UAB
Universitat Autònoma de Barcelona

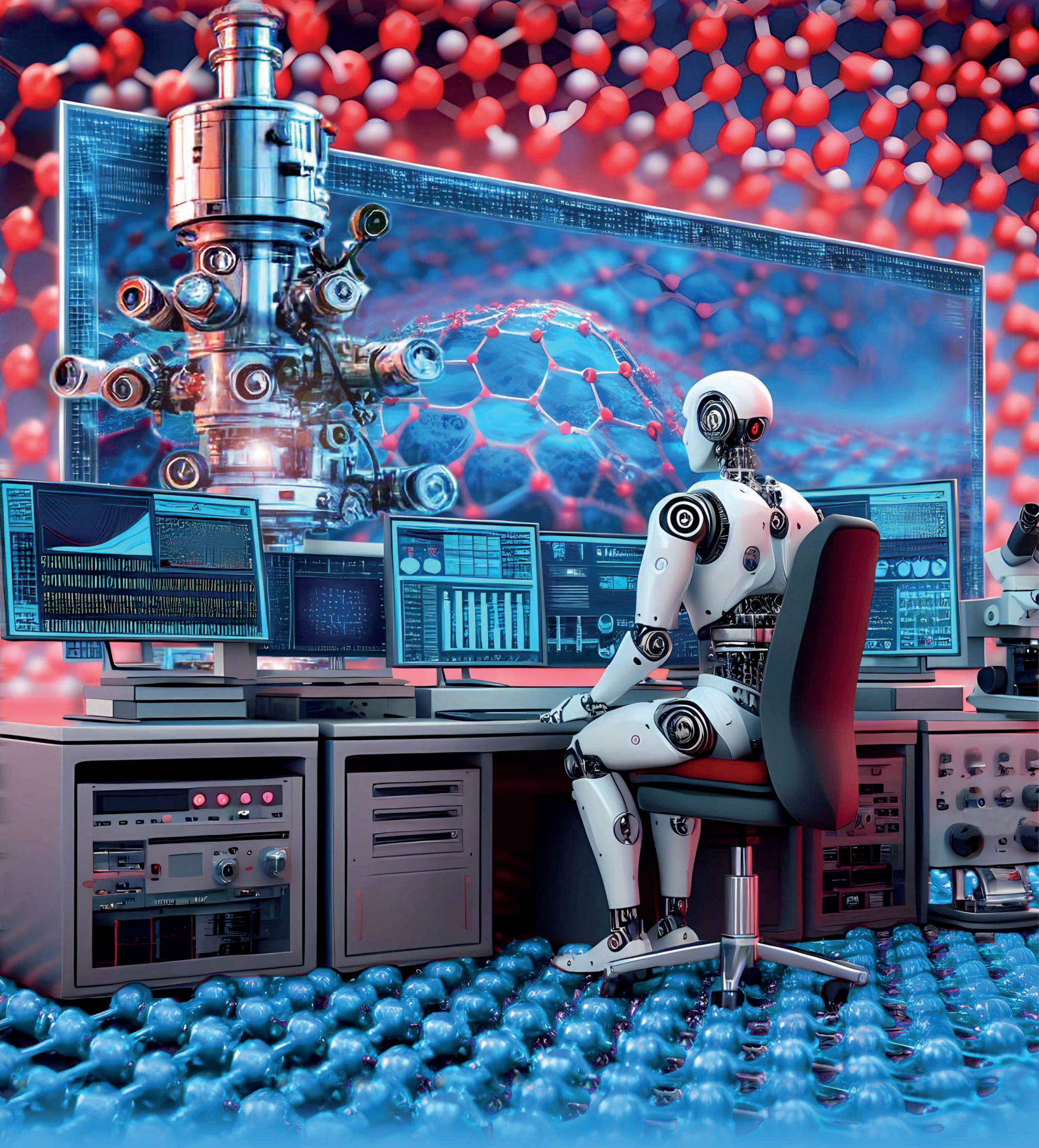
End-to-End Automation in STEM Imaging and Spectroscopy: Revolutionising the Analysis of Quantum Materials and Beyond

Marc Botifoll Moral

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  https://creativecommons.org/licenses/?lang=ca

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  https://creativecommons.org/licenses/?lang=es

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  https://creativecommons.org/licenses/?lang=en



END-TO-END AUTOMATION IN STEM IMAGING AND SPECTROSCOPY: REVOLUTIONISING THE ANALYSIS OF QUANTUM MATERIALS AND BEYOND

MARC BOTIFOLL MORAL

End-to-End Automation in STEM Imaging and Spectroscopy: Revolutionising the Analysis of Quantum Materials and Beyond

by

Marc Botifoll Moral

Supervisor: Prof. Dr. Jordi Arbiol Cobos

Submitted in fulfilment of the requirements for the degree

Philosophiae Doctor in Materials Science

Universitat Autònoma de Barcelona
Catalan Institute of Nanoscience and Nanotechnology
Group of Advanced Electron Nanoscopy

May 2024

The images representing the covers of the thesis dissertation and each chapter were created by prompting tailored commands to a generative artificial intelligence (i.e., Microsoft's Copilot). They aim to highlight the focus of the text they precede while showcasing the astonishing power artificial intelligence has reached, supporting its interest in the present work.

Declaration of Authorship

Jordi Arbiol i Cobos, ICREA Research Professor at Institut Català de Nanociència i Nanotecnologia (ICN2), CSIC and BIST

CERTIFIES

that **Marc Botifoll Moral**, MSc in Multidisciplinary Research in Experimental Sciences, carried out the work entitled *'End-to-End Automation in STEM Imaging and Spectroscopy: Revolutionising the Analysis of Quantum Materials and Beyond'*, under his direction and qualifies for the degree of Doctor in Materials Science. And for that record, sign this certificate.

Prof. Jordi Arbiol i Cobos Marc Botifoll Moral

Bellaterra, Barcelona, Catalonia, Spain, May 2024

Acknowledgements

I would like to start by acknowledging my scientific partners in this wonderful scientific journey at Institut Català de Nanociència i Nanotecnologia (ICN2) which I started back in 2018 with my bachelor's final degree project. I això passa per agrair a l'ICN2 com a institució la seva acollida, des del principi, fent-me sentir com a casa, la meva alma mater en definitiva. En primera instancia, i de forma a vegades inadvertida (aquí radica el seu mèrit), la institució facilitava alleugerir les tasques administratives i burocràtiques per tal de facilitar que em pogues dedicar a la recerca en plenitud. En aquest sentit, agrair al cos administratiu de l'ICN2, en especial a la Inma Caño per la facilitat en gestionar els viatges i el farragós que poden arribar a ser, i l'Andrés Ballesteros per fer el mateix relatiu a totes les gestions vinculades a la beca de doctorat. De la mateixa manera, agrair a l'equip de recursos humans, en primera instancia temporal l'Ainhoa, i continuadament l'Estefania i l'Ana Belén. Aquests són els noms amb els que he tingut un tracte més proper, però sóc plenament conscient que hi ha molts més actors a la rereguarda, els qual també teniu la meva profunda estima i agraïment.

Des del punt de vista científic, l'ICN2 s'ha convertit en la meva alma mater. Més específicament, agraeixo al group of Advanced Electron Nanoscopy (GAeN) per acollir-me i permetre'm l'evolució científica que m'ha transformat a través d'aquests anys. Ha sigut un centre neuràlgic científic immillorable en absolutament tots els sentits. I això passa obligatoriament per agrair-te, Jordi, la teva involucració en el meu desenvolupament. Des de bon principi vas ésser proper, gentil, pacient i sense cap tipus de mala cara, mala paraula o desgana a la hora d'instruir-me i compartir el teu valuós temps i coneixement amb mi. Vaig trovar en tu algú en qui fixar-me tant científicament com humanament i t'has convertit, sens dubte, en el meu mentor i pare científic. Gràcies infinites per tots aquests anys de bona ciència i millors moments, per treballar tant dur per a que nosaltres no ens falti de res. I pels que han de venir! Això ho valoro enormement i així et corresponc amb la meva profunda lleialtat i amistat.

A més, GAeN ha conegut una gran varietat de personalitats, nacionalitats i idiosincràsies, a les quals també em vull referir amb agraïment, ja que ha contributit enormement en la meva transversalitat i tolerancia humana enfront cultures noves que inicialmente xocarien amb la pròpia. Agraeixo la facilitat amb la que tant GAeN com l'ICN2 incentiven viajar, obrir fronteres físiques i mentals, i descubrir nous indrets en els quals establir noves connexions personals i espirituals. Del grup, vull

destacar i agrair-te, Sara, especialmente pels primers mesos que em vas ajudar amb l'adaptació i la instrucción tant de coneixements teòrics com de la operabilitat dels microscopis. Recordo els primers viatges a Saragossa amb molt carinyo i estima, flipant en colors al veure com l'escaneig resultava en unes boletes que resultaven ser àtoms... Aprovecho para agradeceros Rodrigo y Alfonso vuestra ayuda, compromiso y conocimiento en las sesiones en el Titan.

Especial mención para ti Ivan, ya que hemos compartido innumerables momentos de discusión científica pero también encontrábamos fácilmente el espacio de hablar de otras de nuestras pasiones. Pero sobre todo agradecerte por hacer del viaje científico que inicié con el machine learning algo solo y perdido, más ameno y menos solitario, sabiendo que podía contar con una segunda opinión y fuerza motriz para tirar adelante un sueño que ha acabado siendo compartido, y que orgullosamente, parece que absorberá a más mentes pensantes en un futuro cercano.

També a tu Noe, agrair-te la teva ajuda per fer tot més fácil dins de la dinámica administrativa del grup. M'has aportat molta tranquilitat sabent que tot s'està fent correctament en aquest sentit, treient així aquestes cabòries. I no m'oblido d'aquests últims mesos i de l'ajuda i suport tant pràctic com moral en el dipòsit de la tesi.

Agradecimiento especial tambien para los estudiantes que he supervisado tanto en trabajos de máster, en final de grado y ahora en doctorado con Ivan: gracias Carla, Richard, y Chloe. Gracias por poner las cosas tan fáciles y por vuestra luz.

Without forgetting the rest of the members that have been joining the group. Indeed, the friendship of some of them really makes me proud. Xu, Ting, Christian, Chiara, David, Francesco, Helena, Shunrui, Jing, Xinxin, Ying, Zhifu, PengYi, Timm, Jérémy, Ramón, Pol, Meri, Ana, Mar.

Especialment agraït estic també amb la unitat de microscopia electrònica de l'ICN2, els quals sempre han estat, amb un somriure, disponibles per ajudar en qualsevol dubte tant científic com tècnic durant la propia operació dels microscopis. Tanmateix, destaco de la mateixa manera el treball no tan visible de manteniment i cura de l'equipament per trovar-lo en ple funcionament quan el necessitàvem. A més, també hi havia temps per comentar i discutir apassionadament interessos fora de la microscopia i de la ciència. De fet, eren moments molt benvinguts que ajudaven a alliberar la ment en moments d'estrés. Per tot això, moltes gràcies Marcos, Francisco, Belén, Bernat i Kapil.

However, my scientific gratitude is not limited to ICN2, as during these years I have had the opportunity to work with wonderful people and minds.

First of all, I would like to thank you, Quentin, for accepting me in SuperSTEM and within your team. You trained and mentored me with patience and kindness. Your to me unprecedented mastery and understanding of such complex machines

have been truly inspiring and definitely set a reference for what a dreamt human-machine connection can become. Thanks a million for the wonderful discussions and the nice moments with and without the microscopes, as well as your will to share your wisdom in a disinterested way.

Of course, I do not forget about the whole team in SuperSTEM. You helped me to quickly adapt to the place and to fully focus on the science we could share during these three wonderful months. Thank you Demie, Aleksander, Iain, Dorothea, Shihao, Khalil. Also Dani, thanks for the nice moments and for making the SuperSTEM internship nicer. Let us try to park better next time...

Institucionalment, agrair a la Universitat Autònoma de Barcelona, i a l'AGAUR, i en general al sistema públic d'ensenyament i universitats per facilitar la meva instrucció tant mitjançant el programa de doctorat en ciència de materials (així com els coordinadors que hi ha hagut, Eva, Roger, Cristian) com el personal d'administració destinat a que tot rutlli com hauria. A més, gràcies per la beca de doctorat AGAUR de la Generalitat de Catalunya que m'ha permès realitzar els estudis amb moltes més garanties i la possibilitat de gaudir d'activitats extraordinàries en el meu pla formatiu. Gràcies, i seguiu forts i compromesos amb la qualitat de la universitat i ensenyament públic! In that sense, I would also aknowledge Microsoft for its financial support over these years: thanks Emrah, Peter, Philippe.

I would like to extend my thanks to my scientific collaborators during these years. Special emphasis and thanks to Enzo and Vincenzo. You two have been there almost since the very beginning of our dreamt journey of using artificial intelligence in materials science and microscopy. Your contribution has been key since the very beginning, and I would not exaggerate if I said that I would not have been able to finish the thesis in its current form without you. The loop would not be closed and the generalisation capabilities would be much more limited than currently. I know it has been a long journey, and that in some cases the outcomes were not as fast or accurate as desired, but precisely I would like to acknowledge your patience and unconditional help. I believe it has crucially contributed to the current form of the results making the wait worth it. Thanks a million!

The gratitude extends to Prof. Giorgos Katsaros and Prof. Giordano Scappucci, and of course, to their respective teams. It has been an absolute pleasure to work with you and to analyse your devices. Understanding them has been a real challenge, although I firmly believe I have just scratched their surface. Indeed, they have set the basis of a, in my opinion, fascinating journey to come with endless possibilities. Your continuous collaboration and exchange of ideas has been extremely enriching.

Furthermore, special thanks to Prof. Javier García de Abajo and Valerio for their timely support in the development of the theoretical framework for our methodological development. It has been a pleasure to work with you in this, and the theory-experiment correlation is one of the most enriching scientific experiences I

have had so far. Furthermore, I acknowledge Prof. Ernesto Joselevich and Dr. Eitan Oksenberg for their timely contribution as well in the work.

Crucially, a wonderful collaboration with the groups of Prof. Stephan Roche and Prof. Pablo Ordejón raised based on the work presented in the thesis to conduct intriguing and impactful simulations in our systems. I take the chance to acknowledge their timely collaboration, with particular emphasis on Thomas and Catalina. Your contribution has made my thesis a much more complete work. Thanks for pushing your schedule these last months to have these wonderful results. As Humphrey Bogart in Casablanca, I think this is the beginning of a beautiful friendship.

On a smaller scale but not less important, I want to acknowledge the international collaborators with whom I have worked to develop some of the research embedded in the present thesis and in complementary publications as listed at the end of the document. This includes the groups of Prof. Gonzalo Merino, Prof. Anna Fontcuberta i Morral, Prof. Joan Ramón Morante, Prof. Andreu Cabot, Prof. Chen Wang, and their respective teams. Thanks a lot!

Els anys de doctorat no només et revolucionen científicament sinó personal i humanament. Aquesta transformació, d'una crisàlide primerenca a una petita papallona, et prepara per iniciar el desplegament de les ales i estendre-les àmpliament si es desitja suficientment fort. La evolució que he experimentat és enorme, però considero fonamental haver mantingut l'essència dels meus valors. Valors que he anat soldant amb el temps però que he construït amb l'ajut de la meva família i amics. Els meus amics de veritat no disten d'una família també triada. Sempre heu estat allà quan i per si us necessitava, i la vostra lleialtat la valoro com el meu diamant més apreciat. Gràcies pel recolzament, ànims, l'empenta. Gràcies Jose, la teva lleialtat imperiosa sempre m'ha acompanyat. David, tu amistad me ha hecho mejor. Andrés, tu madurez me ha abierto la mente. Adam, tu tenacidad y valentia siempre me han inspirado. Ja sabeu que podeu comptar amb mi pel que sigui i quan sigui.

Aquest treball ha sigut possible gràcies a tots vosaltres, però també en menor mida a molta gent que m'he anat trobant al llarg de la vida i que ha anat influint en els meus pensament i estil de vida. No em vull oblidar dels meus professors, des de ben petit fins al postgrau. No recordo cap professor terriblement terrible, que m'influís negativament o em tallés les ales, cosa que és una sort. Ans al contrari, recordo grans persones influint críticament en ajudar-me a descobrir branques del coneixement apassionants. Així que gràcies a tots per la dedicació, l'esforç i la confiança dipositada en mi.

Ara bé, família, el que sóc jo avui dia és gràcies a vosaltres, pel que us podeu sentir partícips i coautors d'aquest treball. Us estimo a tots infinitament, incondicionalment i em sento extremadament afortunat d'haver tingut la sort de caure en aquesta família. Perquè això si que no es tria, i la meva sort és tan gran com per donar gràcies cada dia. Papa, la teva saviesa, filosofia i profunditat de pensament. Mama, la

teva atenció al detall, la gentilesa, el perdó. Nuria, la teva força, sociabilitat, alegria. Iaia, la teva experiència, saviesa, gentilesa. Quark, la lleialtat incondicional, el compromís i la tranquil·litat. Gràcies per suportar-me tots aquests anys i ajudar-me a créixer, i motivar-me i recolzar-me, esperonar-me, intruir-me, educar-me, filosofar, raonar, cuidar-me, tenir paciència (aquest és el concepte que més s'apropa a infinit de tota la dissertació, comparable a com us estimo) i donar-me les eines per poder fer el meu camí... espero que aquest treball us faci sentir tant orgullosos de mi com jo ho estic de vosaltres. Ainhoa tu també ets de la família, així que gràcies pels bons moments, la paciència, i els moments que sense dubte m'ajudaven a alliberar tensió i facilitar la meva feina. Els bons Mario Karts han ajudat més del que puguis pensar!

Durant aquest període crucial en la meva vida, també he tingut la sort de conèixer a la persona més especial que m'he trobat en aquest camí que és la vida. I en aquest cas, la sort es va convertir en una tria, quan vaig triar ampliar la meva família. Alba, ets el meu amor, t'estimo infinit petardilla. Gràcies per triar-me tu també, per permetre'm poder admirar dia rere dia com d'excepcional ets, per recolzar-me, suportar-me i tenir paciència aquests últims dies de redactat de la tesi que no han sigut fàcils. Però sobretot, gràcies per compartir la teva família i fer-la també meva. Max, t'estimo infinit petardillo. La meva sort de tenir-vos és enorme, i us vull agrair i valorar cada moment amb vosaltres amb el meu amor incondicional.

Moltes gràcies a tots per portar-me a una vida feliç fent el que m'apassiona!

Amb estima i apreci,

Marc

Abstract (English)

The increasing demand for miniaturisation in various fields of materials science, particularly in semiconductor-based devices, requires higher precision and accuracy in characterization. Miniaturisation within the semiconductor platform is exemplified by both classical computing and emerging paradigms like neuromorphic and quantum computing. Quantum computing devices, in particular, require ultrahigh growth quality despite size constraints. However, as devices shrink, correlating materials quality with performance becomes challenging, emphasising the need for detailed nanoscopic characterization. For instance, structural defects can significantly impact functional properties like coherence times in qubits. Therefore, correlating device measurements with structural properties is crucial in the application-materials development cycle. (Scanning) Transmission Electron Microscopy ((S)TEM) serves as a vital tool for linking atomic arrangements with device properties, but challenges persist in enhancing nanodevice characterization.

Among others, we will target some key hurdles accompanying the technique since its inception: (1) lack of statistical significance due to limited areas of analysis, requiring repeated and costly processes, (2) inherent slowness in the characterization loop, from device to analysed results, compromising statistical significance for faster results, (3) still many unmastered knowledge domains, limiting, for instance, retrieving physical insights from spectroscopy, and (4) understanding and gauging the impact of the characterization process on the captured reality. Addressing these challenges is crucial for improving accuracy and efficiency in (S)TEM characterization of nanodevices.

The research in this thesis aims to address these (S)TEM challenges using advanced data analysis and modelling tools, particularly focusing on Machine Learning/Deep Learning (ML/DL) and Artificial Intelligence (AI) algorithms. While traditional computation methods have entailed significant development in (S)TEM, ML/DL and AI offer promising alternatives to automate data analysis and acquisition. These advanced methods can also unveil intricate patterns within data, contributing to unprecedented knowledge discovery, which would be unrealistic to process manually.

Our scientific motivation lies in quantum devices, which as said, would instantly and significantly benefit from successfully addressing the previous challenges. However, the primary focus of the PhD thesis is the use of AI-related methods to tackle these challenges, and while doing so, reaching an impact much beyond the quantum devices that motivate the work. In that sense, the PhD dissertation

is divided into five chapters. The first two chapters settle the motivation and bibliographic foundation (Introduction) and describe the methodological ground which will later enable the findings described throughout (Methodology). However, the gross of the research is presented in the three main results Chapters:

- Chapter 3 focuses on the structural characterisation of quantum devices, particularly for Si/Ge-based devices that can host fault-tolerant and scalable qubits. It aims to structurally evaluate various state-of-the-art Si/Ge devices and correlate them with their performance as operational quantum devices. Additionally, it identifies the characterization requirements that modern transmission electron microscopy must fulfil to provide a comprehensive understanding of these cutting-edge devices. The identified needs will form the foundation for automated workflows that unveil physical knowledge in subsequent chapters.

- Chapter 4 embraces the AI revolution in electron microscopy data analysis in materials science. While current methods are often case-specific and lack generalization, Chapter 4 introduces a groundbreaking end-to-end data analysis workflow to address this limitation. The workflow is presented in two parts: the first automates structural characterization for devices on general material platforms, while the second uses this data to create highly accurate device models that can be the base for unprecedented theoretical simulations relying on experimental knowledge. Throughout the chapter, we extensively prove the validity of the proposed workflow from manifold and diverse device configurations, types and applications, overcoming previous limitations in generalization and applicability.

- Chapter 5 aims to characterize semiconductors' band gap at high spatial resolution, which is a crucial property ruling their functional properties. Despite the capabilities of Electron Energy Loss Spectroscopy, high spatial resolution for band gap mapping remains an open challenge. The chapter proposes a methodology that overcomes previous limitations to understand the spatial distribution of band gaps and its impact on device optoelectronic properties. Besides, a semi-automated solution is introduced to bridge the gap between experimentally measured band gaps and those simulated from models like the ones created in Chapter 4 for devices like the quantum ones in Chapter 3.

Remarkably, the findings from each chapter significantly contribute to the development of the subsequent one, ultimately enhancing electron microscopy in materials science as a whole. This is achieved by strengthening and optimizing the research loop, creating a cumulative impact that reinforces the field's methodologies and understanding.

Resum (Català)

La creixent demanda de miniaturització en diversos camps de la ciència dels materials, especialment en dispositius basats en semiconductors, requereix una major precisió i exactitud en la seva caracterització. Dins la plataforma dels semiconductors, es veu especialment palesa l'alta qualitat requerida tant per computació clàssica com per paradigmes emergents com la computació neuromòrfica i la quàntica. Tanmateix, a mesura que els dispositius es miniaturitzen, correlacionar la qualitat dels materials amb el rendiment es fa cada vegada més desafiant, subratllant la necessitat de nanoscòpia avançada. Per tant, correlacionar les mesures dels dispositius amb les propietats estructurals és crucial en el cicle de recerca enfocat a materials. La microscòpia electrònica de transmissió d'escaneig ((S)TEM) esdevé una eina vital per vincular els arranjaments atòmics amb les propietats dels dispositius, però encara persisteixen reptes pendents de ser resolts.

Entre altres, ens centrarem en els següents reptes clau que acompanyen la tècnica des de la seva creació: (1) la falta de significació estadística a causa d'una àrea d'anàlisi limitada, que requereixen repetició i el seu conseqüent cost, (2) la lentitud inherent del procés global de caracterització, des del dispositiu fins als resultats analitzats, que compromet la significació estadística per a resultats més ràpids, (3) l'existència de molts dominis de coneixement sense dominar, com ara revelar propietats físiques en base a espectroscòpia, i (4) comprendre i mesurar l'impacte del procés de caracterització en com observem la realitat. Abordar aquests reptes és crucial per millorar la precisió i l'eficiència en la caracterització (S)TEM de nanodispositius.

La recerca d'aquesta tesi té com a objectiu abordar aquests reptes de (S)TEM utilitzant eines avançades d'anàlisi de dades i modelització, amb un enfocament particular en l'aprenentatge automàtic/aprenentatge profund (ML/DL) i en algorismes d'intel·ligència artificial (AI). Mentre que els mètodes de càlcul tradicionals han experimentat un desenvolupament significatiu en (S)TEM, ML/DL i AI ofereixen alternatives molt prometedores per automatitzar l'anàlisi i l'adquisició de dades. A més, aquests mètodes avançats també poden revelar patrons ocults en les dades, contribuint a una generació de coneixement sense precedents, inaccessible via intervenció manual.

La nostra motivació científica es basa en els dispositius quàntics, que es beneficiarien immediatament i de manera significativa del fet de superar amb èxit els reptes esmentats. Tanmateix, el focus principal de la tesi doctoral és l'ús de mètodes relacionats amb la IA per abordar aquests reptes, amb la intenció d'assolir un impacte

molt més enllà dels dispositius quàntics que motiven el treball. En aquest sentit, la tesi doctoral es divideix en cinc capítols. Els dos primers capítols estableixen la motivació i els fonaments bibliogràfics (Introducció) i descriuen els pilars metodològics que permetran posteriorment les troballes descrites al llarg de la tesi (Metodologia). Tanmateix, la recerca principal es presenta en els tres capítols de resultats clau:

- El Capítol 3 es centra en la caracterització estructural i correlació funcional de dispositius quàntics, especialment en els dispositius basats en Si/Ge que poden allotjar qubits escalables i resistents a errors. A més, serà clau identificar els requisits de caracterització que la microscòpia electrònica de transmissió moderna ha de complir per proporcionar una comprensió completa d'aquests dispositius punters. Les necessitats identificades en la seva caracterització formaran la base per a fluxos de treball automatitzats per revelar coneixement en els capítols següents.

- El Capítol 4 fa seva la revolució de la intel·ligència artificial en l'anàlisi de dades de microscòpia electrònica en ciència dels materials. Mentre que els mètodes actuals sovint són específics per cada cas i manquen de generalització, el Capítol 4 presenta un procediment innovador d'anàlisi de dades d'extrem a extrem per abordar aquesta limitació. El procediment es presenta en dues parts: la primera automatitza i generalitza la caracterització estructural per a dispositius i tipus de material, mentre que la segona utilitza aquestes dades per crear models precisos dels dispositius, que poden ser la base de simulacions teòriques sense precedents basades en coneixement experimental. Al llarg del capítol, provem extensament la validesa i generalització del flux de treball proposatdes de diverses configurations de dispositius.

- El Capítol 5 té com a objectiu caracteritzar la banda prohibida dels semiconductors amb alta resolució espacial, atès com de crucial que és per les seves propietats funcionals. Malgrat les capacitats de l'Espectroscòpia d'Energia Perduda d'Electrons, la resolució espacial per al mapeig de la banda prohibida segueix sent un repte obert. Per tant, el capítol proposa una metodologia que supera les limitacions anteriors per estudiar la distribució espacial de la banda prohibida i el seu impacte en les propietats optoelectròniques dels dispositius. A més, s'introdueix una solució semiautomatitzada per crear un pont entre les mesures experimentals i les simulades a partir de models com els creats al Capítol 4, per a dispositius com els quàntics del Capítol 3.

Cal destacar que les troballes de cada capítol contribueixen significativament al desenvolupament del següent, millorant, en última instància, la microscòpia electrònica aplicada a la ciència dels materials en el seu conjunt. Això ho aconseguim al reforçar i optimitzar el cicle d'investigació, creant així un impacte acumulatiu que reforça les metodologies i la comprensió del camp.

Resumen (Castellano)

La creciente demanda de miniaturización en diversos campos de la ciencia de materiales, especialmente en dispositivos basados en semiconductores, requiere una mayor precisión y exactitud en su caracterización. Dentro de la plataforma de los semiconductores, es evidente la alta calidad requerida tanto para computación clásica como en paradigmas emergentes como la computación neuromórfica y la computación cuántica. Sin embargo, a medida que los dispositivos se miniaturizan, correlacionar la calidad de los materiales con el rendimiento funcional se vuelve cada vez más desafiante, lo que subraya la necesidad de una nanoscopia avanzada. La microscopía electrónica de transmisión de barrido ((S)TEM) se convierte en una herramienta vital para vincular las distribuciones atómicas con las propiedades de los dispositivos, pero aún persisten desafíos pendientes de resolver.

Entre otros, nos centraremos en los siguientes desafíos clave que acompañan a la técnica desde su creación. (1) La falta de significación estadística debido a un área de análisis limitada, lo que requiere repetición y su consecuente costo, (2) la lentitud inherente del proceso global de caracterización, desde el dispositivo hasta los resultados analizados, lo que compromete la significación y el rendimiento, (3) la existencia de dominios de conocimiento sin dominar, como revelar propiedades físicas en base a espectroscopia, y (4) comprender y medir el impacto del proceso de caracterización en cómo observamos la realidad. Abordar estos desafíos es crucial para mejorar el rendimiento en la caracterización (S)TEM de nanodispositivos.

La investigación de esta tesis tiene como objetivo abordar estos desafíos en (S)TEM utilizando herramientas avanzadas de análisis de datos y modelado, con un enfoque particular en el aprendizaje automático/aprendizaje profundo (ML/DL) y en algoritmos de inteligencia artificial (IA). Mientras que los métodos de cálculo tradicionales han experimentado un desarrollo significativo en (S)TEM, ML/DL y IA ofrecen alternativas muy prometedoras para automatizar el análisis y la adquisición de datos. Además, estos métodos avanzados también pueden revelar patrones ocultos en los datos, contribuyendo a una generación de conocimiento sin precedentes, inaccesible a través de la intervención manual.

Nuestra motivación científica se basa en los dispositivos cuánticos, que se beneficiarían de manera inmediata y significativa de superar con éxito los desafíos mencionados. Sin embargo, el enfoque principal de la tesis doctoral es el uso de métodos relacionados con la IA para abordar estos desafíos, con la intención de lograr un impacto mucho más allá de los dispositivos cuánticos que motivan el trabajo. En este

sentido, la tesis doctoral se divide en cinco capítulos. Los dos primeros capítulos establecen la motivación y los fundamentos bibliográficos (Introducción) y describen los pilares metodológicos que permitirán posteriormente los hallazgos descritos a lo largo de la tesis (Metodología). Sin embargo, la investigación principal se presenta en los tres capítulos de resultados clave:

- El Capítulo 3 se centra en la caracterización estructural y la correlación funcional de dispositivos cuánticos, especialmente en los dispositivos basados en Si/Ge que pueden alojar qubits escalables resistentes a errores. Además, será clave identificar los requisitos de caracterización que la microscopía electrónica de transmisión moderna debe cumplir para proporcionar una comprensión completa de estos dispositivos punteros. Las necesidades identificadas en su caracterización formarán la base para flujos de trabajo automatizados para revelar conocimiento en los capítulos siguientes.
- El Capítulo 4 adopta la revolución de la inteligencia artificial en el análisis de datos de microscopía electrónica en ciencia de materiales. Mientras que los métodos actuales a menudo son específicos y carecen de generalización, el Capítulo 4 presenta un procedimiento innovador de análisis de extremo a extremo para abordar esta limitación. El procedimiento se divide en dos partes: la primera automatiza la caracterización estructural para dispositivos con cualquier tipo de material, mientras que la segunda utiliza estos datos para crear modelos precisos de los dispositivos, que constituyen la base de simulaciones teóricas sin precedentes basadas en conocimiento experimental. A lo largo del capítulo, demostramos extensamente la validez y universalidad del procedimiento propuesto en diversas configuraciones de dispositivos.
- El Capítulo 5 tiene como objetivo caracterizar la banda prohibida de los semiconductores con alta resolución espacial, dado lo crucial que es para sus propiedades funcionales. A pesar de las capacidades de la Espectroscopia de Pérdida de Energía de Electrones, la resolución espacial para el mapeo de la banda prohibida sigue siendo un desafío abierto. Por lo tanto, el capítulo propone una metodología que supera las limitaciones anteriores para estudiar la distribución espacial de la banda prohibida y su impacto en las propiedades optoelectrónicas de los dispositivos. Además, se introduce una solución semiautomatizada para crear un puente entre las mediciones experimentales y las simuladas a partir de modelos como los creados en el Capítulo 4, para dispositivos como los cuánticos del Capítulo 3.

Cabe destacar que los hallazgos de cada capítulo contribuyen significativamente al desarrollo del siguiente, mejorando, en última instancia, la microscopía electrónica aplicada a la ciencia de materiales en su conjunto. Esto lo logramos al reforzar y optimizar el ciclo de investigación, creando así un impacto acumulativo que fortalece las metodologías y la comprensión del campo.

Contents

Declaration of Authorship iv

Acknowledgements vi

Abstract (English) xii

Resum (Català) xiv

Resumen (Castellano) xvii

Contents xviii

List of Figures xxiii

List of Tables xxviii

List of Algorithms xxix

List of Abbreviations xxx

1 Introduction 2

1.1 Motivation 2

1.2 Research objectives 4

1.3 Bibliographic context: Machine learning in transmission electron microscopy for materials science 8

1.3.1 Most important imaging developments: from 2D images to 4D-STEM 8

1.3.1.1 Automation of the electron microscope vs automated data processing 8

1.3.1.2 Unsupervised exploratory and knowledge-revealing routines 12

1.3.1.3 Supervised exploratory and knowledge-revealing routines 16

1.3.1.4 Processing of high dimensionality data by machine learning 19

1.3.1.5 Compressed sensing and machine learning for tomography and in-situ 22

1.3.2 Most important advances in electron spectroscopy 26
1.3.2.1 Unsupervised machine learning 27
1.3.2.2 Supervised machine learning and generation of training sets 33
1.3.2.3 A note on compressed sensing and spectroscopic tomography 38
1.3.3 Machine learning deployment in the electron microscopy community 39
References 43

2 Methodology **59**
2.1 Scope of the chapter 59
2.2 Experimental methods 59
2.2.1 Imaging in the TEM 60
2.2.2 Spectroscopy in the TEM 66
2.2.3 Strain calculations 69
2.3 Computational tools 73
2.3.1 Unsupervised machine learning methods 73
2.3.2 Supervised methods 74
2.3.3 General data analysis for computer vision 76
2.3.4 Programming and software 77
2.4 Experimental equipment and personal expertise 80
References 83

3 Si/Ge-Based Quantum Devices at Atomic Scale: a (Scanning) Transmission Electron Microscopy Approach **90**
3.1 Motivation and context 90
3.2 STEM characterisation of Si/Ge quantum devices 92
3.2.1 Devices towards spin qubits 93
3.2.1.1 Ge quantum wells for spin qubits 94
3.2.1.2 Si quantum wells for spin qubits 108
3.2.2 Devices towards superconducting qubits 113
3.2.2.1 Al as the base of the superconductor 113
3.2.2.2 PtSiGe as the superconductor 117
3.2.3 Additional device morphologies and materials setups 124
3.3 Conclusions 128
3.4 Contributions 129
References 131

4 End-to-end Artificial Intelligence Workflow for the Autonomous Discovery of Physical Knowledge from (S)TEM Data **136**
4.1 Motivation 136
4.2 Workflow description 140

4.2.1 Segmentation . . . 141
4.2.2 Fast Fourier Transform peak finding . . . 141
4.2.3 Phase identification . . . 143
4.2.4 Low-to-high magnification correlation . . . 143
4.2.5 Strain analysis . . . 145
4.2.6 Finite Element Model (FEM) building . . . 146
4.2.7 3D Atomic Model (3DAM) building . . . 146
4.2.8 Strain transfer to atomic model . . . 146
4.2.9 Simulation of functional properties . . . 147

I Comprehensive and General Automation of the Structural Analysis of (Scanning) Transmission Electron Microscopy Data 149
4.I.1 Segmentation . . . 150
4.I.1.1 Methods . . . 150
4.I.1.1.1 Data treatment . . . 150
4.I.1.1.2 Tested algorithms . . . 151
4.I.1.1.3 Metrics . . . 153
4.I.1.2 Results . . . 154
4.I.1.3 Final segmentation model proposed . . . 159
4.I.2 Peak finding in the Fourier spectra . . . 165
4.I.2.1 Machine Learning-based 2D Gaussian fitting . . . 165
4.I.2.2 Encoder-decoder convolutional neural network . . . 167
4.I.2.3 1D profile scanning . . . 172
4.I.2.4 Global combined peak finding model . . . 173
4.I.2.5 Peak finding performance metrics . . . 179
4.I.3 Segmentation - Peak finding interconnection . . . 184
4.I.4 Low-to-high magnification correlation . . . 188
4.I.5 Phase identification . . . 195
4.I.6 Strain analysis: Geometrical Phase Analysis automation . . . 203
4.I.6.1 Definition of the reference region . . . 203
4.I.6.2 Selection of the optimal g vectors pair and mask resolution . . . 204
4.I.6.3 Computation of the rotation angle of the Cartesian axes . . . 206
4.I.6.4 GPA base code . . . 207
4.I.6.4.1 Proofs of the automated GPA . . . 207
4.I.7 Bragg filtering automation . . . 212
4.I.8 Automated phase identification proofs . . . 215

II Automated Creation of Experimentally Realistic 3D Models from the Comprehensive Structural Characterisation of Devices 237
4.II.1 Starting point: Knowledge from Part I . . . 238

4.II.2 Finite Element Modelling automation . . . 240
4.II.2.1 In-(image)-plane alignment of the experimental-model rotation 241
4.II.3 Atomic model building . . . 248
4.II.3.1 ASE adaptation: general crystal surface building . . . 249
4.II.3.2 Virtual unit cell calculation . . . 252
4.II.3.3 Strain transfer from GPA to atomic model . . . 256
4.II.3.4 Compositional information: Spectroscopic mapping and quantification . . . 262
4.II.3.5 Compositional information: Segmentation and symmetry equivalences . . . 266
4.II.4 Proofs of the full workflow . . . 268
4.II.5 Experimental simulation of key functional properties . . . 277
4.II.5.1 STEM simulations . . . 278
4.II.5.2 FEM simulations . . . 279
4.II.5.3 Simulations of atomic forces and relaxations . . . 287
4.II.5.4 Tight-binding simulations . . . 293
4.II.5.4.1 Remarks and considerations of the simulations . . . 293
4.II.5.4.2 Results of the experimental/tight-binding/Keating correlation . . . 298

III Conclusions 305
4.III.1 Conclusions . . . 306
4.III.2 Contributions . . . 308
References . . . 310

5 Sub-Nanometer Mapping of Strain-Induced Band Structure Variations in Semiconductor Heterostructures 322
5.1 Motivation and context . . . 322
5.1.1 Guided-grown ZnSe@ZnTe core-shell nanowires on α-Al_2O_3 . . . 326
5.2 Band gap measurements through Valence EELS . . . 330
5.2.1 Low-loss EELS: theoretical background . . . 330
5.2.2 VEELS band gap mapping methodology . . . 334
5.2.2.1 VEELS simulations . . . 337
5.2.2.2 Theory-based experimental spectrum image cleaning 339
5.2.2.3 Adaptive data fitting and resulting band gap mapping 341
5.2.2.4 Strain considerations . . . 348
5.3 Automation of the methodology . . . 351
5.4 Conclusions . . . 355
5.5 Contributions . . . 359
References . . . 361

6 General conclusions 366

6.1 General conclusions 366
6.2 Future work and outlook 367
References 375

A Supplementary data **376**
A.1 Chapter 4 377
A.1.1 Complementary performance metrics of the peak finding models 377
A.2 Chapter 5 380
A.2.1 Significance tests 380

B List of publications **383**

List of Figures

1.1 ML for data acquisition of STEM 9
1.2 ML for data analysis of HRTEM 12
1.3 ML/DL/AI paradigms for electron microscopy in materials science . 15
1.4 ML for ptychography 20
1.5 ML for tomography 23
1.6 PCA applied to electron spectroscopy 28
1.7 ICA and VCA applied to electron spectroscopy 31
1.8 Unsupervised vs supervised approach in EELS 36
1.9 Physics-aware supervised model on EELS 38
1.10 Hypothetical ML model based on open code and data 41

2.1 Parallel illumination TEM schematics 61
2.2 STEM ray diagram and detection modes 63
2.3 STEM-EELS and monochromator schematics 67
2.4 GPA vs PPA 72
2.5 Supervised ML/DL: U-Net 75
2.6 Electron microscopes used in the thesis 81

3.1 Ge quantum information route 91
3.2 GSQ_1 STEM characterisation 94
3.3 GSQ_2 STEM characterisation 95
3.4 GSQ_3 STEM characterisation 96
3.5 GSQ_1 vs GSQ_2 vs GSQ_3 STEM characterisation 98
3.6 GSQ_4 STEM characterisation 101
3.7 FEM simulations of GSQ_4 for the evaluation of the lamella thinning effect 105
3.8 GSQ_6 STEM characterisation 107
3.9 SSQ_1 vs SSQ_2 vs SSQ_3 STEM characterisation 109
3.10 SSQ_1 vs SSQ_2 vs SSQ_3 thickness comparison (plane counting) . . 111
3.11 GSu_01 vs GSu_02 STEM characterisation 115
3.12 GSu_02 STEM characterisation 116
3.13 GSu_03 vs GSu_06 STEM characterisation 118
3.14 GSu_04 vs GSu_07 STEM characterisation 119
3.15 GSu_05 vs GSu_08 STEM characterisation 121
3.16 GSu_09 STEM characterisation 122

3.17 GSu_10 STEM characterisation . . . 124
3.18 InAs NW @ Al on Si and Ge (@Si) on Si STEM characterisation . . . 126

4.1 Full workflow scheme . . . 139
4.2 Workflow's automated data analysis output . . . 142
4.3 Workflow's automated generation of atomic models . . . 144
4.4 FEM and atomic simulations out of the workflow's outputs . . . 147
4.I.1 HAADF references for segmentation model development and testing 153
4.I.2 Segmentation results for the three stages towards the segmentation model . . . 158
4.I.3 Comparative examples of the segmentation performed with Canny Edge vs SAM (Part 1). . . . 163
4.I.4 Comparative examples of the segmentation performed with Canny Edge vs SAM (Part 2). . . . 164
4.I.5 2D Gaussian fitting machine learning model . . . 166
4.I.6 Training an FFT peak finding model based on U-Net. . . . 168
4.I.7 Generation of training data for CNN-based peak finding model and U-Net training . . . 171
4.I.8 CNN peak finding model performance in FFT spectra . . . 172
4.I.9 Flowchart of the global peak finding model. . . . 174
4.I.10 Performance of the global peak identification model (Part 1). . . . 176
4.I.11 Performance of the global peak identification model (Part 2). . . . 177
4.I.12 Performance of the global peak identification model (Part 3). . . . 178
4.I.13 Histograms of peak finding performance metrics. . . . 180
4.I.14 Performance and metrics of the global peak finding in labelled FFTs (Part 1) . . . 182
4.I.15 Performance and metrics of the global peak finding in labelled FFTs (Part 2) . . . 183
4.I.16 Masking and cropping comparison in the segmentation - peak finding interconnection (Part 1) . . . 185
4.I.17 Masking and cropping comparison in the segmentation - peak finding interconnection (Part 2) . . . 187
4.I.18 Multi-scale template matching chain of GSu_02 . . . 189
4.I.19 Multi-scale template matching chain of GSu_10 . . . 190
4.I.20 Template matching collage of core-shell InAs-Al NW . . . 192
4.I.21 Template matching collage of Ge quantum well with PtSiGe metallic contacts . . . 193
4.I.22 Template matching collage of Ge quantum well with Al-Nb metallic contacts . . . 194
4.I.23 Tolerance in reciprocal space of the spot pairing in the phase identification algorithm . . . 197
4.I.24 Automated GPA g vector and mask selection . . . 205

4.I.25 GPA automation proof: transversal InSb nanowire on InP 208
4.I.26 GPA automation proof: Ge quantum well ($Ge_{0.7}Si_{0.3}$/Ge/$Ge_{0.7}Si_{0.3}$) . 208
4.I.27 GPA automation proof: core-shell InAs - polycrystalline Al nanowire 209
4.I.28 GPA automation proof: longitudinal InSb nanowire on InP 209
4.I.29 GPA automation proof: InAs – $Ga_xIn_{1-x}As$ (compositional gradient) on GaAs 210
4.I.30 GPA automation proof: InP nanowire on InAs 210
4.I.31 GPA automation proof: $CoFe_2O_4$ nanoparticles dispersion 211
4.I.32 GPA automation proof: CuTe nanoparticle dispersion 211
4.I.33 Automated Bragg filtering: transversal cut of an InSb nanowire on InP 212
4.I.34 Automated Bragg filtering: longitudinal cut of a multi-twinned InSb nanowire on InP 213
4.I.35 Automated Bragg filtering: CuTe nanoparticles dispersion 214
4.I.36 Phase identification proof: STEM-HAADF of a longitudinal twinned InSb nanowire on InP 217
4.I.37 Phase identification proof: HRTEM of CuTe nanoparticles (mask intercommunication) 218
4.I.38 Phase identification proof: HRTEM of CuTe nanoparticles (crop intercommunication) 219
4.I.39 Phase identification proof: STEM-HAADF of a transversal cut of an InSb nanowire on InP 221
4.I.40 Phase identification proof: STEM-HAADF of a $Si_{0.3}Ge_{0.7}$/Ge/$Si_{0.3}Ge_{0.7}$ QW 222
4.I.41 Phase identification proof: STEM-HAADF of a $Si_{0.7}Ge_{0.3}$/Si/$Si_{0.7}Ge_{0.2}$ QW 223
4.I.42 Phase identification proof: two STEM-HAADF micrographs of a dispersion of $CoFe_2O_4$ nanoparticles 225
4.I.43 Phase identification proof: HRTEM of CuTe nanoparticles, with key Bragg filtering 226
4.I.44 Phase identification proof: Dataset phase identification of InSb NW on InP 227
4.I.45 Phase identification proof: STEM dataset phase identification of a Ge QW with PtSiGe contacts (Part 1) 229
4.I.46 Phase identification proof: STEM dataset phase identification of a Ge QW with PtSiGe contacts (Part 2) 230
4.I.47 Phase identification proof: STEM dataset phase identification of a core-shell Ge-Si nanowire on Si substrate (Part 1) 232
4.I.48 STEM dataset phase identification of a core-shell Ge-Si nanowire on Si substrate (Part 2) 233
4.I.49 Phase identification proof: STEM dataset phase identification of a core-shell InAs-Al nanowire on Si substrate (Part 1) 235

4.I.50 Phase identification proof: STEM dataset phase identification of a core-shell InAs-Al nanowire on Si substrate (Part 2) 236
4.II.1 Segmentation contour smoothing for automated FEM generation . . 242
4.II.2 Computation of the experimental Miller indices pointing to the cartesian axes for FEM refinement 244
4.II.3 GPA strain transfer into 3D atomic models: effect of the image quality 258
4.II.4 Strain field transfer implications into 3DAM, and example of an InSb NW on InP 261
4.II.5 Spectroscopy-based compositional refinement of 3D atomic models . 264
4.II.6 Compositional refinement of 3D atomic models based on the symmetry of the identified crystal phases 267
4.II.7 Full workflow proof: Ge QW, large FOV 269
4.II.8 Full workflow proof: Ge QW, reduced FOV 271
4.II.9 Full workflow proof: Si QW, reduced FOV 272
4.II.10 Full workflow proof: CuTe nanoparticles 274
4.II.11 Full workflow proof: InSb NW on InP substrate 276
4.II.12 STEM experimental simulations of a 3DAM of a Ge QW 278
4.II.13 Generation of the 3D FEMs from the automated data analysis of the workflow 280
4.II.14 Boundary conditions comparing the thin lamella with the infinitely extended "lamella" in FEM simulations 283
4.II.15 FEM strain relaxation results in a thinned lamella vs unthinned "lamella" vs experimental results 286
4.II.16 Scheme of the bending and stretching in the Keating model 289
4.II.17 Simulation of the forces distribution (Keating model) of a 3DAM of the GSQ_4 device 291
4.II.18 Tight-binding simulations of LDOS through GSQ_4 at full spectral range and band gap zoom-in 299
4.II.19 Tight-binding simulations of GSQ_4's LDOS around the band gap for structural artefact detection 301
4.II.20 Tight-binding simulations of GSQ_4's LDOS around the band gap in a high-quality micrograph 303

5.1 Structure and crystallographic orientations of the sapphire-supported ZnSe@ZnTe core-shell nanowire 325
5.2 Comparison between ZnSe@ZnTe NWs on C and A-plane sapphire . 327
5.3 Atomic modelling and simulated GPA of a cylindrical ZnSe@ZnTe NW 329
5.4 ZnSe and ZnTe complex dielectric functions 331
5.5 Relative thickness map of the ZnSe@ZnTe NW 336
5.6 Workflow of the simulation-based correction of retardation effects for band gap mapping 338

5.7 Correction and relative position finding algorithm 341
5.8 Band gap maps in the ZnSe@ZnTe NW 344
5.9 Pixel-wise analysis of the correction method through a single-pixel interfacial profile 347
5.10 Crossed dielectric effect of ZnSe/ZnTe region-wise hypothetical strain modulations 350
5.11 Semi-automated workflow for band gap mapping through VEELS . . 353
5.12 Automated band gap maps on Ge@Si NW on Si 358

6.1 Cross-fertilization perspectives of data analysis and automation between electron microscopy and other fields 371

List of Tables

2.1 Main Python packages used to develop the workflow in chapter 4, and the methods in chapter 5. 79

3.1 GSQ_1, GSQ_2, and GSQ_3 structural comparison 97
3.2 SSQ_1, SSQ_2 and SSQ_3 strain comparison 110

4.I.1 Performance of Clustering Algorithms in InP-InSb Nanowire 155
4.I.2 Accuracy of clustering algorithms in the nanowire devices. 156
4.I.3 Accuracy of Canny edge detection algorithm in Ge QW devices. . . . 157
4.I.4 Accuracy of Canny edge detection algorithm vs SAM in different nanowires 160
4.I.5 Accuracy of Canny edge detection algorithm vs SAM in different Ge QW devices 161
4.I.6 Noise layers in synthetic data preparation for peak detection CNN . 169
4.I.7 Precision comparison over all the peak finding methods. 181
4.I.8 Recall comparison over all the peak finding methods. 181
4.I.9 F1 score comparison over all the peak finding methods. 181
4.II.1 Keating model constants for stretching in SiGe alloys 290
4.II.2 Keating model constants for bending in SiGe alloys 290

5.1 Band gap average values and type found for the ZnSe@ZnTe NW . . 345

A.1 Performance metrics of 2DGML (full labelled dataset) 377
A.2 Performance metrics of 2DGML (precision > 10%) 377
A.3 Performance metrics of CNN (full labelled dataset) 377
A.4 Performance metrics of CNN (precision > 10%) 378
A.5 Performance metrics of 1DScan (full labelled dataset) 378
A.6 Performance metrics of 1DScan (precision > 10%) 378
A.7 Performance metrics of global peak finding model (full labelled dataset) 379
A.8 Performance metrics of global peak finding model (precision > 10%) 379

List of Algorithms

4.I.1 Crystal phase, zone axis and indexation finding for every spot pair . 196
4.I.2 Recursive global crystal phases and zone axes finding within a single Fast Fourier Transform 198
4.II.1 Programmatic exceptions to compute Miller indices 247

List of Abbreviations

2DHG	**2D H**ole **G**as
3DAM	**3D A**tomic **M**odel
ABF	**A**nnular **B**right-**F**ield
AC	**A**berration **C**orrection
ADF	**A**nnular **D**ark-**F**ield
AFM	**A**tomic **F**orce **M**icroscopy
AI	**A**rtificial **I**ntelligence
API	**A**pplication **P**rogramming **I**nterface
ASE	**A**tomic **S**imulation **E**nvironment
BEM	**B**oundary **E**lement **M**ethod
CBED	**C**onvergent-**B**eam **E**lectron **D**iffraction
CL	**C**athodo**L**uminescence
CMOS	**C**omplementary **M**etal **O**xide **S**emiconductor
CNN	**C**onvolutional **N**eural **N**etwork
CS	**C**ompressive **S**ensing
CTF	**C**ontrast **T**ransfer **F**unction
CV	**C**omputer **V**ision
DBSCAN	**D**ensity-**B**ased **S**patial **C**lustering of **A**pplications with **N**oise
DFT	**D**ensity **F**unctional **T**heory
DF	**D**ark **F**ield
DL	**D**eep **L**earning
DOS	**D**ensity **O**f **S**tates
DPC	**D**ifferential **P**hase **C**ontrast
EDX	**E**nergy **D**ispersive **X**-Ray spectroscopy
EELS	**E**lectron **E**nergy **L**oss **S**pectroscopy
EFTEM	**E**nergy-**F**iltered **T**ransmission **E**lectron **M**icroscopy
ELNES	**E**nergy **L**oss **N**ear **E**dge **S**tructure
FCNN	**F**ully **C**onvolutional **N**eural **N**etwork
FEM	**F**inite **E**lement **M**ethod
FFT	**F**ast **F**ourier **T**ransform
FIB	**F**ocused **I**on **B**eam
FOV	**F**ield **O**f **V**iew
FWHM	**F**ull **W**idth at **H**alf **M**aximum
GAN	**G**enerative **A**dversarial **N**etwork
GDS	**G**raphic **D**ata **S**ystem
GMM	**G**aussian **M**ixture **M**odeling
GPA	**G**eometrical **P**hase **A**nalysis
GP	**G**aussian **P**rocessing
GSQ	**G**ermanium **S**pin **Q**ubit
GSu	**G**ermanium **Su**perconductor **Q**ubit
HAADF	**H**igh-**A**ngle **A**nnular **D**ark-**F**ield
HRTEM	**H**igh-**R**esolution **T**ransmission **E**lectron **M**icroscopy

ICA	**I**ndependent **C**omponent **A**nalysis
iDPC	**i**ntegrated **D**ifferential **P**hase **C**ontrast
iFFT	**i**nverse **F**ast **F**ourier **T**ransform
JDOS	**J**oint **D**ensity **O**f **S**tates
LDOS	**L**ocal **D**ensity **O**f **S**tates
MAADF	**M**iddle-**A**ngle **A**nnular **D**ark-**F**ield
MLP	**M**ulti**L**ayer **P**erceptron
ML	**M**achine **L**earning
MZM	**M**ajorana **Z**ero **M**odes
NMF	**N**on-negative **M**atrix **F**actorization
NN	**N**eural **N**etwork
NW	**N**ano**W**ire
PCA	**P**rincipal **C**omponent **A**nalysis
PIC	**P**unt d'**I**nformació **C**ientífica
PPA	**P**eak **P**airs **A**nalysis
QW	**Q**uantum **W**ell
RL	**R**einforcement **L**earning
RNN	**R**ecurrent **N**eural **N**etwork
rVAE	**r**otationally invariant **V**ariational **A**uto**E**ncoder
SAM	**S**egment **A**nything **M**odel
SIFT	**S**cale-**I**nvariant **F**eature **T**ransform
SNR	**S**ignal-to-**N**oise **R**atio
SPM	**S**canning **P**robe **M**icroscopy
SQUID	**S**uperconducting **QU**antum **I**nterference **D**evice
SSQ	**S**ilicon **S**pin **Q**ubit
STEM	**S**canning **T**ransmission **E**lectron **M**icroscopy
STM	**S**canning **T**unneling **M**icroscopy
SURF	**S**peeded-**U**p **R**obust **F**eature
SVD	**S**ingular **V**alue **D**ecomposition
TEM	**T**ransmission **E**lectron **M**icroscopy
VBM	**V**alence **B**and **M**aximum
VAE	**V**ariational **A**uto**E**ncoders
VCA	**V**ertex **C**omponent **A**nalysis
VEELS	**V**alence **E**lectron **E**nergy **L**oss **S**pectroscopy
VLS	**V**apour-**L**iquid-**S**olid
XRD	**X**-**R**ay **D**iffraction
ZLP	**Z**ero-**L**oss **P**eak

Dedicated to the family I was
given and to the family I have
chosen



CHAPTER 1: INTRODUCTION

Chapter 1

Introduction

In the present Chapter, I introduce the motivation, research objectives, bibliographic context and essentially why the proposed research can be considered as a breakthrough and meaningful at the date of the document release. In summary, the path towards the miniaturisation of devices advances at a vertiginous pace, requiring increasing levels of precision and accuracy in their characterisation. Especially in fields like quantum computing, achieving success in critical applications hinges on maintaining high-quality standards in its materials science, often facilitated by techniques such as transmission electron microscopy. Consequently, there is a pressing demand for high-throughput, atomically precise data analysis paradigms to enhance the process of materials and device discovery and engineering. To address this, we initially identify the characterization requirements in quantum computing devices. Subsequently, we utilize this information to establish a workflow that streamlines data analysis, generates representative 3D models, and pushes the boundaries of measurements achievable within the framework of electron microscopy.

1.1 Motivation

The path towards the miniaturisation of devices advances at a vertiginous pace, requiring increasing levels of precision and accuracy in their characterisation. Remarkably, this is ubiquitous in materials science and applies to every distinct field within it. For instance, the increasing demand to create efficient batteries in devices of continuously decreasing sizes. [1] Even for the devices fed by these batteries, which are mostly based on semiconductors that keep challenging Moore's law. [2, 3] Even beyond this traditional computing wall, on devices that go beyond Moore's law by introducing revolutionary computing paradigms like neuromorphic computing, or quantum computing. [4, 5, 6]

Indeed, quantum computing devices stand as clear examples of this miniaturisation trend and its will towards achieving ultrahigh growth quality despite the size or dimensionality. Not surprisingly, when assessing materials quality and their correlation with device performance, issues start appearing with miniaturisation. As a result, the in-detail advanced nanoscopic characterisation of its constituent materials from the structural point of view turns out to be mandatory and differentiating. For example, if we wish to engineer a qubit with high coherence times, a simple structural defect in a key region of the device could dramatically affect this functional property. [7] Consequently, correlating measurements of the device in its intended operation conditions, with its structural properties is a fundamental research question in any application-materials development loop.

In the realm of experimentation, (Scanning) Transmission Electron Microscopy ((S)TEM) stands as the gold standard for linking the arrangement of the atoms in the devices with the properties they exhibit. [8, 9] Interestingly, the atom-by-atom knowledge revealed through STEM imaging can be complemented with spectroscopy analysis. Thus, spectroscopy in the electron microscope enables accessing supplementary chemical and physical information, which merged with the imaging rewards us with complete information about our system. [10, 11] However, the (S)TEM-based characterisation of nanodevices still implicitly entails important challenges in which it is worth putting effort into improving:

- First and foremost, its tricky path to statistical significance. [12, 13] Thus, (S)TEM can be extremely accurate and representative of the checked system. However, the area in which this accuracy can be claimed is reduced to a maximum of a few hundred nanometers. Therefore, a few micrometre-long devices would inevitably require repeated and spatially periodical analyses throughout their full extension. [14, 15, 16] Also inevitably, it becomes tedious, time-consuming and prohibitive to constantly reproduce this hypothetical thorough analysis on every device. As a result, the most common scenario is jeopardising statistical significance in favour of fast access to results which we generalise with assumptions that are not always applicable.

- Secondly, and closely related to the first point, its reduced throughput. [17, 18] Electron microscopy characterisation is globally a slow process which starts from the sample preparation and thinning, continues with the acquisition of data, and reaches its zenith with the data analysis. Therefore, it is convenient to strategically plan the experiments and select the proper sample preparation on the devices to analyse to avoid bottlenecking the materials research loop. However, this balance is often detrimental to the statistical significance of measurements.

- Particular knowledge domains remain substantial challenges. For instance, the meV low-loss regime of Electron Energy Loss Spectroscopy (EELS) is still

unexplored in most of the systems. Thus, mapping phonons and density of states is close to the current resolution limits of (S)TEM for most study cases. [19, 20, 21] Therefore, there is room to push both the machines and the data analysis methodologies to access this quite unexplored knowledge realm.

- The relation between the measurements and how representative they are of reality. *In-situ* and *operando* microscopy are actively addressing the issue of how the characterisation process affects and alters the reality we capture. [22, 23, 24] For instance, to measure batteries in their charging-discharging cycles, or quantum devices in their working temperature conditions of a few mK. Even more generally, to decipher how the sample thinning impacts measured features like strain and deformations, or how the electron beam damages and changes the sample, especially for beam-sensitive materials. [25, 26] In any case, it is an incipient field which still has more doubts than answers.

Remarkably, when challenges are faced, opportunities arise. That is why, this thesis dissertation aims to provide a complete or partial answer to the listed challenges which hamper a more efficient materials research loop. Our central focus is materials science and electron microscopy. Nonetheless, this focus is tackled through a toolset which at the date of the thesis is incipient and of exponentially increasing importance not only in science but in absolutely every aspect of humankind's daily life.

1.2 Research objectives

As explained, the challenges exposed in the previous section are of huge interest to the materials science community. Therefore, the present thesis dissertation is fully devoted to logically tackling them, and providing everlasting solutions always accessible to its community. The following list orders the main goals of the thesis according to their logical consecution, and the chapters where they are investigated:

1. **To conduct the advanced nanoscopic (S)TEM characterisation of Si/Ge-based devices for quantum computing, to correlate their structural and functional properties.**

   Quantum computing has actively been at the crest of the wave since Richard Feynman thought about the crazy possibility of simulating atoms with themselves. [6] Since then, it has become promising enough to stem multiple practical approaches towards quantum processors (qubits). [27, 28] Remarkably, breakthroughs must still occur to robustly claim quantum supremacy in certain tasks. [29] Interestingly, one of the most promising approaches towards scalable, fault-tolerant qubits is through the semiconductor platform established by Si and Ge heterostructures. [30] Importantly, it has already widely

succeeded in engineering state-of-the-art qubits. In fact, the structural characterisation of some of these successful qubits is presented in the dissertation. [31, 32, 33]

Hence, our objective is to understand, from a structural perspective, why these devices demonstrate exceptional qualities in holding measurable and adjustable quantum states. Given that achieving these attributes necessitates high-quality crystalline structures, our focus lies in providing comprehensive structural characterizations. Moreover, establishing a direct correlation between these functional properties and structural aspects isn't always straightforward, particularly in quantum devices with a tight operating hotspot. Therefore, we propose a workflow aimed at enhancing communication among various data sources and facilitating efficient comparisons between experimental findings from electron microscopy, device measurements during operation, and theoretical simulations of the systems. Consequently, these Si/Ge systems serve as excellent benchmarks for contextualizing and evaluating the requirements that such a workflow must fulfil.

This research is detailed in chapter 3.

2. **To develop a workflow for the automated high-throughput data analysis of STEM micrographs, capable of creating realistic 3D models which are representative of the devices.**

   The second main goal, which constitutes the central objective of the thesis dissertation, is accomplishing the ambitious workflow conceived in objective one in an automated way. First, we identify the characterisation requirements of the quantum devices studied in chapter 3 and the intricacies of correlating the knowledge sources. Then, we transfer this know-how into a workflow that outputs the properties that we want to retrieve, from the experimental (S)TEM-related input data that make sense to provide.

   Ideally, we seek a rapid, automated, and accurate workflow capable of providing comprehensive structural characterization for any device solely from a single (S)TEM image. This means encompassing devices composed of any material or crystal phase, exhibiting any morphology and relative arrangement among their components, and designed for any intended purpose or application. Importantly, a complete structural characterisation ideally involves a total understanding of the arrangement of every atom in 3D. Consequently, the desired workflow must create representative 3D models out of the same input from which the data analysis was carried out. Remarkably, these 3D models must have unprecedented accuracy to represent the experimental device. Thus, they are built out of the experimental information retrieved by the workflow itself. As a result, these 3D models, which must be compatible

inputs for third-party simulation solutions, would represent the exact experimental system in a simulation environment.

The innovative global workflow must act as a link between the device performance, the structural characterisation and the measured physical properties. Therefore, establishing this triple link would speed the research loop up. First of all, the data analysis would no longer bottleneck the process. Related to this, no strategic thinking to adjust to budgets or time frames would be needed anymore. Thus, its high throughput originating from automation would allow the simultaneous analysis of multiple devices and the consequent statistical significance gain. In addition, any research question involving model validation would also be significantly boosted due to the initial obtaining of, at least, base models to work with. The hypothesis testing would be incentivised to efficiently optimise features like the composition and alloying, the morphology, or even to check the effects of sample preparation and how the thinning affects measurements like strain.

In summary, the achievement of this ambitious goal would boost the research loop in any materials science study involving electron microscopy. Particularly, we centre the development on quantum devices as paradigmatic challenging devices. Crucially, the development must parallelly ensure its reproducibility and general character as a priority. Importantly, then, validating the workflow as a general-purpose process is part of the goal, too.

This research is detailed in chapter 4.

3. **To map the band gap with unprecedented spatial and energy resolutions as a key physical property to correlate with the structural properties of semiconducting devices.**

   Achieving the automated workflow to calculate 3D models that match the experimental characterisation would enable the simulation of physical properties in what can be called an "experimental simulation". However, we still miss the direct experimental measurement of the same physical properties computed from these experimental simulations to establish a direct correlation.

   Among all the physical properties which can be simulated, we are particularly interested in the band gap. It is the most representative feature of semiconducting materials. Importantly, the global semiconductor industry sales surpassed $ 525 billion in 2023. [34] These astronomic numbers showcase the huge impact semiconductors have nowadays, and indicate how key it would be to reproducibly and quickly gauge accurate band gap information. In fact, there are already solutions to experimentally measure the band gap. [35] However, we wish to exploit the models generated by the desired workflow of the central goal, and exploit the lateral resolution that the electron microscope can offer. Remarkably, there is almost no available literature achieving band gap

maps with the spatial resolution that (S)TEM can offer. [36] The main reason is the presence of optical excitations hampering the physically meaningful extraction of this property. [36, 37] Thus, the goal is to overcome this limitation and retrieve this knowledge with statistical significance. Although we first aim to break this knowledge barrier, this achievement must be paired with methodological developments. Ideally, then, we pursue a fast, robust, general, significant and if possible, automated methodology to map the band gap at unprecedented spatial and energy resolutions.

This research is detailed in chapter 5.

Bringing all aspects together, the loop ideally closes by integrating the outcomes of each objective. Starting with quantum devices, we address the initial goal by generating 3D models through automated data analysis, as per the second goal. These models allow us to simulate physical properties such as the band gap. Consequently, we aim to experimentally measure the same band gap, aligning with the methodology targeted in the third goal. Ultimately, this comprehensive approach completes the characterization of selected devices outlined in the first goal. It achieves this by correlating their structural characterization with experimentally measured and simulated physical properties, as well as their performance under the intended operational conditions.

The consecution of these ambitious objectives fundamentally relies upon transmission electron microscopy as the central characterisation tool. This statement is supported, detailed and justified in the methodology chapter of this dissertation (chapter 2). The reader will find there all the experimental and computational methods and techniques adopted to pursue these goals. In fact, we seek adequate computational tools to set the basis of the data analysis and modelling required for the second goal.

Thus, there is plenty of literature about methodology development for electron microscopy data. Not surprisingly, most of this bibliography is served by traditional computation and its way of programming. However, new efficient and flexible data analysis paradigms have arisen. Then, to the best of our knowledge, the best candidates to serve as our computational basis are Machine and Deep Learning (ML/DL), and generally Artificial Intelligence (AI) algorithms. They must support the development intended for the central research goals two and three, for which comprehensively reviewing their state of the art is crucial. As a consequence, the next section wraps the literature up and shines light on the most remarkable and recent advances in the field of applying AI-related methods to electron microscopy for materials science. The key lies in transferring this progress to a solution that covers the second and third goals and unlocks the required characterisation of the devices from goal number one. Note that the scientific motivation behind our targeted quantum devices is thoroughly centralised in its dedicated chapter (i.e., Chapter 3).

1.3 Bibliographic context: Machine learning in transmission electron microscopy for materials science

ML, DL and AI have played a pivotal role in the scientific and technological innovations of the past decade across various sectors, from the unexpected realm of finance to the foundational principles of robotics. Its widespread application across a multitude of scientific disciplines has provided essential tools for process automation and knowledge discovery through algorithms that are accessible even to those without an extensive background in computer science. Not surprisingly, electron microscopy and its target in materials science is not an exception to this trend. Interestingly, its benefits widely spread in a synergistic integration with other disciplines which may broaden experimental methodologies into previously untapped areas. In fields like astrophysics, high-energy physics, cryo-TEM for biological samples, geology, or robotics, ML already set a lasting and successful research history. Thus, the potential for electron microscopy, and specifically (S)TEM to reflect the advancements made in these areas to address its own forthcoming challenges is tangible and direct. Importantly, then, the benefits of achieving the major goals listed in the previous section lie beyond the electron microscopy and materials science community. Additionally, it may also offer valuable and expansive approaches for the general microscopy and data analysis sectors of the fields formerly named. Interestingly enough, the final paragraphs of this section pinpoint this cross-fertilisation, providing a clear vision of the trajectory ML in electron microscopy is likely to take in the coming years.

To clarify the current state-of-the-art in the field, we first delineate the progress of ML in imaging techniques, emphasising conventional parallel-beam TEM and STEM, but not forgetting recent innovations such as 4D-STEM. These constitute the role models of what could set the basis for the developments that must lead us to the central goal of the thesis. Later, we delve into STEM spectroscopies and their applications facilitated by ML. In particular, this would be helpful to connect the origin with the cornerstones of what can make us succeed when facing the third goal. In addition, the subsequent section summarises the methodological approaches of the present literature into a practical guide. Pragmatically, it describes the typical computational workflow followed in the field to apply ML to experimental datasets in a broad and less case-specific manner.

1.3.1 Most important imaging developments: from 2D images to 4D-STEM

1.3.1.1 Automation of the electron microscope vs automated data processing

In electron microscopy setups lacking hardware updates, two main strategies are employed to optimize data acquisition: data postprocessing and refining acquisition conditions. Data cleaning or denoising is crucial for ensuring reproducible and statistically significant quantitative analysis in STEM data. [38, 39, 40, 41, 42, 43, 44,

45, 46] Leveraging ML, which has demonstrated proficiency in denoising across various scientific and technical domains, was a natural choice. Initially, unsupervised ML approaches emerged as effective tools for denoising electron microscopy data.

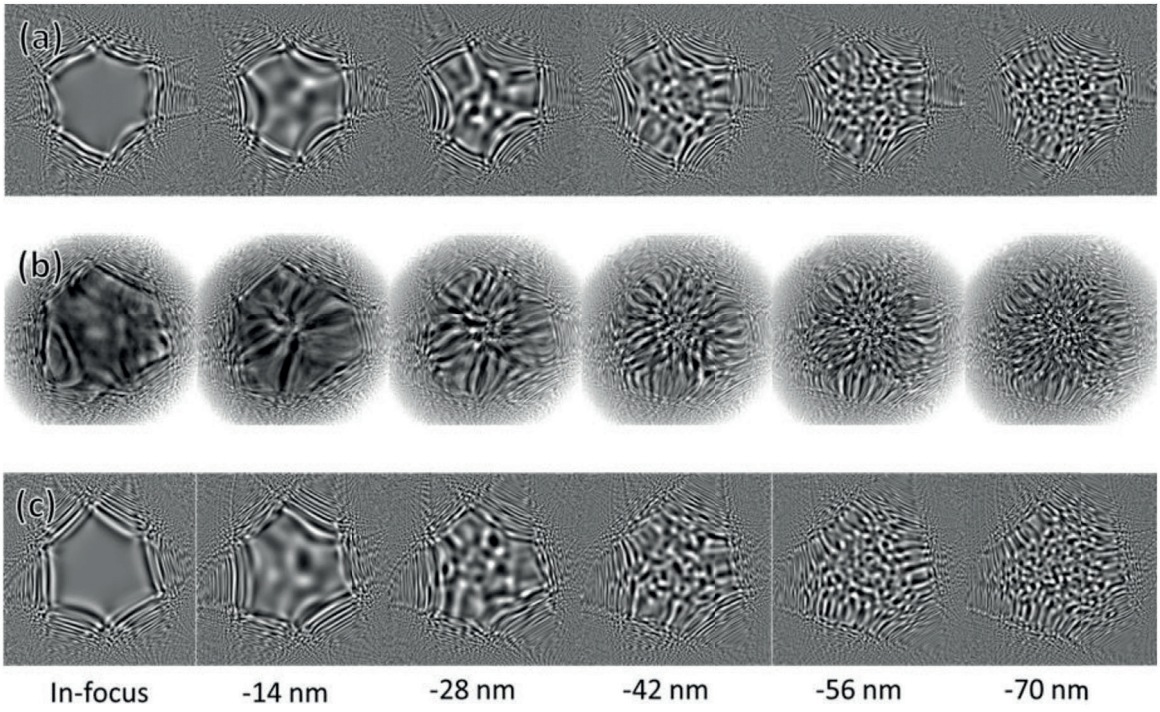


FIGURE 1.1: Complementary to the main trend of advanced data analysis, Machine Learning (ML) revolutionised the acquisition of electron microscopy datasets. The panel represents a research work from JEOL devoted to the automatic measurement and correction of aberrations through ML-based analysis of Ronchigrams. Each column is linked to a different defocus. The top and bottom rows display simulated Ronchigrams with aberrations determined manually (top) or by a Convolutional Neural Network (CNN) (bottom), while the mid row corresponds to the experimentally equivalent ones. [47, 48]

Unsupervised ML methods excel at revealing underlying structures and patterns in data, enabling classification based on intricate criteria. These approaches do not necessitate training and rely solely on the input data, offering versatility and adaptability. The pioneering works of P. Trebbia and N. Bonnet marked a significant milestone by applying unsupervised routines to denoise EELS data using multivariate statistics. This laid the foundation for subsequent advancements in post-processing routines. [49, 50, 51] For instance, the utilization of unsupervised Gaussian Processing (GP) for tasks like cleaner strain mapping showcases the evolution towards more sophisticated denoising techniques in EM data analysis. [49, 50, 52, 53, 54, 55] By embracing ML-driven denoising methods, researchers can enhance the quality, reliability, and interpretability of EM data, paving the way for more accurate and insightful analyses in electron microscopy applications.

Nevertheless, just recently more sophisticated supervised ML methods have emerged, showcasing their potential to advance current techniques. [56, 57, 58] Specifically, DL, a subset of supervised ML, aims to replicate human-like decision-making by identifying intricate patterns in data through complex regressions. In contrast, unsupervised ML is typically favoured for exploratory tasks that focus on uncovering underlying data structures, offering versatility but often with less general performance compared to supervised approaches.

The choice between supervised and unsupervised ML methods depends on the specific requirements of the analysis. While unsupervised ML is valuable for understanding data structure and exploration, supervised ML, particularly DL, excels at learning from labelled data to make predictions and classifications. The advantages and limitations of each approach are extensively discussed in the next few paragraphs.

Interestingly, following the same tendency initiated by unsupervised ML on denoising algorithms, researchers are now exploring denoising techniques from a supervised perspective and beyond. For instance, C. Lee *et al.* successfully enhanced the Signal-to-Noise Ratio (SNR) in an Annular Dark-Field (ADF) STEM image containing single-atom defects. Moreover, they accurately measured the strain field around it by training a Fully Convolutional Neural Network (FCNN) on various defect types. [59] This innovative approach allowed for the extraction of theoretically predicted strain fields that were initially obscured by noise, demonstrating the power of FCNN in identifying shared patterns within image datasets.

The success of initial denoising concepts has spurred the development of additional innovative approaches. Representatively, denoising-noising Generative Adversarial Network (GAN) for active denoising of STEM data and case-independent denoising models that outperformed traditional restoration filters in both TEM and STEM applications (fig. 1.3, GAN). [60, 61, 62, 63, 64] Notably, researchers like J. Vincent *et al.* have explored the latent features learned by DL models to uncover the underlying denoising mechanisms, shedding light on previously opaque processes. [64, 65, 66] Their work revealed that FCNNs can dynamically adjust filtering strategies based on the unique structural characteristics of each image region.

However, the influence of ML has reached further than just the post-processing of data, but its acquisition itself. Automating electron microscope alignments through Computer Vision (CV) routines has become a significant focus for refining the (S)TEM acquisition process. Initially led by microscope manufacturers, these efforts have centred on automating column alignments, including automatic aberration correction and aberration-free convergence angle optimization (fig. 1.1). [47, 67, 68] These routines are available both commercially and in the research literature. Additionally, researchers like K. Roccapriore *et al.* successfully automated dynamic STEM exploration and EELS acquisition by training deep kernels to identify

physically significant regions. [69] This automation extends beyond experiment optimization, enabling novel experimental setups previously unattainable. For instance, E. Rotunno *et al.* trained a Convolutional Neural Network (CNN) to align an orbital angular momentum sorter in beam shaping experiments, showcasing high accuracy and computational speed in estimating induced aberrations. [70, 71] This approach opens avenues for implementing real-time self-alignment across various optical systems and aberration correctors in electron microscopes, and end-to-end complete workflow automation. [72, 73, 74] It is remarkable to note that these ideas are inspired by equivalent developments in sister microscopies such as optical or Scanning Probe Microscopies (SPM). [75, 76] This emphasises how the cross-feeding between disciplines is key to helping each other and also how the collaboration between industry and academia is essential to integrate research community innovations into user-friendly microscope implementations. Additionally, standardizing open and universal data formats is crucial to garner support from companies and propel the field towards more streamlined and advanced microscopy practices. [77] Importantly, this is an idea we will pay special attention to in Chapter 4, to achieve the second goal of the thesis (section 1.2) in the most universal and reproducible manner.

However, the current bottleneck of the workflow lies in the analysis of the acquired data, not in the acquisition itself. Precisely, we envision a data analysis breakthrough as the central goal of the present dissertation (i.e., research goals, section 1.2). Therefore, most of the current efforts are pushing towards automated analysis and ML-based knowledge extraction. The extensive and successful literature in this area provides a solid foundation for developing our urging unified global AI model capable of automating thorough structural characterization tasks without human intervention, as we want to achieve. For instance, CNNs have emerged as powerful tools for defect identification, classification, and quantification in electron microscopy data.

Initial studies focused on the challenges posed by 2D materials in STEM, enabling direct correlations between images and structures. [78, 79, 80] For example, M. Ziatdinov *et al.* trained an encoder-decoder FCNN to detect atomic coordinates in Si-doped graphene, showcasing the model's adaptability to other 2D systems like $Mo_{1-x}W_xSe_2$ with generalized performance. The authors transformed atomic coordinates into graphs for automated Si dopant classification based on bonding patterns with neighbouring atoms (fig. 1.3, CNN). [81] This approach has been widely replicated in further research on similar 2D systems, by taking advantage of the monotonic (with the atomic weight and thickness) dark-field STEM signal. Remarkably, the importance of this dependence is raised in the methodology chapter 2, and will be central to pursuing the goals we set out. In more challenging scenarios implying High-Resolution parallel illumination TEM (HRTEM) , J. Madsen *et al.* employed a similar FCNN architecture on simulated micrographs. They incorporated phase information to identify atomic coordinates in graphene and to count atoms in gold

nanoparticles (fig. 1.2). [82] Importantly, his foundational idea extended the analysis of defects through HRTEM further to surface contaminants on more complex graphene configurations or to quantitative atom counting in Au nanoparticles. [39, 41, 83]

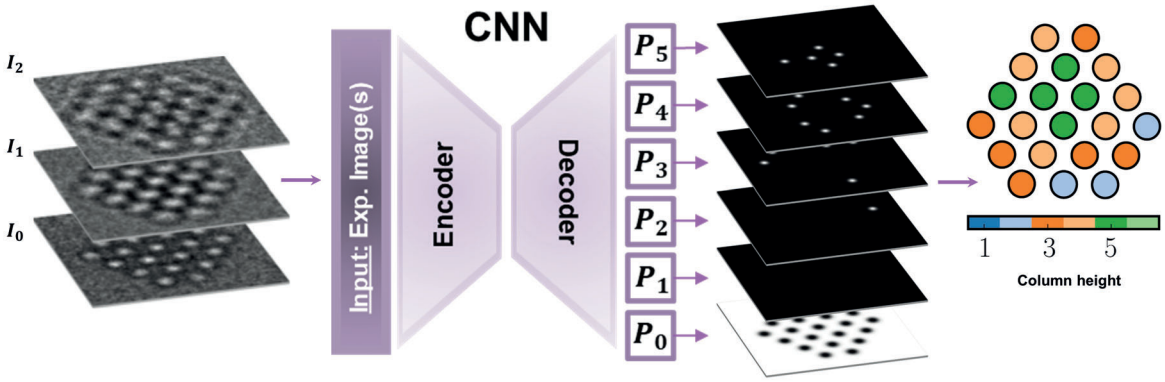


FIGURE 1.2: Together with automated data acquisition, Machine Learning (ML) mostly serves for the advanced analysis of electron microscopy data. We represent the diagram of a CNN trained to count the atoms of an atomic column in gold nanoparticles. The CNN admits both single images and focal series as inputs to classify every atomic column by setting, in the depicted case, a probability (PX) of containing from 1 (P0) to 6 (P5) atoms. [48, 82]

The next step in complexity comprises defect detection in 3D systems. The existing research in this area is limited and primarily theoretical. The proofs of concept are mainly demonstrated in simpler systems like zinc-blende GaAs. As a result, this presents a significant opportunity to advance state-of-the-art defect detection across a wider range of 3D systems. To achieve this, leveraging CNNs in conjunction with unsupervised clustering or anomaly detection methods appears to be a promising approach. [84, 85] The opportunity lies in a comprehensive model that will not only encompass the identification of planar defects and dislocations at an atomic level but also extend to a more macroscopic or industrially-oriented defect analysis in 3D systems. [86, 87, 88] The great potential observed and the perspectives of integrating advanced ML techniques with electron microscopy encourage us to aspire to high hopes. Indeed, we get surprising and intriguing results thanks to this multidisciplinary character (chapter 4).

1.3.1.2 Unsupervised exploratory and knowledge-revealing routines

ML has shown remarkable proficiency in exploring local descriptors and extracting knowledge in electron microscopy data analysis. Not surprisingly, the current research landscape predominantly leans towards utilizing supervised deep neural networks or intricate unsupervised methods for tasks like dimensionality reduction and classification. However, the tendency suggests that unsupervised generative models will play a pivotal role in pushing the boundaries of data science applications in electron microscopy for materials science.

The adoption of unsupervised generative models is anticipated to revolutionize the field. They would significantly streamline the laborious process of generating training sets. Therefore, they would loosen the need for extensive customization and fine-tuning associated with traditional unsupervised techniques. This shift is expected to be a transformative legacy from other scientific domains, offering new avenues for advancements in materials science and electron microscopy. [89]

Nevertheless, despite its future perspectives, the journey of unsupervised ML in electron microscopy commenced with the initial focus on multivariate analysis of EELS spectra, primarily targeting noise reduction. Since the inception of the earliest nearest neighbour algorithm designed to identify the shortest path connecting data points, unsupervised learning has rapidly gained traction across various fields and applications. Its primary strength lies in the absence of a training process, which enhances accessibility and ease of implementation. The general challenge that accessing labelled data entails contributed significantly to its widespread adoption and success. Although unsupervised routines may lack the robustness to generalise models, they provide an excellent platform for uncovering patterns and intricacies of electron microscopy datasets. Remarkably, chapter 4 embraces this particular benefit of using unsupervised models in one of the first steps of the workflow we propose to address our central goal.

Unsupervised ML in imaging techniques has evolved from the initial multivariate analysis of spectra, with Principal Component Analysis (PCA) serving as a fundamental tool for dimensionality reduction and data decomposition. Importantly, PCA enables data cleaning, classification, and can play a crucial role in more complex routines. For instance, it has been instrumental in mapping crystal phases and defects like twins or phase boundaries in micrographs of dichalcogenides such as $MoSe_2$ and WS_2. Interestingly, PCA is perceptive of rotational invariant features and finds them regardless of their in-plane rotation. [78, 90]

Despite its utility, the main limitation of PCA lies in its purely mathematical nature, lacking direct correlation with physical interpretations. This led to the development of alternative methods like Non-negative Matrix Factorization (NMF), which enforces positivity constraints on components. For instance, R. Kannan *et al.* tested NMF in spectral data generated by multimodal STEM and X-ray Diffraction (XRD). [80] The authors generated a hyperspectral dataset from the sliding window Fast Fourier Transform (FFT) of a single atomically resolved STEM image. Then, they mapped the crystallographic phase by assigning a phase to each meaningful NMF component extending the previous PCA approach further. [80, 88, 91, 92] Similarly, B. Martineau *et al.* also applied NMF, combining it with fuzzy c-means clustering (i.e., assigning each datum a probability of being in each cluster), to overcome the effects of sample bending on the diffraction patterns obtained in precession mode while scanning twinned GaAs nanowires (fig. 1.3, C-cluster). [93]

The balance between ease of implementation and remarkable results achieved by PCA or NMF has spurred interest in exploring more complex unsupervised routines. [94, 95, 96, 97] Variational AutoEncoders (VAEs) have emerged as potent tools for dimensionality reduction, capable of extracting physically meaningful information (fig. 1.3, VAEs). While categorized as unsupervised processes, VAEs necessitate manual specification of the data's features of interest for each study. In a crystal phase classification scenario, for instance, the phases across a micrograph must be predefined as input. By delineating these features, the data undergoes compression and decompression within an encoder-decoder architecture, generating a few latent variables that succinctly capture the original image's variability. [98, 99, 100] Eventually, these compressed latent variables can then undergo additional clustering or refinement processes to establish correlations with physically significant features like local crystallography, defects, or sample-specific identifiers. [101, 102]

Notably, S. Kalinin *et al.* utilized rotationally invariant VAEs (rVAEs) to study the evolution of Si on graphene under electron beam irradiation. By encoding the system's variability into three latent variables (i.e., x and y translations, and rotations), the authors demonstrated the robustness of VAEs in tracking time-resolved data. rVAEs proved superior to traditional unmixing methods by efficiently capturing rotation information in a single latent variable and elucidating other sources of variability in independent variables. [103] That is why, rVAEs are particularly valuable for evaluating features susceptible to in-plane rotations, offering processing advantages for polar materials like ferroelectric or ferromagnetic substances. In their work, S. Kalinin *et al.* applied rVAEs to correlate rotation latent variables with ferroic variant orientations. They enabled direct identification of unit cell deformations throughout samples, independently of the structural and chemical variability and under non-ideal imaging conditions. [104]

The attentive reader should have noted that the Si-doped graphene reappears in a more complex computational approach. Therefore, the current status of these advanced ML methods indicates that they are still in the proof-of-concept phase, primarily applied to model systems rather than widely adopted in routine electron microscopy workflows. Moreover, the success so far proven in scenarios with limited variability implies that there is significant potential for their application in more complex systems that offer greater challenges and are inherently more interesting from a materials science perspective.

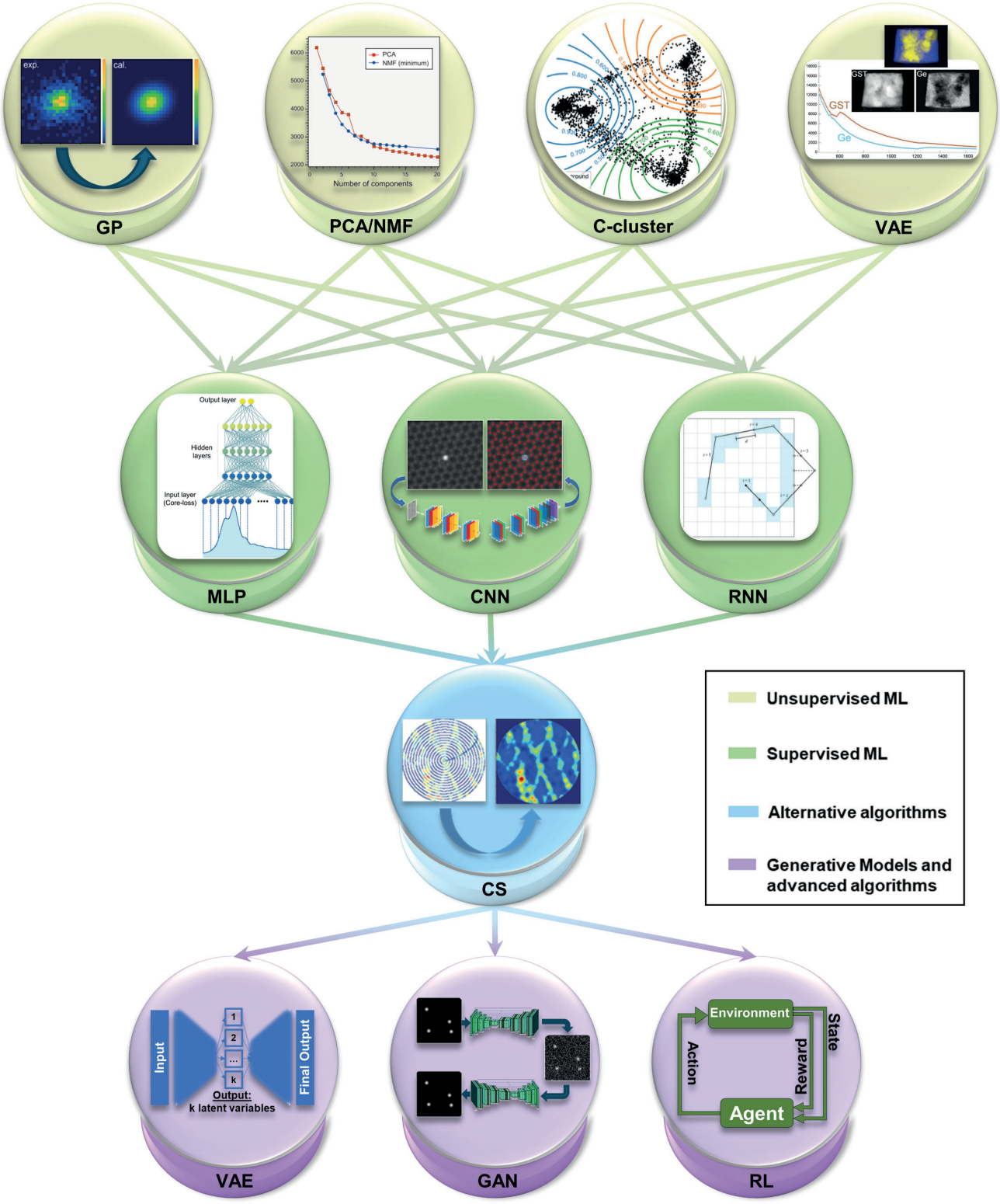


FIGURE 1.3: A selection of the key Machine Learning (ML) methods most recurrently employed in electron microscopy imaging and spectroscopy. The first layer contains the first ML techniques that arose in electron microscopy, unsupervised methods, consisting of Gaussian Processing (GP) [52], Principal Component Analysis and Non-negative Matrix Factorisation (PCA/NMF) [105], fuzzy c-clustering (C-cluster) [93], Vertex Component Analysis (VCA) [106], among others. The evolution in complexity leads to the second layer of supervised techniques, with MultiLayer Perceptron (MLP) [107], and Convolutional Neural Networks (CNN) [81] and Recurrent NN (RNN) [108]. The third layer stands for Compressed Sensing (CS) [109], a group of alternative algorithms with powerful compatibility with ML. Finally, the last layer outputs the future of ML in electron microscopy in the use of generative models like Variational AutoEncoders (VAE) and Generative Adversarial Networks (GAN) [60], and more complex learning paradigms such as Reinforcement Learning (RL). [48].

In summary, to achieve the proposed research goals, it is essential to acknowledge the availability of simple algorithms that can effectively accomplish tasks. However, in cases where challenges arise, more complex algorithms have demonstrated superior performance. Consequently, our starting point in terms of toolbox variability and wealth remarkably sets a strong foundation to work with towards our purposes. In the methodology chapter (chapter 2), one can explore the unsupervised algorithms utilized. While the need for highly complex unsupervised approaches like VAEs may not have been necessary for the specific objectives, the situation may differ when considering supervised approaches. Let us now delve into the same scenarios but through the lens of supervised routines.

1.3.1.3 Supervised exploratory and knowledge-revealing routines

As briefly pointed out before, together with quantum computing, neuromorphic hardware architectures are the latest revolutionary idea in computer science to overcome the limits of Moore's law. [2] On the software and algorithms side, this mimicking trend appeared with the very first perceptron proposed by W. Pitts and W. McCulloch. [110] It continued with the explosion of DL, neural networks and other supervised algorithms, which have systematically outperformed classical classification and regression tasks. [111]

Supervised routines, particularly DL, operate on a training process that necessitates the preparation of a substantial dataset representing the statistical variability inherent in the data under study. The fundamental concept revolves around crafting a model, such as a neural network, capable of discerning the statistical patterns and features within the data through its training process. Once this model is fine-tuned through training, it becomes adept at automatically and robustly analyzing any data falling within these statistical boundaries. [111] In essence, just as a linear regression model adjusts to data exhibiting a linear trend, neural networks can create intricate non-linear multi-dimensional regressions that theoretically can adapt to various data structures. However, the main drawback lies in the resource-intensive nature of setting up the models. It involves generating and preparing (i.e., often manually labelling) the training data while tailoring the model architecture to suit the specific problem at hand. This issue is problematic enough to even encourage the research of automatic AI-driven architecture design and hyperparameter fine-tuning for supervised models. Thus, neural networks that optimise other neural networks. Interestingly, the proof of concept has even been ascertained within our context of electron microscopy and materials science. [112, 113, 114]

Currently, though, the predominant focus of DL in the field has been the localisation of atomic positions and their correlation with mathematical graphs, as stated before, for defect identification and quantification, or image denoising. [59, 64, 65, 66, 80, 81, 82, 83, 115, 116] This trend has seen significant advancements in applications like

atom counting, quantitative TEM, and the eventual extraction of properties with physical significance out of them. [41, 82, 117]

Initially, the identification of atomic positions began with FCNN on 2D systems and later extended to model 3D systems, showcasing their generalization capabilities to handle higher complexity systems. Importantly, efforts aimed at developing more universal models and approaches are an essential spotlight to guide the manner to approach the goals listed in section 1.2. For instance, researchers like M. Ziatdinov *et al.* trained FCNN to detect atomic positions in a La-doped $BiFeO_3$ system to extract local lattice descriptors like polarisation. [118] Subsequently, these local descriptors obtained from their supervised networks were further processed using unsupervised methods to categorize them into physically meaningful groups. Specifically, the authors post-processed with techniques like PCA and k-means clustering to map the distribution of lattice distortions back to the original image. [118, 119]

The previous idea serves as another pillar to achieve our workflow development goals. Multimodal approaches, as exemplified above, play a crucial role in tasks like atomic column finding and phase mapping by overcoming the limitations of standalone models. These methodologies are essential for handling the often complex and challenging output data from neural networks, which can be simplified using classical unsupervised routines or more complex algorithms like VAEs. [104, 120, 121] Beyond multimodal routines, we can exploit the knowledge and state-of-the-art models trained in the literature to fine-tune them to our research needs. We then maximise the reward vs risk or deployed effort by transfer learning of these pre-trained models. For instance, given that atomic columns typically appear as rounded dots in micrographs across DL-based studies, there is potential for developing a universal model through shared transfer learning-based approaches. This strategy could streamline model development by providing a common starting point that can be adapted for specific needs, saving time and resources. Interestingly, we can foresee the mixed contribution of multimodal scopes with transfer learning as the basis of the supervised algorithms deployed to meet our objectives in chapter 4.

The progress of ML in real-space micrographs encouraged targeting the reciprocal space. Remarkably, ML and DL methods demonstrate sensitive geometry awareness, making them well-suited for detecting symmetry constraints in crystals and their reciprocal space representations. Moreover, the task is simplified if we alleviate computational demands by transforming 2D electron diffractograms into 1D spectra. For instance, J. Aguiar *et al.* proposed training a CNN on electron diffraction data and on the FFTs derived from atomic resolution STEM micrographs. The ultimate purpose was to identify the space group of a given unit cell. [122] The fundamental concept involved the azimuthal integration of diffraction patterns or FFTs to generate line profiles encapsulating comprehensive information in a condensed format. Thus, the authors made the most of a substantial dataset comprising 571.340 crystals, achieving a classification confidence of 95% under regular SNR conditions

and 70% in noisier data scenarios. One step further, R. Vasudevan *et al.* developed a CNN capable of determining the Bravais lattice from both experimental and simulated Scanning Tunneling Microscopy (STM) and STEM images directly from 2D diffraction data. Indeed, the authors mapped this Bravais symmetry distribution within a time-resolved set of images of electron beam-damaged WS_2. [115, 123, 124] Interestingly, these constitute preliminary signs highlighting the interest to apply supervised methods, and generally ML, to evolving systems or *in-situ* setups. This idea is reviewed in detail a few paragraphs below.

The versatility of supervised methods has demonstrated their capability to efficiently handle diverse experimental data. As a result, the ML practitioners in the community attempted their implementation in microscopy simulations, too. Thus, simulations play a crucial role in electron microscopy by aiding in the interpretation of micrographs captured under complex conditions and enabling virtual experiments on characterized structures. In that sense, for instance, R. Pennington *et al.* combined Density Functional Theory (DFT) calculations with a neural network-based optimization algorithm to enhance TEM simulations. Beyond this enhancement, they could extract properties along the beam propagation dimension, such as ferroelectric polarisation domains and strain. This methodology was validated using simulated Convergent-Beam Electron Diffraction (CBED) patterns, showcasing its potential applicability to various imaging and spectroscopic modes. [108, 125]

Research has shown that DL models can learn features that extend beyond the training data, even to data seemingly unrelated to the initial training set. However, the exact mechanisms through which DL models acquire these complex routines remain largely unknown. Therefore, understanding how DL models learn is crucial for distilling shared patterns from the task-specific network weights (i.e., coefficients ruling the non-linear transformations of the data within the neural networks) learnt across multiple applications. Thus, studies like that of J. Horwath *et al.*, who examined the learnt convolutional filters of a U-Net architecture for segmenting TEM images of gold nanoparticles, have shed light on the importance of tailoring model architectures for specific tasks. Intriguingly, they learnt that the convolutional kernels could be easily engineered by combinations of traditional filters (i.e., Laplacian, gaussian, etc). Consequently, tailoring the model architecture for a task becomes essential. [64, 126, 127, 128, 129] Not only to elucidate the application of DL in materials science from a computer science perspective, but also essential to identify common patterns in networks with shared objectives and leverage them to streamline specific data processing challenges effectively. In fact, DL on general digital image processing revealed that the first learned kernels of CNNs are those oriented to identify contours in different directions. [130] Not surprisingly then, the idea of contour detection will be key in our workflow, although we exploit it both directly from DL and from its traditional formulation (chapters 2 and 4).

In conclusion, in electron microscopy, the typical research workflow involving supervised ML consists of tackling a specific problem and generating a supervised model that fits into the target information. Maybe without really being aware of it, the electron microscopy community has already benefited from them. Thus, the latest hardware evolutions and their software assistance in the community were based on automation. The clearest example might sit outside materials science, and it is cryo-TEM and its autonomous particle search and analysis. In any case, these task-specific solutions have yet to converge into broader unified models with a customisable or modular nature that gathers all the developed features at once. Moreover, every released model must be open access following the commitment of the community so far. [77, 104, 118] In fact, we pledge to publicly release the code developed to achieve the proposed goals as soon as the content is published in a peer-reviewed journal. Indeed, there is already an open-access publication together with publicly available code partially addressing the third major goal (i.e., research developed in chapter 5). [20, 131]

1.3.1.4 Processing of high dimensionality data by machine learning

ML is driving significant advancements in processing higher dimensionality electron microscopy data. Indeed, the most successful processing routines so far are to an extent based on ML. The latest pixelated detectors are enabling the acquisition of large data volumes, approaching the scale of big data. Consequently, strategies such as processing data in fragments or chunks are becoming essential for the EM community to effectively handle this influx of data. Numerous techniques that fit into the high dimensionality data tag, like Lorentz microscopy or Differential Phase Contrast (DPC) STEM, have benefited from ML to improve the retrieval of the phase information. [132] However, the central focus of the community has primarily shifted towards addressing the management of extensive data streams generated by ptychography and 4D-STEM nanodiffraction.

Ptychography, a promising technique in electron microscopy, involves mathematically reconstructing electron phase information (i.e., the so-called inverse problem) from acquired experimental signals. [133, 134, 135, 136] It has recently demonstrated its value by achieving a spatial resolution record in EM of 14 picometers, limited only by lattice vibrations, and by directly using the atomic orbitals. [137, 138] Moreover, it challenged aberration correction by enabling a resolution of 44 picometers with an uncorrected microscope. [139] Nevertheless, its establishment as a daily-use technique encompasses progress in the computational methods for its reconstructions. To address this, ML applied to ptychography focuses on optimizing phase retrieval routines. [140] For instance, M. Schloz *et al.* combined the multislice formalism (interestingly, implemented with a MultiLayer Perceptron (MLP), which is a fully

connected neural network, fig. 1.3, MLP) with gradient descent regularised optimisation to perform the reconstruction. The multislice formalism accounted for multiple scattering and improved the resolution. On the other hand, the regularisation allowed to loosen the oversampling requirements, still enabling the reconstruction of experimental data under noisy conditions. [141]

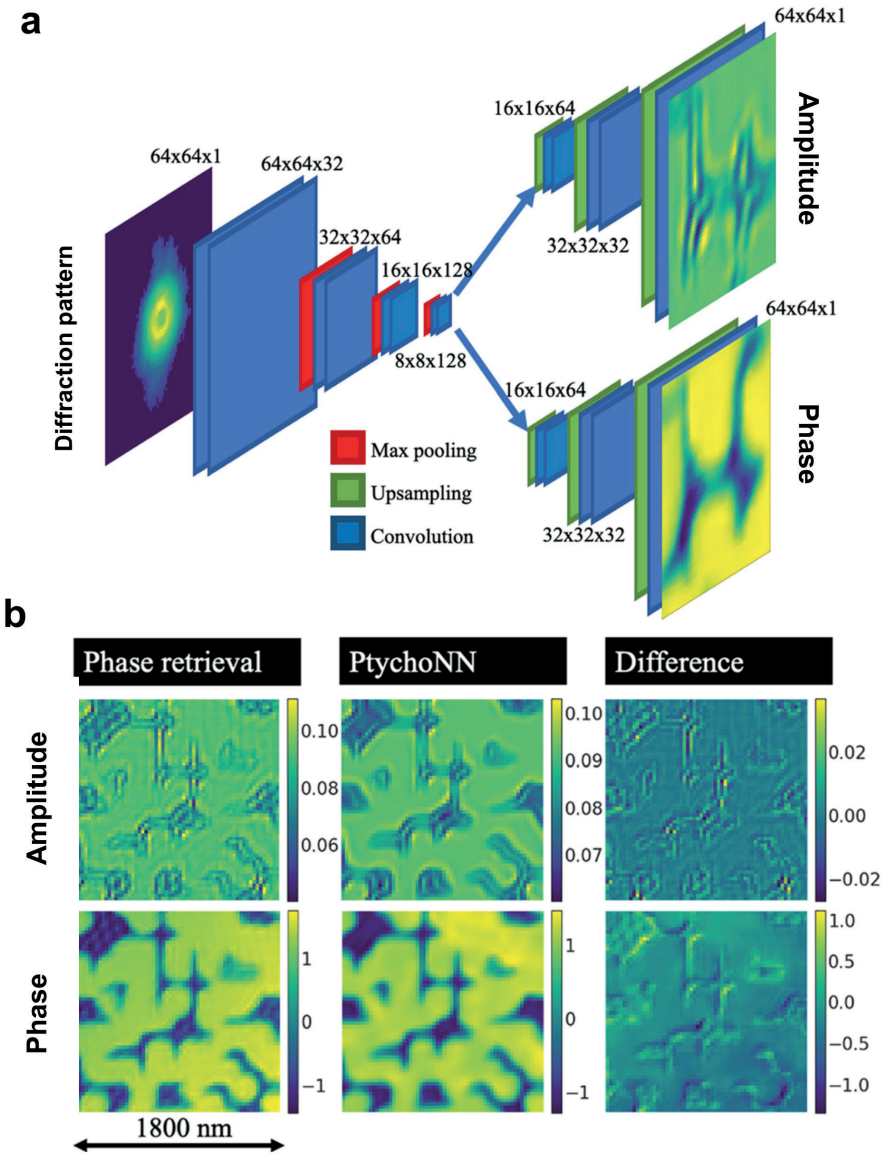


FIGURE 1.4: Machine Learning (ML) applied to electron microscopy high-dimensionality data analysis. (a) An encoder-double decoder convolutional neural network dubbed PtychoNN designed for the ptychography reconstruction of the amplitude and phase information out of single diffraction patterns. PtychoNN allowed (b) to reconstruct the wave information with higher fidelity than traditional iterative phase retrieval algorithms, and up to 300 times faster. [48, 142]

Furthermore, in the realm of ptychography, advancements have been made towards streamlining the process into an end-to-end solution. Note that, generally, the ultimate goal for using ML tends to be a global solution which comprehensively groups all the intermediate steps of a process. Indeed, the major goal of the thesis is to achieve an end-to-end solution (section 1.2). In that sense, but within ptychography, M. Cherukara *et al.* developed a FCNN named PtychoNN, trained to extract both amplitude and phase information from diffraction data. This innovative approach claims to provide a rapid, sub-sampling tolerant, and computationally efficient method for real-time solving of the inverse problem in a single step (fig. 1.4). [142] Interestingly, achieving real-time ptychography opens up possibilities for applications in dose-sensitive and thick samples. In fact, ML has played a crucial role in reducing electron dose requirements in ptychography through advanced methods like deep Reinforcement Learning (RL). RL, which learns by rewarding desired behaviours while penalizing others, has been utilised to generate adaptive scanning paths in real time (fig. 1.3, RL). Consequently, by optimizing scanning paths based on physically meaningful diffraction patterns, RL helps to relax overdetermination constraints for low-dose experiments. The intrinsic time dependence in the scanning process made the authors opt for a Recurrent Neural Network (RNN) which automatically engineered the most optimal scanning path for the reconstructions (fig. 1.3, RNN). [143, 144, 145] Overall, the examples showcased the unlocking potential of ML to enhance imaging modes such as ptychography beyond pure automation.

Nevertheless, the acquisition of diffraction patterns to generate 4D datasets is more general and goes beyond ptychography. In general 4D-STEM nanodiffraction, pixelated detectors are essential for collecting the array of 2D diffraction patterns and have garnered attention from ML practitioners. [16, 146, 147, 148] Notably, G. Correa and D. Muller enhanced detector performance by training a Convolutional Neural Network (CNN) on Monte Carlo electron beam trajectories. This approach aimed to predict experimental beam paths and detector hitting spots to achieve sub-pixel super-resolution, highlighting the potential synergy between software development and hardware advancements in electron microscopy. [149] While such research is currently limited in literature, it exemplifies the evolving landscape of software-hardware integration in electron microscopy research.

ML is increasingly employed in the context of 4D-STEM to efficiently extract physical properties from large datasets. These tasks are challenging for traditional algorithms, primarily due to the enormous data streams associated to 4D datasets. To address this challenge, unsupervised methods have emerged as a common approach, akin to their application in regular atomically-resolved (S)TEM data analysis. These methods aim to attribute a physical interpretation to mathematical components. For instance, associating each NMF or PCA component with a characteristic diffraction pattern. Not surprisingly, the success of unsupervised techniques in processing high-dimensional datasets has established them as a standard practice in

the field. [150, 151, 152] Notably, these techniques have been successfully applied to map crystal phases and their relative rotations in diverse systems like complex oxide structures of similar compositions (e.g., $Ti_{0.87}O_2$ vs Ti_2O_3) and dichalcogenide multilayers such as MoS_2 bilayers. [105, 153] Seemingly, this approach mirrors the multimodal crystal phase mapping seen in conventional STEM applications. Then, the progressive complexity escalation of ML in microscopy is further showcased and encourages us to march toward our ambitions. [154, 155, 156]

Supervised ML has also been proven beneficial for 4D-STEM, particularly in the context of ptychography. However, implementing supervised ML in 4D-STEM can be significantly more time-consuming compared to regular (S)TEM analysis. This increased complexity primarily arises from the higher dimensionality of 4D-STEM data, which necessitates more intricate network architectures and extensive engineering of training data. These challenges can pose obstacles during implementation. [157] As a result, unsupervised methods are often considered a more accessible starting point for newcomers entering the field of smart processing of 4D datasets. Despite these complexities, there have been successful examples of preparing effective training sets for supervised ML in 4D-STEM. For example, consider the training of a 3D Convolutional Neural Network (CNN) using simulated diffraction patterns from a $LaAlO_3$-$SrTiO_3$ heterostructure. Remarkably, this model was capable of discerning whether the interface exhibited atomically sharp steps or a chemically diffuse nature, a task that can be intricate even for experienced microscopists. [158]

Understanding the challenges posed by higher-dimensional data streams remains crucial, even with our clear focus on regular STEM applications. Regardless, it is beneficial to remain mindful of the potential complexities that may arise. In fact, this context will prove fruitful as we tackle the 3D data volumes we anticipate in our research, specifically, spectrum images (in chapter 5), and even stacks of images from full datasets related to a single device (such as the datasets acquired for chapter 3, and subsequently processed in chapter 4). In that sense, let us now explore additional imaging modes that might yield equivalent high-dimensionality data, albeit from different sources from those evaluated thus far.

1.3.1.5 Compressed sensing and machine learning for tomography and in-situ

In the previous paragraphs, we have pointed out that general statistical parameter estimation and DL methods allowed for atom counting and local structure determination. [41, 45, 82, 117, 159, 160, 161] These methods aim for an eventual 3D model representative of the initially unknown nanostructure. Even though these tools can become powerful enough to solve certain case studies, the reproducible extraction of reliable 3D information demands electron tomography. Importantly, our ultimate goal is to skip tomography and reach its typical results without all the intricacies it relates to. Therefore, comprehending its challenges and how ML deals with 3D objects is fundamental.

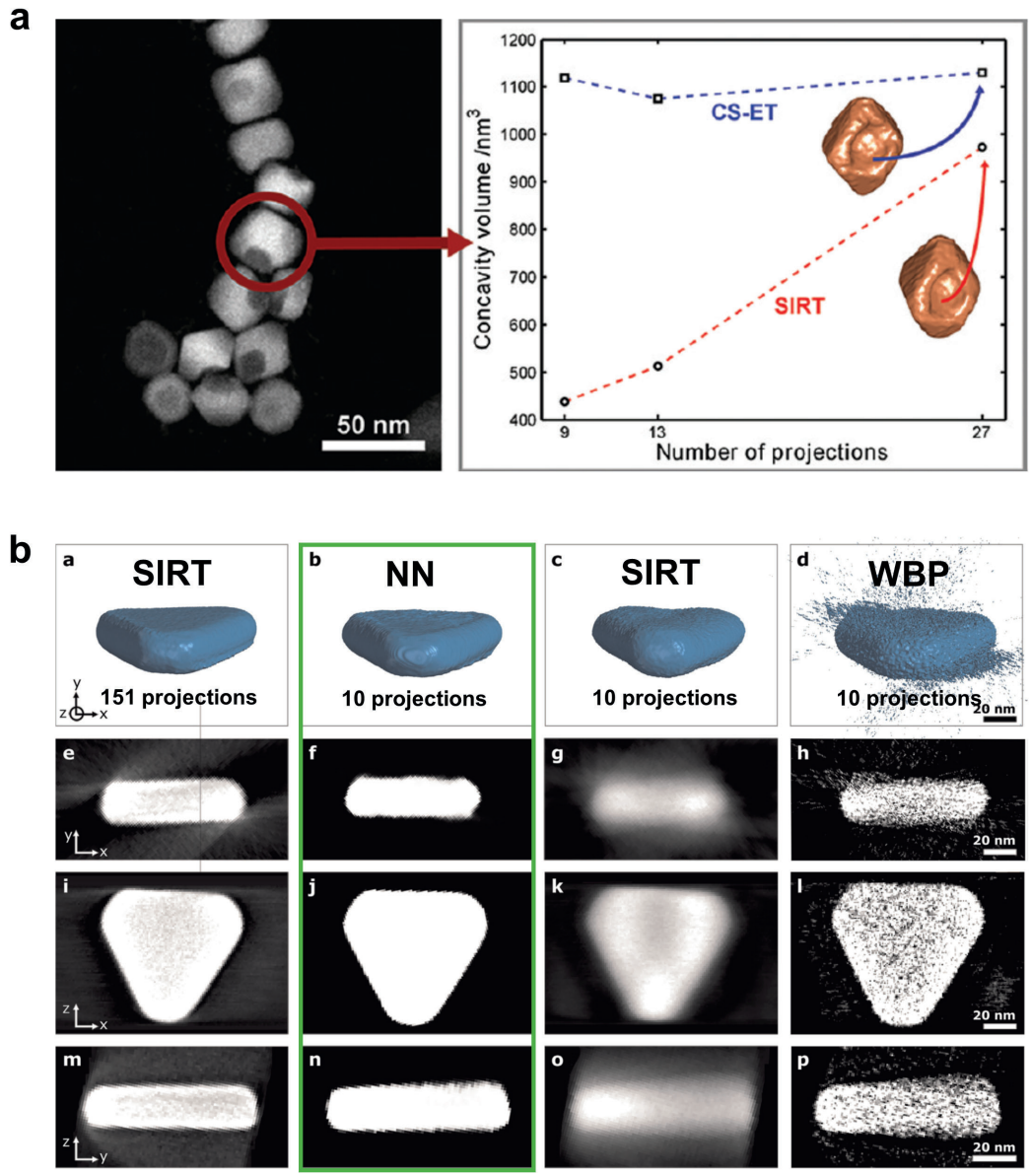


FIGURE 1.5: Machine Learning (ML) applied to electron microscopy high-dimensionality data analysis. (a) Comparison of the accuracy of the electron tomography reconstruction of an iron oxide nanoparticle with a concavity between the SIRT algorithm and a Compressed Sensing (CS)-based one. The CS algorithm reproducibly estimated the concavity size independently of the number of acquired projections while the SIRT approximation was strongly affected by that. [162] (b) Electron tomography reconstruction by, from left to right, the SIRT algorithm with 151 acquired projections, and a Neural Network (NN), SIRT and WBP algorithms with 10 projections, and their orthoslices in the three bottom rows. The NN overwhelmingly improved the 10 projections SIRT and WBP reconstruction, but also the reference 151-projections SIRT. The NN orthoslices clarified its superior performance highlighting the potential for low-dose tomography. [48, 163]

Compressive Sensing (CS) has recently become a keystone used to relax the time-consuming and high-dose experimental setups that tomography usually demands. [164, 165] In just a few words, CS allows to reconstruct a randomly-sampled sparse

signal beyond the Nyquist limit utilizing a mathematical algorithm that contemplates the sampling path. Although strictly speaking, CS does not fall within the ML category, its shared objectives and similar methodologies make it worthwhile to consider as a potentially usable tool.

Research on CS applied to tomography primarily focuses on enhancing current tomogram reconstruction methods. The aim is to reduce the total dose in virtually every experimental setup while achieving dose-reduced, real-time, and shape-independent tomography. [166, 167, 168, 169] This approach not only enhances sensitivity to challenging morphological features like surface rugosity and porosity but also unlocks the imaging of complex shapes. For example, Z. Saghi *et al.* quantitatively evaluated the concavities of iron oxide nanoparticles and reconstructed challenging biological needle-shaped tomograms using CS (fig. 1.5, a). [162, 170]

The success of CS in tomography has led to a trend where ML algorithms seek to replicate these techniques for more efficient implementation. DL has significantly contributed by automating reconstruction processes. [171, 172, 173, 174, 175] For instance, E. Bladt *et al.* surpassed traditional CS methods by introducing a MLP capable of autonomously reconstructing samples from sparse tilt series without the need for user-defined prior knowledge, as typically required in traditional reconstruction algorithms 1.5, b). [163, 176] The integration of ML techniques inspired by CS not only enhances tomographic reconstruction but also signals a shift towards ML superseding any classical implementation, as viewed for DFT and multislice, and virtually for any data science purpose.

The application of CS has extended beyond electron tomography to revolutionize STEM imaging by introducing non-rectangular scanning paths, and promising faster and lower-dose STEM experiments. [177, 178] This innovative approach involves measuring sparse pixels and mathematically inferring the unsampled signal, proving successful in scenarios like undersampled lattice distortions such as point defects in both STEM imaging and EELS. [179] The primary appeal of non-rectangular scanning paths in STEM is not just the speed enhancement but also their utility in mitigating dose limitations for beam-sensitive nanostructures. Researchers like X. Li *et al.* demonstrated real-time CS-based reconstructions on beam-sensitive materials like graphene using non-rectangular spiral and Lissajous scans. [180] Interestingly, this approach has also been applied to SPM, showcasing the versatility of CS in optimizing scanning paths despite their original technique. [181]

Similar to CS in tomography, DL has emulated the benefits of CS in optimizing STEM scanning paths. For instance, J. Ede and R. Beanland modelled a GAN to automatically reconstruct non-rectangular scan paths. Based on this, J. Ede created a sample-aware adaptive scan path learned by reinforcing a recurrent NN that scores the reconstruction of the GAN from the previous study (fig. 1.3, RNN). [182, 183]

This research piece constitutes one of the most sophisticated examples of ML mastery applied to electron microscopy currently available. Therefore, it underscores the importance of a diverse ML toolset for running challenging applications in electron microscopy, such as the ones we propose, which will inevitably require multimodal ML.

CS has also demonstrated its potential in compressing the data volume associated with time series in *in-situ* experiments, offering a valuable tool for managing and analyzing dynamic experimental data efficiently. [184] This application of CS again highlights the significant role that ML can play in addressing challenges posed by *in-situ* experimental setups. [22, 185, 186] At the end, *in-situ* experiments lead to image stacks in which time is the variable correlating frames. In our case, we do not have this dependence, although we will occasionally need to strategically deal with image stacks from the thorough imaging of full devices.

In the previous paragraphs, we have already reviewed some examples in which the impact of the beam, mainly on 2D materials, was carefully examined. Researchers interested in replicating these experiments would benefit from assessing changes in defects or phase distributions. [64, 103, 115, 126, 159] For instance, T. Patra *et al.* conducted a study on MoS_2 across different time scales by integrating ML-enhanced dynamic simulations of defects with time-resolved HRTEM. Their findings revealed that long-time scale displacements of defects could trigger rapid (up to picoseconds) phase transitions from 2H to 1T phases. [187] ML played a crucial role in enabling the necessary time resolution to capture these ultrafast phase transitions, invisible to classical *in-situ* data analysis methods.

Not surprisingly, though, current investigations involving ML in *in-situ* experiments are still in their early stages. They primarily tackle relatively simple systems like nanoparticles and focus on their size and morphology evolution. These initial applications do not require advanced methodologies, making them accessible to newcomers in the field of ML. In fact, some studies in this area combine traditional CV techniques such as thresholding and edge detection with unsupervised clustering to delineate particle contours and central positions. [188, 189] In this context, clustering serves as a form of sample-aware thresholding that can adapt better to changing conditions in time series or dynamic stimuli. For example, Y. Qian *et al.* developed an unsupervised segmentation method that effectively combines k-means clustering for intensity evaluation with edge detection for gradient assessment to extract comprehensive statistics from videos of silica nanoparticles. [190, 191] This approach demonstrated robustness and was successfully applied to more complex scenarios, including the study of the environmental TEM process of Fe nanoparticle formation following the dewetting of a Fe thin film. [192] This research case is inspiring for us, as achieves image segmentation with unsupervised methods. As a consequence, we will try to emulate this accomplishment to segment the STEM images with which

we work. In the Methodology chapter (Chapter 2) we justify this choice, while the results can be checked in Chapter 4.

The previous example remarked that evolving systems may discover their best ally in simple algorithms for capturing their dynamics without the need for complex codes deployment. However, abrupt intensity changes in micrographs, caused by factors such as thickness variations or compositional shifts, can pose challenges for unsupervised routines. In such cases, a supervised approach specifically trained to handle these image variations may be necessary to ensure robustness. Therefore, despite the model design challenges, capturing dynamic changes during *in-situ* experiments would eventually be more plausible. Supervised routines based on CNN architectures like U-Net have proven successful in segmenting and tracking nanoparticle statistics. U-Net, known for its effectiveness in scientific image segmentation tasks such as cell segmentation, has become a foundational model for almost any CNN deployed in microscopy. [193] Before delving into the intricate design of neural network architectures, starting with U-Net as a reliable supervised proof of concept is often recommended. Interestingly, meaningful results have been achieved using U-Net in various conditions such as liquid and gas phases, and under temperature variations, surpassing traditional unsupervised approaches. [194, 195, 196] Not surprisingly, U-Net will be our basis to develop any tailored supervised model (see chapter 2 for further details).

The potential of ML in *in-situ* Electron Microscopy (EM) is vast and is expected to expand significantly as *in-situ* techniques become more widely adopted within the scientific community. Future applications may involve leveraging CS to reduce frame rates for tracking beam-sensitive materials and utilizing VAEs to encode time-dependent descriptors for direct correlation with material dynamics. Furthermore, the combination of ML efficiency and advancements in fast detectors could enable real-time correlation between *in-situ* experiments and parallel simulations. These systems would automatically and comprehensively record and explain the evolution of materials. This seemingly ambitious idea arises from our ultimate goal of 3D model generation (see section 1.2). Thus, if we can generate representative 3D models from devices, and we can do it quickly, the next step is to perform this process live in *in-situ* formats. Whilst efficiency or algorithmic optimisation are welcome, they are not the primary driving factors; our goal is to achieve convergent 3D modelling. Therefore, we will combine the cornerstones of the knowledge developed so far in ML for imaging, as reviewed in this section, to achieve this objective.

1.3.2 Most important advances in electron spectroscopy

Beyond imaging, spectroscopy is also rapidly advancing in its application of ML to address its challenges. While some aspects of spectroscopy will contribute to assembling the workflow that aligns with our primary goal, the comprehensive review of the state-of-the-art ML techniques in spectroscopy will be particularly valuable

for addressing our third objective (see section 1.2). In fact, traditional processing of spectroscopic data is involved in Chapter 3, where it characterises the target devices and establishes the foundation for what needs to be achieved in Chapter 4. However, this process does not entail complex ML processes to retrieve the data we seek. Instead, it relies on classical data analysis routines. Subsequently, this acquired knowledge will be leveraged in Chapter 5, where we address our third major goal and where ML will play a pivotal role in handling spectroscopy data.

Even though we have seen strategies for dealing with high-dimensionality data such as 4D-STEM related techniques, the vast majority of current literature confronts 2D signals (mainly images). This emphasis on 2D data is largely due to the rapid evolution of CV in fields like autonomous driving and macroscopic pattern recognition, where information is primarily conveyed to computers through digital imaging. Consequently, handling more intricate data types, such as those generated by electron spectroscopy, poses a greater challenge compared to the analysis of electron micrographs. Therefore, the next subsections delve into the comparison between advancements in spectroscopy and imaging techniques, exploring situations where the methods employed may overlap in terms of convenience and applicability. Currently, electron spectroscopy has primarily been approached with simpler ML models, although the trend is ascending and aiming to bridge its gap with imaging modes.

1.3.2.1 Unsupervised machine learning

Experimental electron spectroscopy setups often result in either 1D or 3D datasets. While a 1D spectrum can contain invaluable information, leveraging the high spatial resolution of STEM typically leads to the acquisition of 3D datasets. In fact, the primary motivation for choosing electron spectroscopy lies in harnessing its exceptional spatial resolution, which invariably exposes microscopists to high-dimensional data challenges. As a result, creating synthetic high-dimensional data for training supervised models capable of handling such complex data structures poses significant challenges. This task necessitates gathering a large number of similar experiments or resorting to simulation techniques. Fortunately, there exist diverse simulation methods in the literature for generating reliable spectroscopic data. However, these simulation routines are often time-consuming and resource-intensive, making it impractical to generate a sufficiently representative training dataset. [37, 197, 198, 199, 200] Moreover, relying solely on generated data may lead to the development of overly specific models that struggle to generalize beyond the training set. This limitation undermines the justification for investing substantial time and resources in generating training data. Consequently, unsupervised methods have emerged as the preferred approach for handling spectroscopic data due to their efficiency and ability to extract meaningful insights without the need for extensive labelled datasets.

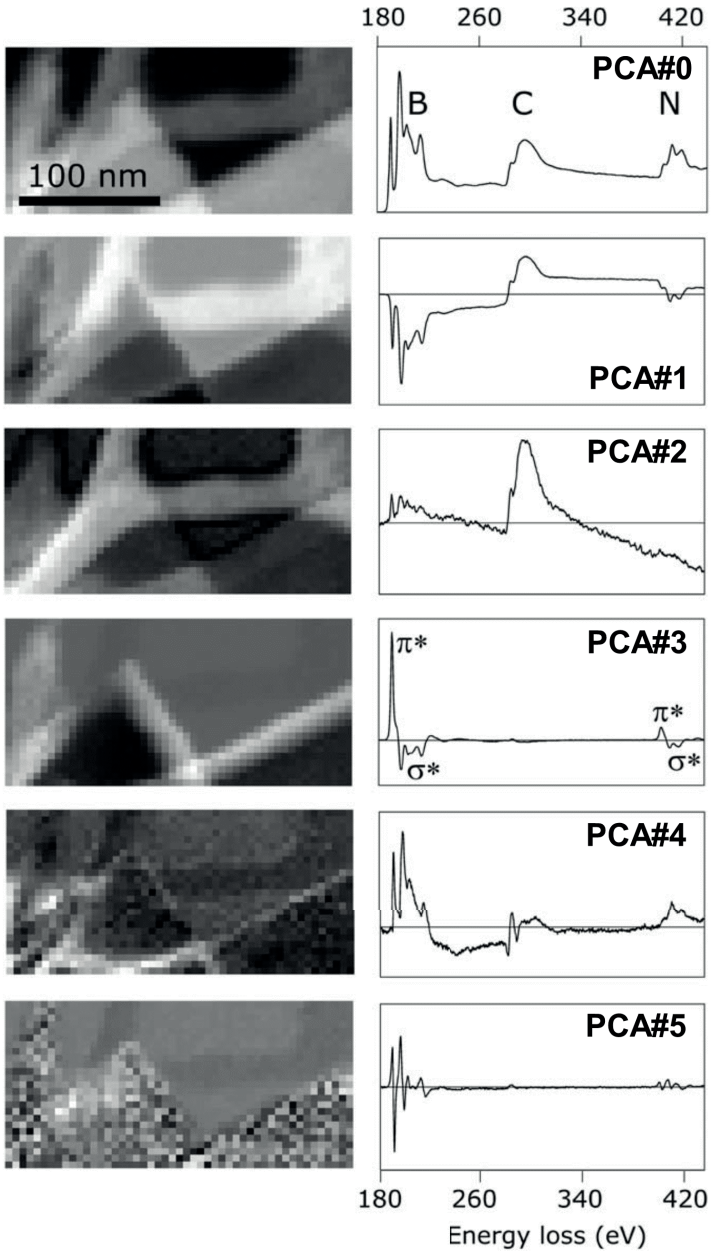


FIGURE 1.6: Unsupervised unmixing methods for hyperspectral analysis: Principal Component Analysis (PCA) of a BN Electron Energy Loss Spectroscopy (EELS) Spectrum Image (SI) and the first six components, from top to bottom, with higher statistical significance. The presence of peaks and features of interest diminishes as going down in significance, where background or noise is the main feature. PCA does not physically constrain the components, as regarded by the negative losses. [48, 51]

In that sense, the evolution of electron spectroscopy mirrors the advancements seen in imaging, where traditional postprocessing techniques have been surpassed by ML methods. This is partly because postprocessing plays a crucial role in spectroscopy, whether it involves enhancing energy resolution or boosting SNR. For instance, these enhancements proved invaluable in emulating the effects of monochromators without the costs and technical challenges of the hardware upgrade. [11, 201,

202, 203, 204] On the other hand, they proved successful in boosting SNR with the non-rigid averaged cross-correlation of single atomic columns applied to Energy-Dispersive X-ray spectroscopy (EDX) and EELS. [205, 206, 207, 208] Additionally, denoising methods utilizing comparisons between experimental data and simulations have been employed, too. [209, 210] However, none of these traditional approaches have matched the straightforward effectiveness of modern post-acquisition operations based on Singular Value Decomposition (SVD) in applications like SNR enhancement.

As mentioned in the very beginning, the initial intersection of ML with electron microscopy, and particularly spectroscopy, primarily occurred in the realm of denoising spectra through unsupervised unmixing methods. This approach has been extensively explored in the literature, showcasing the advantages, limitations, associated challenges, and practical application of PCA and related unmixing techniques to spectroscopic data. [211, 212, 213, 214, 215, 216]

Unmixing aims to simplify spectral features by representing them in terms of key physical properties like material composition or spatial distribution. Many unsupervised unmixing methods are derived from or inspired by SVD (fig. 1.6). [217, 218, 219] This algorithm decomposes data into weighted components, and enables noise removal and extraction of relevant information while preserving spectral fine structures in applications such as cleaning atomic resolution EDX and EELS maps. [220, 221] However, careful consideration is crucial when selecting components for reconstruction to avoid distorting the data. Pragmatically, P. Potapov and A. Lubk published a guide on automatically identifying meaningful PCA components. Their approach involves smoothing the data, assessing the anisotropy of the scree plot and its components, and selecting only those components that do not exhibit noise-like isotropy for reconstruction. [222] Additionally, studies have delved into the statistical biases as a function of the truncated portion of the scree plot, total spatial pixels, and energy channels, aiding in the reproducible implementation of these methods. [217, 218]

PCA, while powerful, can lack physical interpretability in its resulting components. To address this limitation, physically constrained alternatives to PCA have gained popularity and demonstrated enhanced capabilities. The first steps started with NMF, which enforces positivity in components, aligning well with the non-negative nature of measured energies in electron spectroscopy; Independent Component Analysis (ICA), which extracts statistically independent features, assuming that spectral peaks for each element are uncorrelated; with Gaussian Mixture Modeling (GMM), which contrarily assumes that all sources follow Gaussian distributions and separates signals into a finite combination of these distributions; or with Vertex Component Analysis (VCA), which assumes the presence of pure spectral signals from materials and evaluates their distribution within hyperspectral datasets. Mimicking this winning strategy, the incorporation of further physical constraints to ML

is the logical (r)evolution to follow in the coming years, both in the unsupervised and supervised domains. Therefore, physically constrained ML algorithms must certainly intervene to maximise the likelihood of a successful completion of our research objectives. [223, 224]

The literature on unmixing methods for electron spectroscopy is extensive, with numerous articles comparing various methods to determine the best performer for specific applications or purposes. [225, 226, 227, 228, 229, 230, 231] Interestingly, NMF emerged as the preferred approach when the primary objective extended beyond denoising or data inspection. Thus, NMF stands out for its ability to capture the direct physical significance of extracted components and easily accommodate additional constraints. Researchers like M. Shiga *et al.* enhanced NMF by introducing extra physical constraints that better align with the statistical characteristics of EDX and EELS data. These constraints include automatic prior-based component relevance and soft orthogonality determination, enabling a more accurate representation of the underlying physics in spectroscopic datasets. Moreover, these constraints can be customized to suit more complex setups involving simultaneously acquired multimodal spectroscopy, such as multidetector imaging, EELS, EDX, and cathodoluminescence. [232, 233]

Following multi-modality, not only in the technique terms but in the data analysis tools, PCA or SVD actively complemented constrained unmixing methods. ICA has found application in EELS following the pioneering work of N. Bonnet and D. Nuzillard, who applied ICA to simulated EELS data. In practice, ICA is often used in conjunction with PCA rather than as a replacement. Remarkably, Bonnet and Nuzillard highlighted that ICA serves as a complementary technique to standard unmixing methods, given that EELS data does not meet the condition of statistical independence. [234] Interestingly, the synergy between PCA and ICA allows for capitalising on the strengths of each method: PCA helps to prevent overfitting, while ICA can be utilized for spatially mapping crystal phases or elements. [235] Building on this approach, D. Rossouw *et al.* developed a compositional quantification routine based on dual-PCA+ICA and dual-EDX+EELS analyses. They demonstrated the effectiveness of their method on FePt@Fe_3O_4 core-shell nanoparticles. Summarising, they solved the quantification of the spatially overlapping core and shell phases with the complementary spectral unmixing, and the mapping of both light and heavy atoms with the complementary spectroscopy (fig. 1.7, a). [236]

The combined use of PCA/ICA has proven effective in directly distinguishing phases that would typically require laborious least-squares fittings for replication. However, the assumption of independence between spectral peaks in ICA can lead to artefacts when dealing with similar phases that share elements or oxidation states. In such cases, Vertex Component Analysis (VCA) offers a distinct advantage by successfully identifying individual spectral profiles (i.e., combinations of peaks and features) associated with distinguishable phases. Notably, VCA excels

at isolating spectral profiles of physically or chemically independent structures, making it ideal for scenarios where the expected components already are present in isolated hyperspectral pixels. For instance, a comparative study conducted by N. Dobigeon and N. Brun, which confronted PCA, ICA, least-square fitting, and VCA against each other, highlighted the superior interpretability of VCA for analyzing high-complexity sample configurations (fig. 1.7, b). [237] By requiring gentler statistical assumptions, VCA outperformed both PCA and ICA in providing finer analysis of complex sample configurations.

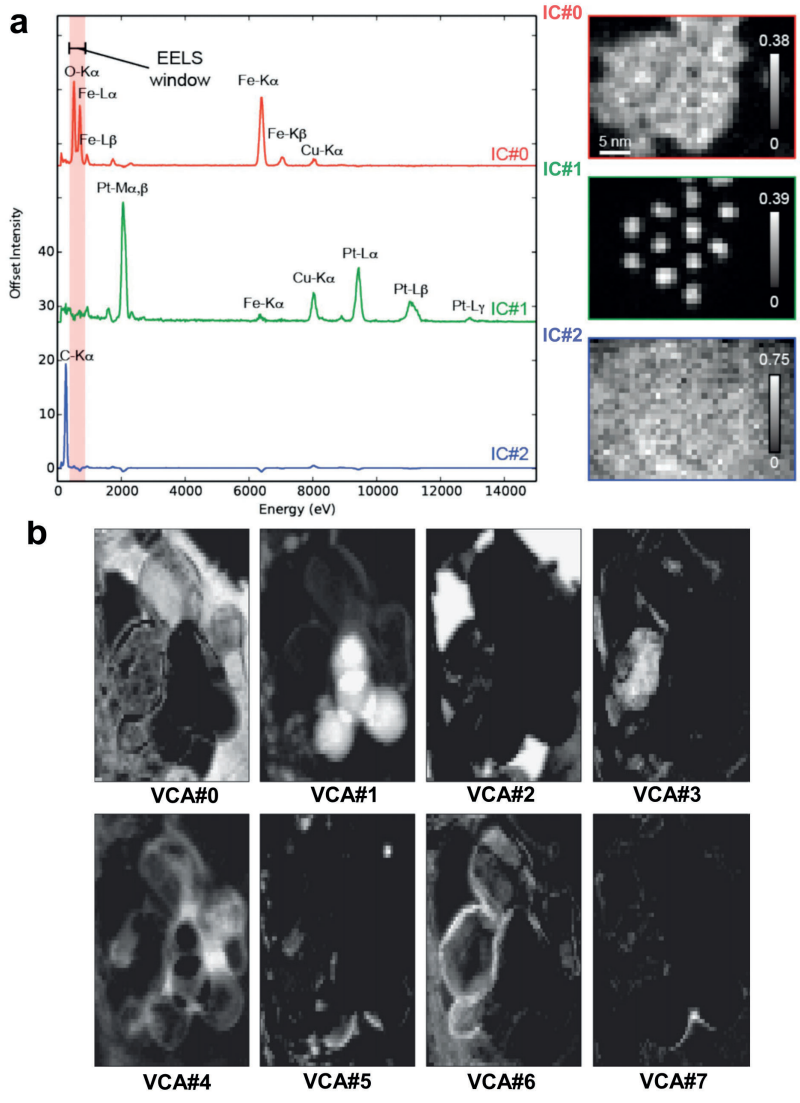


FIGURE 1.7: Unsupervised unmixing methods for hyperspectral analysis: (a) Independent Components Analysis (ICA) of an Energy Dispersive X-Ray (EDX) SI of Fe-based core-shell nanoparticles with the three main independent components. The independence of the components mapped the cores (IC#1), the shells (IC#0), and the C background (IC#2). [236] (b) Vertex Component Analysis (VCA) of an Electron Energy Loss Spectroscopy (EELS) SI of BN/BOx nanoparticles, displaying the eight more important components. As the components were from the original SI, VCA was capable of mapping physically meaningful variations for each component such as particle edges vs particle centres. [48, 237]

This enhanced interpretability allows for the extraction of spatial information through the mapping of spectral components derived solely from the energy domain. Consequently, spatial correlations may be overlooked when focusing solely on spectral similarities without considering the probe location. As a response, and to further account for spatial correlations within the spectrum image, S. Kalinin and colleagues introduced an innovative unmixing method based on spatial Gaussian kernels. These kernels were based on traditionally unmixed components that reduced the complexity of the energy domain and loosened the computational needs of the GP. The methodology can be understood as the unmixing of spectral features across the spatial dimension using these kernels, resulting in enhanced reconstructions and mapping with higher fidelity. [238]

Indeed, one key aspect highlighted by the authors is the flexibility in tuning and adapting these spatial Gaussian kernels to combinations of unmixed components or tailoring them based on specific physical characteristics desired from the hyperspectral dataset. In addition, this adaptability allows for precise engineering of the kernels to capture spatial correlations effectively, making this method particularly suitable for handling highly spatially correlated data such as core-shell nanoparticles, nanoparticle clusters, compositional gradients and others. Importantly, systems like the ones we propose throughout the thesis exhibit high spatial correlation. As a result, spatial information must be handled with care to ultimately achieve the spatial mapping and correlation of physical properties such as the band gap.

As mentioned, unsupervised routines in spectroscopy often rely on unmixing methods for various purposes such as spectra denoising, cleaning, mapping crystal phases, quantifying stoichiometries, and spectral classification. However, the analysis of finer features like Energy Loss Near Edge Structure (ELNES) demands dedicated computing solutions. In addressing these complex tasks, unsupervised clustering methods have emerged as a valuable strategy. [239, 240, 241] Thus, spectral clustering aims to automatically detect subtle spectral variations and group them accordingly, facilitating the identification of distinct spectral features without the need for manual intervention. In spectral clustering, the key aspect involves tuning a hyperparameter related to the desired number of final groups or clusters. It can be, for example, the total number of different crystal phases to expect in a spectrum image. Alternatively, clusters can be set to correlate to distinctive ELNES oxygen K edges, providing a foundation for how more advanced models can follow by progressively considering finer spectral features. For example, S. Kiyohara *et al.* utilized spectral clustering to generate clusters representing ELNES oxygen K edges. [242, 243] Afterwards, they constructed a supervised decision tree which allowed them to, on the one hand, successfully uncover the oxidation nature of oxides based on their spectral characteristics, and on the other hand, to predict the spectral signal of known oxide-based nanostructures.

Interestingly, clustering and unmixing techniques share the common goal of identifying patterns within spectra to enhance material identification in electron spectroscopy applications. As a result, it is welcome to combine both approaches to further refine the process of material identification. This combination can involve applying PCA before clustering to improve the identification of materials, or using PCA to segment individual clusters for more detailed analysis. [244] Moreover, integrating clustering and unmixing with non-linear least squares fitting enables a deeper analysis of ELNES data. This comprehensive approach not only classifies the data but also reveals finer ELNES features such as valence, oxidation states, and coordination states. Indeed, software solutions like WhatEELS offer these capabilities online, providing users with a range of practical tools for advanced spectroscopic analysis. [245] Furthermore, clustering and unmixing techniques are not limited to core losses but can also be applied to low-loss EELS. For example, they can be used to map plasmons or potentially, and thanks to the research work conducted in this thesis, concretely in chapter 5, to accurately map the electronic properties of nanomaterials, such as the energy and type of band gap. [246]

As reviewed, the application of ML in electron spectroscopy has primarily focused on low-complexity methods, indicating that basic approaches are often sufficient for many applications. However, as the field expands and reaches broader audiences, more sophisticated ML routines are emerging and will continue to evolve following the same trend observed in imaging. In that sense, S. Kalinin *et al.* innovatively introduced techniques such as autoencoders from imaging to spectroscopy. [98, 99, 100, 103] In this notable case, 1D convolutional autoencoders were employed to unsupervisedly represent each spectral pixel using only two latent variables. Subsequently, these latent variables were grouped using GMM to distil spectrally distinguishable regions within a heterogeneous array of nanoparticles (e.g., such as fluorine and tin-doped indium oxide structures) (fig. 1.8, c and d). [247, 248] This research exemplifies how autoencoders can revolutionize spectral analysis, potentially paving the way for more complex autoencoder-based routines, soft supervision methods, and the utilization of GAN architectures. [249]

Importantly, this work from S. Kalinin *et al.* highlighted how different unsupervised (i.e., the autoencoder) and semi-supervised approaches can converge into the same spectral analysis. This supervised approach is detailed in the upcoming subsection, together with the pros and cons of using supervised routines for spectroscopy, which have been realised so far in the literature.

1.3.2.2 Supervised machine learning and generation of training sets

As in imaging, the preference for unsupervised ML over supervised in electron spectroscopy data analysis is largely due to the challenging task of generating training data. To make supervised solutions more accessible in spectroscopy, there is a need for a simpler and more resource-efficient method of generating synthetic labelled

data. This labelling typically involves associating full spectra or individual energy channels with specific quantities or properties. This labelling process often requires manual annotation by trained microscopists or the simulation of spectra based on known compositions. Relying on manual labelling is not ideal, and solutions like mechanical Turks are not elegant or efficient for long-term use. Moreover, current EELS databases still fall short in providing a comprehensive solution to the labelling challenge, which varies and should be tailored according to the application of interest. [250, 251]

Fortunately, the literature offers various theoretical frameworks that can aid in generating reliable simulated spectra. [11, 37, 252] The key advantage of using simulated data lies in its ability to generate paired data with labels, aligning with the requirements of supervised models. Balancing the accuracy and efficiency of these simulations is then crucial for effectively planning and implementing supervised ML approaches. The key is ensuring that the generated synthetic data accurately reflects real-world scenarios and can be used effectively for training predictive models.

The most time-effective approximations are based on classical electrodynamics and the dielectric properties of the materials, whose accuracy can be improved by adding quantum corrections (also applicable to Monte Carlo). [253, 254, 255, 256, 257, 258, 259] These quantum calculations can be semi-empirical, utilizing experimental scattering cross-sections, *ab initio* methods, or even estimated with ML itself. [260, 261, 262, 263] For high-fidelity simulations, multiple scattering techniques like the FEFF code or *ab initio* computations of sample potentials in the multislice approximation are commonly employed. [197, 198, 199, 209, 210, 264] Additionally, Bloch wave formulations in thin specimens offer another approach for accurate simulations. [265] Alternatively, to reduce computational costs, traditional pixel-wise simulations of probe displacement can be modelled using scattering matrices. EDX simulations can also be computed if required, which drift from EELS by tailoring the potential for this interaction, modelling the bremsstrahlung background, and considering specific detector geometries. [255, 266] Tools like DTSA-II from NIST provide open-access resources for simulating EDX spectra and analyzing experimental data. [267] The present thesis, however, focuses on EELS as the spectroscopy source. Specifically, in Chapter 5, we utilise some of the simulation tools described in this paragraph. Interestingly, their use will extend beyond ML model training; instead, they will be creatively employed to understand the physical origin of the experimental signal.

Still, simulation solutions can be computationally demanding, making them challenging to use as input for high-dimensional CNNs. However, ML has introduced innovative alternatives to traditional electron microscopy simulations, such as utilizing neural networks for potential propagation through slices via forward propagation. [141] This creative approach involves modelling each slice in a multislice simulation with a layer of the neural network, effectively using the propagated weights as the simulated potential. Importantly, this ML-based acceleration, combined with

efficient simulation techniques, can be sufficient for generating ample data for 1D analyses. Unfortunately, spatial correlations captured in spectrum images would be majorly lost, although the community has already devised ways to circumvent that. [238] For instance, an unsupervised approach proposed by S. Kalinin *et al.* offered a promising solution to bridge the gap between simulated 1D data and the spatial correlations present in spectrum images. [158, 247, 268] Therefore, this approach would dramatically reduce the costs of data engineering and model training, while also achieving the expected spatial correlation typical of 3D CNNs. [238] Looking ahead, the use of generative models like GANs or other complex architectures such as graph NNs is expected to alleviate this pressure, as has been recently demonstrated with success. [269]

As reviewed, the cumbersome generation of training datasets and the adaptability of unsupervised ML methods, have hitherto left supervised approaches on a second plane. However, novel supervised approaches have emerged to address more complex challenges and to simultaneously defy equivalent unsupervised methods. [227, 243] An intriguing example is the work by S. Kalinin *et al.* on autoencoders for categorizing low-loss EELS. [247] In their study, the authors introduced a MLP as a supervised model for mapping spectrally distinct regions in spectrum images. (fig. 1.8, a and b) They achieved results comparable to the autoencoder approach reviewed before (fig. 1.8, c and d). The supervised model was designed to leverage the partial labelling of some pixels in the spectral image, to base its training on these labelled pixels and to infer classifications for the remaining unlabeled ones. By assigning classes (i.e., element type and chemical surroundings) to a fraction of representative pixels, the authors assessed the percentage of labelled pixels required for the accurate supervised classification of the unlabeled ones. Surprisingly, they found that labelling just 0.31% of representative pixels led to results comparable to those achieved by the unsupervised autoencoder. Interestingly, this approach significantly streamlined the training process for supervised models in spectroscopy, providing a practical strategy for microscopists who are considering investing resources in developing such models. By selectively labelling a key fraction of data, researchers can gauge how well a supervised model may generalize without committing extensive resources upfront.

Furthermore, supervised approaches have been utilized in the realm of spectroscopy for various tasks, including automatic mapping and elemental identification, ELNES characterization, and spectral calibration. [268, 270, 271] While these methods tackled more challenging scenarios, their fundamental workflow mirrored that of unsupervised methods. Importantly, the introduction of the additional energy dimension in spectroscopy presents an incommensurable opportunity to actively incorporate physical constraints into ML algorithms.

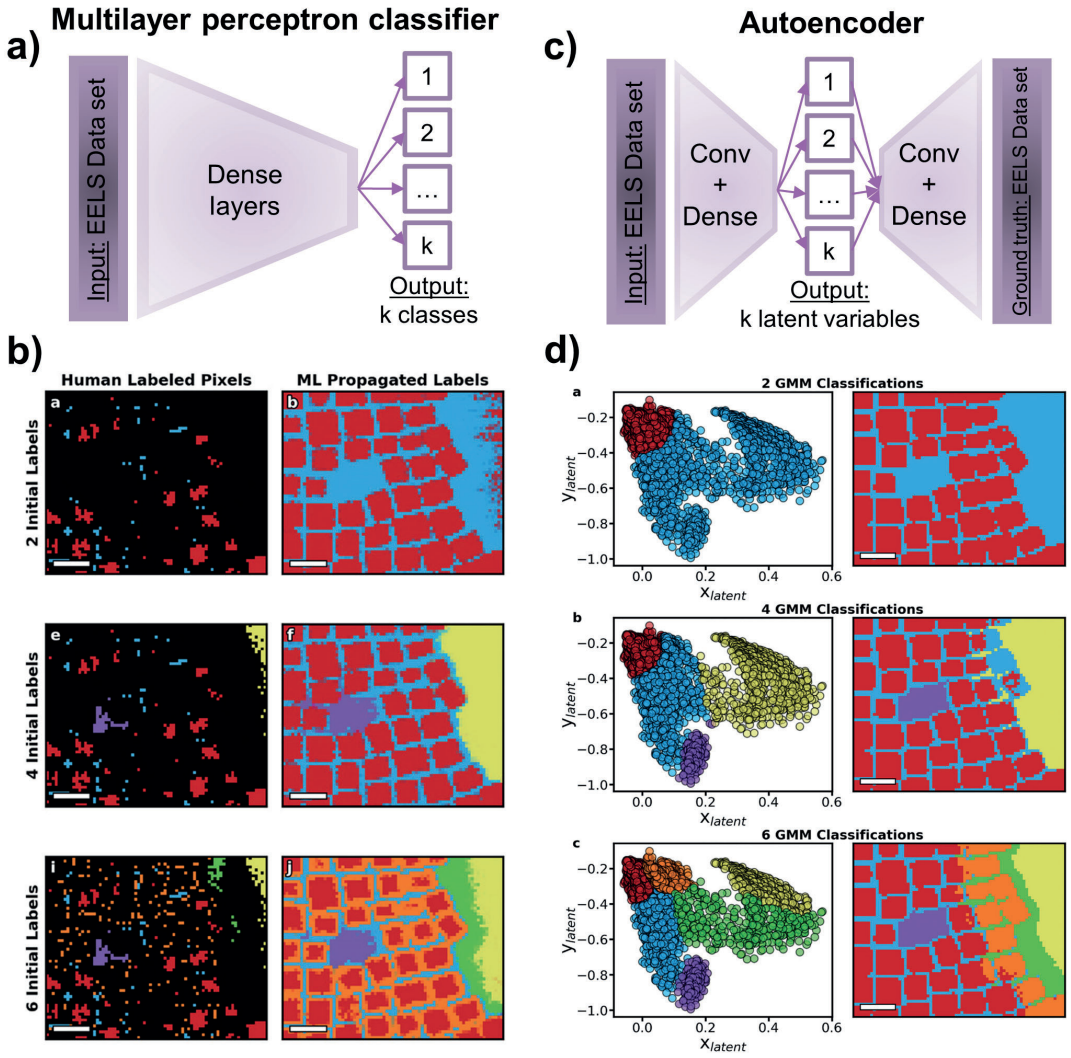


FIGURE 1.8: (a)-(d) Representations of S. Kalinin *et al.*'s work, comparing two learning paradigms to unveil physically-distinguishable Electron Energy-Loss Spectroscopy (EELS) components. [247] The authors suggested: (a) a supervised approach based on a multilayer perceptron capable of (b) extrapolating the partially manually labelled pixels of (b.a,b) two, (b.e,f) four and (b.i,j) six physically-different classes. On the other hand, (c) the authors proposed an unsupervised approach based on an autoencoder. (d) The autoencoder reduced the EELS dataset into two latent variables, which after being Gaussian clustered in (d.a) two, (d.b) four and (d.c) six classes, produced comparable maps to the supervised approach. [48]

Similarly to what NMF posed to PCA, physics-aware ML, is emerging as a promising approach in scientific ML. The idea is that ML models are directly informed by the underlying physical principles they aim to represent. This methodology aims to integrate scientific knowledge into ML models to enhance their interpretability and accuracy. One notable example is the work by L. Roest *et al.*, who developed a physics-aware model for the Zero-Loss Peak (ZLP) using a neural network trained on Monte Carlo simulations. [272, 273] Their model is sensitive to the acquisition conditions, namely the exposure time and the beam energy (fig. 1.9, a and b). [274]

This means that the input data the model receives is the ZLP curve with the exposure time and beam energy, making these variables constitute individual dimensions within the parametric space of the model.

Typically, the science awareness arises from directly inputting the physical conditions into the model, as L. Roest *et al.* did. Nevertheless, similar effects can be achieved with physically-aware labels. For instance, one could use start from a handful of unit cells, jointly simulate spectra and physical properties of interest, and train a model where the input is the spectra and the label (and the model's output) is the simulated physical property. This is what S. Kiyohara *et al.* did with a MLP on core-loss spectra for the extraction of up to six properties of interest (fig. 1.3, MLP). [107] Specifically, the training was performed on 1171 spectra obtained from DFT simulations of 188 different silicon oxides. The labels were also simulated by DFT. Therefore, they eventually engineered a ML-based DFT simulator from experimental EELS spectra. Eventually, from the input spectra, they were capable of extracting their target physical properties which they simulated to train: properties such as average bond length and angle, Voronoi volume, bond overlap population, Mulliken charge and excitation energy.

By now, the physics-aware approaches are essentially based on simulations as the best way to directly link spectra or images with a specific physical property. However, transitioning these approaches to physically-aware experimental datasets necessitates a foundation of reproducible and systematic metadata acquisition alongside the experiments. In microscopy, capturing detailed acquisition conditions and microscope parameters provides a rich starting point for implementing physics-aware ML on experimental data. In that sense, the works of L. Roest *et al.* and S. Kiyohara *et al.* exemplify the potential of integrating scientific constraints into ML models in electron microscopy and spectroscopy research. [107, 274]

This small subset of physics-guided ML models showcases the benefits of incorporating domain knowledge to enhance model generalization and reduce training data requirements. As a result, this concept must be central to building our ML-based solutions, both for the second and third proposed goals. Therefore, metadata handling will be specially addressed to track it throughout the entire developed workflow. In fact, we can assert that the complexity of the global workflow addressing objective number two, and the EELS simulations and data analysis supporting objective number three, would have been unreachable without this approach. Moreover, just as L. Roest *et al.*'s work was inspired by solutions in particle physics, our approach to engineering physically interconnected building blocks into a larger thinking unit also aims to inspire other scientific communities embracing ML.

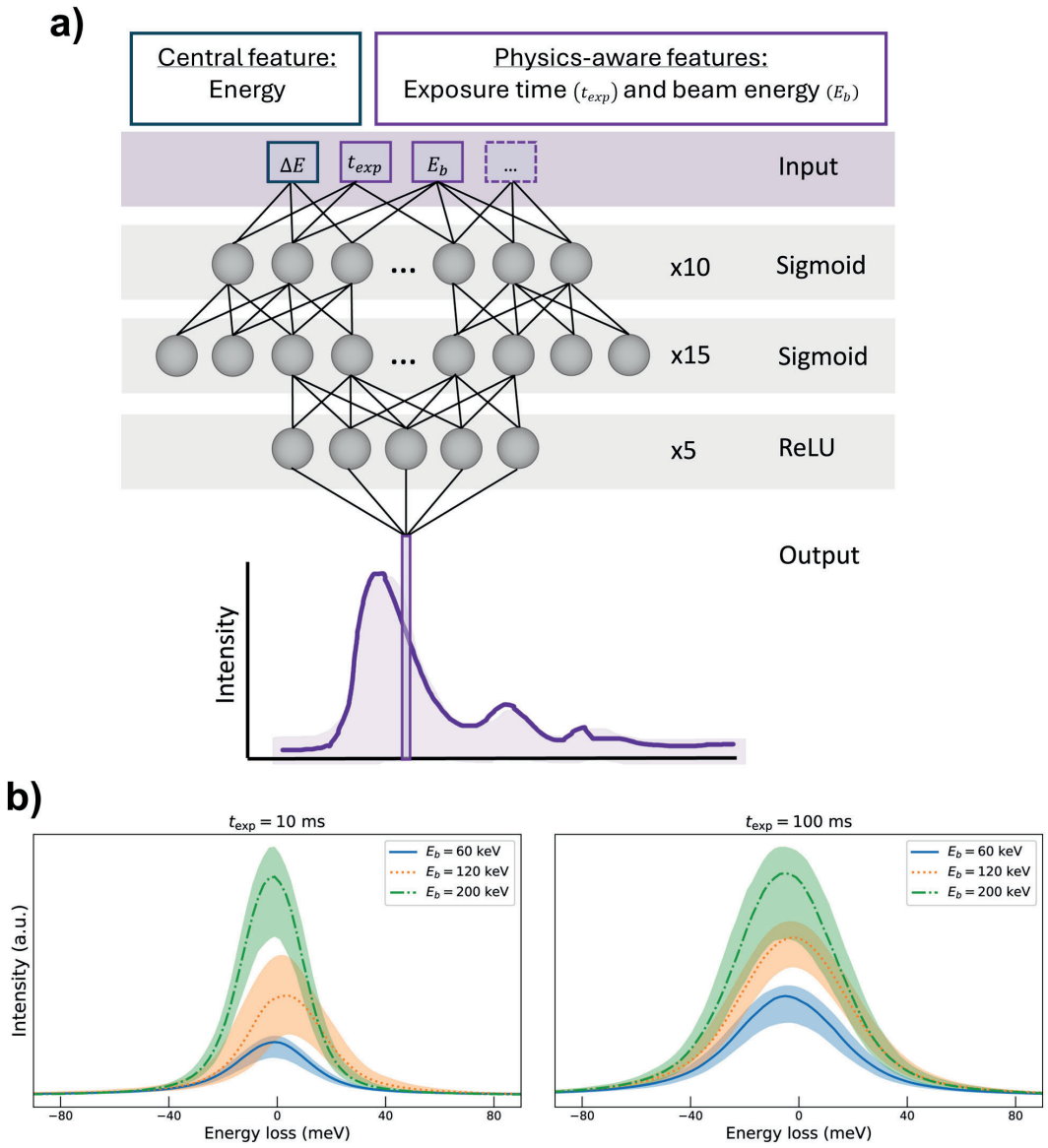


FIGURE 1.9: (a) L. Roest *et al.*'s work, represents the training of a physically-aware multilayer perceptron capable of (b) modelling a zero-loss peak and its uncertainty under different exposure times and beam energies. [48, 274]

1.3.2.3 A note on compressed sensing and spectroscopic tomography

Tomography, by itself, is outside the scope of the present work, as mentioned when evaluating its current status regarding ML. However, since our ultimate goal of obtaining 3D models is shared, we want to highlight an important aspect that shapes our derived workflow. Specifically, while tomography provides the missing structural information in 3D, tomography for spectroscopy completes it in compositional terms. In this context, we want to emphasise the significant role that CS plays and how ML has been utilised to replicate its functionalities. [275, 276]

The key lies in joint imaging plus spectroscopy tomography, which, through multimodal approaches, manages to thoroughly reconstruct structures. [182, 183, 277]

EDX, EELS and Energy-Filtered TEM (EFTEM) tomography successfully combined with imaging mode to improve the reconstruction of nanostructures. [278] For instance, STEM-HAADF and EDX were complemented through simultaneous regularisation to achieve unreachable levels of detail with separate techniques. R. Huber *et al.* explored this method for the 3D reconstruction of nanostructures with both sharp edges and gradual compositional changes. [279] Additionally, unmixing methods (e.g., PCA, ICA and VCA) can be used as a pre-reconstruction step to prepare the data for posterior CS reconstruction. [280] This initial unmixing step reduces the computational cost and complexity of reconstruction by relaxing the prior assumptions, benefiting both EDX and EELS data, especially under challenging sample configurations (fig. 1.3, VCA). [106, 281, 282]

Although there are interesting post-processing approaches, such as DL-based denoising of tomograms, we wanted to emphasize the importance of combining techniques. This synergy is crucial for allowing ML to excel in final 3D reconstruction. [283] In the end, our workflow must incorporate spectroscopic knowledge, even though our primary objective is the analysis of individual micrographs. Ambitiously, our dream goal is to merge the outcomes of the second and third objectives as outlined in (section 1.2), for which this multi-modality approach should always be at the forefront of our minds.

1.3.3 Machine learning deployment in the electron microscopy community

While the specific tools within ML, DL and AI that are employed in this thesis dissertation can be found in the methodology section and within the chapters, our focus lies on describing the underlying philosophy behind their use. This philosophy aligns with the guidance provided by the research community thus far.

The key fact is that this thesis dissertation is not on AI or ML. It is a work devoted to answering materials science questions, which are described in section 1.2. In addition, these are tackled through a major characterisation technique which is TEM, as we justify in Chapter 2. As a result, the primary recipient of knowledge must be materials science and its focus through TEM. Therefore, AI is just a tool or framework we use to answer these materials science questions, for which the direct or on-purpose development of knowledge in AI is not intended.

Remarkably, the AI community is consciously committed to open access and open source, aiming to democratize its use across various knowledge domains. Not surprisingly, the approach resonated in the materials science and electron microscopy communities to develop the toolset we have reviewed so far. Importantly, access to code repositories and databases is commonly provided to facilitate the reproducibility of research findings and enable the creation of new studies.

Consequently, the philosophy underlying AI intervention in this thesis revolves around leveraging pre-built models, packages, and libraries to establish a solid data science foundation to be modified only if it is fully required. We do not consider these tools as black boxes, as understanding their theoretical underpinnings is essential for assessing the strengths of different model types and architectural choices, but more importantly the interpretation of the results and their fine-tuning to specific needs. Thus, if existing ML/DL models demonstrate effectiveness for a task, resources are not going to be allocated to evaluating AI-domain matters like architecture dependencies versus performance. Fortunately, we can assume that this has been addressed already by experts doing focused research on these issues. As a result, we smartly adapt from the retrieved major AI models to stay focused on the primary objective without getting sidetracked.

Remarkably, the process of solving materials science problems without developing any code or creating datasets is indeed possible. For example, we could easily create a simple crystal phase identifier just with resources from literature and repositories. Understanding the building process of such applications is crucial, as they typically follow a three-step approach that needs to be tailored to specific needs. Let us review how it can be done by following along the engineering of this hypothetical application.

In this hypothetical application, open-access data from literature and databases would be utilized. The first step would involve preparing a labelled training set, which can be facilitated by databases like TEM ImageNet (i.e., simulated STEM micrographs). [284] Thus, engineering the labels within the training set is essential to ensure they align with the model's design requirements, focusing on the information to be retrieved. In our hypothetical application, we can easily correlate the images of the training set with the crystal phase and zone axis defined for the simulation. Consequently, our label is a category, where each crystal phase type corresponds to an integer.

The subsequent step, which indeed must be done parallel to the first one, involves designing the model based on the available information and desired output, which in this case is identifying the crystal phase from the images. At this point, extensive literature and repository research are welcome to ensure we filter the most valuable architecture type for the given query. Finally, through iterative training, testing, and fine-tuning, the model aims to achieve the desired performance. In the hypothetical model we suggest, the NN should output the integer representing the category of the crystal phase from which the micrograph was taken. For our convenience, we would utilise a VGG-like architecture, which receives an image as input and outputs integers representing classes. [285] Ultimately, the goal of these models is to generalize to experimental datasets of a similar nature, as they represent the actual targets. To assess the model's behaviour, various metrics can be calculated, providing insights into its overall performance and effectiveness.

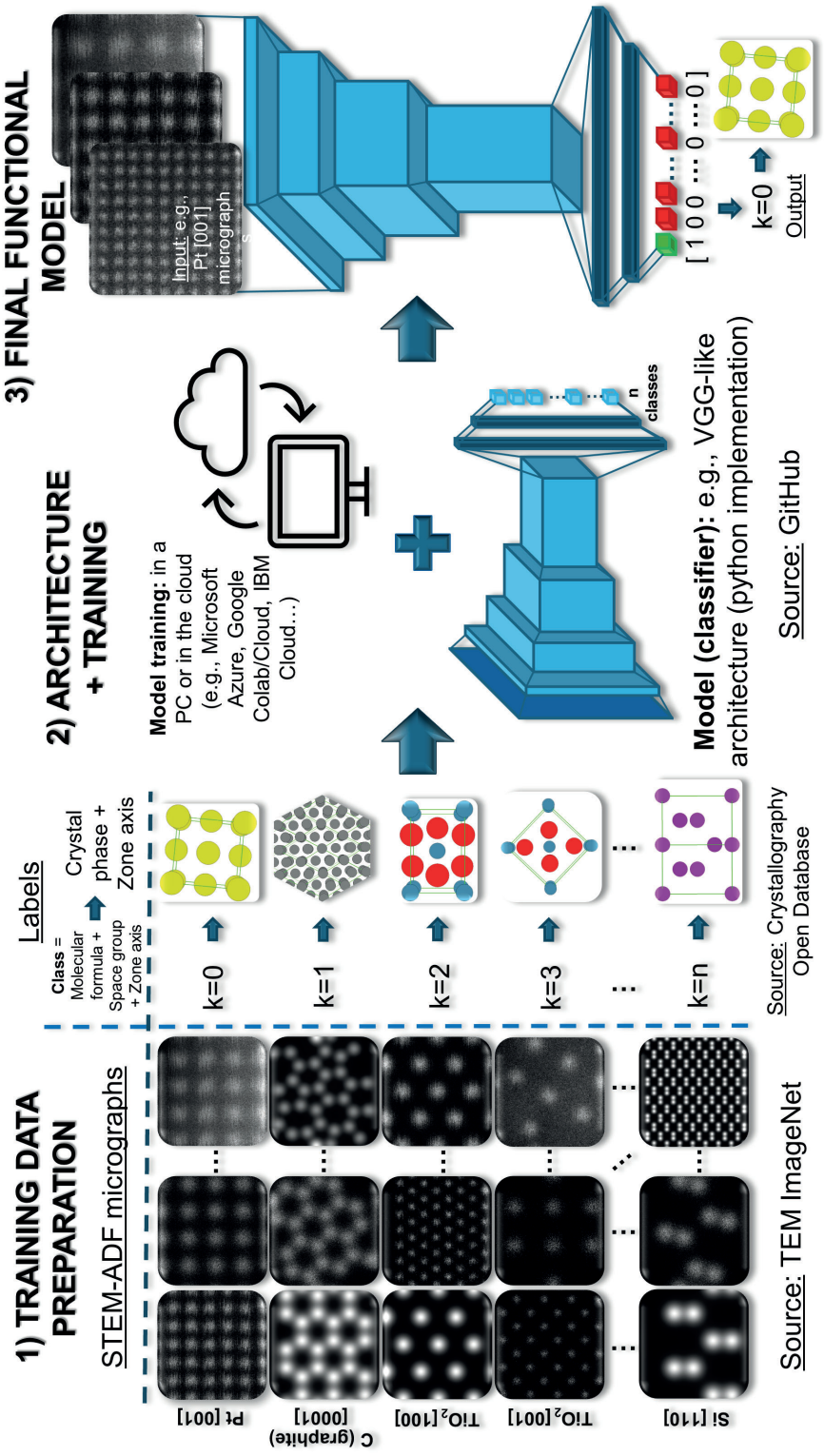


FIGURE 1.10: Example of the hypothetical development of a deep Convolutional Neural Network (CNN) model to correlate Annular Dark-Field Scanning Transmission Electron Microscopy (ADF-STEM) images with the crystal phase and zone axis displayed. The Machine Learning (ML) workflow is thoroughly based on open software and resources, and requires almost zero ML-specific knowledge and funding/infrastructure. The ADF-STEM micrographs are taken from the TEM ImageNet open database, with 14.364 images of multiple crystal structures. [284] The image metadata correlates them with a class defined by the crystal phase and the zone axis (labels). The complete crystallographic information of the crystal phases is extracted from the free Crystallography Open Database [286, 287]. Next, a VGG-like architecture is used from an open-source repository like GitHub, almost eliminating the coding. [285, 288, 289, 290] The training can be done either in a PC or in a cloud service such as Microsoft Azure [291], Google Cloud [292]/Colab [293], IBM Cloud [294], among others. Eventually, the model can be tested and used, allowing the user to identify the crystal phases and orientations of different micrographs. [48]

In summary, this is the general workflow to deploy (supervised) ML models, which adapts well to materials science. On the other hand, the unsupervised ones are much simpler to apply as the tedious infrastructure associated with training can be skipped. Unsupervised models just require smart and case-dependent hyperparameter fine-tuning, which is where the practitioner's creativity lies. In any case, the take-home message is the philosophy of using the resources we find that already perform accurately for our field, and tailoring them if needed. Thus, we will not dive into software or architecture-specific investigations unless it is mandatory to respond to a particular need. Nevertheless, this does not mean that the custom code development is negligible, as remarkably, the total code deployed exceeds 70.000 lines of code. However, every line addresses a so far disregarded question which demanded code from scratch. Particularly, this is well-exemplified by the code that weaves already available packages which provide a solid basis upon which to customise. In fact, these packages are not necessarily AI-related but appropriate for certain steps within the workflow that must answer the third objective, but mainly the second and central objective. The programming environment together with the packages we exploit is discussed in the methodology chapter 2 and throughout the chapters involving methodological development (i.e., chapters 4 and 5).

References

[1] Dawei Yang et al. "NbSe2 Meets C2N: A 2D-2D Heterostructure Catalysts as Multifunctional Polysulfide Mediator in Ultra-Long-Life Lithium–Sulfur Batteries". In: *Advanced Energy Materials* 11.36 (2021), p. 2101250. DOI: https://doi.org/10.1002/aenm.202101250. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aenm.202101250. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/aenm.202101250.

[2] Gordon E. Moore. "Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff." In: *IEEE Solid-State Circuits Society Newsletter* 11.3 (2006), pp. 33–35. DOI: 10.1109/N-SSC.2006.4785860.

[3] David Burg and Jesse H. Ausubel. "Moore's Law revisited through Intel chip density". In: *PLOS ONE* 16.8 (Aug. 2021), pp. 1–18. DOI: 10.1371/journal.pone.0256245. URL: https://doi.org/10.1371/journal.pone.0256245.

[4] John Shalf. "The future of computing beyond Moore's Law". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 378.2166 (2020), p. 20190061. DOI: 10.1098/rsta.2019.0061. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2019.0061. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2019.0061.

[5] "40 years of quantum computing". In: *Nature Reviews Physics* 4.1 (Jan. 2022), pp. 1–1. ISSN: 2522-5820. DOI: 10.1038/s42254-021-00410-6. URL: https://doi.org/10.1038/s42254-021-00410-6.

[6] Richard P. Feynman. "Simulating physics with computers". In: *International Journal of Theoretical Physics* 21.6 (June 1982), pp. 467–488. ISSN: 1572-9575. DOI: 10.1007/BF02650179. URL: https://doi.org/10.1007/BF02650179.

[7] A M Tyryshkin et al. "Coherence of spin qubits in silicon". In: *Journal of Physics: Condensed Matter* 18.21 (May 2006), S783. DOI: 10.1088/0953-8984/18/21/S06. URL: https://dx.doi.org/10.1088/0953-8984/18/21/S06.

[8] María de la Mata and Jordi Arbiol. "High Resolution in STEM Mode: Individual Atom Analysis in Semiconductor Nanowires". In: *Transmission Electron Microscopy Characterization of Nanomaterials*. Ed. by Challa S.S.R. Kumar. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 375–425. ISBN: 978-3-642-38934-4. DOI: 10.1007/978-3-642-38934-4_9. URL: https://doi.org/10.1007/978-3-642-38934-4_9.

[9] Yadira Arroyo Rojas Dasilva et al. "Structural defects in cubic semiconductors characterized by aberration-corrected scanning transmission electron microscopy". In: *Ultramicroscopy* 176 (2017). 70th Birthday of Robert Sinclair and 65th Birthday of Nestor J. Zaluzec PICO 2017 – Fourth Conference on Frontiers of Aberration Corrected Electron Microscopy, pp. 11–22. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2016.09.015. URL: https://www.sciencedirect.com/science/article/pii/S0304399116302121.

[10] David B. Williams and C. Barry Carter. *Transmission Electron Microscopy*. 2nd ed. Springer New York, 2009. ISBN: 978-0-387-76502-0.

[11] R.F. Egerton. *Electron Energy-Loss Spectroscopy in the Electron Microscope*. 3rd ed. Springer, 2011. ISBN: 978-1-4419-9583-4.

[12] Byoungsang Lee et al. "Statistical Characterization of the Morphologies of Nanoparticles through Machine Learning Based Electron Microscopy Image Analysis". In: *ACS Nano* 14.12 (2020). PMID: 33231065, pp. 17125–17133. DOI: 10.1021/acsnano.0c06809. eprint: https://doi.org/10.1021/acsnano.0c06809. URL: https://doi.org/10.1021/acsnano.0c06809.

[13] Biyu Guo et al. "Transmission Electron Microscopy-Based Statistical Analysis of Commercially Available Graphene Oxide Quantum Dots". In: *Crystal Research and Technology* 55.4 (2020), p. 1900231. DOI: https://doi.org/10.1002/crat.201900231. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/crat.201900231. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/crat.201900231.

[14] Frederick Allars et al. "Efficient large field of view electron phase imaging using near-field electron ptychography with a diffuser". In: *Ultramicroscopy* 231 (2021). 80th Birthdays of Colin Humphreys and Knut Urban, 75th Birthdays of Wolfgang Baumeister and John Spence - PICO 2021 – Sixth Conference on Frontiers of Aberration Corrected Electron Microscopy, p. 113257. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2021.113257. URL: https://www.sciencedirect.com/science/article/pii/S0304399121000498.

[15] A Pofelski et al. "Large Field of View Strain Characterization in a Scanning Transmission Electron Microscope Using a Designed Coherent Sampler". In: *Microscopy and Microanalysis* 25.S2 (Aug. 2019), pp. 86–87. ISSN: 1431-9276. DOI: 10.1017/S1431927619001168. eprint: https://academic.oup.com/mam/article-pdf/25/S2/86/48323371/mam0086.pdf. URL: https://doi.org/10.1017/S1431927619001168.

[16] Colin Ophus. "Four-Dimensional Scanning Transmission Electron Microscopy (4D-STEM): From Scanning Nanodiffraction to Ptychography and Beyond". In: *Microscopy and Microanalysis* 25.3 (June 2019), pp. 563–582. ISSN: 1431-9276. DOI: 10.1017/S1431927619000497. eprint: https://academic.oup.com/mam/article-pdf/25/3/563/47996685/mam0563.pdf. URL: https://doi.org/10.1017/S1431927619000497.

[17] Carolin B Wahl, Chad A Mirkin, and Vinayak P Dravid. "Towards Autonomous Electron Microscopy for High-throughput Materials Discovery". In: *Microscopy and Microanalysis* 29.Supplement$_1$ (July 2023), pp. 1913–1914. ISSN: 1431-9276. DOI: 10.1093/micmic/ozad067.988. eprint: https://academic.oup.com/mam/article-pdf/29/Supplement_1/1913/50931398/ozad067.988.pdf. URL: https://doi.org/10.1093/micmic/ozad067.988.

[18] Jörgen Kornfeld and Winfried Denk. "Progress and remaining challenges in high-throughput volume electron microscopy". In: *Current Opinion in Neurobiology* 50 (2018). Neurotechnologies, pp. 261–267. ISSN: 0959-4388. DOI: https://doi.org/10.1016/j.conb.2018.04.030. URL: https://www.sciencedirect.com/science/article/pii/S0959438817303021.

[19] F. S. Hage et al. "Phonon Spectroscopy at Atomic Resolution". In: *Phys. Rev. Lett.* 122 (1 Jan. 2019), p. 016103. DOI: 10.1103/PhysRevLett.122.016103. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.016103.

[20] Sara Martí-sánchez et al. "Sub-nanometer mapping of strain-induced band structure variations in planar nanowire core-shell heterostructures". In: *Nature Communications* 13 (2022), p. 4089. DOI: /10.1038/s41467-022-31778-3. eprint: https://doi.org/10.1038/s41467-022-31778-3. URL: https://doi.org/10.1038/s41467-022-31778-3.

[21] Fredrik S. Hage et al. "Nanoscale momentum-resolved vibrational spectroscopy". In: *Science Advances* 4.6 (2018), eaar7495. DOI: 10.1126/sciadv.aar7495. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.aar7495. URL: https://www.science.org/doi/abs/10.1126/sciadv.aar7495.

[22] Haimei Zheng, Ying Shirley Meng, and Yimei Zhu. "Frontiers of in situ electron microscopy". In: *MRS Bulletin* 40.1 (2015), pp. 12–18. DOI: 10.1557/mrs.2014.305.

[23] Zheng Fan et al. "In Situ Transmission Electron Microscopy for Energy Materials and Devices". In: *Advanced Materials* 31.33 (2019), p. 1900608. DOI: https://doi.org/10.1002/adma.201900608. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201900608. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201900608.

[24] Geoffrey H. Campbell, Joseph T. McKeown, and Melissa K. Santala. "Time resolved electron microscopy for in situ experiments". In: *Applied Physics Reviews* 1.4 (Nov. 2014), p. 041101. ISSN: 1931-9401. DOI: 10.1063/1.4900509. eprint: https://pubs.aip.org/aip/apr/article-pdf/doi/10.1063/1.4900509/14572235/041101_1_online.pdf. URL: https://doi.org/10.1063/1.4900509.

[25] Qiaoli Chen et al. "Imaging Beam-Sensitive Materials by Electron Microscopy". In: *Advanced Materials* 32.16 (2020), p. 1907619. DOI: https://doi.org/10.1002/adma.201907619. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201907619. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201907619.

[26] Daliang Zhang et al. "Atomic-resolution transmission electron microscopy of electron beam–sensitive crystalline materials". In: *Science* 359.6376 (2018), pp. 675–679. DOI: 10.1126/science.aao0865. eprint: https://www.science.org/doi/pdf/10.1126/science.aao0865. URL: https://www.science.org/doi/abs/10.1126/science.aao0865.

[27] Peter Nimbe, Benjamin Asubam Weyori, and Adebayo Felix Adekoya. "Models in quantum computing: a systematic review". In: *Quantum Information Processing* 20.2 (Feb. 2021), p. 80. ISSN: 1573-1332. DOI: 10.1007/s11128-021-03021-3. URL: https://doi.org/10.1007/s11128-021-03021-3.

[28] Sukhpal Singh Gill et al. "Quantum computing: A taxonomy, systematic review and future directions". In: *Software: Practice and Experience* 52.1 (2022), pp. 66–114. DOI: https://doi.org/10.1002/spe.3039. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/spe.3039. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.3039.

[29] Frank Arute et al. "Quantum supremacy using a programmable superconducting processor". In: *Nature* 574.7779 (Oct. 2019), pp. 505–510. ISSN: 1476-4687. DOI: 10.1038/s41586-019-1666-5. URL: https://doi.org/10.1038/s41586-019-1666-5.

[30] Giordano Scappucci et al. "The germanium quantum information route". In: *Nature Reviews Materials* 6.10 (Oct. 2021), pp. 926–943. ISSN: 2058-8437. DOI: 10.1038/s41578-020-00262-z. URL: https://doi.org/10.1038/s41578-020-00262-z.

[31] Daniel Jirovec et al. "A singlet-triplet hole spin qubit in planar Ge". In: *Nature Materials* 20 (2021), pp. 1106–1112. DOI: /10.1038/s41563-021-01022-2. eprint: https://doi.org/10.1038/s41563-021-01022-2. URL: https://doi.org/10.1038/s41563-021-01022-2.

[32] Alberto Tosato et al. "Hard superconducting gap in germanium". In: *Communications Materials* 4 (2023), p. 23. DOI: /10.1038/s43246-023-00351-w. eprint: https://doi.org/10.1038/s43246-023-00351-w. URL: https://doi.org/10.1038/s43246-023-00351-w.

[33] Brian Paquelet Wuetz et al. "Reducing charge noise in quantum dots by using thin silicon quantum wells". In: *Nature Communications* 14.1 (Mar. 2023), p. 1385. ISSN: 2041-1723. DOI: 10.1038/s41467-023-36951-w. URL: https://doi.org/10.1038/s41467-023-36951-w.

[34] Semiconductor Industry Association. *Global Semiconductor Sales Decrease 8.2% in 2023; Market Rebounds Late in Year — semiconductors.org*. https://www.semiconductors.org/global-semiconductor-sales-decrease-8-2-in-2023-market-rebounds-late-in-year/. [Accessed 28-02-2024]. 2024.

[35] Eitan Oksenberg et al. "Surface-Guided Core–Shell ZnSe@ZnTe Nanowires as Radial p–n Heterojunctions with Photovoltaic Behavior". In: *ACS Nano* 11.6 (2017). PMID: 28505415, pp. 6155–6166. DOI: 10.1021/acsnano.7b02199. eprint: https://doi.org/10.1021/acsnano.7b02199. URL: https://doi.org/10.1021/acsnano.7b02199.

[36] Wei Zhan et al. "Band gap maps beyond the delocalization limit: correlation between optical band gaps and plasmon energies at the nanoscale". In: *Scientific Reports* 8.1 (Jan. 2018), p. 848. ISSN: 2045-2322. DOI: 10.1038/s41598-017-18949-9. URL: https://doi.org/10.1038/s41598-017-18949-9.

[37] F. J. García de Abajo. "Optical excitations in electron microscopy". In: *Rev. Mod. Phys.* 82 (1 Feb. 2010), pp. 209–275. DOI: 10.1103/RevModPhys.82.209. URL: https://link.aps.org/doi/10.1103/RevModPhys.82.209.

[38] L. Jones et al. "Smart Align—a new tool for robust non-rigid registration of scanning microscope data". In: *Adv. Struct. Chem. Imaging* 1 (2015), pp. 1–16.

[39] A. De Backer et al. "StatSTEM: An efficient approach for accurate and precise model-based quantification of atomic resolution electron microscopy images". In: *Ultramicroscopy* 171 (2016), pp. 104–116.

[40] Annick De Backer et al. "Experimental reconstructions of 3D atomic structures from electron microscopy images using a Bayesian genetic algorithm". In: *npj Computational Materials* 8.1 (Oct. 2022), p. 216. ISSN: 2057-3960. DOI: 10.1038/s41524-022-00900-w. URL: https://doi.org/10.1038/s41524-022-00900-w.

[41] A. De Backer et al. "Atom counting in HAADF STEM using a statistical model-based approach: Methodology, possibilities, and inherent limitations". In: *Ultramicroscopy* 134 (2013), pp. 23–33.

[42] M. Vatanparast et al. "Strategy for reliable strain measurement in InAs/GaAs materials from high-resolution Z-contrast STEM images". In: *J. Phys. Conf. Ser.* 902 (2017).

[43] J. M. Zuo et al. "Lattice and strain analysis of atomic resolution Z-contrast images based on template matching". In: *Ultramicroscopy* 136 (2014), pp. 50–60.

[44] M. J. Hÿtch, J.-L. Putaux, and J.-M. Pénisson. "Measurement of the displacement field of dislocations to 0.03Å by electron microscopy". In: *Nature* 423 (2003), pp. 270–273.

[45] S. Bals et al. "Statistical estimation of atomic positions from exit wave reconstruction with a precision in the picometer range". In: *Phys. Rev. Lett.* 96 (2006), pp. 1–4.

[46] L. Jones and P. D. Nellist. "Identifying and correcting scan noise and drift in the scanning transmission electron microscope". In: *Microsc. Microanal.* 19 (2013), pp. 1050–1060.

[47] R. Sagawa et al. "Aberration Measurement and Correction in Scanning Transmission Electron Microscopy using Machine Learning". In: *Microsc. Microanal.* 27 (2021), pp. 814–816.

[48] Marc Botifoll, Ivan Pinto-Huguet, and Jordi Arbiol. "Machine learning in electron microscopy for advanced nanocharacterization: current developments, available tools and future outlook". In: *Nanoscale Horiz.* 7 (12 2022), pp. 1427–1477. DOI: 10.1039/D2NH00377E. URL: http://dx.doi.org/10.1039/D2NH00377E.

[49] N. Bonnet. "Multivariate statistical methods for the analysis of microscope image series: Applications in materials science". In: *J. Microsc.* 190 (1998), pp. 2–18.

[50] N. Bonnet et al. "New applications of multivariate statistical analysis in spectroscopy and microscopy". In: *Ultramicroscopy* 40 (1992), pp. 1–11.

[51] M. Bosman et al. "Mapping chemical and bonding information using multivariate analysis of electron energy-loss spectrum images". In: *Ultramicroscopy* 106 (2006), pp. 1024–1032.

[52] K. Aso et al. "Detection of picometer-order atomic displacements in drift-compensated HAADF-STEM images of gold nanorods". In: *Microscopy* 65 (2016), pp. 391–399.

[53] K. Aso et al. "Lattice Tetragonality and Local Strain Depending on Shape of Gold Nanoparticles". In: *Microsc. Microanal.* 25 (2019), pp. 2122–2123.

[54] P. Trebbia. "EELS elemental mapping with unconventional methods I . Theoretical basis: image analysis with multivariate statistics and entropy concepts". In: *Ultramicroscopy* 34 (1990), pp. 165–178.

[55] N. Bonnet, N. Brun, and C. Colliex. "Extracting information from sequences of spatially resolved EELS spectra using multivariate statistical analysis". In: *Ultramicroscopy* 77 (1999).

[56] C. Tian et al. "Deep learning on image denoising: An overview". In: *Neural Networks* 131 (2020), pp. 251–275.

[57] S. V. Kalinin et al. "Lab on a beam-Big data and artificial intelligence in scanning transmission electron microscopy". In: *MRS Bull.* 44 (2019), pp. 565–575.

[58] M. Shen et al. "Multi defect detection and analysis of electron microscopy images with deep learning". In: *Comput. Mater. Sci.* 199 (2021).

[59] Chia-Hao Lee et al. "Deep Learning Enabled Strain Mapping of Single-Atom Defects in Two-Dimensional Transition Metal Dichalcogenides with Sub-Picometer Precision". In: *Nano Letters* 20.5 (2020). PMID: 32243178, pp. 3369–3377. DOI: 10.1021/acs.nanolett.0c00269. eprint: https://doi.org/10.1021/acs.nanolett.0c00269. URL: https://doi.org/10.1021/acs.nanolett.0c00269.

[60] F. Wang et al. "Noise2Atom: unsupervised denoising for scanning transmission electron microscopy images". In: *Appl. Microsc.* 50 (2020).

[61] J. M. Ede and R. Beanland. "Improving electron micrograph signal-to-noise with an atrous convolutional encoder-decoder". In: *Ultramicroscopy* 202 (2019), pp. 18–25.

[62] T. M. Quan et al. "Removing imaging artifacts in electron microscopy using an asymmetrically cyclic adversarial network without paired training data". In: *Proc. - 2019 Int. Conf. Comput. Vis. Work. ICCVW 2019*. 2019, pp. 3804–3813. DOI: 10.1109/ICCVW.2019.00473.

[63] J. P. Buban and S.-Y. Choi. "Auto-encoders for Noise Reduction in Scanning Transmission Electron Microscopy". In: *Microsc. Microanal.* 23 (2017), pp. 130–131.

[64] J. L. Vincent et al. "Developing and Evaluating Deep Neural Network-Based Denoising for Nanoparticle TEM Images with Ultra-Low Signal-to-Noise". In: (2021). DOI: 10.1017/S1431927621012678.

[65] S. Mohan et al. "Deep Denoising For Scientific Discovery: A Case Study In Electron Microscopy". In: (2020).

[66] S. Mohan et al. "Adaptive Denoising via GainTuning". In: (2021).

[67] N. Schnitzer, S. H. Sung, and R. Hovden. "Maximal Resolution from the Ronchigram: Human vs. Deep Learning". In: *Microsc. Microanal.* 25 (2019), pp. 160–161.

[68] C. Zhang et al. "Aberration Corrector Tuning with Machine-Learning-Based Emittance Measurements and Bayesian Optimization". In: *Microsc. Microanal.* 27 (2021), pp. 810–812.

[69] K. Roccapriore, S. V. Kalinin, and M. Ziatdinov. "Physics discovery in nanoplasmonic systems via autonomous experiments in Scanning Transmission Electron Microscopy". In: (2021).

[70] E. Rotunno et al. "Convolutional neural network as a tool for automatic alignment of electron optical beam shaping devices". In: *Microsc. Microanal.* 27 (2021), pp. 822–824.

[71] E. Rotunno et al. "Alignment of electron optical beam shaping elements using a convolutional neural network". In: *Ultramicroscopy* 228 (2021), p. 113338.

[72] Sergei V. Kalinin et al. "Machine learning for automated experimentation in scanning transmission electron microscopy". In: *npj Computational Materials* 9.1 (Dec. 2023), p. 227. ISSN: 2057-3960. DOI: 10.1038/s41524-023-01142-0. URL: https://doi.org/10.1038/s41524-023-01142-0.

[73] E. F. Rauch et al. "Automated nanocrystal orientation and phase mapping in the transmission electron microscope on the basis of precession electron diffraction". In: *Zeitschrift fur Krist.* 225 (2010), pp. 103–109.

[74] Y. Jin et al. "Machine learning guided rapid focusing with sensor-less aberration corrections". In: *Opt. Express* 26 (2018), p. 30162.

[75] O. Dyck, S. Jesse, and S. V. Kalinin. "A self-driving microscope and the Atomic Forge". In: *MRS Bull.* 44 (2019).

[76] S. R. Spurgeon et al. "Towards data-driven next-generation transmission electron microscopy". In: *Nat. Mater.* 20 (2021), pp. 274–279.

[77] S. et al. Somnath. "USID and Pycroscopy: Open frameworks for storing and analyzing spectroscopic and imaging data". In: *Ultramicroscopy* (2021), pp. 1–33.

[78] J. Dan et al. "A hierarchical active-learning framework for classifying structural motifs in atomic resolution microscopy". In: (2020).

[79] H. I. Rasool, C. Ophus, and A. Zettl. "Atomic Defects in Two Dimensional Materials". In: *Adv. Mater.* 27 (2015), pp. 5771–5777.

[80] R. Kannan et al. "Deep data analysis via physically constrained linear unmixing: universal framework, domain examples, and a community-wide platform". In: *Adv. Struct. Chem. Imaging* 4 (2018).

[81] Maxim Ziatdinov et al. "Deep learning of atomically resolved scanning transmission electron microscopy images: Chemical identification and tracking local transformations". en. In: *ACS Nano* 11.12 (Dec. 2017), pp. 12742–12752.

[82] J. Madsen et al. "A Deep Learning Approach to Identify Local Structures in Atomic-Resolution Transmission Electron Microscopy Images". In: *Adv. Theory Simulations* 1 (2018), pp. 1–12.

[83] R. Sadre et al. "Deep Learning Segmentation of Complex Features in Atomic-Resolution Phase-Contrast Transmission Electron Microscopy Images". In: *Microsc. Microanal.* 27 (2021), pp. 804–814.

[84] P. Cho et al. "Defect detection in atomic resolution transmission electron microscopy images using machine learning". In: *Mathematics* 9 (2021), pp. 1–16.

[85] W. Lin et al. "Local crystallography analysis for atomically resolved scanning tunneling microscopy images". In: *Nanotechnology* 24 (2013).

[86] G. Roberts et al. "DefectNet – A Deep Convolutional Neural Network for Semantic Segmentation of Crystallographic Defects in Advanced Microscopy Images". In: *Microsc. Microanal.* 25 (2019), pp. 164–165.

[87] C. Kunka et al. "Decoding defect statistics from diffractograms via machine learning". In: *npj Comput. Mater.* 7 (2021).

[88] J. Dan, X. Zhao, and S. J. Pennycook. "A machine perspective of atomic defects in scanning transmission electron microscopy". In: *InfoMat* 1 (2019), pp. 359–375.

[89] J. Yi, Z. Yuan, and J. Peng. "Adversarial-Prediction Guided Multi-Task Adaptation for Semantic Segmentation of Electron Microscopy Images". In: *Proc. - Int. Symp. Biomed. Imaging*. 2020, pp. 1205–1208.

[90] Q. Li et al. "Quantification of flexoelectricity in PbTiO3/SrTiO3 superlattice polar vortices using machine learning and phase-field modeling". In: *Nature Communications* 8.1 (Nov. 2017), p. 1468. ISSN: 2041-1723. DOI: 10.1038/s41467-017-01733-8. URL: https://doi.org/10.1038/s41467-017-01733-8.

[91] L. Daniel D. and S. H. Sebastian. "Learning the parts of objects by non-negative matrix factorization". In: *Nature* 401 (1999), pp. 788–791.

[92] B. R. Jany, A. Janas, and F. Krok. "Automatic microscopic image analysis by moving window local Fourier Transform and Machine Learning". In: *Micron* 130 (2020), p. 102800.

[93] B. H. Martineau et al. "Unsupervised machine learning applied to scanning precession electron diffraction data". In: *Adv. Struct. Chem. Imaging* (2019). DOI: 10.1186/s40679-019-0063-3.

[94] R. K. Vasudevan et al. "Investigating phase transitions from local crystallographic analysis based on machine learning of atomic environments". In: (2020).

[95] M. E. Winter. "N-FINDR: an algorithm for fast spectral endmember determination in hyperspectral data". In: *Int. Geosci. Remote Sens. Symp.* 1999, pp. 266–275.

[96] Maxim Ziatdinov et al. "Causal analysis of competing atomistic mechanisms in ferroelectric materials from high-resolution scanning transmission electron microscopy data". In: *npj Computational Materials* 6.1 (Aug. 2020), p. 127. ISSN: 2057-3960. DOI: 10.1038/s41524-020-00396-2. URL: https://doi.org/10.1038/s41524-020-00396-2.

[97] R. K. Vasudevan et al. "Phases and Interfaces from Real Space Atomically Resolved Data: Physics-Based Deep Data Image Analysis". In: *Nano Lett.* 16 (2016), pp. 5574–5581.

[98] P. Baldi. "Autoencoders, Unsupervised Learning, and Deep Architectures". In: *ICML Unsupervised Transf. Learn.* 2012, pp. 37–50. DOI: 10.1561/2200000006.

[99] W. H. Lopez Pinaya et al. "Autoencoders". In: *Mach. Learn. Methods Appl. to Brain Disord.* 2019, pp. 193–208. DOI: 10.1016/B978-0-12-815739-8.00011-0.

[100] M. A. Ziatdinov and S. V. Kalinin. "Robust Feature Disentanglement in Imaging Data via Joint Invariant Variational Autoencoders: from Cards to Atoms".

[101] L. Vlcek et al. "Learning from Imperfections: Predicting Structure and Thermodynamics from Atomic Imaging of Fluctuations". In: *ACS Nano* 13 (2019), pp. 718–727.

[102] Mark P. Oxley et al. "Probing atomic-scale symmetry breaking by rotationally invariant machine learning of multidimensional electron scattering". In: *npj Computational Materials* 7.1 (May 2021), p. 65. ISSN: 2057-3960. DOI: 10.1038/s41524-021-00527-3. URL: https://doi.org/10.1038/s41524-021-00527-3.

[103] Sergei V. Kalinin et al. "Exploring order parameters and dynamic processes in disordered systems via variational autoencoders". In: *Science Advances* 7.17 (2021), eabd5084. DOI: 10.1126/sciadv.abd5084. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.abd5084. URL: https://www.science.org/doi/abs/10.1126/sciadv.abd5084.

[104] Mani Valleti et al. *Unsupervised learning of ferroic variants from atomically resolved STEM images*. 2022. arXiv: 2101.06892 [cond-mat.mtrl-sci].

[105] F. Uesugi et al. "Non-negative matrix factorization for mining big data obtained using four-dimensional scanning transmission electron microscopy". In: *Ultramicroscopy* 221 (2021), p. 113168.

[106] M. Jacob et al. "Statistical Machine Learning and Compressed Sensing Approaches for Analytical Electron Tomography - Application to Phase Change Materials". In: *Microsc. Microanal.* 25 (2019), pp. 156–157.

[107] S. Kiyohara et al. "Quantitative estimation of properties from core- loss spectrum via neural network". In: *J. Phys. Mater.* 2 (2019).

[108] R. S. Pennington et al. "Neural-network-based depth-resolved multiscale structural optimization using density functional theory and electron diffraction data". In: *Phys. Rev. B* 97 (2018), pp. 1–10.

[109] Kyle P. Kelley et al. "Fast Scanning Probe Microscopy via Machine Learning: Non-Rectangular Scans with Compressed Sensing and Gaussian Process Optimization". In: *Small* 16.37 (2020), p. 2002878. DOI: https://doi.org/10.1002/smll.202002878. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smll.202002878. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/smll.202002878.

[110] W. S. McCulloch and W. Pitts. "A logical calculus of the ideas immanent in nervous activity". In: *Bull. Math. Biol.* 52 (1990), pp. 99–115.

[111] Y. Lecun, Y. Bengio, and G. Hinton. "Deep learning". In: *Nature* 521 (2015), pp. 436–444.

[112] Jeffrey M Ede. "Deep learning in electron microscopy". In: *Machine Learning: Science and Technology* 2.1 (Mar. 2021), p. 011004. DOI: 10.1088/2632-2153/abd614. URL: https://dx.doi.org/10.1088/2632-2153/abd614.

[113] A. Garcia-Garcia et al. "A Review on Deep Learning Techniques Applied to Semantic Segmentation". In: (2017).

[114] R. M. Patton et al. "167-PFlops deep learning for electron microscopy: From learning physics to atomic manipulation". In: *Proc. - Int. Conf. High Perform. Comput. Networking, Storage, Anal. SC*. 2019, pp. 638–648. DOI: 10.1109/SC.2018.00053.

[115] A. Maksov et al. "Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS2". In: *npj Comput. Mater.* 5 (2019), pp. 1–8.

[116] Sang-Hyeok Yang et al. "Deep Learning-Assisted Quantification of Atomic Dopants and Defects in 2D Materials". In: *Advanced Science* 8.16 (2021), p. 2101099. DOI: https://doi.org/10.1002/advs.202101099. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/advs.202101099. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/advs.202101099.

[117] A. De Backer et al. "Procedure for 3D atomic resolution reconstructions using atom-counting and a Bayesian genetic algorithm". In: (2021).

[118] M. Ziatdinov et al. "Building ferroelectric from the bottom up: The machine learning analysis of the atomic-scale ferroelectric distortions". In: *Applied Physics Letters* 115.5 (Aug. 2019), p. 052902. ISSN: 0003-6951. DOI: 10.1063/1.5109520. eprint: https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/1.5109520/14527440/052902_1_online.pdf. URL: https://doi.org/10.1063/1.5109520.

[119] M. Ziatdinov et al. "Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations". In: *ACS Nano* 11 (2017), pp. 12742–12752.

[120] Sergei V. Kalinin et al. "Unsupervised Machine Learning Discovery of Chemical and Physical Transformation Pathways from Imaging Data".

[121] M. Nord et al. "Atomap: a new software tool for the automated analysis of atomic resolution images using two-dimensional Gaussian fitting". In: *Adv. Struct. Chem. Imaging* 3 (2017).

[122] J. A. Aguiar et al. "Decoding crystallography from high-resolution electron imaging and diffraction datasets with deep learning". In: *Sci. Adv.* 5 (2019), pp. 1–10.

[123] R. K. Vasudevan et al. "Mapping mesoscopic phase evolution during E-beam induced transformations via deep learning of atomically resolved images". In: *npj Comput. Mater.* 4 (2018).

[124] R. K. Vasudevan et al. "Exploring Electron Beam Induced Atomic Assembly via Reinforcement Learning in a Molecular Dynamics Environment".

[125] R. S. Pennington, W. Van den Broek, and C. T. Koch. "Third-dimension information retrieval from a single convergent-beam transmission electron diffraction pattern using an artificial neural network". In: *Phys. Rev. B* 89 (2014), p. 205409.

[126] J. P. Horwath. "Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images". In: *npj Comput. Mater.* (). DOI: 10.1038/s41524-020-00363-x.

[127] G. Benton et al. "Learning invariances in neural networks". In: *Adv. Neural Inf. Process. Syst.* 2020.

[128] M. Ziatdinov et al. "Deep Learning for Atomically Resolved Imaging". In: *Microsc. Microanal.* 24 (2018), pp. 60–61.

[129] M. Finzi et al. "Generalizing convolutional neural networks for equivariance to lie groups on arbitrary continuous data". In: *37th Int. Conf. Mach. Learn. ICML 2020 PartF16814*. 2020, pp. 3146–3157.

[130] Keiron O'Shea and Ryan Nash. *An Introduction to Convolutional Neural Networks*. 2015. arXiv: 1511.08458 [cs.NE].

[131] Marc Botifoll et al. *ZnTe-ZnSe Strain Modulated BG Mapping*. GitHub repository. 2023. URL: https://github.com/mbotifollmoral/ZnTe-ZnSe_Strain_Modulated_BG_Mapping.

[132] T. Zhou, M. Cherukara, and C. Phatak. "High resolution functional imaging through Lorentz transmission electron microscopy and differentiable programming". In: *npj Comput. Mater.* 7 (2021), pp. 1–21.

[133] S. et al. Kandel. "Using automatic differentiation as a general framework for ptychographic reconstruction". In: *Opt. Express* 27 (2019), pp. 18653–18672.

[134] S. Ghosh et al. "ADP: Automatic differentiation ptychography". In: *IEEE Int. Conf. Comput. Photogr. ICCP 2018*. 2018, pp. 1–10. DOI: 10.1109/ICCPHOT.2018.8368470.

[135] W. Hoppe. "Trace structure analysis, ptychography, phase tomography". In: *Ultramicroscopy* 10 (1982), pp. 187–198.

[136] B. A. R. Lupini, M. P. Oxley, and S. V. Kalinin. "Pushing the limits of electron ptychography". In: *Science* 362 (2018).

[137] Zhen Chen et al. "Electron ptychography achieves atomic-resolution limits set by lattice vibrations". In: *Science* 372.6544 (2021), pp. 826–831. DOI: 10.1126/science.abg2533. eprint: https://www.science.org/doi/pdf/10.1126/science.abg2533. URL: https://www.science.org/doi/abs/10.1126/science.abg2533.

[138] Wenfeng Yang et al. "Local-orbital ptychography for ultrahigh-resolution imaging". In: *Nature Nanotechnology* (Jan. 2024). ISSN: 1748-3395. DOI: 10.1038/s41565-023-01595-w. URL: https://doi.org/10.1038/s41565-023-01595-w.

[139] Kayla X. Nguyen et al. "Achieving sub-0.5-angstrom–resolution ptychography in an uncorrected electron microscope". In: *Science* 383.6685 (2024), pp. 865–870. DOI: 10.1126/science.adl2029. eprint: https://www.science.org/doi/pdf/10.1126/science.adl2029. URL: https://www.science.org/doi/abs/10.1126/science.adl2029.

[140] M. et al. Cao. "Machine Learning for Phase Retrieval from 4D-STEM Data". In: *Microsc. Microanal.* 26 (2020), pp. 2020–2021.

[141] M. et al. Schloz. "Overcoming information reduced data and experimentally uncertain parameters in ptychography with regularized optimization". In: *Opt. Express* 28 (2020), p. 28306.

[142] M. J. et al. Cherukara. "AI-enabled high-resolution scanning coherent diffraction imaging". In: *Appl. Phys. Lett.* 117 (2020).

[143] M. Schloz et al. "Adaptive Scanning in Ptychography through Deep Reinforcement Learning". In: *Microsc. Microanal.* 27 (2021), pp. 818–821.

[144] X. et al. Huang. "Optimization of overlap uniformness for ptychography". In: *Opt. Express* 22 (2014), p. 12634.

[145] Katsuaki Nakazawa and Kazutaka Mitsuishi. "Development of temporal series 4D-STEM and application to relaxation time measurement". In: *Microscopy* 72.5 (Jan. 2023), pp. 446–449. ISSN: 2050-5701. DOI: 10.1093/jmicro/dfad006. eprint: https://academic.oup.com/jmicro/

article-pdf/72/5/446/51960184/dfad006.pdf. URL: https://doi.org/10.1093/jmicro/dfad006.

[146] G. W. et al. Paterson. "Fast Pixelated Detectors in Scanning Transmission Electron Microscopy. Part II: Post-Acquisition Data Processing, Visualization, and Structural Characterization". In: *Microsc. Microanal.* 26 (2020), pp. 944–963.

[147] M. et al. Nord. "Fast Pixelated Detectors in Scanning Transmission Electron Microscopy. Part I: Data Acquisition, Live Processing, and Storage". In: *Microsc. Microanal.* 26 (2020), pp. 653–666.

[148] M. et al. Nord. "Strain Anisotropy and Magnetic Domains in Embedded Nanomagnets". In: *Small* 15 (2019).

[149] G. Correa and D. Muller. "Machine Learning for Sub-pixel Super-resolution in Direct Electron Detectors". In: *Microsc. Microanal.* 26 (2020), pp. 1932–1934.

[150] X. et al. Li. "Unsupervised Machine Learning to Distill Structural-Property Insights from 4D-STEM". In: 25 (2021), pp. 2016–2017.

[151] C. Zhang et al. "Denoising atomic resolution 4D scanning transmission electron microscopy data with tensor singular value decomposition". In: *Ultramicroscopy* 219 (2020), p. 113123.

[152] Xin Li et al. "Manifold learning of four-dimensional scanning transmission electron microscopy". In: *npj Computational Materials* 5.1 (Jan. 2019), p. 5. ISSN: 2057-3960. DOI: 10.1038/s41524-018-0139-y. URL: https://doi.org/10.1038/s41524-018-0139-y.

[153] A. et al. Nalin Mehta. "Unravelling stacking order in epitaxial bilayer MX2 using 4D-STEM with unsupervised learning". In: *Nanotechnology* 31 (2020), p. 445702.

[154] P. Cueva, E. Padget, and D. A. Muller. "A Natural Basis for Unsupervised Machine Learning on Scanning Diffraction Data". In: *Microsc. Microanal.* 24 (2018), pp. 490–491.

[155] C. Shi et al. "Rapid and Semi-Automated Analysis of 4D-STEM data via Unsupervised Learning". In: *Microsc. Microanal.* 27 (2021), pp. 58–59.

[156] F. I. et al. Allen. "Fast Grain Mapping with Sub-Nanometer Resolution Using 4D-STEM with Grain Classification by Principal Component Analysis and Non-Negative Matrix Factorization". In: *Microsc. Microanal.* 27 (2021), pp. 794–803.

[157] C. Zhang et al. "Atomic resolution convergent beam electron diffraction analysis using convolutional neural networks". In: *Ultramicroscopy* 210 (2020), p. 112921.

[158] M. P. et al. Oxley. "Deep learning of interface structures from simulated 4D STEM data: cation intermixing vs. roughening". In: *Machine Learning: Science and Technology* 1 (2020), 04LT01.

[159] R. Manzorro et al. "A Deep Learning Approach to Retrieving 3D Structure Information from High Resolution Time-Resolved TEM Images". In: *Microsc. Microanal.* 27 (2021), pp. 464–465.

[160] S. Van Aert et al. "Maximum likelihood estimation of structure parameters from high resolution electron microscopy images. Part II: A practical example". In: *Ultramicroscopy* 104 (2005), pp. 107–125.

[161] L. C. Gontard et al. "Photogrammetry of the three-dimensional shape and texture of a nanoscale particle using scanning electron microscopy and free software". In: *Ultramicroscopy* 169 (2016), pp. 80–88.

[162] Z. et al. Saghi. "Three-dimensional morphology of iron oxide nanoparticles with reactive concave surfaces. A compressed sensing-electron tomography (CS-ET) approach". In: *Nano Lett.* 11 (2011), pp. 4666–4673.

[163] E. Bladt et al. "Electron tomography based on highly limited data using a neural network reconstruction technique". In: *Ultramicroscopy* 158 (2015), pp. 81–88.

[164] J. M. Thomas et al. "A new approach to the investigation of nanoparticles: Electron tomography with compressed sensing". In: *J. Colloid Interface Sci.* 392 (2013), pp. 7–14.

[165] R. Leary et al. "Compressed sensing electron tomography". In: *Ultramicroscopy* 131 (2013), pp. 70–91.

[166] L. Staniewicz and P. A. Midgley. "Machine learning as a tool for classifying electron tomographic reconstructions". In: *Adv. Struct. Chem. Imaging* 1 (2015), pp. 1–15.

[167] A. Béché et al. "Development of a fast electromagnetic beam blanker for compressed sensing in scanning transmission electron microscopy". In: *Appl. Phys. Lett.* 108 (2016).

[168] L. Kovarik et al. "Implementing an accurate and rapid sparse sampling approach for low-dose atomic resolution STEM imaging". In: *Appl. Phys. Lett.* 109 (2016), pp. 0–5.

[169] B. Goris et al. "Electron tomography based on a total variation minimization reconstruction technique". In: *Ultramicroscopy* 113 (2012), pp. 120–130.

[170] Zineb Saghi et al. "Compressed sensing electron tomography of needle-shaped biological specimens – Potential for improved reconstruction fidelity with reduced dose". In: *Ultramicroscopy* 160 (2016), pp. 230–238. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2015.10.021. URL: https://www.sciencedirect.com/science/article/pii/S0304399115300565.

[171] M. et al. López-Haro. "A Macroscopically Relevant 3D-Metrology Approach for Nanocatalysis Research". In: *Part. Part. Syst. Charact.* 35 (2018), pp. 1–11.

[172] J. M. et al. Muñoz-Ocaña. "Optimization of STEM-HAADF Electron Tomography Reconstructions by Parameter Selection in Compressed Sensing Total Variation Minimization-Based Algorithms". In: *Part. Part. Syst. Charact.* 37 (2020), pp. 1–9.

[173] A. Rakowski et al. "Learning Frame Interpolation for Tilt Series Tomography". In: *Microsc. Microanal.* (2020), pp. 928–930. DOI: 10.1017/S1431927620016360.

[174] Y. et al. Zhao. "Five-second STEM dislocation tomography for 300 nm thick specimen assisted by deep-learning-based noise filtering". In: *Sci. Rep.* 11 (2021), pp. 1–12.

[175] A. A. Hendriksen, D. M. Pelt, and K. J. Batenburg. "Noise2Inverse: Self-supervised deep convolutional denoising for linear inverse problems in imaging". In: *arXiv* 6 (2020), pp. 1320–1335.

[176] D. M. Pelt and K. J. Batenburg. "Fast Tomographic Reconstruction from Limited Data Using Artificial Neural Networks". In: *IEEE Trans. Image Process.* 22 (2013), pp. 5238–5251.

[177] A. Stevens et al. "The potential for bayesian compressive sensing to significantly reduce electron dose in high-resolution STEM images". In: *Microscopy* 63 (2014), pp. 41–51.

[178] N. Browning et al. "The Potential Benefits of Compressed Sensing and Machine Learning for Advanced Imaging and Spectroscopy in the Electron Microscope". In: *Microsc. Microanal.* 26 (2020), pp. 2458–2460.

[179] D. et al. Mucke-Herzberg. "Practical Implementation of Compressive Sensing for High Resolution STEM". In: *Microsc. Microanal.* 22 (2016), pp. 558–559.

[180] Xin Li et al. "Compressed Sensing of Scanning Transmission Electron Microscopy (STEM) With Nonrectangular Scans". In: *Microscopy and Microanalysis* 24.6 (2018), pp. 623–633. DOI: 10.1017/S143192761801543X.

[181] X. et al. Li. "Manifold learning of four-dimensional scanning transmission electron microscopy". In: *npj Comput. Mater.* 5 (2019), pp. 1–8.

[182] J. M. Ede. "Deep learning in electron microscopy". In: *Mach. Learn. Sci. Technol.* 2 (2021).

[183] Jeffrey M. Ede and Richard Beanland. "Partial Scanning Transmission Electron Microscopy with Deep Learning". In: *Scientific Reports* 10.1 (May 2020), p. 8332. ISSN: 2045-2322. DOI: 10.1038/s41598-020-65261-0. URL: https://doi.org/10.1038/s41598-020-65261-0.

[184] S. Zheng et al. "Super-compression of large electron microscopy time series by deep compressive sensing learning". In: *Patterns* 2 (2021), p. 100292.

[185] M. D. Sangid. "Coupling in situ experiments and modeling – Opportunities for data fusion, machine learning, and discovery of emergent behavior". In: *Curr. Opin. Solid State Mater. Sci.* 24 (2020), p. 100797.

[186] K. et al. Higgins. "Exploration of Electrochemical Reactions at Organic–Inorganic Halide Perovskite Interfaces via Machine Learning in In Situ Time-of-Flight Secondary Ion Mass Spectrometry". In: *Advanced Functional Materials* 30 (2020).

[187] T. K. Patra et al. "Defect dynamics in 2-D MoS2 probed by using machine learning, atomistic simulations, and high-resolution microscopy". In: *ACS Nano* 12 (2018), pp. 8006–8016.

[188] Xingzhi Wang et al. "AutoDetect-mNP: An Unsupervised Machine Learning Algorithm for Automated Analysis of Transmission Electron Microscope Images of Metal Nanoparticles". In: *JACS Au* 1.3 (2021). PMID: 33778811, pp. 316–327. DOI: 10.1021/jacsau.0c00030. eprint: https://doi.org/10.1021/jacsau.0c00030. URL: https://doi.org/10.1021/jacsau.0c00030.

[189] N. M. Schneider et al. "Automated analysis of evolving interfaces during in situ electron microscopy". In: *Adv. Struct. Chem. Imaging* 2 (2016).

[190] Y. Qian et al. "Robust nanoparticles detection from noisy background by fusing complementary image information". In: *IEEE Trans. Image Process.* 25 (2016), pp. 5713–5726.

[191] Y. Qian, J. Z. Huang, and Y. Ding. "Identifying multi-stage nanocrystal growth using in situ TEM video data". In: *IISE Trans.* 49 (2017), pp. 532–543.

[192] A. A. Ezzat and M. Bedewy. "Machine learning for revealing spatial dependence among nanoparticles: Understanding catalyst film dewetting via gibbs point process models". In: *J. Phys. Chem. C* 124 (2020), pp. 27479–27494.

[193] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4.

[194] L. Yao et al. "Machine Learning to Reveal Nanoparticle Dynamics from Liquid-Phase TEM Videos". In: *ACS Cent. Sci.* 6 (2020), pp. 1421–1430.

[195] K. Faraz et al. "A Machine Learning pipeline to track the dynamics of a population of nanoparticles during in situ Environmental Transmission Electron Microscopy in gases". In: *Microsc. Microanal.* 27 (2021), pp. 2236–2237.

[196] Xin Li et al. "Statistical learning of governing equations of dynamics from in-situ electron microscopy imaging data". In: *Materials Design* 195 (2020), p. 108973. ISSN: 0264-1275. DOI: https://doi.org/10.1016/j.matdes.2020.108973. URL: https://www.sciencedirect.com/science/article/pii/S0264127520305074.

[197] M. P. Prange et al. "Simulation of STEM-EELS Including Diffraction and Solid-State Effects I: Mixed Dynamic Form Factor beyond the Dipole Approximation". In: (2021), pp. 808–809.

[198] M. P. et al. Oxley. "Simulation of Probe Position-Dependent Electron Energy-Loss Fine Structure". In: (2014), pp. 784–797. DOI: 10.1017/S1431927614000610.

[199] A. Eljarrat et al. "Density Functional Theory Modeling of Low-Loss Electron Energy-Loss Spectroscopy in Wurtzite III-Nitride Ternary Alloys". In: (2016), pp. 706–716. DOI: 10.1017/S1431927616000106.

[200] S. et al. Chen. "Electron energy loss spectroscopy and ab initio investigation of iron oxide nanomaterials grown by a hydrothermal process". In: (2009), pp. 1–10. DOI: 10.1103/PhysRevB.79.104103.

[201] A. Gloter et al. "Improving energy resolution of EELS spectra: an alternative to the monochromator solution". In: *Ultramicroscopy* 96 (2003), pp. 385–400.

[202] O.L. Krivanek et al. "Progress in ultrahigh energy resolution EELS". In: *Ultramicroscopy* 203 (2019). 75th Birthday of Christian Colliex, 85th Birthday of Archie Howie, and 75th Birthday of Hannes Lichte / PICO 2019 - Fifth Conference on Frontiers of Aberration Corrected Electron Microscopy, pp. 60–67. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2018.12.006. URL: https://www.sciencedirect.com/science/article/pii/S0304399118303668.

[203] K. Kimoto and Y. Matsui. "Software techniques for EELS to realize about 0 . 3 eV energy resolution using 300 kV FEG-TEM". In: *J. Microsc.* 208 (2002), pp. 224–228.

[204] P. Cueva et al. "New Approaches to Data Processing for Atomic Resolution EELS". In: *Microscopy and Microanalysis* 18.S2 (2012), pp. 970–971. DOI: 10.1017/S1431927612006708.

[205] L. Jones et al. "Improving the SNR of Atomic Resolution STEM EELS EDX Mapping while Reducing Beam-damage by using Non-rigid Spectrum-image Averaging". In: *Microsc. Microanal.* 21 (2021), pp. 1215–1216.

[206] J. S. Jeong and K. A. Mkhoyan. "Improving Signal-to-Noise Ratio in Scanning Transmission Electron Microscopy Energy-Dispersive X-Ray (STEM-EDX) Spectrum Images Using Single-Atomic-Column Cross-Correlation Averaging". In: *Microsc. Microanal.* 22 (2016), pp. 536–543.

[207] L. Jones et al. "Managing dose-, damage- and data-rates in multi-frame spectrum-imaging". In: *Microscopy* (2018), pp. 98–113. DOI: 10.1093/jmicro/dfx125.

[208] K. Sader et al. "Smart acquisition EELS". In: *Ultramicroscopy* 110 (2010), pp. 998–1003.

[209] P. Torruella et al. "Assessing Oxygen Vacancies in Bismuth Oxide Through EELS Measurements and DFT Simulations Assessing Oxygen Vacancies in Bismuth Oxide through EELS Measurements and DFT Simulations". In: (2017). DOI: 10.1021/acs.jpcc.7b06310.

[210] I. Arslan et al. "Comparison of simulation methods for electronic structure calculations with experimental electron energy-loss spectra". In: *Micron* 34 (2003), pp. 255–260.

[211] N. Mevenkamp et al. "Multi-modal and multi-scale non-local means method to analyze spectroscopic datasets". In: *Ultramicroscopy* (2019), p. 112877. DOI: 10.1016/j.ultramic.2019.112877.

[212] S. Jesse and S. V. Kalinin. "Principal component and spatial correlation analysis of spectroscopic-imaging data in scanning probe microscopy". In: *Nanotechnology* 085714 (2009).

[213] H. Chang and D. Yeung. "Robust locally linear embedding". In: *Pattern Recognit.* 39 (2006), pp. 1053–1065.

[214] B. J. Rodriguez et al. *Dynamic and Spectroscopic Modes and Multivariate Data Analysis in Piezoresponse Force Microscopy*. 2010. DOI: 10.1007/978-1-4419-7167-8.

[215] A. A. Varambhia, L. Jones, and A. London. "Determining EDS and EELS partial cross-sections from multiple calibration standards to accurately quantify bi-metallic nanoparticles using STEM". In: *Micron* (2018). DOI: 10.1016/j.micron.2018.06.015.

[216] A. Genç et al. "Tuning the Plasmonic Response up: Hollow Cuboid Metal Nanostructures". In: *ACS Photonics* 3 (2016), pp. 770–779.

[217] J. Spiegelberg and J. Rusz. "Can we use PCA to detect small signals in noisy data?" In: *Ultramicroscopy* 16 (2016).

[218] S. Lichtert and J. Verbeeck. "Statistical consequences of applying a PCA noise filter on EELS spectrum images". In: *Ultramicroscopy* 125 (2013), pp. 35–42.

[219] P. Potapov. "Why Principal Component Analysis of STEM spectrum-images results in " abstract ", uninterpretable loadings?" In: *Ultramicroscopy* 160 (2016), pp. 197–212.

[220] P. Potapov, P. Longo, and E. Okunishi. "Enhancement of noisy EDX HRSTEM spectrum-images by combination of filtering and PCA". In: *Micron* 96 (2017), pp. 29–37.

[221] P. Cueva et al. "New Approaches to Data Processing for Atomic Resolution EELS". In: 18 (2012), pp. 970–971.

[222] P. Potapov and A. Lubk. "Optimal principal component analysis of STEM XEDS spectrum images". In: *Adv. Struct. Chem. Imaging* (2019). DOI: 10.1186/s40679-019-0066-0.

[223] A. Tharwat. "Independent component analysis: An introduction". In: *Appl. Comput. Informatics* 17 (2018), pp. 222–249.

[224] J. M. P. Nascimento and J. M. B. Dias. "Vertex component analysis: A fast algorithm to unmix hyperspectral data". In: *IEEE Trans. Geosci. Remote Sens.* 43 (2005), pp. 898–910.

[225] T. Blum et al. "Machine Learning Method Reveals Hidden Strong Metal-Support Interaction in Microscopy Datasets". In: *2100035* (2021), pp. 1–7.

[226] T. Blum et al. "Machine Learning for Challenging EELS and EDS Spectral Decomposition". In: *25* (2021), pp. 180–181.

[227] Y. Suzuki et al. "Automated estimation of materials parameter from X-ray absorption and electron energy-loss spectra with similarity measures". In: *npj Comput. Mater.* (2019), pp. 1–7. DOI: 10.1038/s41524-019-0176-1.

[228] R. D. Leapman and C. R. Swyt. "Separation of Overlapping Core Edges in Electron Energy Loss Spectra by Multiple-Least-Squares Fitting". In: *Ultramicroscopy* 26 (1988), pp. 393–404.

[229] R. Hovden et al. "The Open-Source Cornell Spectrum Imager". In: *Micros. Today* (2013). DOI: 10.1017/S1551929512000995.

[230] M. Ziatdinov et al. "Robust multi-scale multi-feature deep learning for atomic and defect identification in Scanning Tunneling Microscopy on H-Si(100) 2x1 surface". In: (2020).

[231] L. Yedra et al. "Oxide Wizard: An EELS Application to Characterize the White Lines of Transition Metal Edges". In: *Microsc. Microanal.* (2014), pp. 1–8. DOI: 10.1017/S1431927614000440.

[232] M. Shiga et al. "Sparse modeling of EELS and EDX spectral imaging data by nonnegative matrix factorization". In: *Ultramicroscopy* 170 (2016), pp. 43–59.

[233] S. Muto and M. Shiga. "Application of machine learning techniques to electron microscopic - spectroscopic image data analysis". In: (2020), pp. 110–122. DOI: 10.1093/jmicro/dfz036.

[234] N. Bonnet and D. Nuzillard. "Independent component analysis: A new possibility for analysing series of electron energy loss spectra". In: *Ultramicroscopy* 102 (2005), pp. 327–337.

[235] F. de la Peña et al. "Mapping titanium and tin oxide phases using EELS: An application of independent component analysis". In: *Ultramicroscopy* 111 (2011), pp. 169–176.

[236] D. Rossouw et al. "A New Method for Determining the Composition of Core – Shell Nanoparticles via Dual-EDX + EELS Spectrum Imaging". In: *Part. Part. Syst. Charact.* (2016). DOI: 10.1002/ppsc.201600096.

[237] N. Dobigeon and N. Brun. "Spectral mixture analysis of EELS spectrum-images". In: *Ultramicroscopy* 120 (2012), pp. 25–34.

[238] S. V. Kalinin et al. "Automated and Autonomous Experiments in Electron and Scanning Probe Microscopy". In: *ACS Nano* 15 (2021), pp. 12604–12627.

[239] N. Creange et al. "Propagation of priors for more accurate and efficient spectroscopic functional fits and their application to ferroelectric hysteresis". In: *Mach. Learn. Sci. Technol.* 2 (2021).

[240] S. Yang et al. "Deep Learning-Assisted Quantification of Atomic Dopants and Defects in 2D Materials". In: (2021), pp. 2101099, 1–10.

[241] M. Pfannm et al. "Visualizing a Homogeneous Blend in Bulk Heterojunction Polymer Solar Cells by Analytical Electron Microscopy". In: *Nano Lett.* 11 (2011), pp. 3099–3107.

[242] S. Kiyohara et al. "Data-driven approach for the prediction and interpretation of core-electron loss spectroscopy". In: *Sci. Rep.* (2018), pp. 1–4. DOI: 10.1038/s41598-018-30994-6.

[243] T. Mizoguchi and S. Kiyohara. "Machine learning approaches for ELNES / XANES". In: *Microscopy* 69 (2020), pp. 92–109.

[244] P. Torruella et al. "Clustering analysis strategies for electron energy loss spectroscopy (EELS)". In: *Ultramicroscopy* 185 (2018), pp. 42–48.

[245] J. Blanco-Portals et al. "WhatEELS . A python-based interactive software solution for ELNES analysis combining clustering and NLLS". In: *Ultramicroscopy* 232 (2022).

[246] J. Ryu et al. "Dimensionality reduction and unsupervised clustering for EELS-SI". In: *Ultramicroscopy* 231 (2021), p. 113314.

[247] S. V Kalinin et al. "Separating Physically Distinct Mechanisms in Complex Infrared Plasmonic Nanostructures via Machine Learning Enhanced Electron Energy Loss Spectroscopy". In: *Adv. Opt. Mater.* 2001808 (2021), pp. 1–13.

[248] J. Hachtel et al. "Beyond NMF: Advanced Signal Processing and Machine Learning Methodologies for Hyperspectral Analysis in EELS". In: *27* (2021), pp. 322–324.

[249] C. M. Pate, J. L. Hart, and M. L. Taheri. "RapidEELS: machine learning for denoising and classification in rapid acquisition electron energy loss spectroscopy". In: *Sci. Rep.* 11 (2021), pp. 1–10.

[250] P. Ewels et al. "A Complete Overhaul of the Electron Energy-Loss Spectroscopy and X-Ray Absorption Spectroscopy Database: eelsdb . eu". In: *Microsc. Microanal.* (2017), pp. 717–724. DOI: 10.1017/S1431927616000179.

[251] H. L. Xin et al. "One Million EEL Spectra Acquisition with Aberration-Corrected STEM: 2-D Chemical Investigation of a Statistically Significant Ensemble of Nanocatalysts". In: *Microsc. Microanal.* 16 (2010), pp. 2009–2010.

[252] E. J. Kirkland. *Advanced Computing in Electron Microscopy*. 1998. DOI: 10.1007/978-1-4757-4406-4.

[253] F. J. García de Abajo and A. Howie. "Retarded field calculation of electron energy loss in inhomogeneous dielectrics". In: *Phys. Rev. B* 65 (2002), pp. 1–17.

[254] L. Kiewidt and M. Karamehmedovi. *The Generalized Multipole Technique for the Simulation of Low-Loss Electron Energy Loss Spectroscopy*. 2018. DOI: 10.1007/978-3-319-74890-0.

[255] M. P. Oxley and S. J. Pennycook. "Image simulation for electron energy loss spectroscopy". In: *Micron* 39 (2008), pp. 676–684.

[256] A. Mitsutake, Y. Sugita, and Y. Okamoto. "Replica-exchange multicanonical and multicanonical replica-exchange Monte Carlo simulations of peptides. I. Formulation and benchmark test". In: *J. Chem. Phys.* 118 (2003), pp. 6664–6675.

[257] A. Desalvo et al. "Analysis of light elements in superposed layers by Monte Carlo simulation of EELS spectra". In: *Mikrochim. Acta* 114–115 (1994), pp. 267–275.

[258] M. Attarian Shandiz, F. Salvat, and R. Gauvin. "Detectability Limits in EELS by Monte Carlo Simulations". In: *Microsc. Microanal.* 18 (2012), pp. 998–999.

[259] M. Attarian Shandiz, F. Salvat, and R. Gauvin. "Fine Structure of Core Loss Excitations in EELS by Monte Carlo Simulation". In: *Microsc. Microanal.* 19 (2013), pp. 366–367.

[260] J. Verbeeck and S. Van Aert. "Model based quantification of EELS spectra". In: *Ultramicroscopy* 101.2 (2004), pp. 207–224. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2004.06.004. URL: https://www.sciencedirect.com/science/article/pii/S0304399104001433.

[261] J. Verbeeck, S. Van Aert, and G. Bertoni. "Model-based quantification of EELS spectra: Including the fine structure". In: *Ultramicroscopy* 106.11 (2006). Proceedings of the International Workshop on Enhanced Data Generated by Electrons, pp. 976–980. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2006.05.006. URL: https://www.sciencedirect.com/science/article/pii/S0304399106000945.

[262] X. et al. Liu. "Machine learning approach for the prediction of electron inelastic mean free paths". In: *Phys. Rev. Mater.* 033802 (2021).

[263] E. et al. Quattrocchi. "The deep-DRT: A deep neural network approach to deconvolve the distribution of relaxation times from multidimensional electrochemical impedance spectroscopy data". In: *Electrochim. Acta* 392 (2021), p. 139010.

[264] M. P. Oxley et al. "Probing atomic-scale symmetry breaking by rotationally invariant machine learning of multidimensional electron scattering". In: *npj Comput. Mater.* 7 (2021), pp. 1–6.

[265] T. Morimura and M. Hasaka. "Bloch-wave-based STEM image simulation with layer-by-layer representation". In: *Ultramicroscopy* 109 (2009), pp. 1203–1209.

[266] F. Eggert. "EDX-spectra simulation in electron probe microanalysis. Optimization of excitation conditions and detection limits". In: *Microchim. Acta* 155 (2006), pp. 129–136.

[267] NIST. *DTSA-II Microscopium*. https://cstl.nist.gov/div837/837.02/epq/dtsa2/. 2021.

[268] M. Chatzidakis and G. A. Botton. "Towards calibration-invariant spectroscopy using deep learning". In: *Sci. Rep.* 9 (2019), pp. 1–10.

[269] Daniel del-Pozo-Bueno et al. "Comparative of machine learning classification strategies for electron energy loss spectroscopy: Support vector machines and artificial neural networks". In: *Ultramicroscopy* 253 (2023), p. 113828. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2023.113828. URL: https://www.sciencedirect.com/science/article/pii/S0304399123001456.

[270] A. Y. Borisevich et al. "Using Neural Network Algorithms for Compositional Mapping in STEM EELS". In: *Microsc. Microanal.* 15 (2009), pp. 50–51.

[271] D. del Pozo-Bueno, F. Peiró, and S. Estradé. "Support vector machine for EELS oxidation state determination". In: *Ultramicroscopy* 221 (2021).

[272] P. J. Thomas and P. A. Midgley. "Image-spectroscopy – I . The advantages of increased spectral information for compositional EFTEM analysis". In: *Ultramicroscopy* 88 (2001), pp. 179–186.

[273] P. J. Thomas and P. A. Midgley. "Image-spectroscopy – II . The removal of plural scattering from extended energy-filtered series by Fourier deconvolution". In: *Ultramicroscopy* 88 (2001), pp. 187–194.

[274] L. I. Roest et al. "Charting the low-loss region in electron energy loss spectroscopy with machine learning". In: *Ultramicroscopy* 222 (2021).

[275] A. et al. Stevens. "Compressive STEM-EELS". In: *Microsc. Microanal.* 22 (2016), pp. 560–561.

[276] E. et al. Monier. "Fast reconstruction of atomic-scale STEM-EELS images from sparse sampling". In: *Ultramicroscopy* (2020). DOI: 10.1016/j.ultramic.2020.112993.

[277] J. et al. Schwartz. "Recovering Chemistry at Atomic Resolution using Multi-Modal Spectroscopy". In: *Microsc. Microanal.* 27 (2021), pp. 1226–1228.

[278] S. M. et al. Collins. "Scan Strategies for Electron Energy Loss Spectroscopy at Optical and Vibrational Energies in Perylene Diimide Nanobelts". In: *Microsc. Microanal.* 25 (2019), pp. 1738–1739.

[279] R. Huber et al. "Total generalized variation regularization for multi-modal electron tomography". In: *Nanoscale* 11 (2019), pp. 5617–5632.

[280] L. et al. Yedra. "EEL spectroscopic tomography: Towards a new dimension in nanomaterials analysis". In: *Ultramicroscopy* 122 (2012), pp. 12–18.

[281] Z. et al. Saghi. "Compressed sensing electron tomography of needle-shaped biological specimens - Potential for improved reconstruction fidelity with reduced dose". In: *Ultramicroscopy* 160 (2016), pp. 230–238.

[282] P. et al. Torruella. "3D visualization of iron oxidation state in FeO/ Fe3O4 core-shell nanocubes from electron energy loss tomography". In: *Nano Lett.* 16 (2016), pp. 5068–5073.

[283] A. Skorikov et al. "Deep learning-based denoising for improved dose efficiency in EDX tomography of nanoparticles". In: *Nanoscale* 13 (2021), pp. 12242–12249.

[284] R. Lin et al. "TEMImageNet training library and AtomSegNet deep-learning models for high-precision atom segmentation, localization, denoising, and deblurring of atomic-resolution images". In: *Sci. Reports 2021 111 11* (2021), pp. 1–15.

[285] K. Simonyan and A. Zisserman. "Very deep convolutional networks for large-scale image recognition". In: *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.* 2015, pp. 1–14.

[286] S. et al. Gražulis. "Crystallography Open Database (COD): An open-access collection of crystal structures and platform for world-wide collaboration". In: *Nucleic Acids Res.* 40 (2012), pp. 420–427.

[287] S. et al. Gražulis. "Crystallography Open Database - An open-access collection of crystal structures". In: *J. Appl. Crystallogr.* 42 (2009), pp. 726–729.

[288] machrisaa. *GitHub - machrisaa/tensorflow-vgg: VGG19 and VGG16 on Tensorflow — github.com*. https://github.com/machrisaa/tensorflow-vgg. [Accessed 07-03-2024].

[289] rcmalli. *GitHub - rcmalli/keras-vggface: VGGFace implementation with Keras Framework — github.com*. https://github.com/rcmalli/keras-vggface. [Accessed 07-03-2024].

[290] GitHub. *GitHub: Let's build from here — github.com*. https://github.com/. [Accessed 07-03-2024].

[291] Microsoft. *Cloud Computing Services | Microsoft Azure — azure.microsoft.com*. https://azure.microsoft.com/en-us/. [Accessed 07-03-2024]. 2024.

[292] Google. *Cloud Computing Services | Google Cloud — cloud.google.com*. https://cloud.google.com/. [Accessed 07-03-2024]. 2024.

[293] Google. *Google Colaboratory — colab.research.google.com*. https://colab.research.google.com/. [Accessed 07-03-2024]. 2024.

[294] IBM. *IBM Cloud — ibm.com*. https://www.ibm.com/es-es/cloud. [Accessed 07-03-2024]. 2024.



CHAPTER 2: METHODOLOGY

Chapter 2

Methodology

The ambitious goals presented in the introduction require both specific existing methodologies and the development of new ones. Chapters 4 and 5 focus on the new workflows we propose for addressing these challenges. Thus, the final purpose is to tackle the introduced scientific questions, and the pioneering or improvement of the methodologies is our tool to face them. However, the solutions we defend are based on existing experimental and computational resources. Therefore, this chapter aims to justify the choices of the techniques and methods that are deployed throughout the thesis.

2.1 Scope of the chapter

An attentive reading of this chapter is rewarded with the necessary knowledge to logically follow the thesis dissertation. We present the methods, techniques, algorithms, software packages and tools we have used to go through the science contained in the next chapters. However, the text is written as a rational argumentation that justifies the choices we have made when selecting the most appropriate toolbox for our ends. Therefore, particular fundamentals of the techniques are described to guarantee the required knowledge to understand the thesis. Nevertheless, the attentive reader is invited to check the references we pinpoint for the detailed and exhaustive study of the introduced techniques. As a result, we do not overlap the missions of each dedicated document, and we can focus on the key aspects we exploit.

2.2 Experimental methods

Our ultimate goal is to understand how the functional properties of the devices correlate with their structural properties. This knowledge is fundamental to iteratively

improve the design of the devices towards the attainment of these functional properties. Thus, to unveil the structural properties of the devices we are targeting we need a tool capable of offering enough spatial resolution to identify its building blocks. If an atom can be approximated to the Angstrom realm, our characterisation tools must go beyond. This constraint already rejects plenty of far-away candidates like conventional optical microscopy, and others close to the edge like atom probe tomography. Consequently, the best options left on the table are Transmission Electron Microscopy (TEM) and Scanning Probe Microscopies (SPM). Nevertheless, SPMs like Scanning Tunneling Microscopy (STM) or Atomic Force Microscopy (AFM) limit the interaction with the sample to its surface. 2D materials and devices with surface effects would greatly benefit from this approach, but its extent would not reach much further. That is why, our preferred choice is TEM. Note that SPMs and complementary techniques would surely add value in future correlative experiments, but the presented research is adequately based solely on TEM.

TEM can reach spatial resolutions of up to 14 pm and energy resolutions below 5meV. [1, 2, 3, 4] These capabilities enable a wide toolbox for complete experiments that perfectly answer our needs. Nevertheless, the main limitations of TEM are the sample preparation and its requirement of electron transparency, and its low throughput or, in other words, its limited local reach. The present work will just live along with the first limitation. On the other hand, we will use the second one in our favour to propose a path towards a high-throughput TEM that no longer supposes the traditional bottleneck it typically is, and converts the TEM into a viable technique for seeking statistical significance. The TEM pathway typically follows the steps of 1) imaging, 2) spectroscopy, 3) strain analysis, and 4) modelling. Then, let us present the TEM toolset we have used in the thesis, classified according to the type of data they generate.

2.2.1 Imaging in the TEM

Imaging is perhaps the most paradigmatic feature of microscopy. Usually, images are representations of the reality which we, as humans, feel comfortable with. Interestingly, we have also made computers feel comfortable with them, generating the nowadays vital field of computer vision, somehow justifying the present work. [5]

TEM has evolved so much that manifold imaging modes and approaches can be easily accessed with reasonably modern equipment. [6] From regular parallel illumination mode to scanning mode and the broad range of detected signals they can lead to. All of them offer distinctive features that make them preferential over the others for certain purposes. Interestingly, the questions we need to answer for our target systems just require the use of regular detection modes enabling atomic resolution. That is, the currently hotter imaging modes like 4D-Scanning Transmission Electron Microscopy (STEM) and relatives like ptychography would mostly be overkill for our needs. [7] With the proper processing, they would surely suppose a resolution

improvement and an alternative, for instance, in strain measurements. However, the associated data processing is much more complex due to the high dimensionality of the data and its consequent huge data sizes (several gigabytes). As a result, as mentioned, we will see that regular TEM and STEM are the perfect trade-off between resolution and data complexity. This applies to the direct and manual advanced characterisation of nanodevices, as has been regularly done for many years, and as it is presented in chapter 3.

In addition, the complexity vs knowledge access trade-off also applies to the workflow for automated data analysis and modelling, and remarkably, to its democratisation. The proposed workflow, thoroughly explained in chapter 4, relies upon operational modes which are relatively simple to access. Almost any materials scientist has access to the modes we have used, as state-of-the-art equipment of 2010 already offered the possibilities we exploit. The idea of democratisation is utterly important, as with this approach, we allow the whole materials science community to the immediate access to our proposed solution and its application to their daily life research problems. A solution that not only provides access to unprecedented levels of modelling accuracy but also enhances the throughput of advanced characterization for their devices.

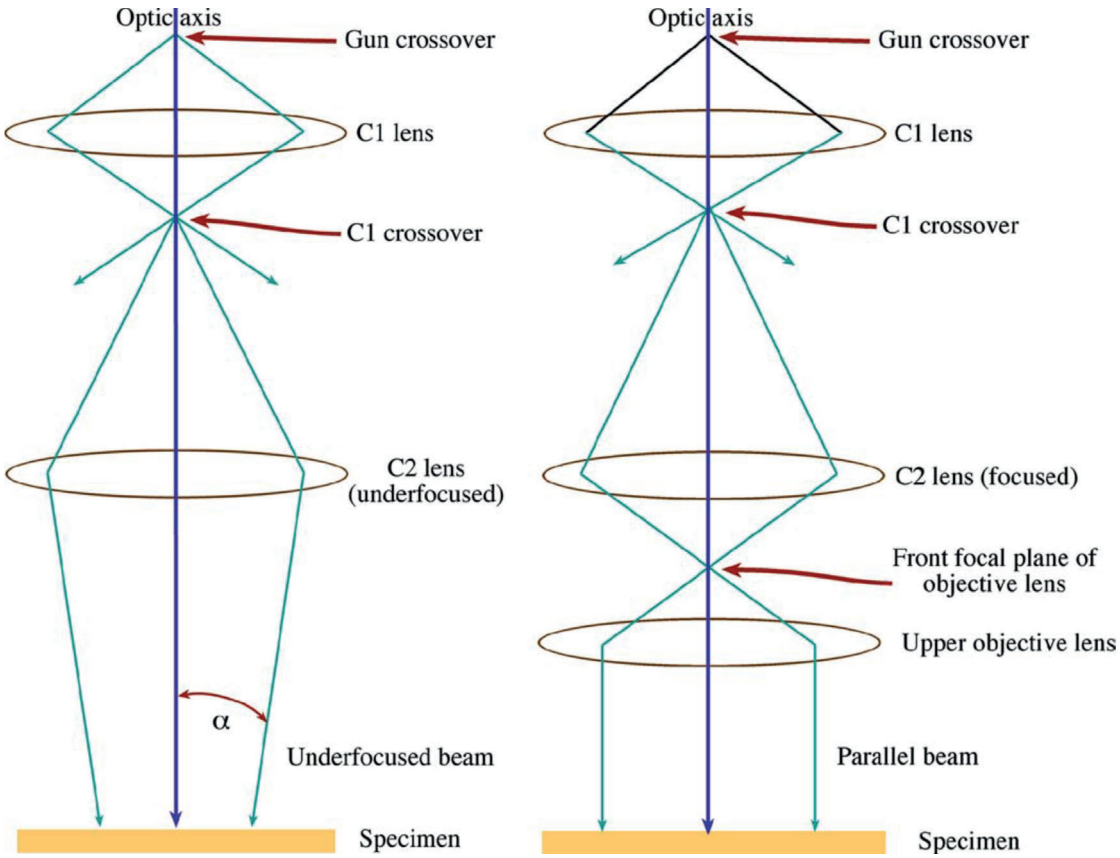


FIGURE 2.1: Ray diagrams of a Transmission Electron Microscope (TEM) operating at parallel illumination conditions. At left, the condenser lenses C_1 and C_2 cannot produce a parallel beam unless, as in the right diagram, an objective lens is introduced before the specimen. Adapted with permission from [6].

The operational modes we exploit have a resolution limit which perfectly fits the needs of the advanced characterisation of materials through imaging. That is, a regular Aberration Corrected (AC) TEM with parallel illumination leads to state-of-the-art resolutions of 80 pm and 50 pm at 200 and 300kV of accelerating voltage, respectively. [8] To achieve this, we use correctors that mostly affect the spherical aberration, C_s, which stands for the lenses that have a distribution of focal planes as a function of the rays' crossing points within the lens. C_s is the major source of image blurring in the electron microscope, and its correction meant a breakthrough in electron microscopy. Generally speaking, the improvement enabled unprecedented resolution achievements. Aberration correctors do not only affect C_s rounded aberrations but also enable to correct the remaining non-rounded aberrations like coma or astigmatism (especially in high orders). In fact, the apparition of aberration correctors was so crucial that there is still a huge joint scientific and engineering effort to continuously enhance the traditional approach of electromagnetic lens-based aberration correction, but also the emergence of new promising and democratising perspectives like phase plates. [6, 9, 10, 11, 12, 13, 14]

In particular, in parallel illumination mode, the lens contributing the most to the signal is the objective lens (i.e., the main image-forming lens) (Figure 2.1). Therefore, the aberration correction compensates the aberrations of the objective lens, in the so-called image-corrected microscopes. Moreover, the use of image correctors helps in reducing phase delocalisation and attaining sharper interfaces. Alternatively, this phase-contrast governed mode can also offer atomic resolution without aberration correctors, for which its reach is even major. Of course, as mentioned, the image quality will not generally be comparable.

In any case, high-resolution TEM mode allows us to quickly access atomic resolution and unveil the structural features of our electron transparent samples. Importantly, the outcoming signal, governed by Bragg diffraction, considers the full thickness of the thin sample, averaging the information throughout. Furthermore, the interpretation of the contrast captured by high-resolution TEM micrographs is not obvious and requires more than a trained eye. What we acquire is the convolution of the beam phase with the crystalline lattice. Consequently, the signal is tightly bound to the fine-tuning of our beam conditions. In other words, the appearance of distinctive features like the atoms themselves will shift its contrast from dark to bright following a Contrast Transfer Function (CTF) which oscillates from negative to positive values as a function of the defocus. [6] As a result, the correct interpretation of the images demands a non-intuitive simulation considering the nature of the imaging system and conditions, its CTF and its convolution with the candidate model representing the structure we need to validate. Noticeably, this idea brings to the table a valuable characterisation and model validation routine. Nevertheless, it might not be the best candidate for the automation of certain data analysis routines implying,

for instance, atomic position detection and contrast-based analysis. [15, 16] Fortunately, our automated workflow does not rely on these principles. Although it is developed and built with STEM mode as the central cornerstone (i.e., as the main source of information), we managed to also successfully generalise it to TEM and the *a priori* harder to process contrast it offers (see chapter 4).

The access to atomic resolution allows to bypass going through diffraction mode. That is, instead of changing the way we project the diffraction pattern to acquire it directly, we take an image and we process its discrete Fourier transform. This process is typically done through the algorithm of the Fast Fourier Transform (FFT). [17] Strictly speaking, the diffraction pattern and the FFT are not the same, although they both provide crystallographic information. Diffraction mode is preferred when the crystal phase is unknown and we want to determine its structure. On the other hand, the FFT is more pragmatic when we require the crystal phase identification from a database. Thus, comparing the experimental FFT with candidate crystal phases which allow us to explain our data and fit their models. In fact, this idea is one of the cornerstones of chapter 4. In summary, to reduce the complexity of the associated data processing, we prefer a single image and its processed FFT rather than acquiring an image plus a diffraction pattern containing almost the same information.

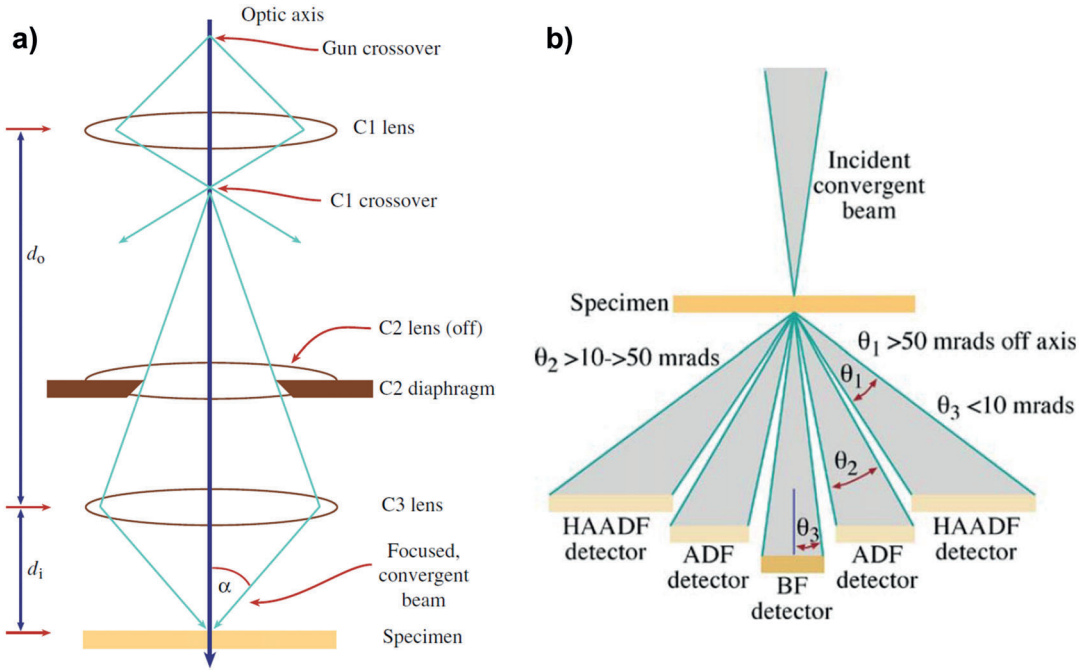


FIGURE 2.2: Schematics representing the fundamentals of Scanning Transmission Electron Microscopy (STEM). a) Ray diagram of the formation of a probe converging on the sample because of the action of the condenser lenses. b) The spatial distribution of the main available STEM detectors and their characteristic angular ranges defining the scattered electrons that will be detected by each of them. Adapted with permission from [6].

On the other hand, instead of shaping the beam to illuminate the sample parallelly, we can create a convergent probe that scans the sample laterally, producing a signal

at each beam position. This mode is STEM, and it is the main imaging mode upon which we rely to retrieve the results described throughout the thesis. Both for the manual advanced characterisation (i.e., chapter 3), and as the source of data with which to feed the automated workflows we propose (described in chapters 4 and 5). The main reason is the contrast generation mechanism of the base detection mode we use. Thus, through the High-Angle Annular Dark-Field (HAADF) detector, the signal we get is approximately proportional to the square of the atomic number, Z of the imaged element, and the thickness of the sample. Thus, we capture the electrons elastically scattered at a high angle through Rutherford scattering, which is directly dependent on the mass of the atoms' nucleus. [6, 18] For example, if we can resolve the atoms, then we can distinguish their element type if the contrast difference in intensity is captured within our dynamic range. In fact, atomic resolution in STEM is more demanding than in TEM. Although STEM can achieve atomic resolution without correctors, requiring optimised narrow pole pieces, or cold field emission guns reducing the chromatic blurring, as well as careful alignments (or a combination of all of them), aberration correction revolutionised STEM and has converted it into the nowadays leading electron microscopy characterisation technique. That is, aberration correctors allow the formation of an aberration-free probe size down to 50 pm at 300kV, therefore systematically enabling a high-resolution analysis almost unattainable without the correctors or through advanced imaging approaches such as ptychography. [19, 20] In the present case, we refer to probe correction, as the aberration correction compensates for the aberrations of the main probe-forming lenses, the condenser lenses. That is why, accessing high-resolution STEM might be logistically more challenging than for TEM, as aberration correction is still costly.

As briefly indicated at the beginning of the section, manifold detection approaches can be accessed with STEM. As seen with HAADF, they are paired with the detector from which they arise, and have minor differences from each other. The Z-contrast differences between device regions provided by our preferred mode, HAADF, will be a major key to making our image analysis automated workflow successful. On the other hand, we can find complementary detectors. For instance, the Annular Bright-Field (ABF) detector is more prone to capturing lighter atoms like oxygen or nitrogen due to its detection of electrons scattered at low angles, which are sensitive to phase contrast. [21] Consequently, it complements HAADF with further details about the in-depth nature of the sample. In fact, although essential for HAADF, the concept of depth of focus (i.e., sampled real space along the optical axis) dilutes for bright-field imaging due to the low-angle scattering and the phase contrast. That is, ABF can be seen as an intermediate mode between STEM-HAADF and parallel illumination mode. Interestingly, both correction paradigms can be merged to build the so-called double-corrected electron microscopes, which gather in one single machine both the probe and image correctors. Generally, having two microscopes with separated image and probe corrections is more robust and convenient than merging both corrections in a single microscope. Thus, image correction in STEM can cause

unwanted artefacts in the scattered signal. Indeed, image correctors are placed after the sample, therefore interacting with the scattered electrons and potentially distorting the physics to be found from the specimens. In this case, we want the output to be as raw as possible from the specimen, while image correctors interfere with this will. On the other hand, probe correction will add no value to parallel beam mode, as its purpose of reducing probe size is no longer required.

To reproduce our results and to make use of the proposed automated workflow we essentially exploit the imaging benefits of probe aberration-corrected STEM-HAADF and its capacity to correlate intensities with atoms, to distinguish structural defects, and to more intuitively interpret the overview of the devices. This idea is perfectly exemplified in chapter 3, when distinctive features of the micrographs are highlighted. Additionally, chapter 4 showcases how certain knowledge could only be retrieved thanks to the nature of the mentioned signal.

In TEM, one must always consider that the imaged objects represent 2D projections of a 3D reality. This holds true unless an electron tomography experiment is specifically conducted. In such cases excepting tomography, our goal might be to capture an image of a material along a certain zone axis. When imaging along a zone axis, we obtain not merely a single atom but rather the projection of the entire atomic column within the depth of focus of the operational mode. [6] This concept bears significant importance in advanced characterisation workflows, ensuring that our conclusions remain accurate. However, this consideration becomes critical when generating 3D models with our automated data analysis workflow (Chapter 4). If the projection appears homogeneous across the entire image, we can assume it represents the whole thickness of our device. Consequently, we invoke the principle of translational invariance. By "homogeneous", we mean the absence of abrupt intensity variations between atomic columns. Such variations could indicate uneven atom counts, defects, vacancies, or similar features. Naturally, we must be careful when accepting this assumption in these challenging scenarios. The morphology and symmetry of the device play a crucial role in validating our hypothesis. In practice, translational invariance typically holds true for samples presented as thin lamellae - electron-transparent slices of a device - prepared using techniques like Focused Ion Beam (FIB). These lamellae typically have a thickness ranging from 20 to 50 nm.

As we briefly pinpointed before, from the data complexity perspective, the regular TEM and STEM modes are also justified. They release images, which are just 2D matrices with a single complex number in every pixel (i.e., they typically are real numbers, but in a Fast Fourier Transform (FFT), they can be complex). The processing of these images or 2D matrices can reach unimaginable levels of complexity, but it can also be simple and fast. Nonetheless, higher dimensionality generally implies that the simplest available processing methods will inherently be tedious and slow.

Furthermore, the abundance of open-access and open-source tools designed for image processing positions them as the optimal candidates for handling software implementations that are customized and fine-tuned to meet our specific requirements. [22]

2.2.2 Spectroscopy in the TEM

TEM imaging is commonly not enough to completely resolve the characterisation inquiries of our devices. The complete structural characterisation compels a minimum of compositional information. As we have seen, the most immediate source of compositional information is brought by the *Z*-contrast retained by the HAADF detector. However, it is more sensitive to large atomic numbers than to small ones. As a result, if atoms like N or O are present in the sample, or more generally, if two different elements with a significant atomic mass difference are imaged simultaneously, the compositional information and even the pure resolving of the atoms can be lost. A paradigmatic example of this effect can be noticed when imaging certain 2D systems like transition metal dichalcogenides, such WS_2. Interestingly, it is likely that the heavy W atom, (i.e., the transition metal) shifts the dynamic range to an extent that the light S atom is no longer distinguishable beyond the noise levels under certain imaging conditions. [23, 24] In these cases, though, smaller and less chromatically blurred probes can circumvent this limitation by separating the scattering from neighbouring W and S atoms and consequently separate them in the micrograph, still in HAADF. For instance, employing a cold field emission source would aid in reducing the chromatic blur associated with the low-voltage imaging necessary to prevent knock-on damage in such samples.

Outside HAADF domains, we can reduce the detected scattering angles and choose alternative imaging modes like Integrated Differential Phase Contrast (iDPC), which does not have a square but a linear dependence on the atomic number, and therefore gentler conditions for imaging light elements. [25]. However, although imaging modes like iDPC can provide considerable value, especially in thin samples, the general answer lies in spectroscopy.

Spectroscopy essentially means to measure the energy of the electrons (or complementary signals), rather than where their scattering leads them (i.e., imaging). Excitingly, spectroscopy opens a wide toolbox which facilitates the measurement of compositional information, but also of physical properties which directly rule the behaviour of macroscopic devices. In fact, we use electron spectroscopy as the main source of compositional knowledge for chapter 3, as a complementary compositional refiner together with the *Z*-contrast of HAADF in chapter 4, and as a resource to access physical properties like the bandgap or the plasmons, as explored in Chapter 5. The combination of imaging and spectroscopy allows us to comprehensively characterise the structural properties and even go further providing direct functional

properties. This thorough characterisation enabled by aberration correctors, state-of-the-art energy resolutions and new-generation detectors has popularised the term "synchrotron in a microscope", specially referred to when the TEM operates in scanning mode. [3, 4, 7, 25, 26, 27]

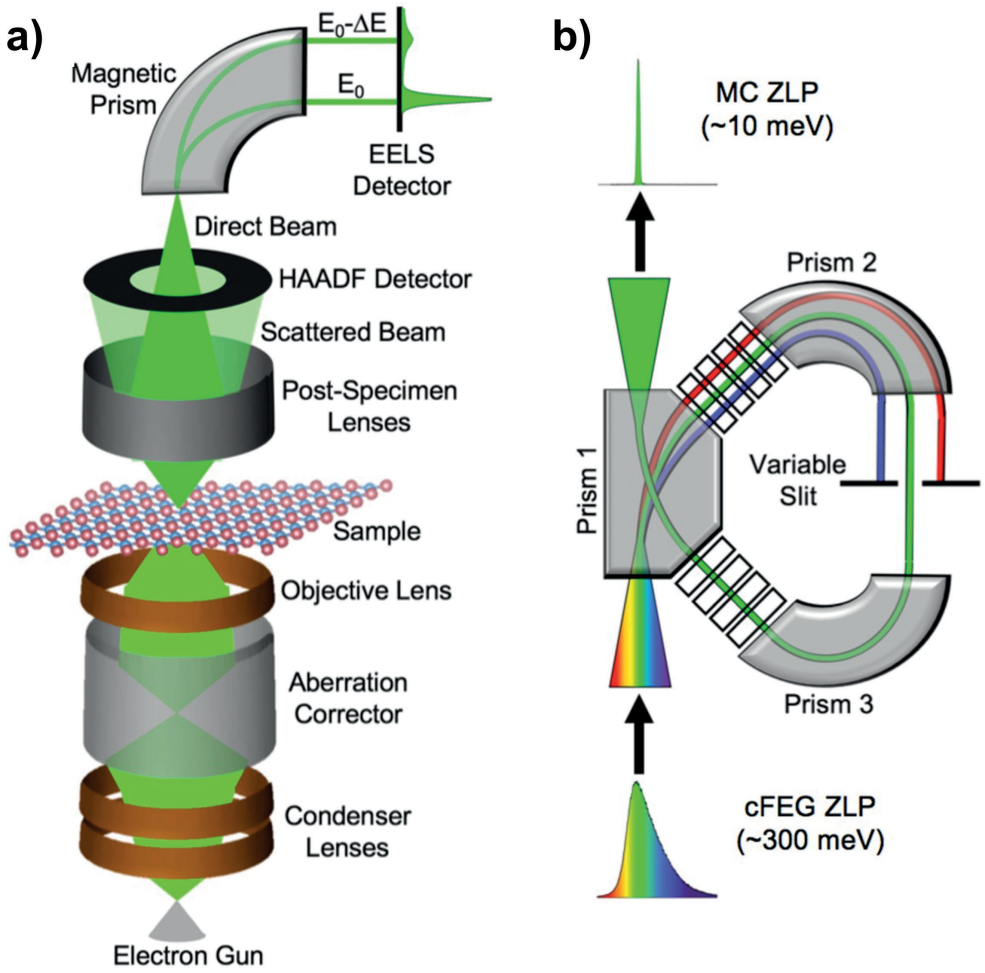


FIGURE 2.3: Schematics of an a) Electron Energy Loss Spectroscopy (EELS) acquisition when operating in Scanning Transmission Electron Microscopy (STEM) mode. Panel b) showcases the working principle of the monochromator and how it enhances the resolution of a cold field emission gun. Beware the electron source being at the bottom of the microscope in the depicted machine. Adapted from [28].

The electron spectroscopies are EFTEM (Energy-Filtered TEM), Energy Dispersive X-ray spectroscopy (EDX), CathodoLuminescence (CL), and Electron Energy Loss Spectroscopy (EELS), Auger Spectroscopy and Electrons Backscattered Diffraction (EBSD). They all rely upon the same physical origin, which is the inelastically scattered electrons traversing the sample. Thus, their differences lie in the electron/photon/X-ray optics detecting the signal, and the nature of the signal. For instance, in both EFTEM and EELS, we directly measure the electrons which have lost a certain amount of energy, ΔE, although EFTEM generally works with parallel illumination and EELS with scanning mode (fig. 2.3,a). EDX, CL and EELS work in scanning mode. In CL and EDX we detect photons in the visible and

X-ray spectra respectively. In fact, the detected X-rays arise from the relaxation of the electronic states after they are excited by an incoming electron. In fact, this electron will lose energy in this interaction, which correlates with the energy we measure for the electron in EELS. The exact energies and their relation will depend on the characteristic binding energies of each excited and relaxed energy level in the interacting atom. On the other hand, CL gauges excitations from valence to conduction band and their relaxation arising from secondary excitations from the primary inelastic excitations that generate the EDX and EELS signals. As secondary events, their probability to occur is lower and therefore CL is a typically lower signal technique. [6]

In the present thesis dissertation, we only use STEM-EELS as the source of spectroscopic information. The arguments behind this choice are manifold. First and foremost, the joint access to both low-loss (i.e., 0-100 eV) and high-loss (i.e., $\geq$ 100 eV) features. In other words, the simultaneous access to physical and chemical information, respectively. In addition, EELS provides the best joint spatial-energy resolution. That is, EDX is spatially limited by the volume of X-ray emission, while EFTEM and EDX are spectrally limited to elemental composition maps, but not oxidation states. In the case of CL, we can retrieve bandgap and other optical and photonic effects, although the limiting fact is the intensive dose required which limits the resolution and the susceptible samples. EELS, on the other hand, offers these functionalities at a spatial resolution essentially limited by the probe size, together with more complex phenomena like channelling. Moreover, unlike EFTEM, we can locally map a whole spectrum of signals simultaneously, allowing a more flexible data analysis approach. [29, 30, 31]

Indeed, from the data point of view, there are important differences concerning the images. Spectroscopy can generate single 1D spectra, 2D spectral profiles, and 3D spectral images. The latter ones are the most informative, yet the most complex to process. These 3D data volumes have a full 1D spectrum at every pixel or beam position, which allows us to map the distribution of spectral features along the image. For Chapters 3 and 4, the features representing compositional information are the edges generated when exciting the electrons of the different energy levels of the atoms with which they scatter (K edges for principal quantum number n=1 electrons, L for n=2, M for n=3, and so on). [6, 31] Particularly, in this thesis, we only consider the overall morphology of these edges for the regular elemental mapping. Thus, near-edge features accounting for oxidation state and chemical environment are beyond our needs and the scope of Chapters 3 and 4. What is more, for Chapter 5, the features we exploit are physical excitations like bulk and surface plasmons, phonons, and density of states. They induce inelastic scattering with electrons, which appear as edges or peaks since the very few meV ranges. [6, 31, 32, 33, 34, 35]

With spectral images (or spectrum images), we thoroughly exploit the lateral resolution of the microscope and its spectral resolution. Interestingly enough, the spatial

resolution will be defined by how we map the real space. In other words, by the spacing or jumps between probe positions. In other words, by our pixel size. On the other hand, spectral resolution is initially defined by the nature of our electron gun. For instance, a cold field emission gun will have a better default energy spread than a thermionic gun (i.e., 0.3eV vs 1.1eV, respectively) (figure 2.3, b). [36, 37] However, the central player in defining the spectral resolution will be the monochromator. Technically, there is not a unique approach for monochromating an electron beam. Nevertheless, it essentially consists of a mechanical slit which opens and closes to filter and permit the electrons within the desired energy spread. Although it depends on the microscope-specific energy dispersion along the slit, as a rule of thumb we consider that the energy spread is reduced by the same factor we reduce the total current through the slit (fig. 2.3, b). This can lead to the state-of-the-art resolutions we mentioned before. [3, 4, 11] In Chapters 3 and 4, our demands for energy resolution are soft, as we just search for the presence of intense compositional edges. However, in Chapter 5 we pushed the resolution far enough to resolve bandgaps in the low-loss regime, at energy domains between 0.5 and 10 eV. Remarkably, to attain our goals, we reached energy resolutions of 6 meV. [38] Note that the energy resolution is convened as the full width at half maximum of the zero-loss peak acquired in vacuum. Finally, we briefly indicate that an additional benefit of monochromating the electron beam is the reduction of the chromatic blur. This is complementary to the correction of the chromatic aberration, C_c, although monochromation always implies intensity loss. Consequently, by monochromating we improve the resolution at the expense of signal-to-noise ratio, because of the forced reduction in the total current not undergone when just correcting C_c. [6] Few microscopes are equipped with C_c correctors as their main chromatic blur reduction factor when imaging. Particularly, they are useful for EFTEM imaging experiments where losing counts by monocrhomating is not an option.

2.2.3 Strain calculations

Once we obtain the description of the crystalline arrangements within the analysed devices, our goal is to compute the local strain distribution that corresponds to the observations. To achieve this, we compare the expected atomic arrangement in a completely relaxed material with the actual measured arrangement. If no differences are detected, it implies that no strain is present. It's crucial to recognize that strain inherently implies change. It can be viewed either as a temporal evolution (before and after a stimulus) or as a condition change (bulk material versus material in a heterostructure). The significance lies in the fact that TEM imaging inherently misses this dynamic process. When we capture an image, it represents an instant in time, unless *in-situ* experiments are carried out. This instant corresponds uniquely to the final state of the material after the influence of specific driving forces, such as lattice mismatch in a heterostructure. Therefore, the strain we measure is relative to the information contained within the image, which serves as our sole source of

data. To convert this relative or image-specific strain into absolute strain, we must characterise the previous or original state of the materials. Reasonably, we assume that every material was originally in its bulk and fully relaxed state, as reported in the literature for its corresponding crystal phase. By balancing both states, we can quantify the absolute strain as a measure of the material's temporal evolution.

In Chapter 3, we delve into the concept of absolute strain. Here, our focus lies on understanding how the growth of heterostructures impacts the material. Strain plays a key role in shaping the functional properties of the devices characterised in Chapter 3, specifically influencing the splitting of their quantised energy levels. [39] On the other hand, we also consider the concept of relative strain, particularly the mapping of displacements. This approach proves sufficient for accurately modelling absolute strain within the atomic models generated by our automated data analysis workflow. Interestingly, this relies on the automated retrieval of crystal phases, providing context for the relaxed state of these phases. More details can be regarded in Chapter 4.

There are multiple ways to retrieve the relative strain in TEM. For instance, 4D-STEM and convergent beam electron diffraction have emerged as powerful tools to map the strain even in large fields of view. [40, 41] Nevertheless, as happened when evaluating imaging, the data processing is no longer tackling a 2D data element but a 4D one. To circumvent the processing complexity, one can instead rely on the regular TEM imaging described in section 2.2.1 and its post-processing to compute strain. Geometrical Phase Analysis (GPA) and Peak Pairs Analysis (PPA) are the most common algorithms to locally map the strain components of the 2D linear strain tensor (equation 2.1).

$$\varepsilon_{ij} = \begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} \\ \varepsilon_{yx} & \varepsilon_{yy} \end{pmatrix} \tag{2.1}$$

These two methods are similar despite their particularities. Their first requirement, is, of course, atomic-resolution images from which to extract the crystalline lattice information. As in TEM we compute relative strain or displacements, they both rely on defining a reference from which to compute these relative displacements. As a result, the final result is the mapping of how different the lattice is when compared to the reference. The main difference then, lies in how each method computes these differences. GPA maps the geometrical phase variations from reciprocal space. In other words, it sets a reference interplanar distance in reciprocal space (i.e., reciprocal space vector, $\vec{g}_i$) and computes the location and degree of the deviations from its central values. [42, 43] Finally, the reciprocal distances are mapped back to real space, where the relative displacements can be converted into absolute strain if the prior information is known. On the other hand, PPA does not need to step into the reciprocal space. PPA evaluates the real space coordinates of the resolved atoms and

maps their positional differences concerning the reference atomic lattice. [44, 45, 46] As with GPA, we retrieve relative strain, which again, requires prior knowledge to be converted into absolute values.

The post-processing of both GPA and PPA is straightforward. They both are few-step processes which quickly end up in favourable results after a manageable parameter tuning. Moreover, the interpretability of the results is direct as soon as its core idea of reference setting is understood. Interestingly, this is a welcome consequence of setting off from a 2D image instead of a 4D data object. Therefore, we reinforce opting for GPA and PPA instead of 4D-STEM.

When comparing GPA and PPA, the main difference is the field of view in which they excel. PPA is better considered for small fields of view with reduced pixel sizes that sample every atom with several pixels. This allows the proper identification of the atomic coordinates and a precise mapping of the relative displacements (in principle infinite resolution with an infinite sampling of the real space) (fig. 2.4, PPA). Consequently, nanoparticles and snippets of 2D materials are outstanding candidates for applying PPA. [45]

On the other hand, GPA shines in larger and intermediate fields of view, as soon as atomic resolution is accomplished and a big enough reference can be delimited. This broadens the reach of GPA to larger devices. Larger devices like those we are targeting in the present dissertation, which although they can present few-nm features, might extend for several microns. For instance, a 4K micrograph could enable the strain mapping of a maximum FOV of 512 nm, considering a lattice spacing of 2.5 angstroms and the minimum Nyquist sampling. That is why GPA ends up being our preferential tool for the strain calculation of samples which will be typically presented as thin lamellae. However, GPA steps into reciprocal space for its analysis, making the definition of its precision less obvious than in PPA. This concept will be paramount, especially in Chapter 4. It essentially reduces to defining a mask around the FFT peaks of interest ($\vec{g}_i$), whose size will be inversely proportional to the achieved precision in the real space strain maps (fig. 2.4, GPA, b, red and blue circumferences around $\vec{g}_1$ and $\vec{g}_2$). In other words, opening a wider mask captures a larger portion of reciprocal space, unlocking then a better resolution while simultaneously increasing the noise. Contrarily, shrinking the mask reduces the noise as most considered reciprocal space pixels arise from the considered crystal planes, although finer details scattered outside our mask and outside our reference $\vec{g}_i$ will be lost.

Interestingly enough, we will apply GPA in all the devices we analyse in Chapter 3, and applies for most of the development in Chapter 4. Remarkably, Chapter 4 showcases the validity of GPA when applied to tiny fields of view, and even glimpses the

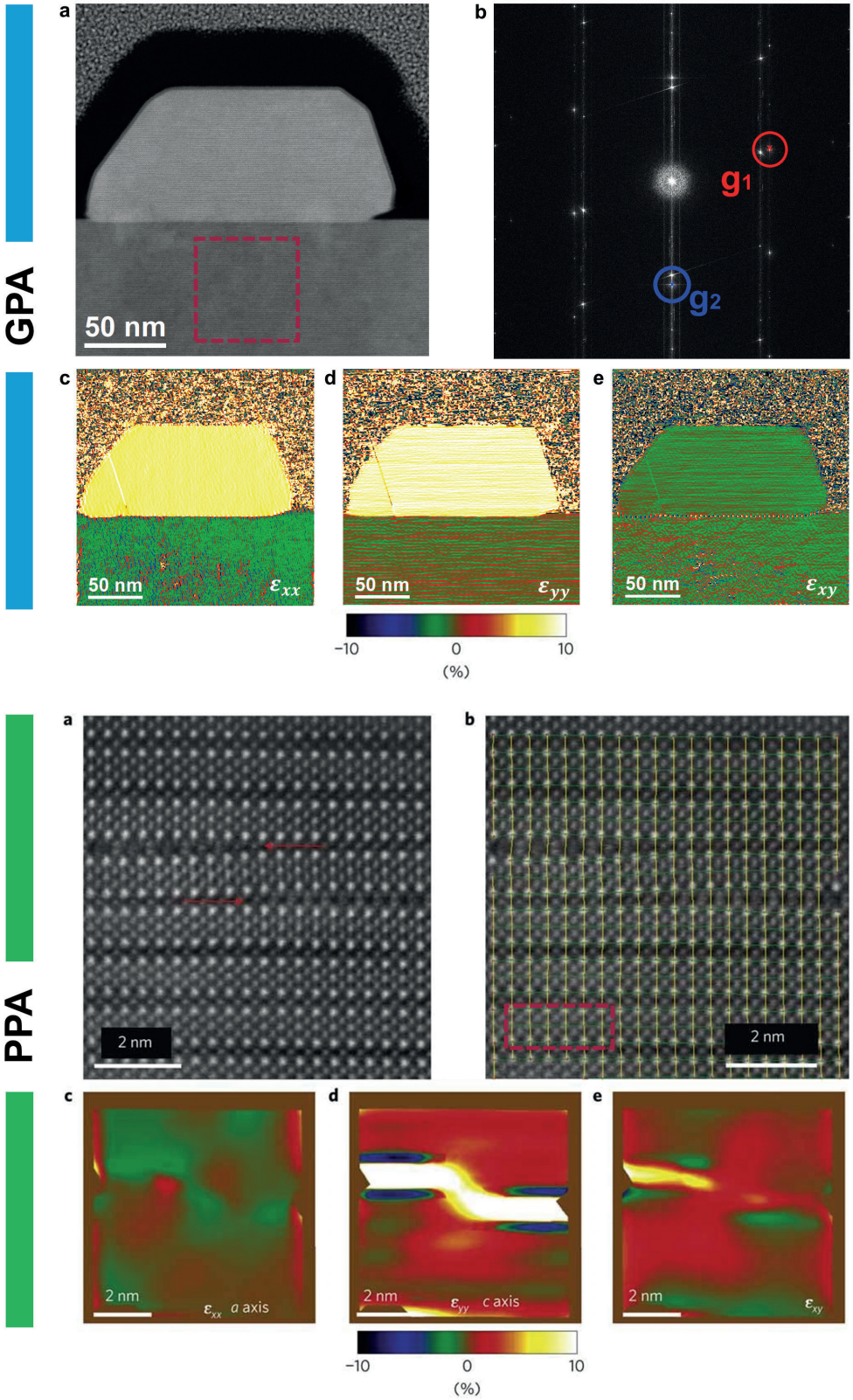


FIGURE 2.4: Comparison between Geometrical Phase Analysis (GPA) and Peak Pairs Analysis (PPA). The top panel represents the GPA analysis in a (a) 200 nm wide InSb nanowire grown on InP representing large fields of view and the reference region in red, (b) the choice of the reciprocal space vectors and the resulting (c) ε_{xx}, (d) ε_{yy} and (e) ε_{xy} components of the linear strain tensor. The bottom panel shows the same analysis through (b) PPA, on a (a) 10 nm crop of YBCO, and the resulting (c) ε_{xx}, (d) ε_{yy} and (e) ε_{xy} strain components. Adapted with permission from [46].

extension of GPA to unexpectedly large regions. Furthermore, in Chapter 4 we detail the achievement of a thorough automation of GPA for its accounting of the type and arrangement of the materials and their specific parametric needs. The mentioned differences are highlighted in two paradigmatic examples in Figure 2.4, with GPA applied in a 200 nm wide InSb nanowire grown on InP representing large fields of view, and a 10 nm crop of YBCO analysed through PPA.

GPA is a sensitive computational method capable of offering high spatial resolutions. Consequently, it is also sensitive to noise. Therefore, special attention must be paid to the data acquisition process to ensure its noise levels are kept to a minimum. For instance, STEM acquisitions should prioritise the reduction of scanning-related noise and artefacts through, for example, fast serial acquisitions, or sampling rates the larger than Nyquist the better. The effect of the noise seen by GPA is described in detail in chapter 4, as it can seriously affect the generation of the ultimate atomic models we pursue.

2.3 Computational tools

The longest and most innovative chapter of the dissertation is chapter 4. It focuses on developing a workflow for the automated generation of representative 3D models of experimental devices. To achieve this, we acquire experimental data using the methods described in section 2.2. Subsequently, we process this data to obtain our automated data analysis results and create valuable 3D models. Implicitly, this process incorporates regular data analysis routines, Machine Learning (ML) and Deep Learning (DL), and Computer Vision (CV), collectively forming a functional Artificial Intelligence (AI). Let's delve into the tools we utilize and explore why we choose them over alternative approaches. In the following subsections, we classify these tools based on their implementation characteristics and purpose. Specifically, within machine learning, we distinguish between supervised and unsupervised methods. This distinction relates to the underlying principle of the methods and whether they rely on prior knowledge.

2.3.1 Unsupervised machine learning methods

Unsupervised methods do not need any prior knowledge of the data structure. Thus, unsupervised algorithms do not require training nor labelled data, and just work and find patterns in the input data as a standalone entry.[22, 47, 48] Their main purpose is the classification of the data points of the input according to these complex patterns identified according to the nature of the specific algorithm. As a result, we employ unsupervised routines when the access to labelled data is not realistic nor reproducible. In Chapter 4, unsupervised algorithms are mainly used for the segmentation of electron micrographs, a purpose for which obtaining labelled data is not trivial. Manual labelling would involve painstakingly drawing the device's

regions, a time-consuming process that adds little value to our research. It is essential to emphasise that our research scope extends beyond workflow development; we aim to extract knowledge and scientific insights during the process. We intend to share this technical knowledge with the materials science community. However, brute-force manual labelling is not a strategy of general interest. Instead, we explore an alternative approach in the section on supervised methods (subsection 2.3.2) and in Chapter 4. Here, we discuss how scientific value lies in circumventing label generation through physical models that automatically annotate data.

Multiple unsupervised methods are used to segment the STEM-HAADF micrographs. The purpose is to correlate the clusters found with the segments within the image. We evaluated the performance of three clustering algorithms: K-means clustering [49], Agglomerative clustering [50], and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [51]. Each algorithm has its particularities, which are discussed in chapter 4 in line with their performance in the segmentation application and how they deal with the contrast distribution of HAADF STEM images.

In addition, we have used unsupervised ML to denoise and cluster EELS spectrum images. Both in Chapter 3 for manual denoising and data smoothing, and in Chapter 5 to unveil the characteristic spectral features composing the EELS maps. As mentioned in the Introduction, the spectral unmixing techniques that originally launched the use of ML within electron microscopy were: Principal Component Analysis (PCA) as the main denoising algorithm, together with Non-negative Matrix Factorisation (NMF), Independent Component Analysis (ICA) and Vertex Component Analysis (VCA) for clustering and endmember retrieval. This decision is reinforced by the fact that they have become synonymous with spectral unmixing since their initial application.

2.3.2 Supervised methods

Supervised methods do require training with (typically plenty of) labelled training data which resembles but is not the target data piece. [22, 47, 48, 52] By labelled data, we understand data pairs which include the data itself together with a description of the information we want our method to extract from it. For instance, if we label a picture of a dog so our method finds dogs in images, our label can be a box containing the dog. If our goal is to find atoms from an electron micrograph, our labels can be boxes containing the atoms. Hence, our labels must tell the model what we want to extract from our data. That is why, it is typically challenging to find labelled data that adjusts to our necessities.

As we briefly mentioned in subsection 2.3.1, our proposed goals are not just getting an application done. Not surprisingly, in most cases, the attainment of the requested application would just demand the brute-force labelling of a huge amount of data.

Nevertheless, it would contribute with no new knowledge beyond the pure working application. This way of proceeding may be acceptable in productivity-oriented environments, where the process is dismissable in comparison with getting functional applications, but it is not a winning strategy for science. Not being scaleable in time or reach, it would demand unaffordable human capital and prohibitive resources to end up with black boxes, which are operational, but in the end still black boxes.

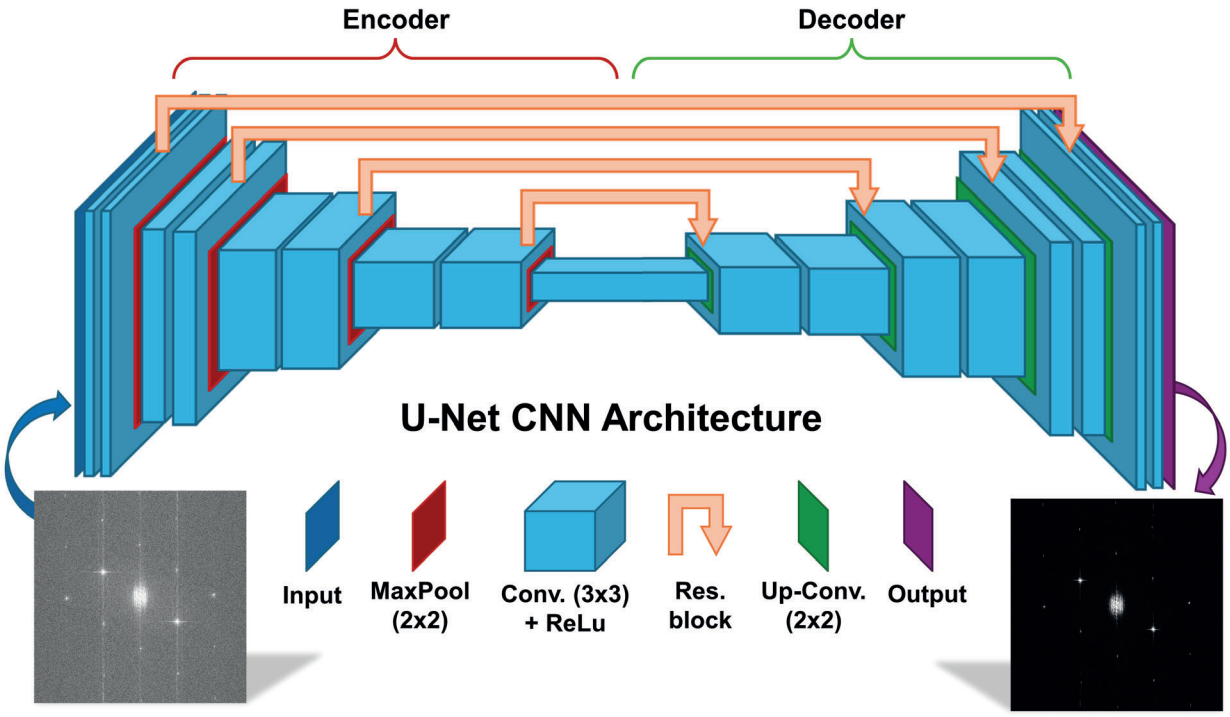


FIGURE 2.5: Schematic representation of the well-known encoder-decoder convolutional neural network U-Net for image segmentation. It is composed of convolutional filters (light blue), downsampling filters (red), upsampling filters (green), and residual blocks (orange). Its encoder-decoder architecture ensures its input (dark blue) and output (purple) remain the same size. In the depicted example, the peaks from a fast Fourier transform are highlighted because of the neural network. [53].

As a result, in Chapter 4, we implement supervised methods in the workflow segments where access to labelled data is intuitive, automatic, high-throughput and model-based. Particularly, we used it in two scenarios. Firstly, we used a traditional DL solution to obtain a peak-finding model. The reason why we need a peak finding model is described in detail in the core of chapter 4. In any case, the solution is based on a Convolutional Neural Network (CNN) with an encoder-decoder architecture which is based on the well-known U-Net architecture (fig. 2.5). [53] We opt for this kind of architecture because we want both our inputs and outputs to be images. We input an image in which to find peaks, and we output an image of the same size with the peaks as its distinctive features. We can graphically see this concept in Figure 2.5, where U-Net is used to highlight the intensity peaks of an FFT. Remarkably, the DL model is trained upon labelled data obtained from a physical model, specifically the kinematic diffraction model. Therefore, we are aware of the physical

base the neural network relies on. In addition, it is a flexible training approach as we can fine-tune anytime through the use of more complex complementary models, or even through manually annotated data. However, the use of this hypothetical annotated data would not constitute the core of the DL model, which is a substantial difference concerning the scenario we were condemning before.

The second use of supervised ML also targets peak finding, but not in a customary manner. We used a well-established data analysis routine (i.e., 2D Gaussian fitting) which performs satisfactorily in specific cases but requires manual adjustment of hyperparameters. [54] Then, the idea lies in manually tailoring these hyperparameters in different situations to cover a broad spectrum of situations that resemble our experimental conditions. Afterwards, we fit a simple polynomial supervised ML model to capture how these hyperparameters vary with experimental conditions, achieving automation in the process. Interestingly enough, we chased manual labelling of very few regression points because of the nature of these points. Importantly, each hyperparameter set successfully encompasses a wide compendium of experimental contexts. Consequently, this justifies the manual labelling in a high-reward/low-risk (i.e., quick labelling) situation that easily generalises, in combination with a simple ML model that adequately balances over and underfitting.

2.3.3 General data analysis for computer vision

Most of the methods described in sections 2.3.1 and 2.3.2 can be grouped as computer vision routines when they are applied to images or visual inputs. In the present section, though, we differentiate a handful of methods we employed, which despite being designed specifically for images, do not fit within ML or DL. As a result, no training is required. Here we just mention the methods used to automate processes in Chapter 4, not including pre-processing like filters or denoisers for image enhancement. This does not mean we do not use them: for instance, we employ a Wiener filter to denoise the GPA maps in the proposed workflow and smooth the computed strain fields down.

Our general data analysis routines are essentially two: contour detection and template matching. Contour detection, as the name suggests, consists of highlighting the contrast changes within images. It is based on the application of convolutional kernels or filters that highlight these changes. As with the unsupervised ML methods described in section 2.3.1, contour detection aims to segment STEM-HAADF micrographs. Its simple tuning with few hyperparameters makes it an outstanding candidate to accommodate the variability found in electron micrographs. Interestingly enough, the explicit differences between general digital images (i.e., a picture from a cell phone) and electron micrographs make contour detection a fascinating tool to analyse within the context we present. The results of the evaluation of the algorithm can be found in chapter 4.

On the other hand, template matching is used to correlate pairs of images. While we won't delve into the intricate details explained in Chapter 4, we'll provide a concise overview. Our goal, as outlined in Chapter 1, is to automate data analysis. To achieve this, we must have a tool capable of identifying features in one micrograph and mapping them to corresponding features in another. In the realm of feature detection, various solutions exist. However, multiscale template matching strikes a balance between computational efficiency, performance, and flexibility. [55, 56] This method proves advantageous when dealing with micrographs that may differ in pixel size (calibration). By adopting a multiscale approach, we optimize efficiency by leveraging relative positioning, ultimately saving time. Compared to more complex algorithms like Scale-Invariant Feature Transform (SIFT) or Speeded-Up Robust Feature (SURF), multiscale template matching stands out due to its simplicity. [57, 58] These advanced algorithms account for intricate rotations and inversions in features, but unless we specifically anticipate such scenarios (e.g., abrupt changes in zone axis), they become overkill. Chapter 4 provides further details into the purpose and performance of this algorithm within the context of electron microscopy.

2.3.4 Programming and software

In the scientific endeavours taken into account in the present PhD dissertation, we engage in various forms of manual analysis related to quantum devices (Chapter 3), automated workflows (Chapter 4), and bandgap mapping (Chapter 5). Even if some intend to reach automation, their development involves, to some extent, manual meticulous examination and evaluation. To achieve this, we rely on a trusted tool: the Gatan Microscopy Suite (GMS), formerly known as DigitalMicrograph.

In the following, we list the key advantages, which set GMS apart and establish it as the gold standard for microscopists:

1. **Intuitive Interface:** GMS offers an interface that is both user-friendly and intuitive. Researchers can swiftly access the necessary features without unnecessary complexity.

2. **Robust Toolset:** Within GMS, a plethora of tools specifically designed for image analysis and hyperspectral data exploration await. Whether it is quantifying nanoscopic details or assessing material properties, GMS provides the arsenal needed for comprehensive analysis.

3. **Customization:** GMS does not stop at pre-defined functionalities. Researchers can create custom and tailored routines using its built-in programming language.

However, when it comes to developing intricate solutions such as the complex artificial intelligence discussed in Chapter 4, we turn to a different language: Python. Python empowers us to craft personalized algorithms and solutions that align precisely with our research objectives. In summary, while GMS remains our go-to for

standard microscopy analysis, Python serves as our creative canvas for pushing the boundaries of scientific exploration. We develop our Python codes outside GMS to benefit from intuitive development environments, although the proposed data analysis routines depicted throughout the thesis can indeed be used within GMS to further integrate them within its environment.

Python stands out as the optimal choice for developing tailored data science solutions. [59] This preference is primarily attributed to its dedicated community devoted to open knowledge and open source principles, consistently striving for transparent and reproducible scientific practices. As a result, a vast array of open-source tools, packages, Application Programming Interfaces (APIs), repositories, tutorials and courses are readily available for self-training purposes, catering to both novice and advanced users in ML/DL. This abundance renders Python an exceptional platform for data scientists who view programming not merely as a means of investigation and development, but as a conduit for exploring broader inquiries. Crucially, leveraging these pre-existing tools expedites research processes, enabling scientists to focus on their specific research queries. Simultaneously, Python offers sufficient flexibility to extend methodological boundaries beyond what is achievable with commercial solutions. Our own requirements for automation and methodological advancement required the development of custom routines reliant on Python's core data science packages.

For instance, the NumPy package saves us from self-scripting basic mathematical arithmetics by providing a well-optimised and intuitive use of its functions and classes. [60] Like NumPy, we used several packages and APIs to accelerate the methodological development needed for the research aspirations pointed out in the introduction (Chapter 1). The main packages and libraries we have used in Chapters 4 and 5 are presented in table 2.1.

Interestingly, we have conducted our coding endeavours using these packages on a commercially available personal computer. While this might seem like a trivial detail, it holds significant importance in our will towards democratisation. It ensures that the processing can be handled and reproduced by a personal machine accessible to anyone. The laptop, equipped with 16 GB of RAM, an Intel Core i7-8750H processor clocked at 2.20 GHz, and an Nvidia RTX 2060 graphics card, proved adequate for all the manual data analyses outlined in Chapter 3, as well as for the development, testing and validation of the workflow described in Chapter 4, and similarly in Chapter 5. However, an interesting observation was made during the execution of Chapter 4, where the computation of extensive 3D atomic models exceeded the laptop's RAM capacity. Consequently, we migrated to a workstation for these computations. They never exceeded 20 GB of RAM usage across all the research cases studied, despite dealing with more than half a million atoms. Remarkably, the creation of these atomic models is facilitated by the Python package known as the Atomic Simulation Environment (ASE), leveraging multiple pre-built

functions and classes to streamline the manipulation and representation of atoms within symmetric cells. [61]

Package name	Purpose	References
ASE	Atomic modelling and simulation environment	[61]
AtomAI	AI tools for microscopy analysis	[62]
ctypes	Python-C/C++ integration	[63]
HyperSpy	Hyperspectral data analysis	[64, 65]
math	Maths and arithmetics programming	[66]
matpotlib	Figure plotting	[67]
mendeleev	Periodic table information	[68]
NumPy	Array programming	[60]
OpenCV	Computer vision algorithms	[55]
pandas	Tabular data processing	[69]
PIL	Image processing	[70]
Pycroscopy	Microscopy related algorithms	[71]
PyTorch	Low-level ML/DL	[72]
Scikit-image	Scientific image processing	[73]
Scikit-learn	Scientific machine learning (mostly unsupervised)	[74]
SciPy	Scientific computing	[75]
STEMTooL	STEM data processing algorithms	[76]
SymPy	Symbolic maths programming	[77]
tkinter	Interface design	[78]

TABLE 2.1: Main Python packages used to develop the workflow in chapter 4, and the methods in chapter 5.

Furthermore, the workstation was used in various scenarios, not solely due to a requirement for computational power but also to expedite computationally intensive calculations. This workstation, specifically designed for AI development, boasted 256 GB of RAM, an Intel Xeon W-2295 CPU processor running at 3 GHz, and an Nvidia Quadro RTX A6000 GPU. Furthermore, we leveraged the collaborative resources of a high-performance computing centre like Punt d'Informació Científica

(PIC), specialising in AI, to access GPUs such as the RTX 2080 Ti and V100. [79] Specifically, this high-end equipment was employed to iteratively refine the training of the U-Net model with our target training data (fig. 2.5). Although memory-efficient methods such as batch training could have been implemented on the personal computer mentioned earlier, the additional computational resource greatly expedited the training process and enabled iterative hyperparameter tuning until the desired model was achieved, as detailed in Chapter 4.

Additionally, our research involved using software dedicated to data acquisition, serving as an intermediary between the microscopes and us, the microscopists. Software suites such as FEI TIA, TUI, and Velox for Thermo Fisher Scientific (formerly FEI) microscopes, or Nion Swift for Nion microscopes, were used. The following section concludes the methodology chapter by justifying the selection of the experimental equipment, which, despite facing geographical and institutional constraints, ultimately proved to be ideal for achieving our research objectives.

2.4 Experimental equipment and personal expertise

Achieving the development objectives outlined in subsection 2.3.4 necessitates expertise from the author in software utilization and programming libraries. However, the author's proficiency extends beyond scripting to encompass the experimental techniques detailed in section 2.2. Interestingly, access to these experimental techniques was facilitated through the equipment belonging to the author's institution, the Catalan Institute of Nanoscience and Nanotechnology, and their research group, the Group of Advanced Electron Nanoscopy.

Most of the STEM images in the thesis dissertation were taken using a probe-corrected TITAN low-base operating at 300 kV. Specifically, the instrument available at the Laboratorio de Microscopías Avanzadas, within the Instituto de Nanociencia y Materiales de Aragón, at the Universidad de Zaragoza. [80] As disclosed in section 2.2, probe-correction enabled atomic resolution STEM imaging, primarily in HAADF mode, which was the primary focus of our investigations. The author attained autonomous expertise in operating this machine, either personally acquiring data or overseeing the acquisitions, particularly during periods coinciding with the COVID-19 pandemic when logistical constraints necessitated the involvement of group colleagues for data collection.

The other main microscope used is the uncorrected Tecnai F20, operating at 200 kV and available within the Catalan Institute of Nanoscience and Nanotechnology. [81] The Tecnai F20 was mainly employed to generate EELS compositional maps of samples previously imaged using the TITAN. It's worth noting that EELS analysis was not conducted on the TITAN, as its operating time was dedicated solely to imaging with lateral resolutions beyond the capabilities of the Tecnai. Consequently, while

high spatial resolutions for compositional mapping were neither feasible nor a priority (given that atomic resolution EELS would be excessive), the Tecnai served as an ideal complement to TITAN imaging. Furthermore, the Tecnai was used to perform high-resolution imaging through parallel-illumination TEM mode on various complementary devices for diverse applications. Although these applications are omitted from the dissertation due to their functional properties diverging from the primary scope outlined in the introduction, they provide valuable insights for generalising the data analysis workflow developed in Chapter 4. In summary, the author reached complete autonomous expertise in operating this machine.

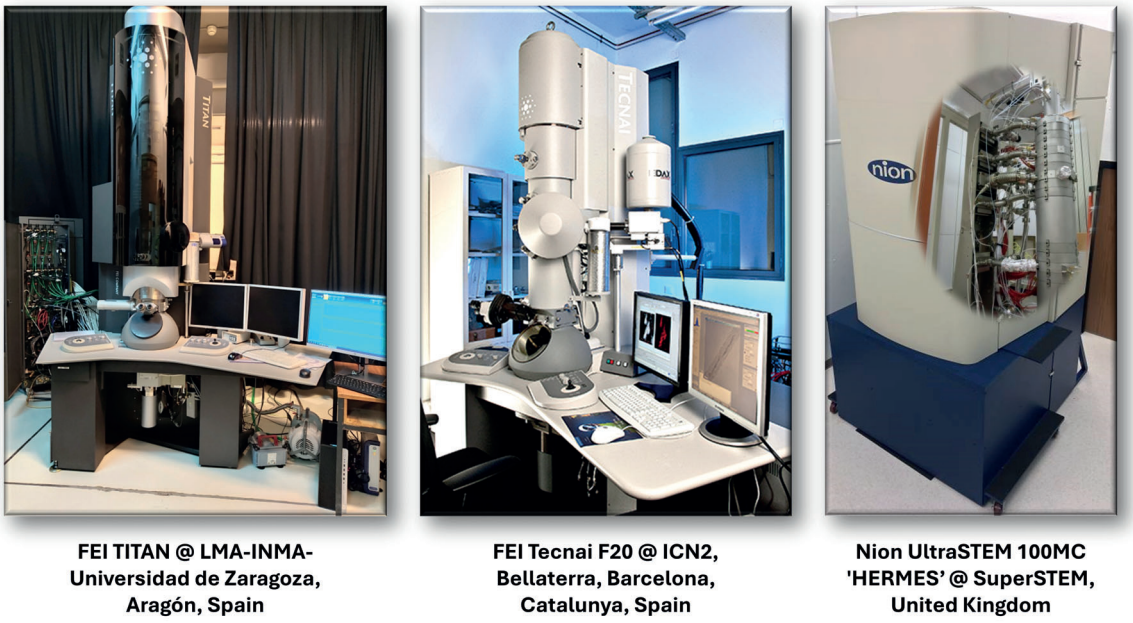


FIGURE 2.6: The three electron microscopes used to acquire the data presented in the thesis dissertation: an FEI TITAN [80], an FEI Tecnai F20 [81], and a Nion UltraSTEM 100MC 'HERMES' [38].

Crucially, an additional microscope was indispensable to fulfil our goal of low-loss EELS data analysis. The maximum energy resolution attainable with the TITAN in EELS is approximately 140 meV, insufficient to capture the relative spatial changes in the density of states we sought. [80] It is important to note that while these energies may fall within the sub-eV range, achieving the necessary resolution to map local changes requires pushing beyond the apparent starting point. As a result, we employed a monochromated STEM-dedicated microscope to achieve state-of-the-art energy resolutions and access fine features, such as the bandgap, in the low-loss regime. This microscope is the Nion UltraSTEM 100MC 'HERMES', located at SuperSTEM, capable of reaching energy resolutions as low as 5 meV when combining its cold field emission gun and monochromator at 60 kV. [38] Hyperspectral data was

compiled collaboratively with partners from SuperSTEM and the author. Following a 3-months research internship at the microscopy centre, the author gained autonomous proficiency in operating the machine, achieving energy resolutions down to 9 meV in EELS, together with atomic-resolution STEM imaging.

Lastly, the lamellae used in our analyses were prepared by the technical support staff in Laboratorio de Microscopías Avanzadas. While the author possesses basic expertise in FIB, none of the analysed lamellae were prepared by the author personally.

References

[1] Wenfeng Yang et al. "Local-orbital ptychography for ultrahigh-resolution imaging". In: *Nature Nanotechnology* (Jan. 2024). ISSN: 1748-3395. DOI: 10.1038/s41565-023-01595-w. URL: https://doi.org/10.1038/s41565-023-01595-w.

[2] Zhen Chen et al. "Electron ptychography achieves atomic-resolution limits set by lattice vibrations". In: *Science* 372.6544 (2021), pp. 826–831. DOI: 10.1126/science.abg2533. eprint: https://www.science.org/doi/pdf/10.1126/science.abg2533. URL: https://www.science.org/doi/abs/10.1126/science.abg2533.

[3] O.L. Krivanek et al. "Progress in ultrahigh energy resolution EELS". In: *Ultramicroscopy* 203 (2019). 75th Birthday of Christian Colliex, 85th Birthday of Archie Howie, and 75th Birthday of Hannes Lichte / PICO 2019 - Fifth Conference on Frontiers of Aberration Corrected Electron Microscopy, pp. 60–67. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2018.12.006. URL: https://www.sciencedirect.com/science/article/pii/S0304399118303668.

[4] Ondrej L Krivanek et al. "Towards sub-10 meV energy resolution STEM-EELS". In: *Journal of Physics: Conference Series* 522.1 (June 2014), p. 012023. DOI: 10.1088/1742-6596/522/1/012023. URL: https://dx.doi.org/10.1088/1742-6596/522/1/012023.

[5] Jean Ponce David A. Forsyth. *Computer Vision: A Modern Approach*. 1st ed. Prentice Hall Professional Technical Reference, 2002. ISBN: 0130851981.

[6] David B. Williams and C. Barry Carter. *Transmission Electron Microscopy*. 2nd ed. Springer New York, 2009. ISBN: 978-0-387-76502-0.

[7] Colin Ophus. "Four-Dimensional Scanning Transmission Electron Microscopy (4D-STEM): From Scanning Nanodiffraction to Ptychography and Beyond". In: *Microscopy and Microanalysis* 25.3 (2019), pp. 563–582. DOI: 10.1017/S1431927619000497.

[8] *CETCOR | Corrected Electron Optical Systems — ceos-gmbh.de*. https://www.ceos-gmbh.de/en/produkte/temcorrectors/cetcor. [Accessed 19-02-2024].

[9] Crispin Hetherington. "Aberration correction for TEM". In: *Materials Today* 7.12 (2004), pp. 50–55. ISSN: 1369-7021. DOI: https://doi.org/10.1016/S1369-7021(04)00571-1. URL: https://www.sciencedirect.com/science/article/pii/S1369702104005711.

[10] Stephanie M Ribet et al. "Design of Electrostatic Aberration Correctors for Scanning Transmission Electron Microscopy". In: *Microscopy and Microanalysis* 29.6 (Oct. 2023), pp. 1950–1960. ISSN: 1431-9276. DOI: 10.1093/micmic/ozad111. eprint: https://academic.oup.com/mam/article-pdf/29/6/1950/54730716/ozad111.pdf. URL: https://doi.org/10.1093/micmic/ozad111.

[11] Peter W. Hawkes and Ondrej L. Krivanek. "Aberration Correctors, Monochromators, Spectrometers". In: *Springer Handbook of Microscopy*. Ed. by Peter W. Hawkes and John C. H. Spence. Cham: Springer International Publishing, 2019, pp. 625–675. ISBN: 978-3-030-00069-1. DOI: 10.1007/978-3-030-00069-1_13. URL: https://doi.org/10.1007/978-3-030-00069-1_13.

[12] Simon Hettler and Raul Arenal. "Aberration-corrected transmission electron microscopy with Zernike phase plates". In: *Ultramicroscopy* 239 (2022), p. 113564. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2022.113564. URL: https://www.sciencedirect.com/science/article/pii/S030439912200095X.

[13] Francisco Vega Ibáñez, Armand Béché, and Johan Verbeeck. "Can a Programmable Phase Plate Serve as an Aberration Corrector in the Transmission Electron Microscope (TEM)?" In: *Microscopy and Microanalysis* 29.1 (Jan. 2023), pp. 341–351. ISSN: 1431-9276. DOI: 10.1017/S1431927622012260. eprint: https://academic.oup.com/mam/article-pdf/29/1/341/49034109/s1431927622012-260jra.pdf. URL: https://doi.org/10.1017/S1431927622012260.

[14] Vincenzo Grillo et al. "Towards a holographic approach to spherical aberration correction in scanning transmission electron microscopy". In: *Opt. Express* 25.18 (Sept. 2017), pp. 21851–21860. DOI: 10.1364/OE.25.021851. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-25-18-21851.

[15] Maxim Ziatdinov et al. "Deep learning of atomically resolved scanning transmission electron microscopy images: Chemical identification and tracking local transformations". en. In: *ACS Nano* 11.12 (Dec. 2017), pp. 12742–12752.

[16] M. Ziatdinov et al. "Building ferroelectric from the bottom up: The machine learning analysis of the atomic-scale ferroelectric distortions". In: *Applied Physics Letters* 115.5 (Aug. 2019), p. 052902. ISSN: 0003-6951. DOI: 10.1063/1.5109520. eprint: https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/1.5109520/14527440/052902_1_online.pdf. URL: https://doi.org/10.1063/1.5109520.

[17] K.R. Rao, D.N. Kim, and J.-J. Hwang. *Fast Fourier Transform: Algorithms and Applications*. Signals and Communication Technology. Springer, 2010. ISBN: 978-1-4020-6628-3.

[18] Shunsuke Yamashita et al. "Atomic number dependence of Z contrast in scanning transmission electron microscopy". In: *Scientific Reports* 8.1 (Aug. 2018), p. 12325. ISSN: 2045-2322. DOI: 10.1038/s41598-018-30941-5. URL: https://doi.org/10.1038/s41598-018-30941-5.

[19] *CESCOR | Corrected Electron Optical Systems — ceos-gmbh.de*. https://www.ceos-gmbh.de/en/produkte/stemkorrektoren/cescor. [Accessed 19-02-2024].

[20] Kayla X. Nguyen et al. "Achieving sub-0.5-angstrom–resolution ptychography in an uncorrected electron microscope". In: *Science* 383.6685 (2024), pp. 865–870. DOI: 10.1126/science.adl2029. eprint: https://www.science.org/doi/pdf/10.1126/science.adl2029. URL: https://www.science.org/doi/abs/10.1126/science.adl2029.

[21] María de la Mata and Jordi Arbiol. "High Resolution in STEM Mode: Individual Atom Analysis in Semiconductor Nanowires". In: *Transmission Electron Microscopy Characterization of Nanomaterials*. Ed. by Challa S.S.R. Kumar. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 375–425. ISBN: 978-3-642-38934-4. DOI: 10.1007/978-3-642-38934-4_9. URL: https://doi.org/10.1007/978-3-642-38934-4_9.

[22] Marc Botifoll, Ivan Pinto-Huguet, and Jordi Arbiol. "Machine learning in electron microscopy for advanced nanocharacterization: current developments, available tools and future outlook". In: *Nanoscale Horiz.* 7 (12 2022), pp. 1427–1477. DOI: 10.1039/D2NH00377E. URL: http://dx.doi.org/10.1039/D2NH00377E.

[23] Sajedeh Manzeli et al. "2D transition metal dichalcogenides". In: *Nature Reviews Materials* 2.8 (June 2017), p. 17033. ISSN: 2058-8437. DOI: 10.1038/natrevmats.2017.33. URL: https://doi.org/10.1038/natrevmats.2017.33.

[24] Shanshan Wang, Alex Robertson, and Jamie H. Warner. "Atomic structure of defects and dopants in 2D layered transition metal dichalcogenides". In: *Chem. Soc. Rev.* 47 (17 2018), pp. 6764–6794. DOI: 10.1039/C8CS00236C. URL: http://dx.doi.org/10.1039/C8CS00236C.

[25] Emrah Yücelen, Ivan Lazić, and Eric G. T. Bosch. "Phase contrast scanning transmission electron microscopy imaging of light and heavy atoms at the limit of contrast and resolution". In: *Scientific Reports* 8.1 (Feb. 2018), p. 2676. ISSN: 2045-2322. DOI: 10.1038/s41598-018-20377-2. URL: https://doi.org/10.1038/s41598-018-20377-2.

[26] Quentin M. Ramasse. "Twenty years after: How "Aberration correction in the STEM" truly placed a "A synchrotron in a Microscope"". In: *Ultramicroscopy* 180 (2017). Ondrej Krivanek: A research life in EELS and aberration corrected STEM, pp. 41–51. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2017.03.016. URL: https://www.sciencedirect.com/science/article/pii/S030439911730116X.

[27] John M. Rodenburg, ed. *Electron Microscopy and Analysis 1997, Proceedings of the Institute of Physics Electron Microscopy and Analysis Group Conference, University of Cambridge, 2-5 September 1997*. Institute of Physics Conference Series. CRC Press, 1997. ISBN: 978-0-7503-0441-2. DOI: 10.1201/9781003063056.

[28] Jordan A. Hachtel, Andrew R. Lupini, and Juan Carlos Idrobo. "Exploring the capabilities of monochromated electron energy loss spectroscopy in the infrared regime". In: *Scientific Reports* 8.1 (Apr. 2018), p. 5637. ISSN: 2045-2322. DOI: 10.1038/s41598-018-23805-5. URL: https://doi.org/10.1038/s41598-018-23805-5.

[29] GW Bailey et al. "Comparison of STEM EELS Spectrum Imaging vs EFTEM Spectrum Imaging". In: *Microscopy and Microanalysis* 5.S2 (Aug. 1999), pp. 616–617. ISSN: 1431-9276. DOI: 10.1017/S1431927600016408. eprint: https://academic.oup.com/mam/article-pdf/5/S2/616/48284874/mam0616.pdf. URL: https://doi.org/10.1017/S1431927600016408.

[30] *Why EELS? — eels.info*. https://eels.info/why-eels/overview. [Accessed 20-02-2024].

[31] R.F. Egerton. *Electron Energy-Loss Spectroscopy in the Electron Microscope*. 3rd ed. Springer, 2011. ISBN: 978-1-4419-9583-4.

[32] Sara Martí-sánchez et al. "Sub-nanometer mapping of strain-induced band structure variations in planar nanowire core-shell heterostructures". In: *Nature Communications* 13 (2022), p. 4089. DOI: /10.1038/s41467-022-31778-3. eprint: https://doi.org/10.1038/s41467-022-31778-3. URL: https://doi.org/10.1038/s41467-022-31778-3.

[33] Ming-Wen Chu et al. "Probing Bright and Dark Surface-Plasmon Modes in Individual and Coupled Noble Metal Nanoparticles Using an Electron Beam". In: *Nano Letters* 9.1 (2009). PMID: 19063614, pp. 399–404. DOI: 10.1021/nl803270x. eprint: https://doi.org/10.1021/nl803270x. URL: https://doi.org/10.1021/nl803270x.

[34] F. J. García de Abajo. "Optical excitations in electron microscopy". In: *Rev. Mod. Phys.* 82 (1 Feb. 2010), pp. 209–275. DOI: 10.1103/RevModPhys.82.209. URL: https://link.aps.org/doi/10.1103/RevModPhys.82.209.

[35] F. S. Hage et al. "Phonon Spectroscopy at Atomic Resolution". In: *Phys. Rev. Lett.* 122 (1 Jan. 2019), p. 016103. DOI: 10.1103/PhysRevLett.122.016103. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.016103.

[36] ThermoFisher Scientific. *Accelerate Materials Insights with Our New Cold Field Emission Electron Gun for Atomic-Scale High-Resolution TEM*. https://www.thermofisher.com/blog/materials/cold-field-emission-electron-gun-for-atomic-scale-high-resolution-tem/. [Accessed 21-02-2024]. 2022.

[37] L. Piazza et al. "Design and implementation of a fs-resolved transmission electron microscope based on thermionic gun technology". In: *Chemical Physics* 423 (2013), pp. 79–84. ISSN: 0301-0104. DOI: https://doi.org/10.1016/j.chemphys.2013.06.026. URL: https://www.sciencedirect.com/science/article/pii/S0301010413002826.

[38] SuperSTEM. *Uncovering atomic secrets - Instrumentation — superstem.org*. https://www.superstem.org/facility/instrumentation. [Accessed 21-02-2024].

[39] Daniel Jirovec et al. "A singlet-triplet hole spin qubit in planar Ge". In: *Nature Materials* 20 (2021), pp. 1106–1112. DOI: /10.1038/s41563-021-01022-2. eprint: https://doi.org/10.1038/s41563-021-01022-2. URL: https://doi.org/10.1038/s41563-021-01022-2.

[40] Christoph Mahr et al. "Accurate measurement of strain at interfaces in 4D-STEM: A comparison of various methods". In: *Ultramicroscopy* 221 (2021), p. 113196. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2020.113196. URL: https://www.sciencedirect.com/science/article/pii/S0304399120303363.

[41] Sihan Wang et al. "AutoDisk: Automated diffraction processing and strain mapping in 4D-STEM". In: *Ultramicroscopy* 236 (2022), p. 113513. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2022.113513. URL: https://www.sciencedirect.com/science/article/pii/S0304399122000481.

[42] M.J. Hÿtch, E. Snoeck, and R. Kilaas. "Quantitative measurement of displacement and strain fields from HREM micrographs". In: *Ultramicroscopy* 74.3 (1998), pp. 131–146. ISSN: 0304-3991. DOI: https://doi.org/10.1016/S0304-3991(98)00035-7. URL: https://www.sciencedirect.com/science/article/pii/S0304399198000357.

[43] Martin J. Hÿtch, Jean-Luc Putaux, and Jean-Michel Pénisson. "Measurement of the displacement field of dislocations to 0.03 Å by electron microscopy". In: *Nature* 423 (2003), pp. 270–273. DOI: /10.1038/nature01638. eprint: https://doi.org/10.1038/nature01638. URL: https://doi.org/10.1038/nature01638.

[44] Pedro L. Galindo et al. "The Peak Pairs algorithm for strain mapping from HRTEM images". In: *Ultramicroscopy* 107.12 (2007), pp. 1186–1193. ISSN: 0304-3991. DOI: https://doi.org/10.

1016/j.ultramic.2007.01.019. URL: https://www.sciencedirect.com/science/article/pii/S0304399107000204.

[45] D L Sales et al. "Critical strain region evaluation of self-assembled semiconductor quantum dots". In: *Nanotechnology* 18.47 (Oct. 2007), p. 475503. DOI: 10.1088/0957-4484/18/47/475503. URL: https://dx.doi.org/10.1088/0957-4484/18/47/475503.

[46] A. Llordés et al. "Nanoscale strain-induced pair suppression as a vortex-pinning mechanism in high-temperature superconductors". In: *Nature Materials* 11.4 (Apr. 2012), pp. 329–336. ISSN: 1476-4660. DOI: 10.1038/nmat3247. URL: https://doi.org/10.1038/nmat3247.

[47] Mohamed Alloghani et al. "A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science". In: *Supervised and Unsupervised Learning for Data Science*. Ed. by Michael W. Berry, Azlinah Mohamed, and Bee Wah Yap. Cham: Springer International Publishing, 2020, pp. 3–21. ISBN: 978-3-030-22475-2. DOI: 10.1007/978-3-030-22475-2_1. URL: https://doi.org/10.1007/978-3-030-22475-2_1.

[48] Michael W. Berry, Azlinah Mohamed, and Bee Wah Yap. *Supervised and Unsupervised Learning for Data Science*. Cham: Springer International Publishing, 2020. ISBN: 978-3-030-22475-2. DOI: 10.1007/978-3-030-22475-2. URL: https://doi.org/10.1007/978-3-030-22475-2.

[49] Lingfeng Wang and Chunhong Pan. "Robust level set image segmentation via a local correntropy-based K-means clustering". In: *Pattern Recognition* 47.5 (2014), pp. 1917–1925. ISSN: 0031-3203. DOI: https://doi.org/10.1016/j.patcog.2013.11.014. URL: https://www.sciencedirect.com/science/article/pii/S0031320313005001.

[50] Toufiq Parag et al. "A context-aware delayed agglomeration framework for electron microscopy segmentation". en. In: *PLoS One* 10.5 (May 2015), e0125825.

[51] Randy M Sterbentz, Kristine L Haley, and Joshua O Island. "Universal image segmentation for optical identification of 2D materials". en. In: *Sci. Rep.* 11.1 (Mar. 2021), p. 5808.

[52] Jeffrey M Ede. "Deep learning in electron microscopy". In: *Machine Learning: Science and Technology* 2.1 (Mar. 2021), p. 011004. DOI: 10.1088/2632-2153/abd614. URL: https://dx.doi.org/10.1088/2632-2153/abd614.

[53] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4.

[54] Debangshu Mukherjee et al. "mpfit: a robust method for fitting atomic resolution images with multiple Gaussian peaks". In: *Advanced Structural and Chemical Imaging* 6.1 (Jan. 2020), p. 1. ISSN: 2198-0926. DOI: 10.1186/s40679-020-0068-y. URL: https://doi.org/10.1186/s40679-020-0068-y.

[55] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. " O'Reilly Media, Inc.", 2008.

[56] *The OpenCV Reference Manual*. 2.4.13.7. OpenCV. Apr. 2014.

[57] D.G Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". In: *International Journal of Computer Vision* 60 (2004), pp. 91–110. DOI: /10.1023/B:VISI.0000029664.99615.94. URL: https://doi.org/10.1023/B:VISI.0000029664.99615.94.

[58] Herbert Bay et al. "Speeded-Up Robust Features (SURF)". In: *Computer Vision and Image Understanding* 110.3 (2008). Similarity Matching in Computer Vision and Multimedia, pp. 346–359. ISSN: 1077-3142. DOI: https://doi.org/10.1016/j.cviu.2007.09.014. URL: https://www.sciencedirect.com/science/article/pii/S1077314207001555.

[59] Guido Van Rossum and Fred L Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

[60] Charles R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2.

[61] Ask Hjorth Larsen et al. "The atomic simulation environment—a Python library for working with atoms". In: *Journal of Physics: Condensed Matter* 29.27 (June 2017), p. 273002. DOI: 10.1088/1361-648X/aa680e. URL: https://dx.doi.org/10.1088/1361-648X/aa680e.

[62] M. Ziatdinov. *pycroscopy/atomai: Deep and Machine Learning for Microscopy*. https://github.com/pycroscopy/atomai. 2022.

[63] Python Software Foundation. *ctypes — A foreign function library for Python — docs.python.org*. https://docs.python.org/3/library/ctypes.html. [Accessed 25-02-2024].

[64] Francisco de la Peña et al. *hyperspy/hyperspy: v2.0*. Version v2.0. Dec. 2023. DOI: 10.5281/zenodo.10412190. URL: https://doi.org/10.5281/zenodo.10412190.

[65] Francisco de la Pena et al. "Electron Microscopy (Big and Small) Data Analysis With the Open Source Software Package HyperSpy". In: *Microscopy and Microanalysis* 23.S1 (Aug. 2017), pp. 214–215. ISSN: 1431-9276. DOI: 10.1017/S1431927617001751. eprint: https://academic.oup.com/mam/article-pdf/23/S1/214/48301678/mam0214.pdf. URL: https://doi.org/10.1017/S1431927617001751.

[66] Python Software Foundation. *math — Mathematical functions — docs.python.org*. https://docs.python.org/3/library/math.html. [Accessed 25-02-2024].

[67] John D. Hunter. "Matplotlib: A 2D Graphics Environment". In: *Computing in Science Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.

[68] Lukasz Mentel. *Welcome to mendeleev (2019); documentation 2014; mendeleev 0.15.0 documentation — mendeleev.readthedocs.io*. https://mendeleev.readthedocs.io/en/stable/. [Accessed 25-02-2024]. 2015.

[69] Jeff Reback et al. *pandas-dev/pandas: Pandas 1.3.5*. Version v1.3.5. Dec. 2021. DOI: 10.5281/zenodo.5774815. URL: https://doi.org/10.5281/zenodo.5774815.

[70] pillow. *Reference — pillow.readthedocs.io*. https://pillow.readthedocs.io/en/stable/reference/index.html. [Accessed 25-02-2024].

[71] Rama Vasudevan et al. *A processing and analytics system for microscopy data workflows: the Pycroscopy ecosystem of packages*. 2023. arXiv: 2302.14629 [physics.ins-det].

[72] Adam Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Curran Associates, Inc., 2019. URL: https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92-f2bfa9f7012727740-Paper.pdf.

[73] S. van der Walt et al. "scikit-image: image processing in Python". In: *PeerJ* 2 (2014), e453. DOI: 10.7717/peerj.453.

[74] Fabian Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12.85 (2011), pp. 2825–2830. URL: http://jmlr.org/papers/v12/pedregosa11a.html.

[75] Pauli Virtanen et al. "SciPy 1.0: fundamental algorithms for scientific computing in Python". In: *Nature Methods* 17.3 (Mar. 2020), pp. 261–272. ISSN: 1548-7105. DOI: 10.1038/s41592-019-0686-2. URL: https://doi.org/10.1038/s41592-019-0686-2.

[76] Debangshu Mukherjee and Raymond Unocic. "STEMTooL: An Open Source Python Toolkit for Analyzing Electron Microscopy Datasets". In: *Microscopy and Microanalysis* 26.S2 (Aug. 2020), pp. 2960–2962. ISSN: 1431-9276. DOI: 10.1017/S143192762002334X. eprint: https://academic.oup.com/mam/article-pdf/26/S2/2960/48324548/mam2960.pdf. URL: https://doi.org/10.1017/S143192762002334X.

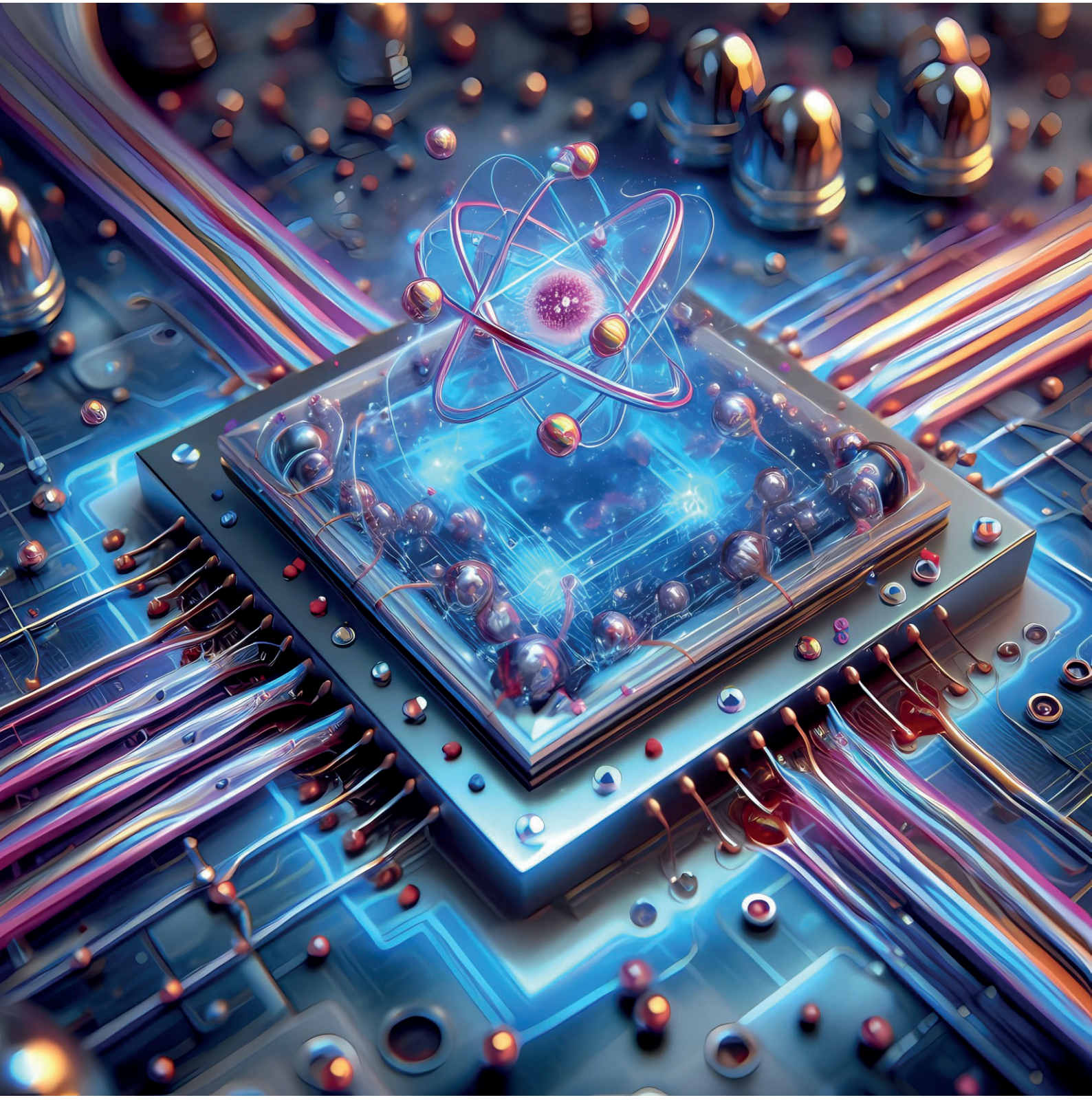
[77] Aaron Meurer et al. "SymPy: symbolic computing in Python". In: *PeerJ Computer Science* 3 (Jan. 2017), e103. ISSN: 2376-5992. DOI: 10.7717/peerj-cs.103. URL: https://doi.org/10.7717/peerj-cs.103.

[78] Fredrik Lundh. "An introduction to tkinter". In: (1999). URL: www.pythonware.com/library/tkinter/introduction/index.%20htm.

[79] PIC. *PIC — pic.es*. https://www.pic.es/. [Accessed 26-02-2024].

[80] Laboratorio de Microscopías Avanzadas Universidad de Zaragoza. *TITAN Analítico*. https://lma.unizar.es/titan-analitico/. [Accessed 26-02-2024].

[81] Catalan Institute of Nanoscience and Nanotechnology. *FEI Tecnai G2 F20 S-TWIN HR(S)TEM - ICN2 — icn2.cat*. https://icn2.cat/en/emu-equipments/fei-tecnai-g2-f20-s-twin-hr-s-tem. [Accessed 26-02-2024].



CHAPTER 3: SI/GE-BASED QUANTUM DEVICES AT ATOMIC SCALE: A (SCANNING) TRANSMISSION ELECTRON MICROSCOPY APPROACH

Chapter 3

Si/Ge-Based Quantum Devices at Atomic Scale: a (Scanning) Transmission Electron Microscopy Approach

Quantum devices represent outstanding systems to investigate the modern challenges of advanced structural characterisation using (scanning) transmission electron microscopy. As technology continues to push beyond Moore's law, the miniaturisation of these devices becomes crucial for characterising their optimum operational conditions and functional properties. Interestingly, Si/Ge-based devices have emerged as a promising platform for hosting fault-tolerant and scalable qubits. This chapter aims to evaluate various state-of-the-art Si/Ge devices designed to explore varied computing paradigms and achieve viable qubits. Furthermore, we analyse how their structural properties correlate with their performance as fully operational quantum devices. Lastly, we systematically assess the characterisation requirements that modern transmission electron microscopy must fulfil to provide a comprehensive understanding of these cutting-edge devices. These identified needs will form the foundation for the automated workflows that unveil physical knowledge in subsequent chapters.

3.1 Motivation and context

Quantum computing presents a groundbreaking opportunity to enhance algorithm efficiency across various fields such as numerical optimization, cryptography, data searching, finances, or drug and materials development. [1, 2, 3, 4, 5] Nowadays, researchers are challenging the finding of scenarios in which quantum computers do present a benefit (i.e., quantum advantage). However, quantum supremacy has

already been proven to outperform classical computers. [6] Outperforming classical computers entails leveraging the quantum phenomena of superposition and entanglement. Unlike classical bits, which are limited to 0 or 1 states, quantum bits (qubits) can exist in infinite combinations of 0s and 1s through superposition. In addition, qubits exhibit entanglement, where the state of one qubit affects its entangled counterparts and vice versa.

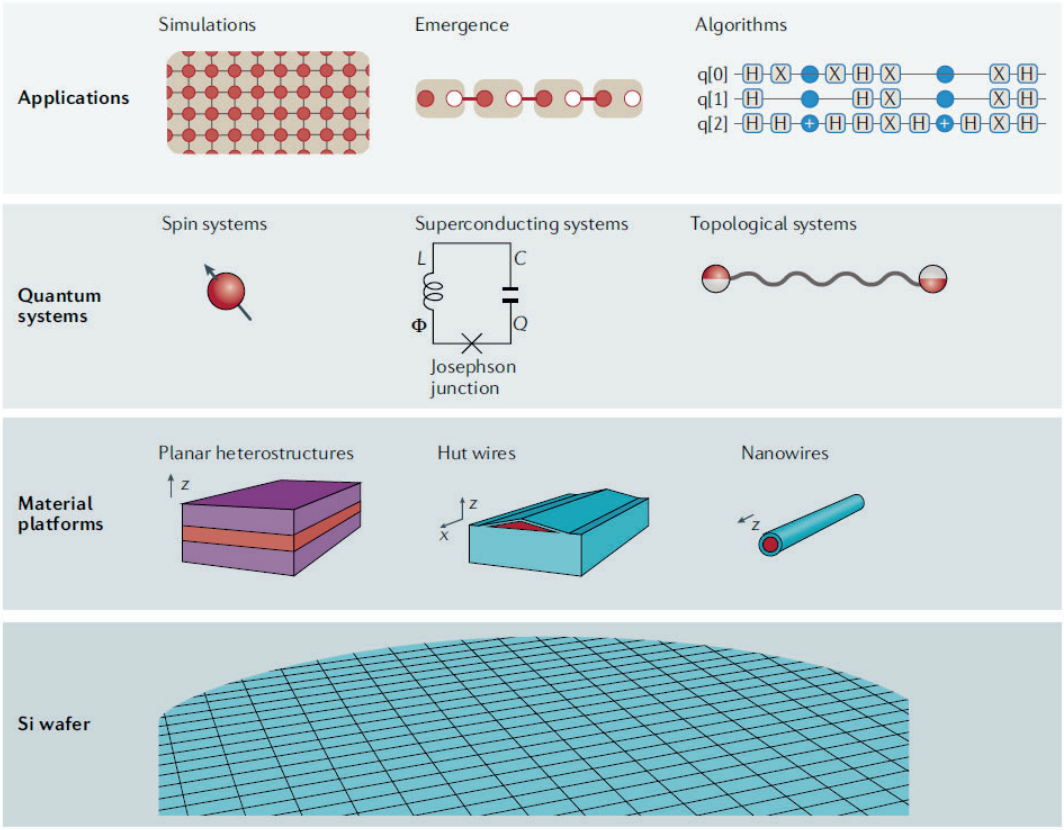


FIGURE 3.1: The three technologies of the germanium quantum information route. Ge serves as a versatile platform supporting the advancement of spin, superconducting, and topological systems. The integration of these diverse systems in Ge offers a promising pathway towards scalable quantum technology on silicon wafers. Hybrid quantum systems provide a unique environment for simulating complex Hamiltonians, and allow for the controlled emulation of emergent phenomena like Majorana zero modes, which can be complemented with spin qubits to define universal quantum gate sets. Adapted with permission from [7].

The primary challenge in quantum computers development is identifying a qubit platform that can deliver high-fidelity measurements, coherent control, and robust isolation from external factors. Various physical systems like superconducting circuits, trapped ions, and semiconductors have been explored for qubit implementation. [8, 9, 10] Among these options, silicon (Si) and germanium (Ge), but especially Ge, emerged as leading contenders due to their scalability potential. [7, 11]

The Complementary Metal Oxide Semiconductor (CMOS) technology, which underpins the whole semiconductor industry, is based on Si. Its high-quality (native) silicon oxide, has enabled a level of miniaturisation that has continuously pushed the

boundaries of Moore's law. [12, 13] However, in order to extend beyond the realm of Moore's law, CMOS technology has had to adopt high dielectric constant materials. Interestingly, Ge has emerged as a material offering mature compatibility with high dielectric materials and outstanding integrability on Si wafers. In addition, Ge boasts the highest hole mobility among semiconductors at room temperature. These unique properties and compatibility attributes highlight Ge's potential to not only advance classical computers beyond miniaturisation but also to serve as a platform for storing and interconnecting qubits. [7]

Importantly, the Ge platform can consistently hold various quantum computing paradigms. Manifold characteristic properties of this material make it possible: high hole mobilities, low effective mass carriers, which are controllable through confinement and strain, the absence of hyperfine interaction, the possibility to isotopically purify it, and its large and tuneable g-factor and spin-orbit interaction. For instance, high hole mobility enables well-controlled quantum dots with large energy-level spacing that relax the fabrication constraints in favour of scalability. As a result, all-electrical spin-qubits, gate-controlled superconducting qubits (i.e., gatemons), and Majorana zero modes topologically protecting quantum states, can emerge from Ge. [7]

Each technological paradigm has a preferential device morphology. Spin qubits are achieved through planar heterostructures, superconducting qubits with superconductor-semiconductor-superconductor planar devices, and topological states with hybrid nanowires. Particular challenges apply for each materials platform, for which the target of the characterisation must adapt to the key information to retrieve. For instance, planar heterostructures need to be defect-free in the central layer representing the quantum well. Consequently, structural characterisation is key to unveiling whether defects propagate from the Si wafer to the top layers of the device. On the other hand, for hybrid nanowires and super-semi-super structures, the interfaces are the most revealing features which need careful evaluation. Therefore, in the following section, we present the advanced (Scanning) Transmission Electron Microscopy ((S)TEM) characterisation of multiple devices representing these three material platforms, with special emphasis on planar heterostructures for spin qubits, and superconducting qubits for Josephson junctions also based on planar Ge.

3.2 STEM characterisation of Si/Ge quantum devices

The devices presented in this section showcase the path from the intermediate development stages to the final working device. Thus, STEM characterisation comes into play majorly when device and transport measurements reach a minimum viability of foreseen quantum performance. We intervene in optimising the devices

by correlating structural features with the functional properties they exhibit. Importantly, the iterative and progressive process of characterization of these devices has resulted in the development of devices with exceptional functional properties and operational conditions, to the point of redefining the state of the art.

However, each material technology or platform entails its growth strategies and characteristic device morphologies. Consequently, in this section, we introduce devices representing the three quantum technologies that Si and Ge can handle, and the key parameters we pinpointed when characterising their viability as operational quantum devices. Indeed, identifying these essentials was fundamental to fine-tuning the device properties to the desired performance based on growth condition shifts acting back on the structural properties. Not surprisingly, these key aspects drift from platform to platform. Generally, planar heterostructures demand a defect-free quantum well, while superconducting devices and nanowires care about interface sharpness between the superconductor and semiconductor. The details can be found in the next subsections, which are classified by the material platform. [7]

We pay special descriptive emphasis on pairing the structural properties with the experimental techniques. Thus, being aware of the intrinsic challenges linked to retrieving properties like the crystal phase and its spatial orientation, the relative and absolute strain (see the nuance in the Methodology chapter, section 2.2), the local composition, and even missing but thoroughly desired information, will guide the backbone of the automated workflow developed in the next chapters (i.e., chapter 4 and 5). Thus, the central destination of this thesis dissertation must always be kept in mind: ultimately, the workflow should be capable of automatically reproducing the manual characterisation undergone in the following devices. First, it would skyrocket the efficiency of extracting conclusions. More importantly, it would unlock a deeper layer of comprehension of any targeted system through the intervention of experimental simulations of their representative 3D models.

3.2.1 Devices towards spin qubits

The spin qubits utilize the two spin states (up and down) of single electrons or holes as the logical states 1 and 0. Spin qubits were first proved in GaAs/AlGaAs planar heterostructures, which used top metallic gates to confine quantum dots with discrete quantum states in which the electron remains. [8] Their main limitation was hyperfine interaction (i.e., electron-nuclei spin interaction), which was overcome with Si. Thus, Si contains only a 4.7% of non-zero nuclear spin isotopes (i.e., ^{29}Si), which reduces the nuclear spin noise down to negligible if isotopically purified Si is obtained. [13] As a result, Si spin qubits exhibit state-of-the-art coherence, despite presenting scalability challenges because of fabrication constraints. [12, 14, 15, 16] Remarkably, these challenges can be overcome by Ge thanks to its higher mobility and integrability with the high dielectric constant materials. Therefore, we pay

special attention to spin qubits held by Ge quantum wells, although we do not fully forget about Si quantum wells. [17, 18]

3.2.1.1 Ge quantum wells for spin qubits

In Si/Ge planar heterostructures where the quantum well is pure Ge, the band-edge alignment between compressively strained Ge and relaxed $Si_{1-x}Ge_x$ results in a type I band alignment that confines holes in the growth direction, forming a 2D hole gas. [19] Indeed, the compressive strain in the epitaxial Ge quantum well is determined by the in-(image)-plane lattice parameter of the underlying relaxed $Si_{1-x}Ge_x$ buffer layer. Practically, these $Si_{1-x}Ge_x$ buffers necessitate a high Ge content ($0.6 \leq x \leq 0.9$) to support strained Ge epilayers of a thickness ranging from 10 to 30 nm. As a result, we ensure they are smaller than the critical thickness for initiating plastic relaxation. [20, 21] The process by which the Ge quantum well is populated with holes and subsequently modulated falls outside the scope of the present work and belongs to the field of devices and transport. [7] However, the technological challenge of not reaching the critical thickness presents an opportunity for investigation in materials science. Consequently, the proposed structural characterisation focuses on assessing the concentration of defects or misfit dislocations at the interfaces of the quantum wells. This underlying cause of device malfunction must be considered alongside the growth conditions that give rise to it, guiding iterative adjustments towards achieving the ideal growth settings.

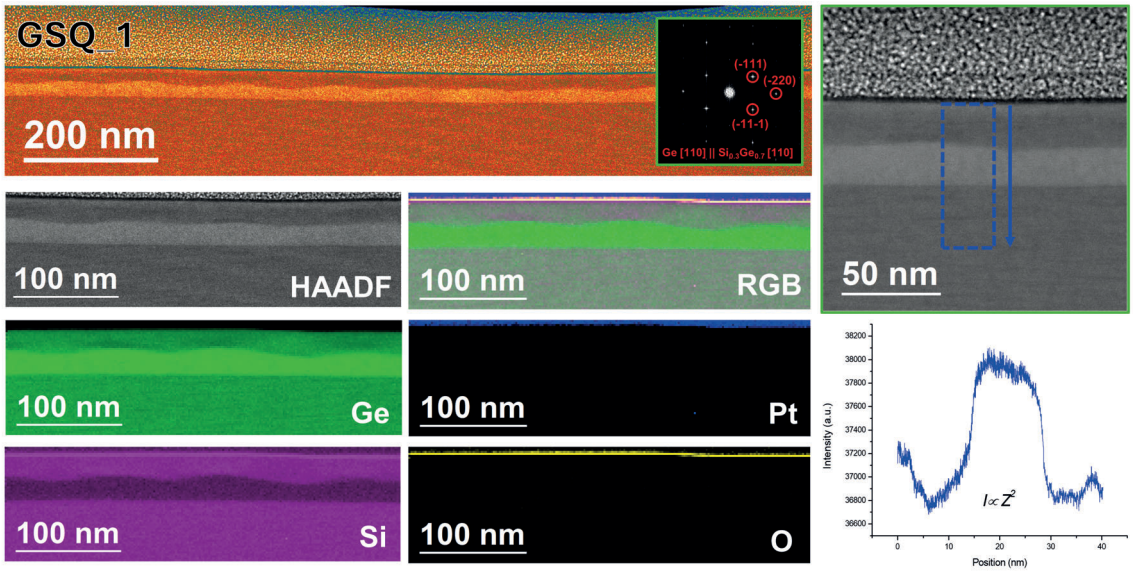


FIGURE 3.2: GSQ_1 Scanning Transmission Electron Microscopy (STEM) characterisation. The elongated High-Angle Annular Dark-Field (HAADF) overview on top depicts the wavy interface of the quantum well, oriented along the [110] zone axis, below which qualitative compositional maps can be found. On the right, a HAADF STEM intensity profile and its proportionality with the atomic number Z.

Other features can alter the optimal performance of spin qubits in Ge quantum wells. We identified these features in our initial devices, which mark the first steps towards determining the optimal growth conditions and device characteristics. GSQ_X, standing for Ge Spin Qubit, is the designation we use to refer to the Ge quantum wells utilized for spin qubits.

The first batch of three analysed devices contains GSQ_1, GSQ_2, and GSQ_3 (table 3.1). Nominally, we expected a 15 nm pure Ge Quantum Well (QW) embedded between two SiGe buffer layers. The intended composition of these buffer layers aimed for $Si_{0.3}Ge_{0.7}$. This thickness falls below the critical thickness of a Ge thin film grown on top of a fully relaxed $Si_{0.3}Ge_{0.7}$. [20, 21] Consequently, we anticipate no misfit dislocations at the interface between the QW and the bottom buffer layer. In addition, achieving suitable device performance requires a well-defined and straight interface between layers to confine energy levels and finely adjust them with the expected strain induced by a sharp and mismatched interface. The theoretical mismatch between Ge and $Si_{0.3}Ge_{0.7}$ is 1.231 %, which is sufficiently small to accommodate a dislocation-free 15 nm layer.

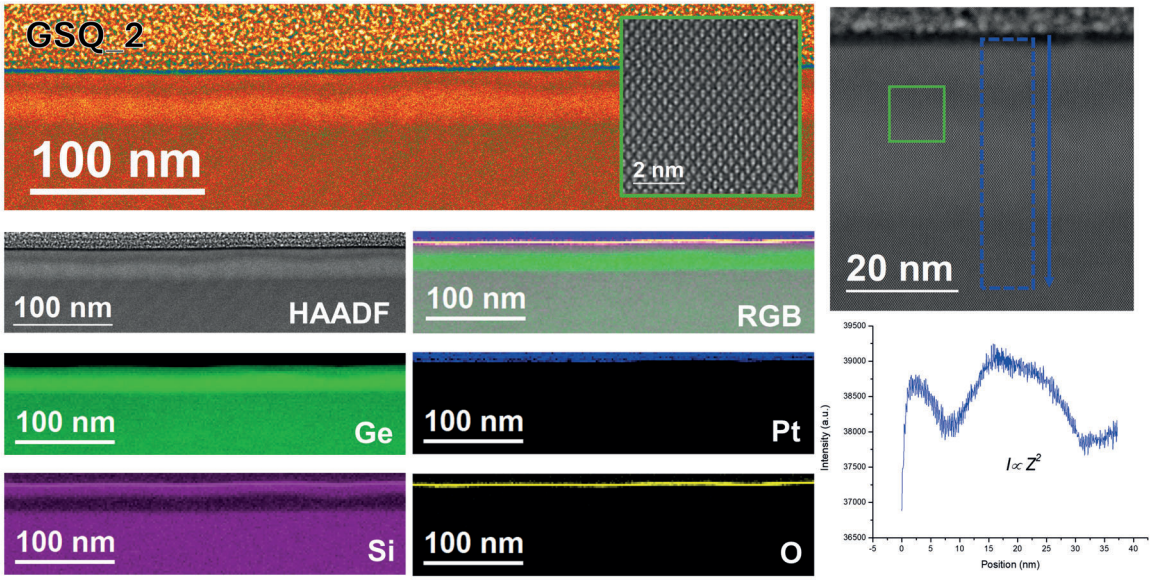


FIGURE 3.3: GSQ_2 Scanning Transmission Electron Microscopy (STEM) characterisation. The elongated High-Angle Annular Dark-Field (HAADF) overview on top depicts the diffuse interface of the quantum well, oriented along the [110] zone axis as seen in the atomic resolution inset. At the bottom, qualitative compositional maps can be found. At right, a HAADF STEM intensity profile and its proportionality with the atomic number Z, highlighting the diffuse top interface.

The device growers and designers conducted complementary characterisation of the surface morphology of the chips containing GSQ_1, GSQ_2, and GSQ_3, to evaluate their flatness. AFM was employed to identify a certain waviness on the surfaces of

GSQ_1 and GSQ_2. Conversely, GSQ_3 exhibited a flat surface consistent with the ideal scenario.

The role of the (S)TEM characterisation is to evaluate the morphology of the QW and the buffer layers and correlate it with the AFM measurements. First of all, the atomic resolution STEM indicates we have Ge and, with no compositional information yet, $Si_{1-x}Ge_x$ in the cubic diamond-like crystal phase (i.e., space group 227, Fd-3m). They are imaged along the [110] zone axis, with the epitaxial planes (i.e., vertical planes) being the (-220). We use the crystallographic information encoded in the FFT of the atomically-resolved micrograph to identify the crystal phase. This can be observed in figure 3.2, where we extract the FFT of the HAADF micrograph and add it as an inset on the right side (green squared pattern). The resulting pattern is the prototypical reciprocal space representation of the [110] zone of a cubic material. The confirmation encompasses the goodness of fit evaluation of model unit cells adjusting to the experimental interplanar distances and separating angles represented by the reflections arising in the FFT. Thus, our prior knowledge constrains our fitting of crystal phases to those containing just Si and Ge. In this case, the diamond cubic structure is the winning fitting. The result can be double-checked by plotting the characteristic crystallographic facet exposed by the [110] zone axis, with the atomic dumbbells (i.e., pairs of atoms) as can be seen in the image insets (squared in green) in figures 3.3 and 3.4.

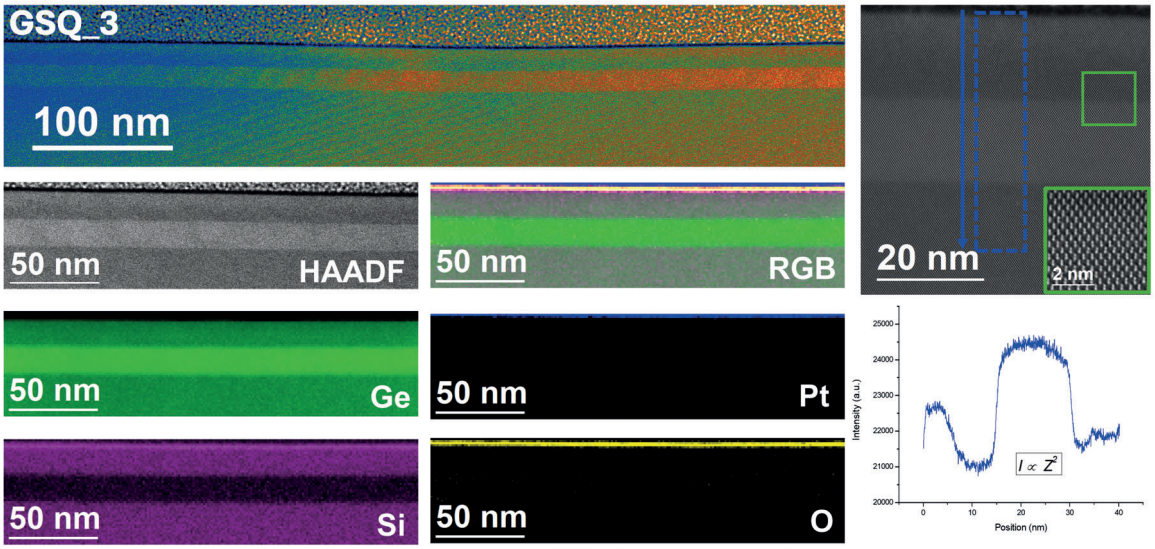


FIGURE 3.4: GSQ_3 Scanning Transmission Electron Microscopy (STEM) characterisation. The top panel shows an elongated High-Angle Annular Dark-Field (HAADF) overview, revealing the sharp and straight interface of the quantum well, despite the bent lamella. The atoms are imaged along the [110] zone axis, as seen in the green squared atomic resolution inset. At the bottom, qualitative compositional maps are presented. Additionally, on the right, an HAADF STEM intensity profile emphasizes the sharp interface thanks to its Z-contrast.

This methodology for phase identification is of utmost importance for the present thesis dissertation. Importantly, every phase identification presented in it is based on model fitting and database collation. Furthermore, it will set the basis of how phase identification is handled autonomously, requiring model fitting and database filtering. Further details can be found in chapter 4.

We make use of the relation between pixel intensity and atomic number to test interface flatness and sharpness (see Methodology, chapter 2). This idea is depicted individually per device in figures 3.2, 3.3 and 3.4, and globally as a direct comparison in figure 3.5. The temperature colour scale was applied to the HAADF micrographs as a way to emphasise contrast shifts. We observe three combinations of diverse sharpness and flatness in the devices. First, GSQ_1 exhibits waviness while maintaining sharpness. Second, GSQ_2 presents a more diffuse interface despite being flatter than GSQ_1. On the other hand, GSQ_3 emerges as the best candidate in terms of morphology, as it combines both flatness and sharpness in the QW. Consequently, the conditions necessary to replicate GSQ_3 would allow for reproducible engineering. The flatness and, more importantly, the sharpness pattern are evident in figures 3.2, 3.3 and 3.4, where we plot the pixel intensity profile against the position. Notably, the sudden intensity ramp is unique to GSQ_3. In contrast, GSQ_1 and especially GSQ_2 exhibit gradual intensity changes due to the combined influence of waviness and diffuse interfaces.

In thin films, waviness arises due to the growth process not allowing full relaxation of the growing mismatched layers. This phenomenon is akin to applying horizontal compressive mechanical loading, which can lead to local surface wrinkles or global buckling in the thin film. [22, 23] In the present case, this loading effect is equivalently induced by the mismatched heterostructure and the larger cell parameter of Ge (with a bulk value of 5.6580 Å) compared to $Si_{0.3}Ge_{0.7}$ alloy (with an estimated bulk value of 5.5899 Å). Typically, cell parameters of binary compounds are computed using Vegard's law (as expressed in equation 3.1). However, it is well-known that a slightly more precise trend can be obtained by following Equation 3.2 (which provides cell parameters in Å). [24] The compressive stress experienced by Ge due to the mismatch results in the observed features, which would not arise under tensile stresses that tend to relax and flatten the layers further. [23]

	GSQ_1	**GSQ_2**	**GSQ_3**
Top buffer layer	$Si_{31\pm3}Ge_{69\pm3}$	$Si_{31\pm3}Ge_{69\pm3}$	$Si_{32\pm3}Ge_{68\pm3}$
Quantum well	$Si_{5\pm4}Ge_{95\pm4}$	$Si_{6\pm4}Ge_{94\pm4}$	$Si_{4\pm4}Ge_{96\pm4}$
Bottom buffer layer	$Si_{30\pm3}Ge_{70\pm3}$	$Si_{30\pm2}Ge_{70\pm2}$	$Si_{30\pm3}Ge_{70\pm3}$
Interface	Sharp	Diffuse	Sharp
Waviness	Pronounced	Intermediate	None
Thickness (nm)	[20, 23]	[20, 23]	15.5

TABLE 3.1: GSQ_1, GSQ_2, and GSQ_3 structural comparison.

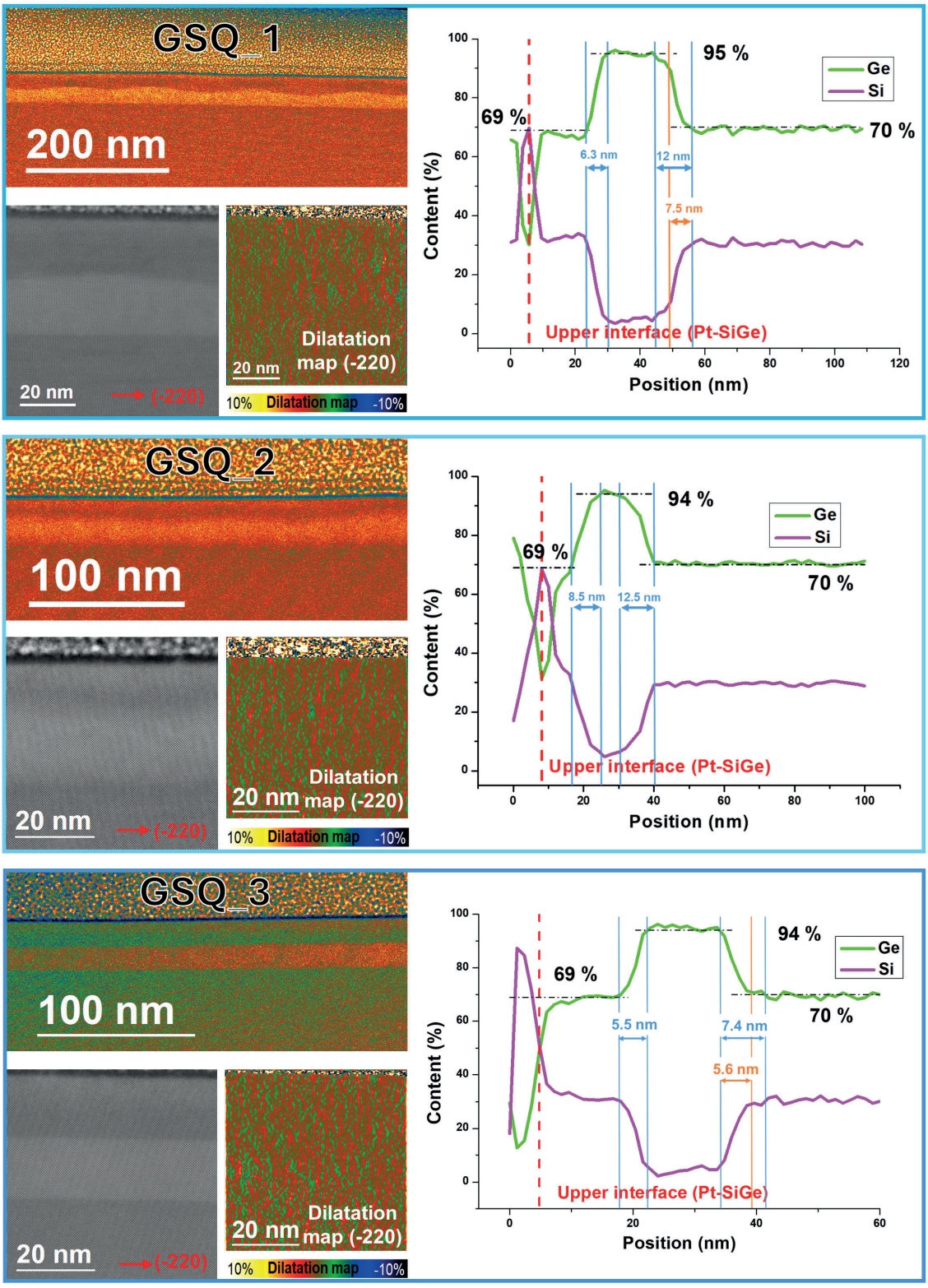


FIGURE 3.5: Comparative Scanning Transmission Electron Microscopy (STEM) characterisation of devices GSQ_1, GSQ_2 and GSQ_3. On top, the morphology of the thin film is compared through High-Angle Annular Dark-Field (HAADF) imaging. In the middle, the Geometrical Phase Analysis (GPA) highlights the homogeneity of the interplanar spacing of the epitaxial planes in the three devices. Lastly, the quantitative Electron Energy Loss Spectroscopy (EELS) profiles showcase the average stoichiometry per region.

Surface roughening occurs due to surface diffusion, often promoted by subsequent annealing processes during overall growth. This surface diffusion acts as a stabilising mechanism under conditions of low strain. [23, 25] Together with The key difference between GSQ_1 and GSQ_2, and GSQ_3 lies in their growth characteristics. The first two exhibit a faster growth range (higher temperature or pressure), which hinders strain accommodation. Additionally, GSQ_1 and GSQ_2's greater thickness, closer to the critical thickness, results in the observed wavy patterns, where valleys often align with anticipated dislocation spots (table 3.1). To mitigate this effect, pursuing the growth observed in GSQ_3 is recommended.

$$a_{A_xB_{1-x}} = xa_A + (1-x)a_B \tag{3.1}$$

$$a_{Ge_xSi_{1-x}}(angstroms) = 5.431 + 0.2x + 0.027x \tag{3.2}$$

Interestingly, no significant differences in the dilatation maps computed by GPA can be observed. Indeed, the values are homogeneous throughout the full width and height of the QWs. The dilatation maps indicating the dilatation of the (-220) planes ruling the epitaxy entail a QW perfectly adapting to the cell parameters of the buffer layer. Consequently, the absolute strain directly correlates with the mismatch of the Ge and $Si_{0.3}Ge_{0.7}$ substrate with Ge being the cell parameter of reference that changes.

The concept of morphology limitation in terms of intensity defined by HAADF STEM is crucial. It will enable assessing the sharpness of interfaces within more complex interface heterostructures encountered in this chapter. Importantly, it will establish the foundation for clear segmentation and region definition in the next chapters. For instance, we have manually or visually evaluated the waviness and sharpness of interfaces. However, in automated workflows, special attention will be required to identify regions of interest for further quantitative analysis. Retrieving this information quantitatively is essential. Furthermore, it is worth noting that HAADF STEM intensity is significantly influenced by thickness or effective thickness. In Figure 3.4, the transition from blue to orange in the HAADF STEM image occurs due to the bending of the lamellae. This bending introduces thickness variations that affect the interaction with the electron beam. Analogous to the effect of atomic mass, this leads to a shift in electron scattering at high angles.

Additionally, EELS was essential to verify whether the nominal values align with the actual composition, as well as disclosing misleading features observed through HAADF STEM. For example, a brighter intensity is observed immediately below the SiO_x protective layer, suggesting a potential Ge accumulation. Note that this hypothesis is based solely on pixel intensity, as depicted in the dark blue plots in figures 3.2, 3.3, and 3.4. Remarkably, quantitative EELS profiles revealed that this

effect is not due to actual Ge accumulation but rather a thickness modulation caused by the protective layer, without any real compositional shift. The qualitative maps presented in Figures 3.2, 3.3, and 3.4 show consistent homogeneity in composition across all cases, indicating successful controlled growth. The waviness is also clearly depicted in the qualitative maps, although their most representative feature is how they showcase if the interface is sharp or diffuse.

Quantitative profiles are also depicted in figure 3.5. First, the three coincide in the absence of the Ge accumulation immediately below the protective oxide layer, labelled with a vertical dashed red line as "Upper interface Pt-SiGe". Secondly, we computed the average of the Si/Ge stoichiometry relations in every region. The values define the reference stoichiometry the change of which supposes the start of an interface. The interface extends until it reaches the average composition of the next area. Interestingly, we can reproducibly compare the interfaces and quantify their extent. GSQ_2 stands as the most diffuse or gradual interface, as observed by naked eye, and GSQ_3 is by far the sharpest interface.

Importantly, quantitative maps are obtained from a known reference. In other words, prior knowledge provides a reliable compositional reference for precise quantitative analysis. For instance, this prior knowledge can originate from a complementary characterisation technique like XRD or from a calibrated growth chamber. Interestingly, we use this reference value to centre the average within the region where the reference is established. Subsequently, we iteratively adjust the fitting of the background and edges of interest to align the retrieved average with that external reference value. As a result, we bypass the strict procedure of plural scattering deconvolution and its additional layer of data complexity (such as a longer spectral range or a dual EELS configuration to capture the Zero-Loss Peak (ZLP) and the edge of interest simultaneously). For the devices GSQ_1, GSQ_2, and GSQ_3, we set the reference to be the pure Ge within the QW. While we consistently match the nominal composition in the buffer layers, the values in the QW are not precisely 100%. No model fitting to the edges yielded a perfect 100% but rather 95% or 94%. The expected experimental uncertainty for these deviations is approximately 3-5%. Interestingly, contamination from Si in the growth chamber may impact the purity of the Ge growth, contributing to these deviations.

The previous quantification strategy is followed throughout the thesis, and although it will not be directly automated, its integration and output will be crucial for achieving the central goal. Retaining this information while maintaining data simplicity is an essential requirement for the convenience of the data analysis routine, a philosophy kept since the early stages of development.

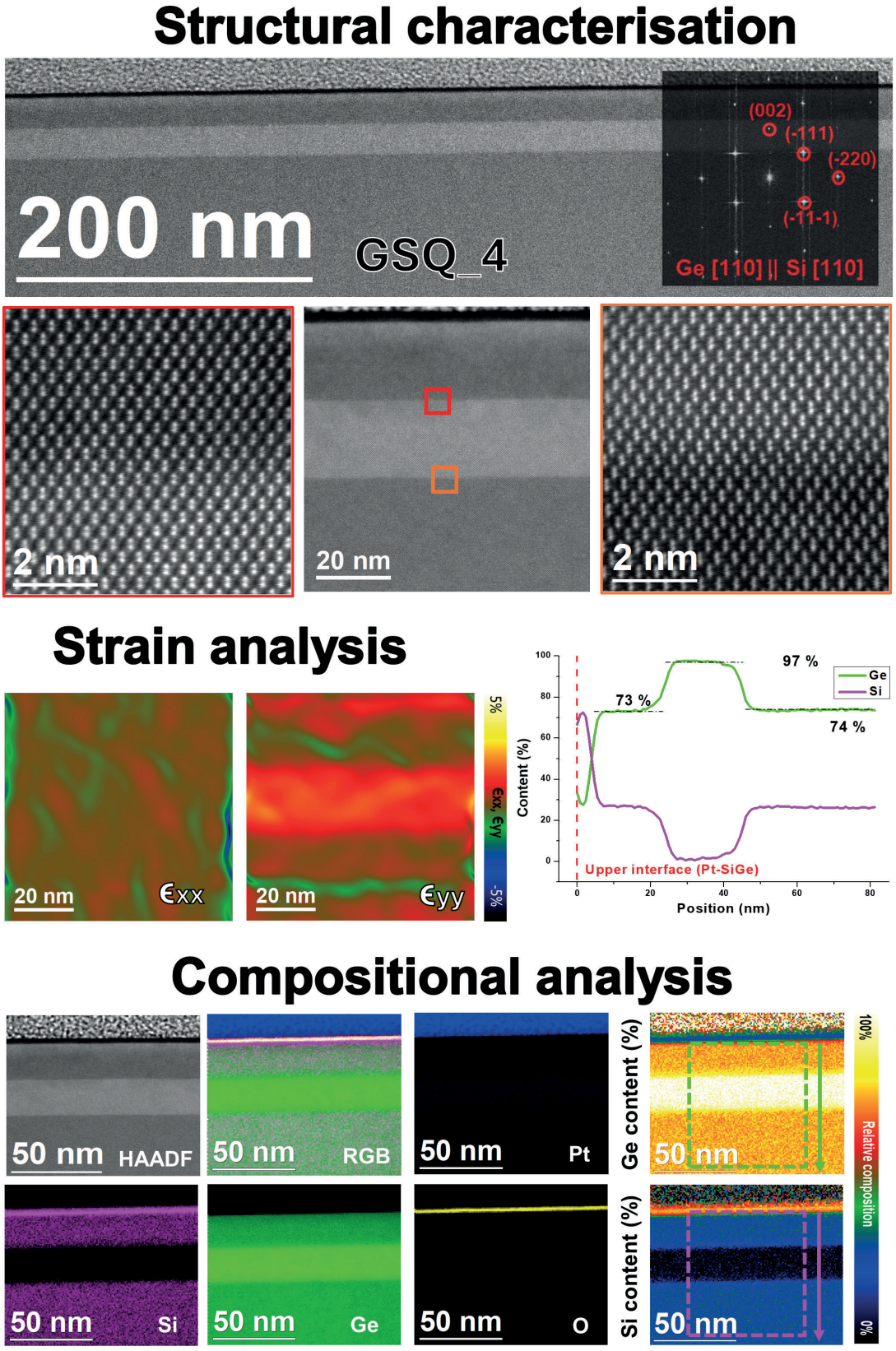


FIGURE 3.6: GSQ_4 Scanning Transmission Electron Microscopy (STEM) characterisation. The crystal phase identification and structural analysis through High-Angle Annular Dark-Field (HAADF) showcases the absence of defects or misfit dislocations in the layers. In the middle, the Geometrical Phase Analysis (GPA) computes the horizontal and vertical components of the in-(image)-plane strain. Lastly, Electron Energy Loss Spectroscopy (EELS) qualitative and quantitative maps agree with sharp, straight and compositionally homogeneous materials. [26]

GSQ_1, GSQ_2 and GSQ_3 provided valuable insights into the path towards optimal transport and quantum properties. These findings emphasized alignment with the results from GSQ_3. The key lies in achieving a sharp and straight interface, an essential requirement for the ultimate device. As a result, the growth process was iterated until conditions converged in GSQ_4. This novel approach surpassed the state-of-the-art performance with a singlet-triplet spin qubit contained in two quantum dots confined by electrostatic potential. [26] The successful growth strategy employed low-energy plasma-enhanced chemical vapour deposition. Starting from a pure Si (001) wafer, a graded layer (approximately 10 μm thick) with increasing Ge content was grown at a high deposition rate and temperature (5-10 nm/s and 760ºC) until reaching 70% of Ge content. The threading dislocations are kept in this specific layer. Subsequently, the layer with 70% Ge composition was extended by an additional 2 μm at 550ºC. The final growth segment involved a $Si_{0.3}Ge_{0.7}/Ge/Si_{0.3}Ge_{0.7}$ structure, grown at a lower deposition rate and temperature (0.5 nm/s and 350ºC) to mitigate the effects of surface diffusion and wrinkling. Finally, a layer of native SiO_x, was deposited to complete the process. The observed performance boost underscores the superiority of the present growth approximation over traditional thermal chemical vapour deposition methods.

Additional devices were engineered and characterised in the process of obtaining GSQ_4. For instance, GSQ_5 was grown similarly but with a shift in the hydrogen pressure when acting as a carrier gas in the growth chamber. The changes were detrimental to device performance. Not surprisingly, given the structural characterisation, we would attribute the effect to decreased interface sharpness. Indeed, the sharpest and most straight device coincides with the best performing one, GSQ_4. Its thorough STEM characterisation is detailed in figure 3.6. The lamella is cut to expose the [110] zone axis with the spatial orientation found so far. Remarkably, the high-magnification insets reveal an atomically sharp interface reinforced by the EELS maps. Its homogeneity is thorough and the Si and Ge quantification deviates from the nominal range only within the experimental uncertainty (i.e., 3%).

The paragraphs above highlight the importance of strain in controlling quantum devices and the splitting of quantum states, such as heavy holes and light holes. From GPA, we calculate the relative strain by considering the $Si_{0.3}Ge_{0.7}$ as a reference for the relaxed region. Importantly, when examining STEM micrographs, the information is projected and averaged along the axis parallel to the electron beam. This allows us to consistently retrieve the in-(image)-plane components of the strain tensor. The conversion from relative strain to absolute strain was done according to the following criteria.

GPA considers the relative displacement from a reference, which is generally assigned to the substrate because of its generally relaxed status. [27, 28] Conveniently, computing strain by considering a reference area for the relative strain calculation helps minimize the undesirable effects drift can have on the results. Then the strain

of a region like the QW, computed along an in-(image)-plane axis, i, with a characteristic plane having an interplanar distance, d_i, is defined by equation 3.3 (in percentage). Note that the calculation stands for the elastic component of strain.

$$\epsilon^{rel}_{i,QW} = 100\frac{d_{i,QW} - d_{i,subs}}{d_{i,subs}} \tag{3.3}$$

This relative measurement is directly retrieved from GPA maps. To convert relative strain to absolute strain, compositional knowledge is essential for assessing the initial configurations of the materials involved. In simpler terms, this refers to the theoretical mismatch within the heterostructure. There is no intervening strain when the theoretical mismatch aligns with the dilatation. However, if they diverge, we must evaluate the experimental value of the interplanar distance measured along the i axis:

$$d^{exp}_{i,QW} = \frac{\epsilon^{rel}_{i,QW} d^{theo}_{i,subs}}{100} + d^{theo}_{i,subs} \tag{3.4}$$

Finally, we compare the experimental value for that interplanar distance with the theoretical relaxed value:

$$\epsilon^{abs}_{i,QW} = -100(1 - \frac{d^{exp}_{i,QW}}{d^{theo}_{i,QW}}) \tag{3.5}$$

which is the real percentage of change the material perceives because of the mismatched heterojunction with respect to its fully relaxed status.

The strain values we obtain are consistently small, hovering around 1 %. The theoretical mismatch of 1.231 % induces perfect epitaxy of the vertical (-220) planes. Consequently, the (-220) planes of Ge compress to accommodate the smaller interplanar distance of the buffer layer below. Specifically, they experience compression by 0.9602 %. Despite deviating slightly from the total mismatch relaxation, this strain serves a crucial purpose: it facilitates the coherence of the crystal planes and impedes the emergence of misfit dislocations. Interestingly, the strain that horizontally adapts the epitaxial planes finds spatial compensation through the expansion of the horizontal (002) planes. Intuitively, the tensile strain experienced by the (002) planes aligns with the spatial arrangement of the compressed atoms, following Poisson's ratio (as expressed in Equation 3.6). In particular, the tensile strain of the horizontal planes of GSQ_4 measures 0.2689 %, defining a Poisson ratio consistent with the theoretical value for Ge, namely 0.28.

$$\nu = -\frac{\epsilon_{trans}}{\epsilon_{long}} \tag{3.6}$$

Indeed, the transversal strain accommodation of the longitudinal inception is the main compensation mechanism, but not the only one. Sample preparation and lamellae thinning create an additional axis from which to expect strain release. Thus, in the absence of thinning, the effective infinity in the direction parallel to the electron beam hinders the displacement of the atoms in that same direction. The effectively infinite number of atoms to displace, infinitely enlarges the required energy, making it unfavourable and simultaneously promoting the release in the remaining free axis. Contrarily, by cropping the sample we add an interface and we remove this infinity. Consequently, the displacement is plausible in that direction too, and the strain release will be a combination of two transversal components. The weight of these two components relates to the type of material and thickness of the electron-transparent sample.

In the context of (S)TEM-based strain analysis, it is crucial to account for the influence of the lamella when correcting for single-axis longitudinal strain compensation. Failing to do so can lead to a significant underestimation of the global relaxation mechanism, unless there is evidence of its independence from the sample thickness. Surprisingly, this critical step is often overlooked by the scientific community, despite its clear implications for method validation. The complexity of evaluating this cross-check may explain why it is frequently missed. To address this, one possible approach involves mapping multiple lamella thicknesses and analysing the strain using consistent terms. The goal is to identify a pattern that accounts for the in-depth strain component. However, implementing this approach presents several experimental challenges. First, accurately measuring the lamellae thickness is essential. Additionally, characterising multiple lamellae in detail is a resource-intensive task.

Alternatively, relying on simulations can loosen these experimental restraints. Working with Finite Element Simulations (FEM) is convenient due to their capability to handle systems of any size (i.e., suitable to model mechanical deformations in space rockets and in nanodevices like our quantum wells). The underlying strategy is to simulate the release of the strain of the device when embedded in a thin lamella, with or without periodic boundary conditions in the surface cut. Practically, this informs us about whether the thinning changes the strain release mechanism when compared to the original full uncut device.

Interestingly enough, we followed the previous strategy to unveil whether the provided absolute biaxial strain values can be trusted. The results are presented in figure 3.7. An unthinned system and a thinned one are simulated to address the effect of cutting the lamellae. In both, we set the bottom surface as fixed (fixed null displacement) and the top surface as free (fixed zero stress). The difference stands in the facets of the lamella. In the unthinned version, cyclic boundary conditions are defined to account for a virtually infinite scenario. On the other hand, we define two free surfaces of null stress for the cut version.

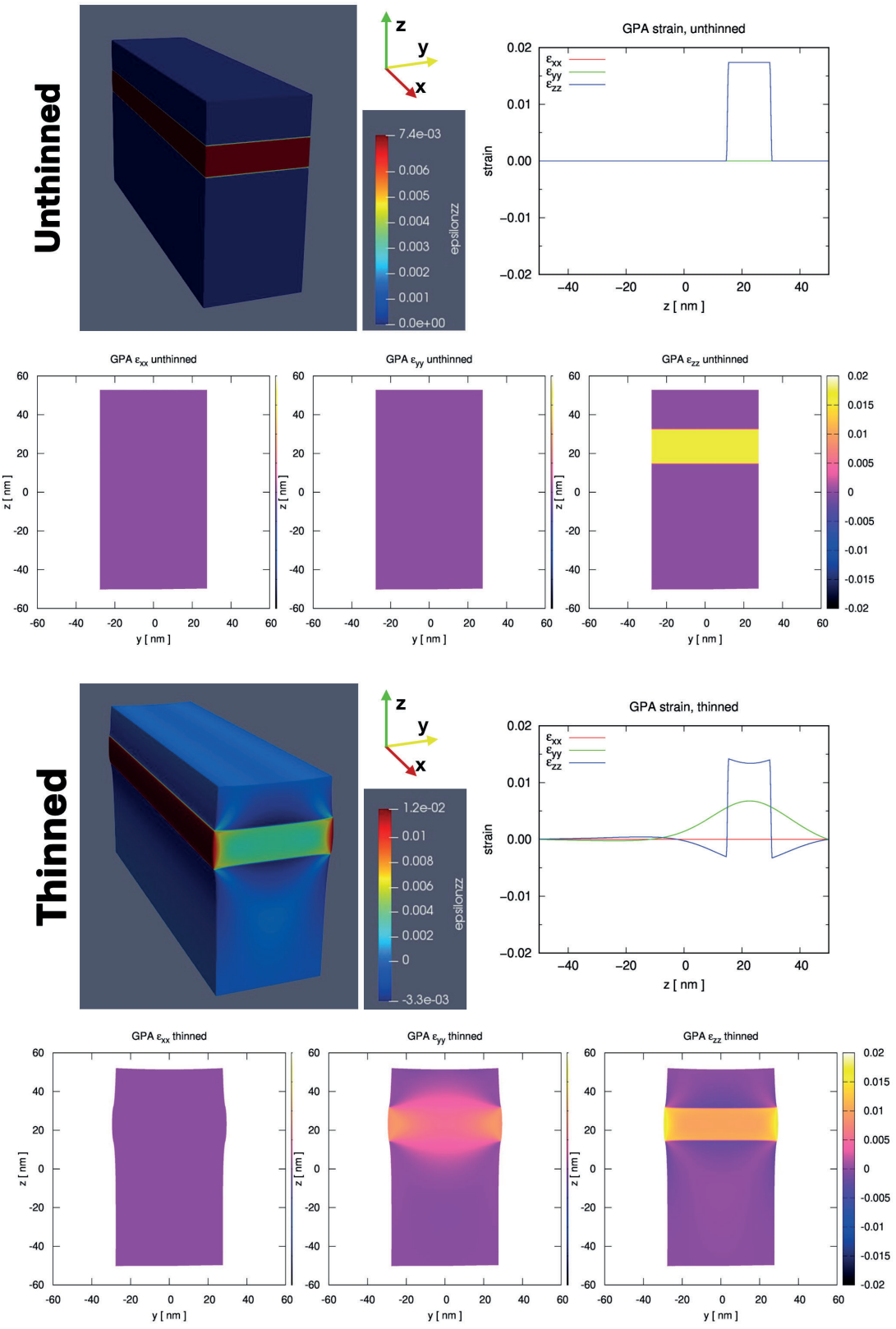


FIGURE 3.7: Effect of lamella thinning in the calculation of the strain components through Geometrical Strain Analysis of the device GSQ_4. Finite Element Simulations (FEM) compare an unthinned (i.e., infinite) and thinned system, indicating a strain release along the y axis (i.e., parallel to the electron beam) induced by the thinning.

Note that the graphs express relative strain for the direct comparison with GPA. The thickness of the lamella is limited to the 50 nm directly reported from the FIB preparation. The simulations ran in the open-source FEM software OpenFoam. [29]

As expected, the unthinned original device only presents a single transversal component. The relative strain is null along x, in agreement with GPA, and dilatation is observed only along the confinement direction (y in the experiment, z in the simulation). However, if we consider the thinned version, there is an additional component of strain release along the axis of the cut (y in the simulation). Thus, dilatation is partially contained in the confinement direction and in the cut one due to the additional degree of freedom the cut provides. As a result, the amount of strain released along the confinement dimension is smaller than what we would observe with no cut, as part of it is released in the axis parallel to the electron beam, being therefore invisible to us. In other words, we measure a smaller quantity of relative transversal strain than the expected along that dimension. Significantly, then, a correction considering sample preparation must be taken into account if further conclusions are extracted from the provided GPA analysis, at least in this particular system.

Notably, the central goal defined in the goals section 1.2 branches out to systematically provide models of the devices to quickly address concerns like the lamella thinning effect. As the answer might be device-dependent and must entail a case-specific simulation, skipping the manual modelling leads to a strategic advantage. In addition, these models can potentially lead to scientific questions never asked before due to the possibilities born from the direct use of experimental knowledge to build them. Remarkably, our workflow successfully answers this need, as we explain in detail in chapter 4.

To wrap the present section up, we exhibit the starting steps towards improving the quantum properties of the already outstanding GSQ_4 device. The next step encompasses the reduction of noise coming from the nuclei spin affecting the carriers (i.e., heavy holes of Ge) by isotopically purifying the Ge in the QW and surrounding layers. Device GSQ_6 does not yet present isotopic purification but guides the path towards a growth process that would eventually incorporate this refinement. In other words, GSQ_6 is a trial towards optimising the growth recipe to, once optimised, apply it to the more expensive purified material. The purification would lead to zero nuclear spin ^{72}Ge, ^{74}Ge or a combination of both isotopes, as their sum is more than 50 % of the abundance of Ge. The exact purifying strategy would need further investigation, as the isotopic purification of Ge is an absent concern in the literature. [30, 31]

The growing team was split in two to deliver a two-step growth. First, the growing team in the Laboratory for Nanostructure Epitaxy and Spintronics on Silicon (L-NESS) from the Politecnico di Milano, in Como, Italy, grew the device in the same way as explained for GSQ_4. Nevertheless, their contribution ends after growing

the 2 μm of constant $Si_{0.3}Ge_{0.7}$ buffer. Afterwards, the growth was transferred to the Chinese Academy of Science (CAS) in Beijing, China, to add 250 nm of constant composition $Si_{0.3}Ge_{0.7}$, 20 nm of strained Ge, 20nm of the $Si_{0.3}Ge_{0.7}$ top barrier, and 2 nm of Si to induce the protective native oxide. Interestingly, the latter step would ultimately imply isotopic purification thanks to the expertise of the conducting laboratory in China. Importantly, the growth of the first segment is thoroughly optimised in Como and directly calibrated and correlated to the electrical and transport measurements carried out at the Institute of Science and Technology Austria (ISTA). This justifies splitting the growth process, forcing it to only re-optimise its final stage.

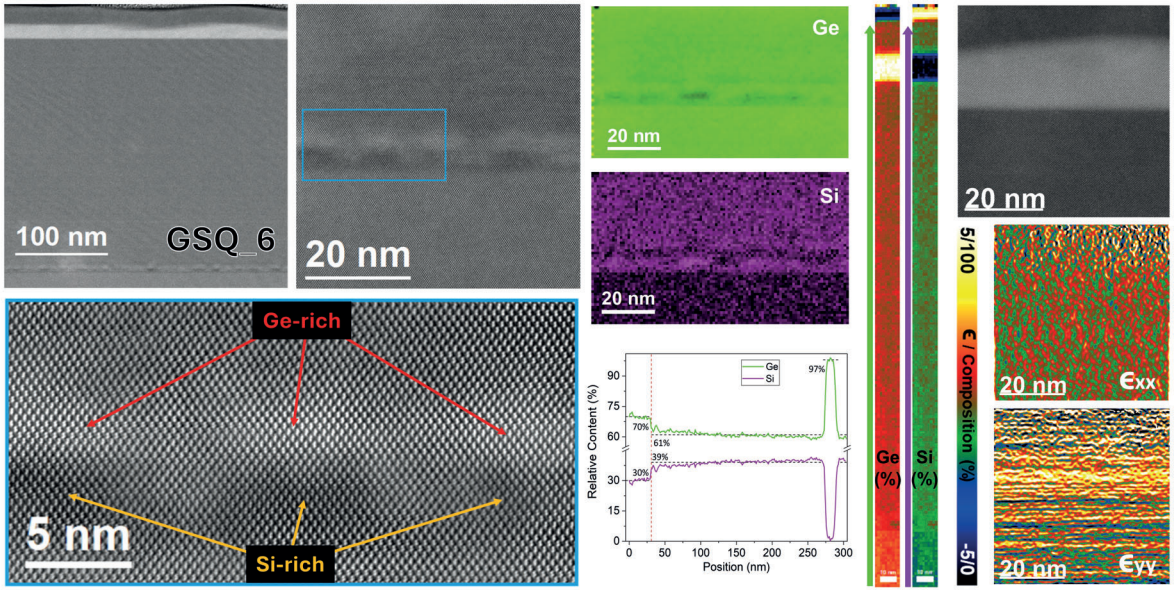


FIGURE 3.8: GSQ_6 Scanning Transmission Electron Microscopy (STEM) characterisation. High-Angle Annular Dark-Field (HAADF) overview of the wavy and diffuse upper interface of the quantum well. Special emphasis on the intensity modulations in the interface where the growth paradigm changes. Electron Energy Loss Spectroscopy (EELS) proves they are Ge-rich and Si-rich islands. The strain analysis showcases uniformity in the epitaxial planes, total adaptation and the absence of misfit dislocations or defects in the quantum well.

Figure 3.8 showcases the STEM characterisation of GSQ_6. The main feature of interest that we emphasize is the interface between the two growth steps. This interface exhibits distinguishable contrast blobs in the HAADF STEM micrographs, indicating compositional shifts from the expected stoichiometry. Interestingly, EELS further confirms the presence of Ge and Si-rich accumulations at this interface. These artefacts suggest an unwanted bumpy transition that requires additional smoothing and refinement. However, this is not the only issue revealed by the STEM characterisation. We obtained a quantitative EELS profile covering the full extension growth during the second growth stage and the last nanometers of the first. Surprisingly, the quantitative profile shows an unexpected compositional discontinuity, once again coinciding with the interface between the growth steps. Specifically, the

expected 70% of Ge content, which was well-calibrated in the initial steps (highlighting the importance of maintaining the original growth conditions), transitions to 61 % of Ge when initiating the second growth interval. This miscalibration results in a $Si_{0.39}Ge_{0.61}/Ge/Si_{0.39}Ge_{0.61}$ heterostructure, which likely contributes to the formation of a wavy and diffuse QW, as observed in devices GSQ_1 and GSQ_2.

Despite not meeting the expected quality standards and desired compositional profile, GSQ_6 constitutes a proper starting point to push GSQ_4 and its performance. [26] Despite the unexpected compositional misadjustment summing on the strain and causing wrinkling in the QW, promising quantum measurements were already obtained. Notably, the accumulation of holes in the QW in the absence of bias is a convenient characteristic to eventually confine them as liked. In fact, traditional schemes of reversed graded growth do not accomplish unbiased hole localisation. [32] Thus, these results suggest that the growth inception might be influential to hole confinement despite its long-range action, tracing a path to future devices.

3.2.1.2 Si quantum wells for spin qubits

The attainment of a reproducible large-scale platform of Ge-based spin qubits is promising, as evidenced in the previous subsection. However, its progress is mimicking the also promising path carved by silicon. [13, 16, 33] However, its main underlying challenges are encountered when tackling scalability and the fabrication constraints that hamper it. [34, 35] For instance, as happened in regular CMOS technology, its integration with the dielectric materials. [13, 36] Therefore, the current research on Si-based spin qubits diverged from the demonstration of the small-scale building units to the increasing interest in scalability towards many-qubits systems.

Precisely, the devices we are about to introduce are based on a well-established concept for housing spin qubits. Their purpose is to enhance one or more properties that significantly impact their performance and scalability. These devices are labeled as SSQ_X, representing Silicon Spin Qubits. These devices consist of three Si quantum wells with varying thicknesses. SSQ_1, SSQ_2 and SSQ_3 follow the same growth pattern but have nominal thicknesses of 8, 6 and 4 nm, respectively (figure 3.9). However, it is important to note that the nominal thickness tends to underestimate the experimental measurements.

Importantly, HAADF STEM was used once again to exploit the intensity dependence on atomic number, allowing us to detect intensity shifts at the interfaces. It is essential to note the contrast inversion compared to previous devices with Ge QWs: now, the smallest atomic number resides within the QW. The nominal composition of the buffer layers is $Si_{0.7}Ge_{0.3}$, resulting in a negative mismatch of -1.24 %. We meticulously double-checked the nominal composition using quantitative EELS measurements. The obtained profiles were traced along the full device to assess compositional homogeneity. Remarkably, we retrieved homogeneous profiles,

and their average values align closely with the nominal composition, falling within the uncertainty intervals (approximately 3 %) for all three devices. Interestingly, we even resolved a 1 % increase in Si composition in the graded buffer layer immediately below the buffer layer in contact with the QW, at a distance of 350 nm from the SiO_x capping layer. Notably, SSQ_1, SSQ_2 and SSQ_3 consist of nearly perfect QWs. They exhibit a straight horizontal extension without any signs of diffusion. In contrast, the Ge-based spin qubit platform is relatively young, which justifies evaluating imperfect devices like GSQ_1, GSQ_2, and GSQ_6. Meanwhile, the Si platform has already progressed beyond this stage and now faces additional challenges.

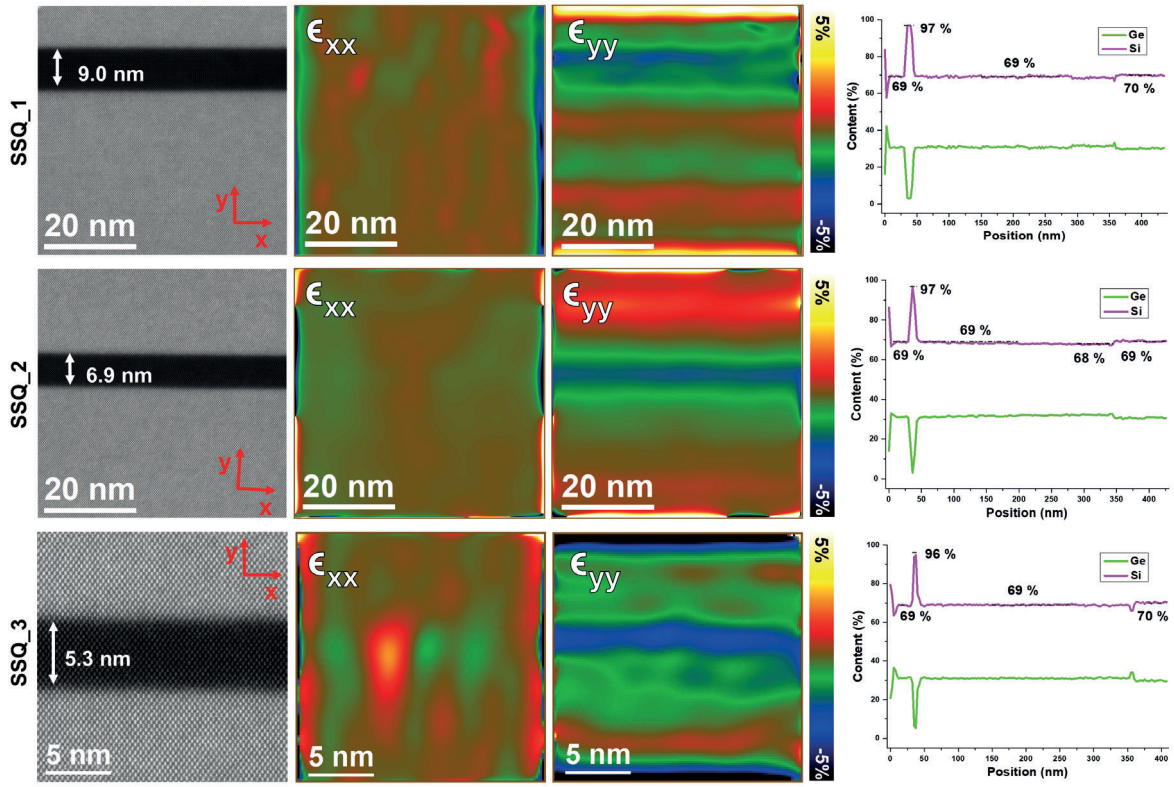


FIGURE 3.9: SSQ_1 vs SSQ_2 vs SSQ_3 Scanning Transmission Electron Microscopy (STEM) characterisation. The High-Angle Annular Dark-Field micrographs depict the thickness of the layer and enable the strain analysis through Geometrical Phase Analysis (GPA). The quantitative Electron Energy Loss Spectroscopy (EELS) confirms the expected $Si_{0.7}Ge_{0.3}/Si/Si_{0.7}Ge_{0.3}$ relation, well-relaxed due to the 2.4 μm $Si_{0.7}Ge_{0.3}$ buffer layer.

The growth strategy is similar to the aforementioned devices. It uses a low-pressure chemical vapour deposition process. Starting from a Si (001) wafer, a 3 μm graded $Si_{1-x}Ge_x$ layer is grown in four discrete steps (x = 0.07, 0.14, 0.21, and 0.3). 2.4 μm of $Si_{0.7}Ge_{0.3}$ follow before the QW of varying thickness. Remarkably, the Si in the QW is isotopically purified. Therefore, we have ^{28}Si, which directly removes the hyperfine contribution from the equation. Finally, the upper buffer layer of 30 nm is added, and 5 nm of sacrificial Si are passivated at 550 °C, below an AlO_x layer. As with GSQs, the graded growth aspires to completely remove any threading dislocation

	SSQ_1	**SSQ_2**	**SSQ_3**
Thickness (nm)	9.0±0.5	6.9±0.5	5.3±0.5
Relative ϵ_{xx} (%)	-0.03269±0.00019	-0.1026±0.0002	-0.027±0.010
Relative ϵ_{yy} (%)	-1.5990±0.0004	-1.7005±0.0003	-1.716±0.008
Absolute ϵ_{xx} (%)	1.04±0.05	0.97±0.05	1.01±0.06
Absolute ϵ_{yy} (%)	-0.54±0.05	-0.64±0.05	-0.69±0.05
Plane counting ϵ_{xx} (%)	1.04±0.09	1.00±0.07	0.96±0.12

TABLE 3.2: SSQ_1, SSQ_2 and SSQ_3 relative and absolute strain comparison. The absolute strain calculation is computed from the compositional values defined in figure 3.9.

rising from the very first grown layers. As a result, the buffer layers confining the QW can be perfectly relaxed and the Si in the thin film can allocate the desired strain.

As part of our analysis, we employed GPA to compute the strain. Additionally, we leveraged compositional information to complement the GPA results, allowing us to translate relative strain measurements into absolute values. This workflow was consistent with the approach used for GSQ_1-6. The computed relative and absolute strain values for the three SSQ devices are summarized in table 3.2. The relative strain values directly reflect the raw measurements obtained from the GPA maps. In contrast, the absolute strain accounts for the compositional variations detailed in figure 3.9. Interestingly, there are no significant differences in the elastic strain components related to the epitaxial planes between these devices. The vertical adaption is nearly perfect, yielding consistent strain values even when comparing the GSQ devices with inverted stoichiometry. This outcome is not surprising, given that the thickest SSQ QW is approximately half of the height of GSQ_4. [26] However, when it comes to longitudinal strain, ϵ_{yy}, we observe an increase in compression (in absolute value) as the QW becomes thinner. This trend is reasonable because a thicker layer experiences less distortion due to the greater number of atoms that need to be compressed. As the layer thickness approaches infinity, theoretically, the distortion should approach zero, as depicted in figure 3.7. Nevertheless, it is essential to handle the exact values with care, considering the complexities and nuances associated with the thinning effect of the lamellae.

Interestingly, the thickness of the QWs, which as mentioned, plays a key role in defining the properties of the spin qubit, was calculated through plane counting. For that, the characterised strain in table 3.2 is crucial for the accurate determination of the thickness. Of course, the counted planes are the horizontal (002) planes, which are perfectly parallel to the interface. The workflow operates under the assumption that we have a relaxed buffer layer. It is a reasonable assumption considering the growth process we are pursuing. Calculating the interplanar distance, denoted as d_{buff}, is straightforward and follows equation 3.7. This calculation holds because both Si and Ge exhibit cubic crystal structures.

$$d_{hkl} = \frac{a_{cell}}{\sqrt{h^2+k^2+l^2}} = \frac{a_{cell}}{\sqrt{0^2+0^2+2^2}} = \frac{a_{cell}}{2} \tag{3.7}$$

In addition, GPA outputs the average dilatation of the QW planes from the buffer ones, δ. As a result, we can compute the experimental average interplanar distance of the (002) planes inside the QW, d_{QW} (equation 3.8).

$$d_{QW} = d_{buff}(1+\delta) \tag{3.8}$$

Finally, the thickness of the layer is obtained by counting the number of planes inside the QW, for which the clear intensity shift from the HAADF STEM signal is crucial. Then, the thickness t_{QW} of the QW is directly calculated by the total number of planes inside, n_{QW}: $t_{QW} = n_{QW} d_{QW}$.

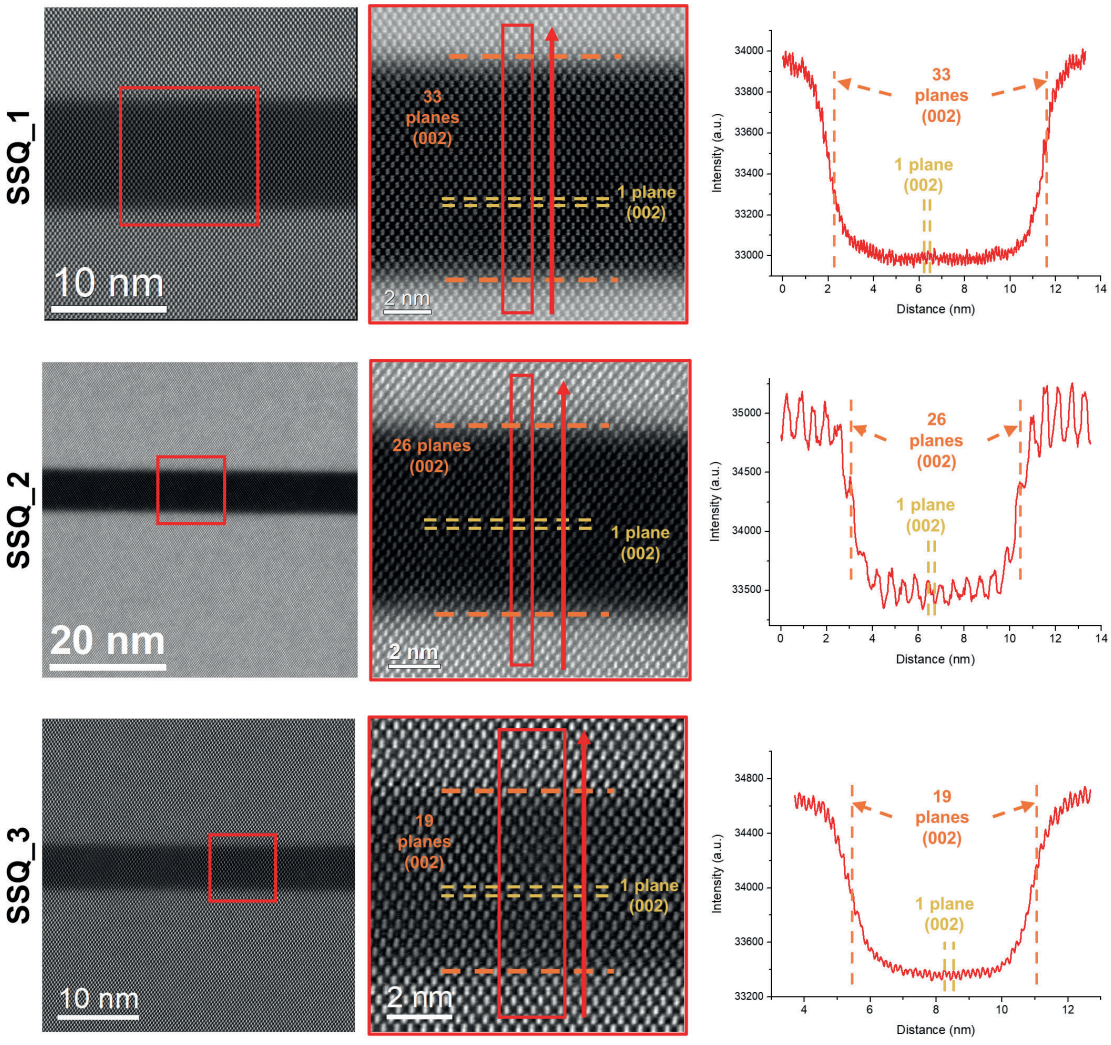


FIGURE 3.10: SSQ_1, SSQ_2 and SSQ_3 thickness comparison by plane counting, as a workflow to reduce and quantify the uncertainty of the measurement.

The uncertainty of the measurement comes from the number of planes. Thus, the measurements are replicated to improve the statistical significance, as the bias of the value comes from whether to include the first and last planes. In that sense, the

standard deviation of t_{QW} is $\sigma = 2d_{QW}$. Indeed, we quantify and consistently define the bias from a measure whose bias is typically undetermined.

These rules were applied to the three Si QWs, as depicted by figure 3.10. The graphs showcase the HAADF STEM intensity profile and how the (002) planes sit between an intensity peak and its consecutive valley. Four measurements were done per device, averaging the 9.0±0.5, 6.9±0.5 and 5.3±0.5 nm for the devices SSQ_1, SSQ_2 and SSQ_3, respectively (fig. 3.9 and table 3.2). Note that the measurement discrepancy stems from the slightly varying contrast along the horizontal axis in both interfaces. This variation can locally affect the inclusion or exclusion of planes, leading to different counts in other parts of the interface. Therefore, consistently addressing this uncertainty is crucial for accurate estimation, regardless of the individual microscopist's perception.

Computing the thickness by plane counting opens up the possibility of estimating strain by adapting the workflow. Our main motivation lies in devices SSQ_2 and SSQ_3, which have relatively few planes due to their thickness. In these cases, the associated error when averaging the GPA maps within the exact area of the QW can be significant. The limited number of planes contributes to deviations in the g vectors within that region. As a solution, we directly sample plane stacks within the QW, allowing for a pixel-precise measurement of the average interplanar distance. We extract the distance for a group of planes and divide it by the number of planes forming that stack. The retrieved interplanar distance, based on several replicas, is then compared with the expected relaxed interplanar distance, as defined in equation 3.5. It is important to note that our interest lies in the vertical planes forming the epitaxy, as this allows us to retrieve the longitudinal strain ϵ_{xx}. The results are presented in table 3.2 and are remarkably consistent with the GPA-based measurements. Interestingly, the uncertainties associated with this method are larger than those obtained using GPA for all three devices. The images where GPA is computed are high-magnification ones, implying that the well-sampled QWs occupy a significant portion of the image. Consequently, the increment of the g vector is clearly visible in reciprocal space and effectively represented by GPA. In any case, this method represents a significant improvement for lower magnification acquisitions. Additionally, it serves as a cross-check for the perfect epitaxy below the critical thickness, providing two different yet equivalent paths for validation.

Remarkably, the three devices have proven exceptional quantum and transport properties to enhance the scalability of Si-based spin qubits. SSQ_1 and SSQ_3, and especially SSQ_3, improved the scattering and uniformity of the 2D electron gas over a 100 mm wafer, significantly reducing charge noise. Importantly, charge noise degrades the performance of spin qubits hampering their scalability. The disorder arises from impurities and dislocations within the QW, the semiconductor barrier, its interface with the dielectrics and the dielectrics themselves. On the one hand, SSQ_1

and SSQ_3 focused on removing impurities from the semiconductor-dielectric interface. On the other hand, the absence of dislocations, and the outstanding crystal quality and sharpness of the QW improved an order of magnitude the dephasing times in gate-defined Si-quantum dots. [37]

Moreover, SSQ_2 was used to challenge the balance between disorder, charge noise, and valley splitting in quantum dot qubits. The contribution of this device lies in its design based on the simultaneous and dependent tackling of disorder and valley splitting. A high valley splitting separates the quantum states holding the spin qubit. However, common strategies to increase it are detrimental to disorder, and vice-versa. SSQ_2 benchmarked the path towards its mutual enhancement, also aiming for large-scale integration of multiple spin qubits within the Si platform. [38]

3.2.2 Devices towards superconducting qubits

Superconducting qubits offer a wide and promising platform to realise scalable quantum processors. [39] The platform relies on contacting semiconductors (Sm) with superconductors (S) to exploit the proximity effect of inducing superconductivity on the Sm because of proximity. [40, 41] Interestingly enough, we can exploit the same advantageous characteristics of Ge that enabled the successful spin qubits of the previous section. [7] Therefore, taking advantage of the hole mobility in Ge would unlock Superconducting QUantum Interference Devices (SQUIDs) and topological quantum computing. [40, 42] Indeed, the devices presented in this section have been successfully proven in building SQUIDs.

The major objective is obtaining a platform with a S-Sm contact with low disorder, high superconducting gap and high critical magnetic field. Aluminum (Al) extensively served as the launching point to create transparent S-Sm interfaces for the supercurrent, although limited by their small critical magnetic field. [40, 43] This forces the operation at low temperatures and low magnetic fields. To allow further magnetic resilience, we propose two strategies. First, using Al as the starting point and combining it with elements whose properties can complement the weak points of Al. Secondly, directly accessing other materials, either by exploiting already known superconductors or by discovering brand-new ones. The devices in this section are encoded as GSu_X, standing for Ge SUperconducting qubits.

3.2.2.1 Al as the base of the superconductor

We propose a S-Sm-S junction on planar Ge. Therefore, the junction uses as Sm the Ge QWs that confine the transport carriers in the layer itself. The superconductivity was initially tested using Al independently, serving as a preliminary cross-check of the reproducibility for building and integrating other materials within that device geometry. The morphology of the device is a double superconducting barrier

enabling supercurrent reflections known as Andreev reflections. The preliminary device integrating the Al contacts, GSU_01, is depicted in figure 3.11.

At first glance, asymmetry is the main distinguishable feature. Thus, the right and left metallic Al contacts do not present the same morphology as can be seen in the HAADF STEM overview at the top of figure 3.11. In the overview, the left contact is distorted by anomalous growth of the dielectric layer of HfO_x (see the compositional distribution in the qualitative EELS maps). The cause lies in the intrinsic challenge of growing conformal HfO_x on top of the Al. Thus, the high affinity with the metal can turn into the reported distortion. Like GSQ_X devices, the SiGe alloy and Ge feature a cubic diamond crystal structure, although in the present case, the buffer layers are fixed to nominally be $Si_{0.2}Ge_{0.8}$. The change is justified because the difference in energy splitting is not as crucial as for spin qubits, because the quantum states ruling the qubit are not encoded in this manner.

The strain we measure also decreases with respect GSQ device due to the closer composition between layers. It is -0.71 % (i.e., compressive) for the vertical planes and 0.40 % (i.e., tensile) for the horizontal ones. Note that the horizontal component in the GSQ devices is around -1 % of compressive strain. Importantly, the direct comparison of the composition and strain between the GSQ and GSu devices, and its justification in terms of band splitting and confinement of quantum states exposes an exciting research question mark. To the best of our knowledge, the quantitative correlation of composition, strain and band structure is unaddressed in literature. Even more so in systems where these changes are subtle and can locally evolve across nanometric spatial ranges, like the targets of the present chapter. The main obstacle is the absence of a robust way to spatially map the band structure or its most representative characteristic, the band gap energy and type. As an answer, in chapter 5 we develop a workflow to map the band gap, and to potentially answer this kind of inquiry directly linked to the foreseen third objective in our list of goals for the present dissertation (section 1.2).

The Al contact is polycrystalline, as showcased by the HRTEM wide detail of the left contact. Its polycrystalline nature exposes varied zone axes, such as the [110] highlighted in the right contact, orange inset of figure 3.11. The same orange inset showcases a feature of discussion in the right contact. There is a focus shift in the fragment of intermediate intensity outlining the curvy Al - Ge interface. Waviness parallel to the electron beam direction is likely causing this local intensity change. As a result, the expected interface is not as sharp and continuous as desired. Remarkably, the desired overall result can be found in device GSu_02. This optimised device manages without HfO_x in favour of a gentler AlO_x. This deliberate change contributes to the overall symmetry of the device. Remarkably, the metallic contact is now split into two regions, as observed in the bottom half of figure 3.11 and in figure 3.12.

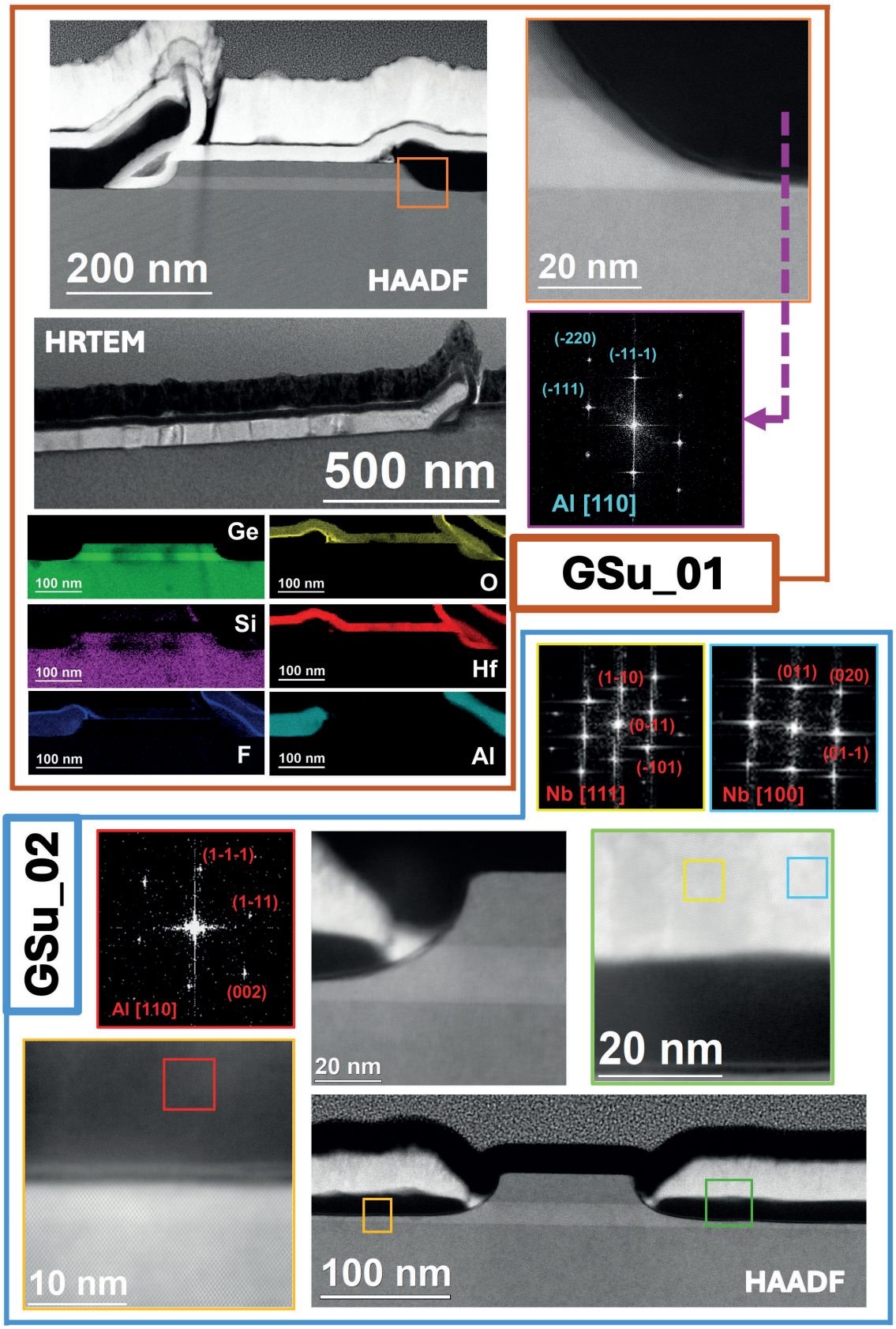


FIGURE 3.11: Complete (Scanning) Transmission Electron Microscopy (S)TEM characterisation of GSu_01 and GSu_02. Their direct comparison demonstrates the major structural quality of GSu_02, mainly highlighted by the symmetry and the absence of metallic contact deformation.

The compositional EELS maps and HAADF STEM show that the superconducting contact is built upon two metals, Al and niobium (Nb). The Al crystallises in a face-centred cubic crystal phase, space group number 225, Fm-3m, while Nb does it in a body-centred cubic phase, space group 229, Im-3m. Note that again, these crystal phases were identified by model comparison through the FFT-based approach detailed previously. This strengthens its reproducibility towards general crystal phase identification and supports its central role in the workflow detailed in Chapter 4.

The device was grown by combining reduced chemical vapour deposition for the Ge QW, and evaporation for the metallic contacts. [44] The device was fabricated using a 100-keV electron beam lithography system, where the SiGe/Ge is etched through reactive ion plasma to create room for the superconductor and ensure its contact with the Ge QW. It is plasma-based passivated to reduce contact resistance. Afterwards, the Al and Nb layers are evaporated. Finally, the thick AlO_x layer is deposited at 150ºC with plasma atomic layer deposition. [45]

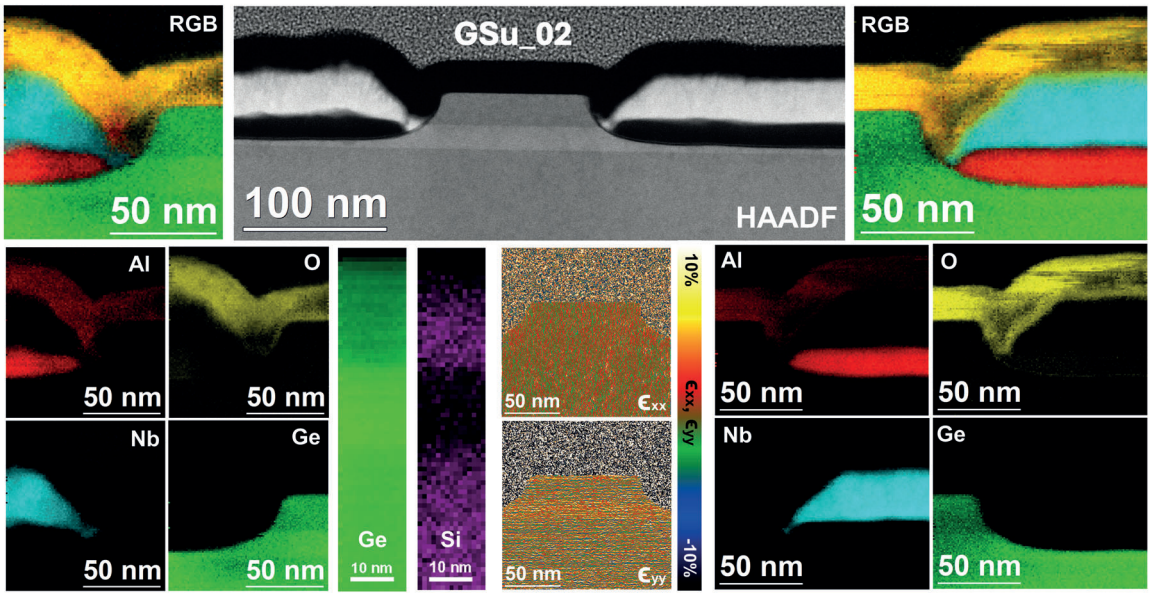


FIGURE 3.12: GSu_02 Scanning Transmission Electron Microscopy STEM characterisation. The partial qualitative Electron Energy Loss Spectroscopy (EELS) elemental maps showcase the sharp interface between Al and $Ge/Si_{0.2}Ge_{0.8}$, and Al and Nb. Interestingly, the conformal growth of the AlO_x layer does not abruptly alter the contact morphology as in GSu_01 (fig. 3.11), but slightly oxidises the last 3-4 nm of the metallic Nb. Furthermore, the Geometrical Phase Analysis (GPA) maps of the in-(image)-plane components of the strain reveals the perfect epitaxy due to a gentle elastic strain induced by the stoichiometric proximity of the implied layers. [45]

The junction implying these up to three crystallographic phases offers a rich landscape of contact types and interfaces. First, the monocrystalline Al directly contacts the $Si_{0.2}Ge_{0.8}$ with tiny (i.e., 1-1.5 nm) oxidation and amorphisation of the Al just in between, probably due to the effect of the plasma-based passivation step. Although it impedes the epitaxial relation, it promotes the preferential growth of the

Al along the [110] axis, where the pseudo-epitaxy is driven by the 111 planes. On the other hand, Nb is polycrystalline, for which epitaxy is discarded everywhere (fig. 3.11). The polycrystalline nature is proved in the blue and yellow insets of figure 3.11, in which two Nb grains are oriented along their [100] and [111] axes, respectively. Nb is also contacting the Ge QW through a small channel which at the very contact with Ge gets amorphised. In any case, the transport measurements indicated that the supercurrent advances through the Ge-Al interface and not through the Nb, as the Al defines the transparency and low-disorder of the junction. Moreover, devices made just of Nb showed no signs of supercurrent. Interestingly, the superconducting gap of the system is significantly increased because of the by default high superconducting gap of Nb. Therefore, the junction gets the best of the two worlds by exploiting the high transparency of Al and the high gap of Nb. [45] In fact, the remarkable improvement of these two functional properties permit up to fifth-order multiple Andreev reflections. Even beyond, the device proved successful as a Josephson junction-based transistor for SQUIDs.

3.2.2.2 PtSiGe as the superconductor

The other alternative is leaving Al behind and exploring materials with potential superconducting properties and ease of integrability to S-Sm-S junctions. The next eight devices, GSu_03-10 show the use of a material whose superconducting properties had never been exploited before. The only prior evidence referred to its binary counterparts PtGe and PtSi. [46] The GSu_03-10 devices represent the progress from the early stages of discovering and gathering evidence of the superconducting $PtSi_{1-x}Ge_x$, till its successful implementation and the measurement of quantum properties. [47] To report for the first time the superconductivity of $PtSi_{1-x}Ge_x$ alloys, the investigation went through the material science research loop detailed in the introduction: multiple device types are schematised, developed, grown and characterised iteratively until exploitable features are retrieved. Indeed, this case study illustrates the need for a workflow capable of loosening the resource-consuming demands of the (S)TEM characterisation. The iterative investigation of the eight-device batch presented in this section took around one year to ultimately click with the optimal device configuration (i.e., GSu_10, fig. 3.17).

The research loop starts with two three-device batches. Their mission is to evaluate the superconductor growth and its contact with the Ge QW. Importantly, the Ge QW is again the source of carriers as in the previous subsection regarding Al-based contacts. The main difference between the two batches (1^{st} batch: GSu_03, GSu_04 and GSu_05; 2^{nd} batch: GSu_06, GSu_07 and GSu_08) is the vertical position of the QW. In other words, the thickness of the upper buffer layer (1^{st} batch: around 20 nm; 2^{nd} batch: around 50 nm), which nominally is $Si_{0.2}Ge_{0.8}$. Within the batch, the three devices evolve progressively adding more complex features until the realisation of the ideal device configuration. The progression is seen in figures 3.13, 3.14 and 3.15,

which follow the same pattern of showing the same device type from the first batch on the top half, and the second batch on the bottom half. The semiconductor also adapts its thickness to be in contact with the Ge, for which a larger mass of PtSiGe alloy is expected in the second batch.

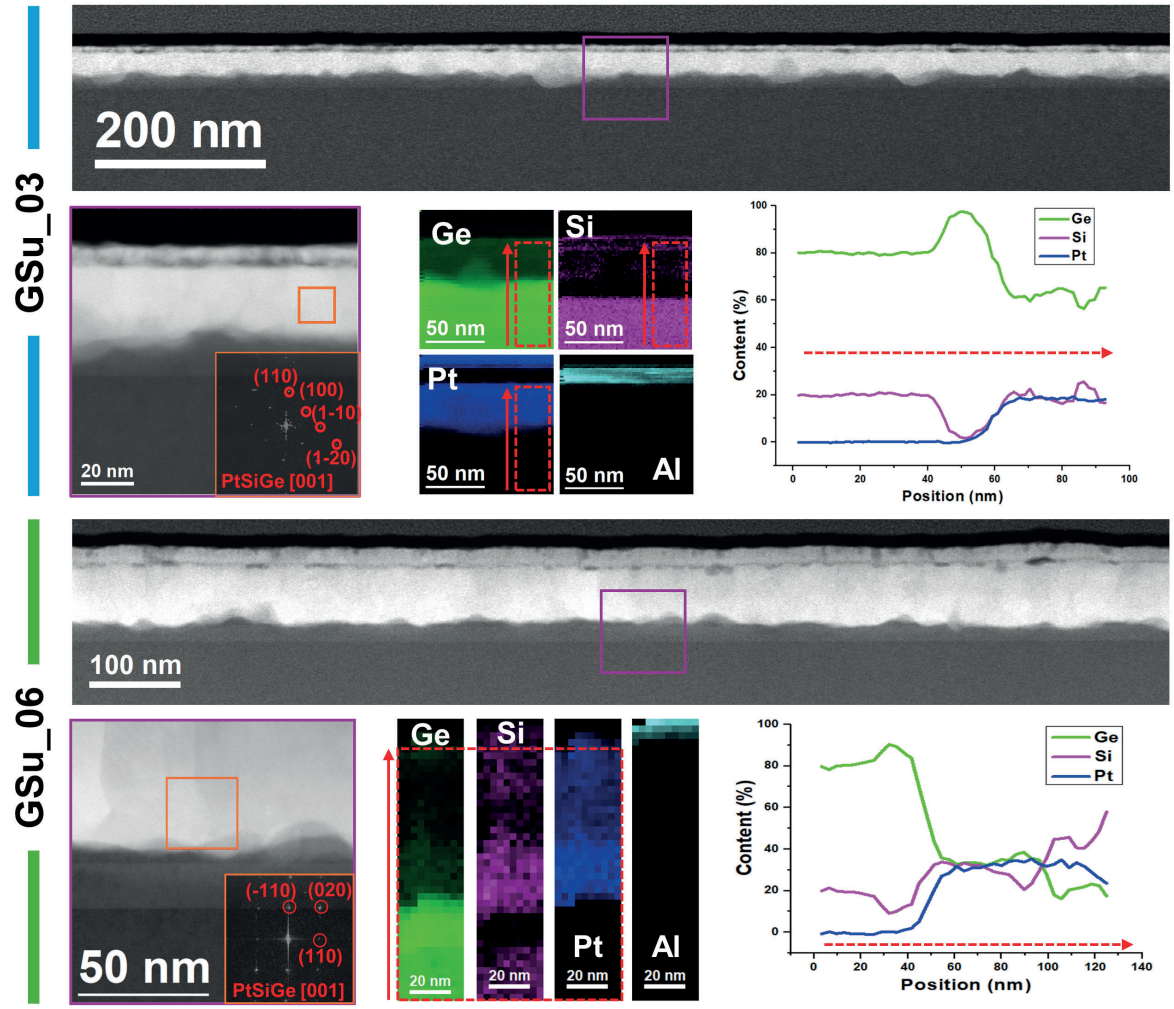


FIGURE 3.13: GSu_03 and GSu_06 Scanning Transmission Electron Microscopy (STEM). GSu_06 lays the Ge QW deeper than GSu_03 (i.e., thicker upper buffer layer for GSu_06). Qualitative and quantitative Electron Energy Loss Spectroscopy (EELS) showcase the Pt:Si:Ge stoichiometry of a polycrystalline layer validated by its Fast Fourier Transform.

Figure 3.13 showcases the first device design. It focuses on the direct contact between the Ge QW and the PtSiGe. The morphology to attain is the thorough and uninterrupted contact with the Ge all along the QW. Indeed, the most obvious difference between GSu_03 and GSu_06 is the total Ge QW-PtSiGe contact in the GSu_03 device. Thus, this condition is not met in GSu_06, as the upper buffer layer is occasionally exposed, leaving a $Si_{0.2}Ge_{0.8}$ gap in between.

The $PtSi_{1-x}Ge_x$ is structurally characterised through both the analysis of local FFTs and the Bragg filtering of isolated spots in global FFTs. Similarly to Nb in device

GSu_2, its polycrystalline nature forces a local approach in seeking symmetric frequency arrangements. Bragg filtering is useful in this case as it allows us to map the spatial distribution of crystallographic planes of interest, pointing them to the crystals where they come from and their zone axis. With this combination, we managed to identify the crystallographic phase of PtSiGe. Not surprisingly, it crystallises in an orthorhombic crystal system, specifically in space group Pbnm, number 62, as both binary compounds PtSi and PtGe do so. In fact, knowing the $PtSi_{1-x}Ge_x$, we can approximate the cell parameters from the direct application of Vegard's law from the binary compounds. [46]

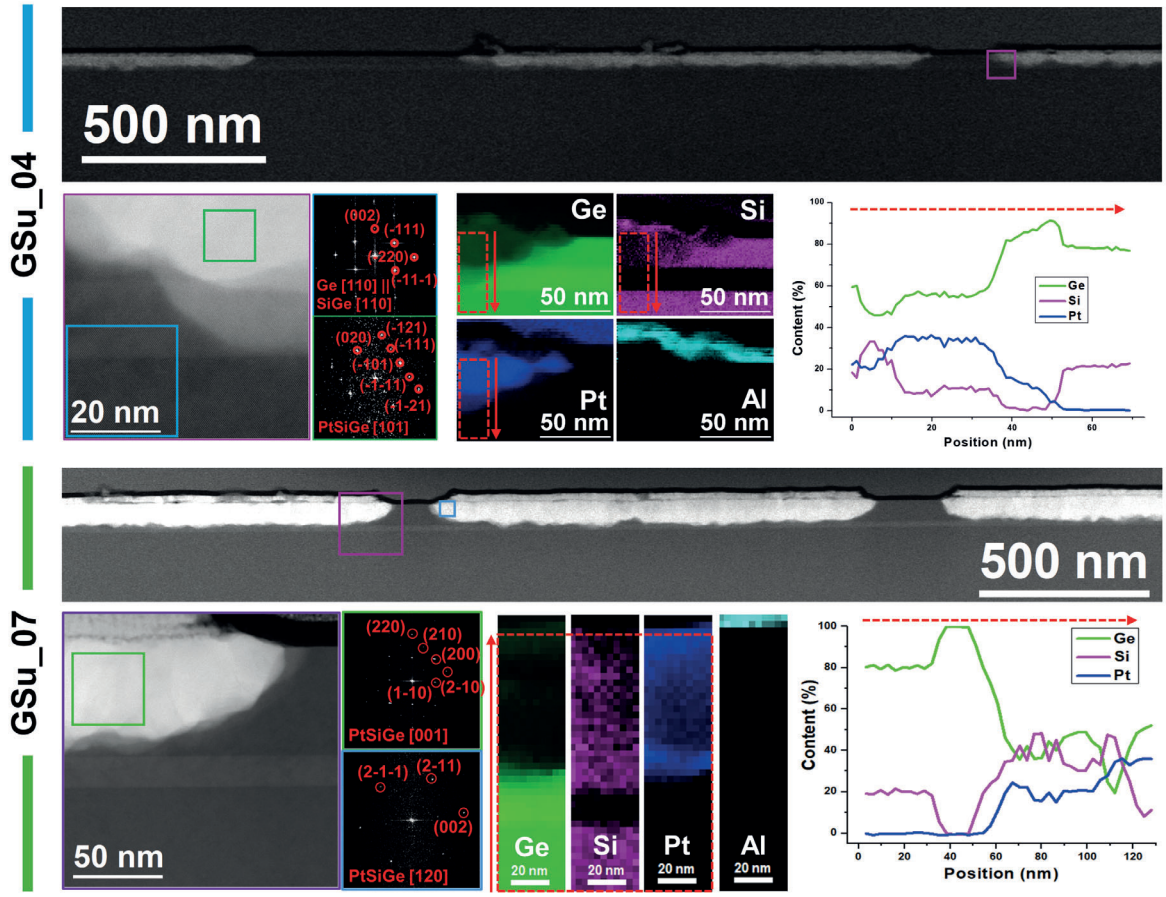


FIGURE 3.14: GSu_04 vs GSu_07 Scanning Transmission Electron Microscopy (STEM). GSu_07 lays the Ge QW deeper than GSu_03 (i.e., thicker upper buffer layer for GSu_04). Qualitative and quantitative Electron Energy Loss Spectroscopy (EELS) showcase the Pt:Si:Ge stoichiometry of a polycrystalline layer of the orthorhombic PtSiGe validated by its Fast Fourier Transform in multiple grains and zone axes. The superconductor-semiconductor-superconductor junction is added to the configuration of GSu_03 and GSu_06 (fig. 3.13).

Both qualitative and quantitative EELS are acquired first, to validate the presence of the Pt and its homogeneity as an alloy, and second, to define its stoichiometry. Due to its polycrystalline nature, slight variations can be observed in the contrast of the main three elemental maps. These local changes are more prominent close to the QW, concentrated at the very bottom of the contact. This would enable the

local divergent displacement of the Ge and Si from grain to grain. Therefore, the Ge-richer grains might originate from the alloying of the last few nanometers which come into contact with a higher default concentration of Ge from the QW itself. Importantly, the principal quantitative conclusion is reproducible for the three pairs of devices comparing the two QW depths. The superconductor of the QW starting at 20 nm from the top typically contains a comparable amount of Pt and Si (i.e., Pt > Si) which is considerably lower than the Ge. On the other hand, for the Ge QW at 50 nm, the three elements are more equilibrated in composition. Thus, the limited area analysed makes us think that the further the QW is, the less Ge will be contained in the grains. Although local variations are grain-dependent, the first batch will experience an overall higher Ge content because of the QW proximity. Of course, the average stoichiometry should respect the $Si_{0.2}Ge_{0.8}$ relation from the original buffer layer, but with this added particularity.

GSu_04 and GSu_07 add the S-Sm-S junction to the device by leaving an area without the metallic contact. The open channel is where the proximity and the multiple Andreev reflections should arise. Different widths of the channel are evaluated, too. Nonetheless, the most important feature to asses is the quality of the interfaces controlling the supercurrent exchange. These interfaces are highlighted in the HAADF STEM insets (i.e., purple windows). As in the GSu_03 vs GSu_06 comparison, the GSu_07 device is not fully contacting the PtSiGe with the Ge QW. Consequently, the transport channel would be malformed hindering the supercurrent formation. Contrarily, GSu_04 showcases a more promising interface to sustain the supercurrent. In fact, the interface inset mostly showcases a brighter [110] zone axis of diamond structure with the same cell parameters as Ge, suggesting PtSiGe grains outside the realms of the depth of focus. The granular nature of PtSiGe is again supported by the FFTs and the varying zone axes we see. Furthermore, the compositional conclusions are consistent with the previous device collation (fig. 3.13).

GSu_05 and GSu_08 add the metallic gates on top of the semiconductor to essentially the same configuration observed in GSu_04 and GSu_07. As a result, GSu_05 and GSu_08 approximate the configuration of a functional device. For instance, the first top gate aims to extend throughout the exposed semiconductor part of the junction. In the HAADF micrographs of figure 3.15, the top gate does so regularly for GSu_05, although it is less conformal for GSu_08. The complementary observations regarding phase identification, compositional analysis and most importantly, interface quality and PtSiGe-Ge distribution, align with the comments for GSu_04 and GSu_07.

The device configuration of both GSu_03 and GSu_06 impeded measuring any strain because the full Ge QW is covered by the alloy. However, it could be computed for the GSu_04-GSu_07 and GSu_05-GSu_08 device pairs. Not surprisingly, the results match the trend of the $Si_{0.2}Ge_{0.8}/Ge/Si_{0.2}Ge_{0.8}$ heterostructure initially found for the most similar device, the GSu_02. [45] The absolute strain values exhibit oscillations

due to the use of experimental stoichiometry directly from the quantitative EELS maps. The theoretical mismatch between the QW and bottom $Si_{0.2}Ge_{0.8}$ buffer is 0.81 %. Specifically, the ϵ_{xx} values fall within the range of -0.82 and -0.75 %. Considering the experimental uncertainty, these values essentially indicate that the mismatch is fully elastically adapted by the QW, resulting in a perfect epitaxy. Furthermore, the vertical (-220) planes exhibit a misfit dislocation-free adaptation, which leads to an expansion of the (002) planes. The (ϵ_{yy}) values range from 0.61 to 0.41 %. It is important to note that, as mentioned before, the lamellae thinning must be taken into account to reveal the underlying source of the calculated ϵ_{yy} values.

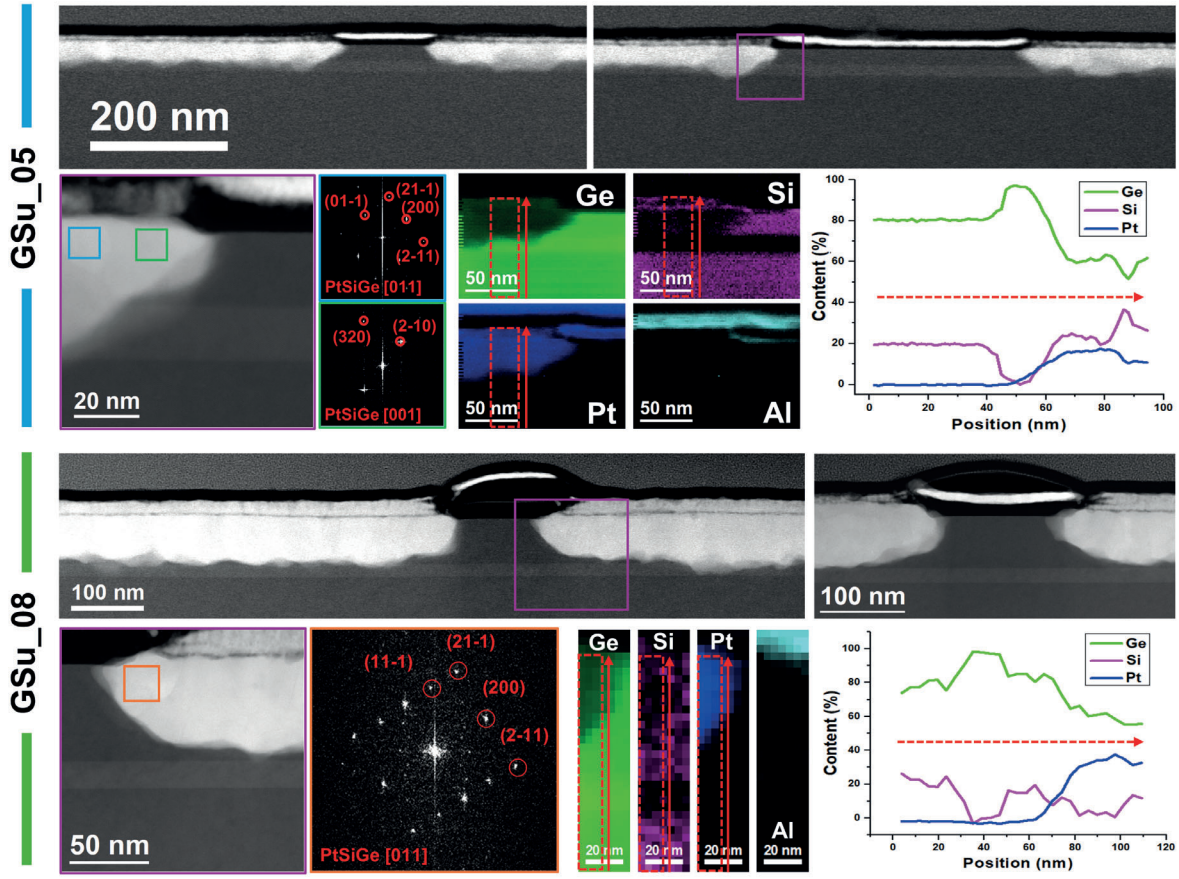


FIGURE 3.15: GSu_05 vs GSu_08 Scanning Transmission Electron Microscopy (STEM). GSu_08 lays the Ge QW deeper than GSu_04 (i.e., thicker upper buffer layer for GSu_04). Qualitative and quantitative Electron Energy Loss Spectroscopy (EELS) showcase the Pt:Si:Ge stoichiometry of the polycrystalline PtSiGe layer. GSu_05 and GSu_08 already display the main features the ultimate superconductor-semiconductor-superconductor junction must present.

Growing a thicker upper buffer is advantageous due to promising potential landscapes. [48] However, for the second batch of devices, the overall comment is that the total covering of the top buffer layer would require further filling to accomplish the S-Sm-S junction. Indeed, at this point, it is interesting to describe the growth process to pinpoint direct causes and their possible fixes. Remarkably, the fabrication process does not require etching or special vacuum conditions. In addition, the

known compatibility of Pt with existing Si CMOS technology makes the system an excellent candidate for large-scale integration of semiconductor-superconductor hybrids. The QW is grown by the same reduced-pressure chemical vapour deposition explained for SSQ devices, with a reverse gradient from a Si (001) wafer. The process, though, continues with the growth of the superconducting contact by the electron beam-induced evaporation of 15 nm of Pt after wet etching of the Si capping layer. The most important step is the alloy-forming step, which consists of rapid thermal annealing at 400 ºC for 15 min, in an argon atmosphere. The atomic layer deposition at 300 ºC of 10 nm of AlO_x ends the process. [47] Therefore, to achieve a deeper PtSiGe layer fully covering the deeper QW, the annealing should entail a thicker Pt initial layer, more annealing time or temperature, or a combination of the three. The temperature should never surpass 500 ºC as the Si and Ge from the layers would start to diffuse. As a result, the strategy would probably shift towards thicker Pt, further annealing time and lower temperature.

The research loop continued with the last steps towards building the final functional device. That is, the progressive inclusion of further top gates. The top gates consist of Pd and Ti alloys separated by AlO_x layers. The deposition conditions coincide with those of the Pt layer. [47] The overall growth process is schematised in figure 3.17. The device with a single Pd/Ti contact is GSu_09 (fig. 3.16), while the double one is GSu_10 (fig. 3.17).

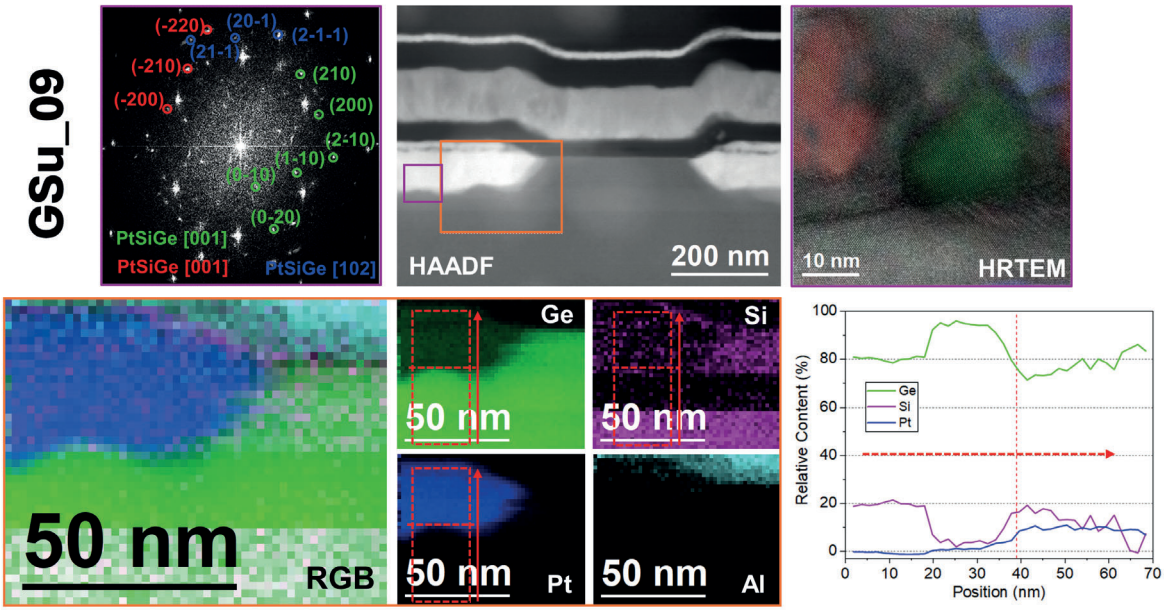


FIGURE 3.16: GSu_09 Scanning Transmission Electron Microscopy (STEM) characterisation. The central High-Angle Annular Dark-Field (HAADF) overview showcasing the metallic top gates is surrounded by a high-magnification High-Resolution TEM (HRTEM) image of the PtSiGe grains. The micrograph is Bragg-filtered following the zone axes identified in its Fast Fourier Transform (FFT). Qualitative and quantitative Electron Energy Loss Spectroscopy (EELS) showcase the Pt:Si:Ge stoichiometry of the polycrystalline PtSiGe layer next to the junction.

The winning strategy, at least at the current point of the investigation, is to keep the QWs on the shallow side. Thus, we stick with the appropriate covering of the shallow ones to improve the deep QWs in the future. Therefore, aligning with the first batch of GSu_03, GSu_04 and GSu_05, the Ge in the QW is fully covered by the PtSiGe. The HAADF STEM contrast of both figures 3.16 and 3.17 showcase the same grain-dependent compositional distribution. It is especially clear for device GSu_10 in figure 3.17, in the HAADF STEM overview, how the pixel intensity within the superconducting layer varies locally. For instance, the quantitative EELS profiles along the grains neighbouring the S-Sm-S interfaces exhibit equivalent compositional relations of Pt:Si:Ge. Importantly, these S-Sm interfaces are the sharpest of all the GSu devices analysed so far. They compare to the also outstanding GSu_02 device. [45] However, there is an unintended Si accumulation in the AlO_x domain, as well as a Ge accumulation on the top edge of the top buffer layer. This uncontrolled diffusion is likely coming from the growth of the top gate, as it is the only differentiating process compared to the two initial batches.

Interestingly, in figure 3.16 we present the Bragg filtering approach mentioned a few paragraphs above. The global FFT from a high-magnification HRTEM micrograph simultaneously contains up to three zone axes of the same PtSiGe orthorhombic phase. Consequently, the Bragg filtering procedure unlocks mapping its spatial distribution back to the original image. Thus, we identify the FFT frequencies per zone axis, and colour the grains from the independent representation of each zone axis.

Regardless of whether local or global FFTs are used, the initial step involves model matching and the identification of a crystallographic phase along with its spatial orientation. In the present case, we recognise orthorhombic PtSiGe, whose conventional unit cell is sketched in figure 3.17. Generally, by successfully identifying the spots corresponding to a crystal phase in a specific orientation, we unlock a Bragg filtering process capable of mapping the spatial distribution of these phases. Remarkably, this procedure provides an excellent complementary approach for intuitively interpreting electron micrographs and understanding their physical content. Interestingly, it is equally applicable in both HRTEM and STEM operating modes. As a consequence, this concept will also be exploited in the workflow deployed in Chapter 4.

The combination of devices GSu_09 and GSu_10 was crucial to push the state-of-the-art of S-Sm-S quantum devices. The devices exhibited large supercurrents and strong suppression of subgap states. Although the magnitude of the induced superconducting gap is not as large as in Al-based devices, the retrieved hard superconducting gap had never been reported before for Ge/SiGe. A hard superconducting gap relies on the low disorder leading to the absence of subgap states. This is obtained and reproduced in up to six similar devices with the same growth recipe and results. Despite improving the conductance, it did not reach the expected and ideal

one for this device type. [47] Most likely, the interface is not fully transparent, being detrimental to the general conductivity of the junction. Moreover, the joint leakage of Si in the AlO_x layer and the Ge accumulation might also play a role. Thus, the electrical isolation from AlO_x would be hindered by these contaminants. In addition, the heterogeneous distribution of grains within the superconducting layer might also hamper the conductance and reduce the ideal value. Therefore, the grain formation process would require further investigation to produce homogeneous grains and better interfaces, for example, by altering the annealing as described before.

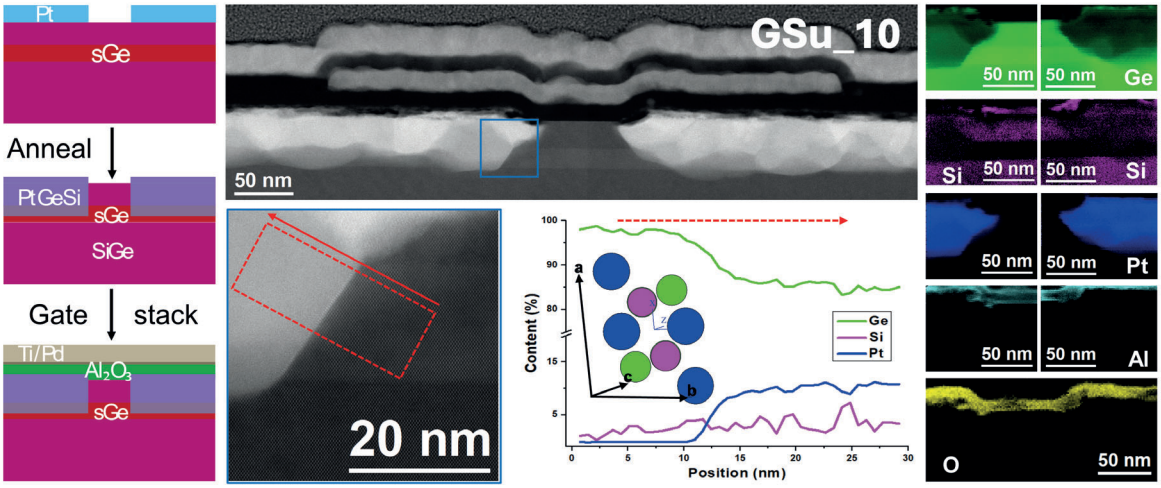


FIGURE 3.17: GSu_10 Scanning Transmission Electron Microscopy (STEM) characterisation. The overview highlights the complete configuration of the top gates. To the left, the schematics of the lithographic growth process illustrate the conversion of the metallic Pt layer into the superconducting PtSiGe. To the right, the Electron Energy Loss Spectroscopy (EELS) compositional maps complement the quantitative profile drawn along the sharp superconductor-semiconductor left interface of the device. The inset in the EELS profile corresponds to the conventional unit cell of the orthorhombic PtSiGe crystal phase (space group Pbnm, number 62). [47]

The reported hard gap opens a wide range of possibilities. Not only to pursue superconducting qubits based on Josephson transistors but to push for topological quantum computing with Majorana Zero Modes (MZM). Thus, to enable and open the possibility to merge the three technological platforms described before for the ultimate quantum processor. [7] Interestingly, the next subsection targets complementary devices not incorporating a QW. However, they are engineered either with Si, Ge or both, and present an alternative morphology or device dimensionality to stimulate the rising of profitable exotic quantum phenomena.

3.2.3 Additional device morphologies and materials setups

The analysed devices in the previous sections hopefully clarified the enormous potential of devices based on Si and Ge QWs. They create 2D electron or hole gases to confine quantum states and exploit outstanding transport properties. However, limiting the investigations to QWs would torpedo the overall progress in the field,

as certain phenomena can only arise in different device morphologies. For instance, nanowires are thought to be an ideal platform for allocating topologically-protected fault-tolerant quantum computing through MZM. [40] According to Kitaev's model, MZM theoretically arise at the edges (along the translational invariant direction) of proximitised semiconductor-superconductor interfaces. [49] These MZM would encode quantum states based on their occupation of no (logic 0) or one electron (logic 1). These quantum states are not local but dependent on the whole proximitised nanowire (NW). As a result, they are robust against local disorder or perturbations. This is why we refer to them as topologically protected, for which they are promising candidates for large-scale quantum processing without logical errors. [50]

The main problem is that these theoretical modes are shy experimentally, and their evidence in the literature is scarce. Finding and reproducibly obtaining them is the longing of many in the academic quantum computing community, but also within its industry. [51, 52] For instance, Microsoft is allocating significant resources to find and exploit them in a quantum processor that would provide a strategic advantage over other companies that address errors with error-correcting qubits. [41, 53] Ideally, This achievement would end up with quantum processors in which every qubit is logical and operates without errors. However, gathering evidence of their traces can be misleading as they can easily be confused with other excitations. For instance, zero bias peaks of electrical conductance are the clearest evidence of MZM. However, they can also occur for Andreev bound states or Kondo effect, for which distinguishing their origin is key to univocally point at MZM for quantum computing.

The first device we analyse in this section successfully answered this concern. In fact, the central development is an experimental workflow based on simultaneous tunnelling and Coulomb blockade spectroscopy. This method helps to eliminate uncertainties in Majorana detection by comparing the outputs of both spectroscopies in different conditions. [54] We used a semiconducting InAs NW with a superconducting Al shell to validate this methodology. The first spotlight is the perfect hexagonal symmetry showcased in the HAADF STEM overview of figure 3.18. Not surprisingly, the phase identification revealed the InAs wurtzite polytype (space group $P6_3mc$, 186) and the characteristic FFT of the [0001] zone axis. The crystal quality is outstanding, shaping a perfectly regular hexagonal NW that exposes the main facets corresponding to the main lower Laue zone reflections (i.e., 1-100). The Al phase coincides with the customary one also found for devices GSu_01 and GSu_02. In addition, the qualitative EELS maps feature the homogeneous distribution of every involved element. Of course, the InAs NW is a perfect monocrystal. On the other hand, the Al is polycrystalline, although large single grains covering the extension of a single facet are present. Interestingly, the zoom-in stresses the right-most corner of the hexagonal NW, for which we provide an atomic model that schematically informs on the atomic disposition in that region. [55]

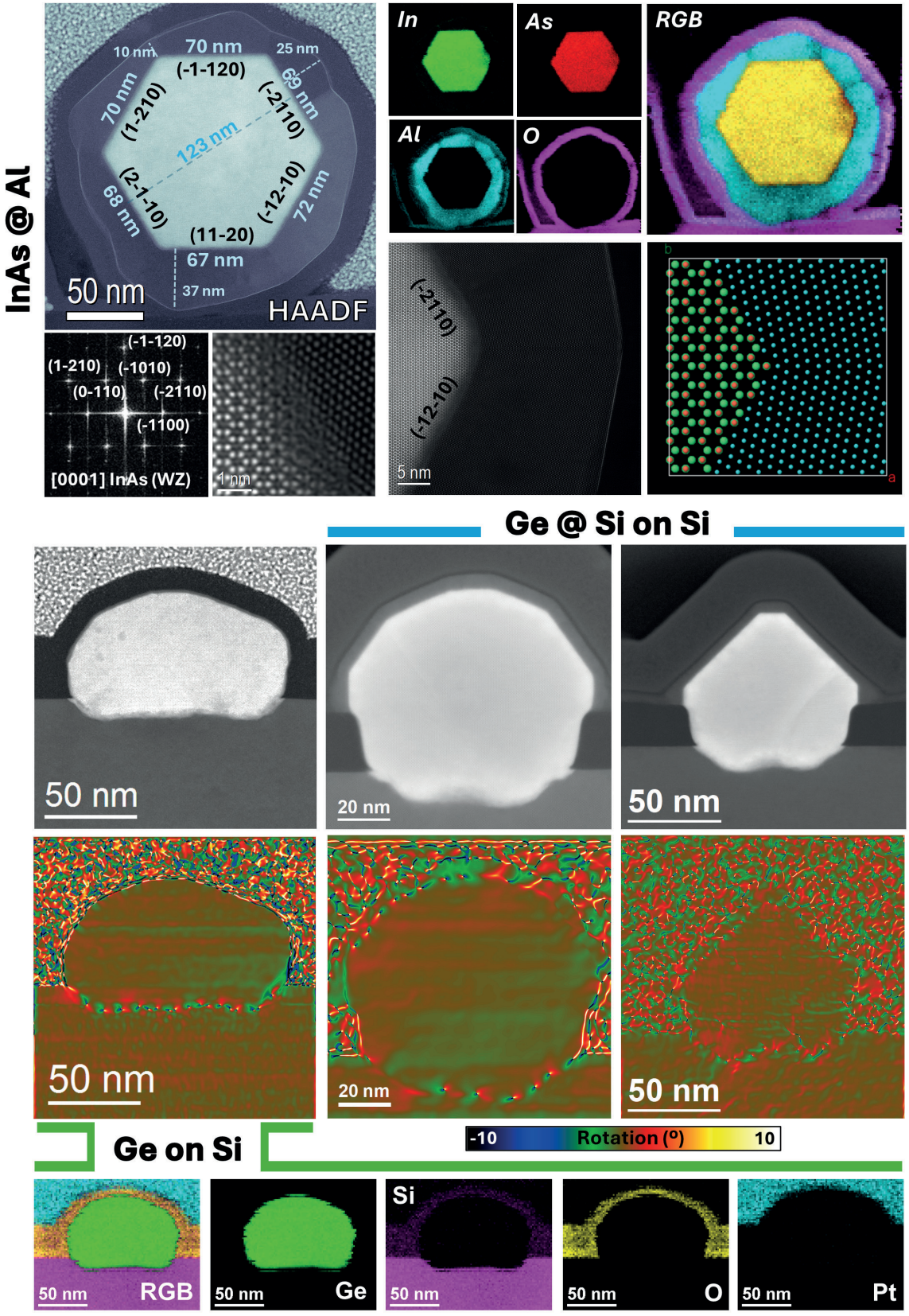


FIGURE 3.18: On the top half: comprehensive Scanning Transmission Electron Microscopy (STEM) characterisation of an InAs nanowire with a superconducting Al shell. The detail on the right corner of the wire is manually modelled with a representative atomic model of the InAs @ Al contact. [54] On the bottom half: Overview micrographs of Ge (@Si) on Si nanowires and their rotation maps computed by Geometrical Phase Analysis (GPA) to highlight the array of dislocations plastically relaxing the mismatch between the Si substrate and the Ge wire.

Remarkably, the provided atomic model is built by manually defining the crystallographic facets as collected from the micrograph. Then, the areas within the virtually cut supercell are filled with the crystallographic phase identified for InAs and Al. The final output is a list of atoms and their 3D coordinates. Thus, their arrangement structurally represents the analysed device. Remarkably, it allows a comprehensive understanding of the device's properties, both by checking further the structural and compositional details, and even beyond, by simulating functional properties and their performance in their operating context. The vision of ultimately engineering a 3D model out of experimental data and its analysis is the ultimate pillar of the present thesis. Thus, the central goal outlined in section 1.2 intends to push the workflow to also retrieve these 3D models automatically. Indeed, the automation would unlock more realistic and reliable 3D models than any atomic model that has been manually created in the past.

Secondly, we present a batch of three nanowires representative of a larger batch but narrowed down for concision (fig. 3.18, bottom half). Grown by metalorganic vapour phase epitaxy, the NWs fully rely on the Si and Ge platforms for their main focus on spin qubits and topological quantum computing. [56] As a result, further research stages should aim for the creation of a proximitised NW contacting a semiconductor, the Si or Ge, with a superconductor. Nevertheless, the stage of this investigation is the earliest of all the device types analysed. Therefore, they are solely focused on establishing the ideal growth settings and pinpointing the challenges of the targeted morphology. Thus, the present stage only involves the development of the semiconductor, for which no superconducting shell is expected.

The Ge NWs grow on a Si substrate. The high mismatch of 4.2 % makes us expect plastic relaxation mechanisms. Indeed, the observations coincide and an array of misfit dislocations is featured in the interface between the substrate and the NW. Mainly two device types are studied. First, the basic configuration of the Ge NW on top of the Si substrate (fig. 3.18, bottom half, left column and EELS maps). On the other hand, an additional layer of complexity stacks a Si shell on top of the same Ge wire (fig. 3.18, bottom half, mid and right columns). In the latter case, the stack of dislocations is also visible between the Ge QW and the Si shell. The best way to highlight the dislocation array is with the GPA rotation maps. In figure 3.18, the HAADF STEM micrographs are paired with their GPA rotation maps, where the features that abruptly grade from negative to positive rotations correspond to individual misfit dislocations. Interestingly, the early stages of development are noticeable as we observe the first structural defects of all the structures that have been scrutinized thus far. The source of the observed stacking faults and twin boundaries is the high mismatch of a non-optimised growth mechanism of a highly mismatched heterostructure. Indeed, optimising the semiconductor means a defect-free nanostructure, as these structural defects are detrimental to the transport properties. In fact, in this case, a flawless NW must be the already demanding starting point of

seeking the elusive MZM. In that sense, transport measurements are still lacking to experimentally prove the beneficial theoretical properties they could present. [56, 57]

3.3 Conclusions

Chapter 3 has offered an overview of the current state of the art of Si but mostly Ge devices for quantum computing applications. The central scope is their structural characterisation and their improvement across the research loop. Thus, certain devices like those presented within the GSQ batch evolved from advanced but still defective devices until a quantum well with outstanding properties to realise the first validation of a singlet-triplet spin qubit on Ge. [26] On the other hand, the SSQ devices directly exhibited exploitable properties to act as spin qubits, but in the framework of a much more mature technology (i.e., Si platform). As a result, although SSQ pushed the state of the art, it was by enclosing a particular riddle for one step further in the route towards the scalable and fault-tolerant quantum processor. [37, 38]

The knowledge of homogenous QWs without defects and their realization through graded growth approaches that mitigate defects with an extensive buffer have been further tested on more complex configurations. The idea was to add complementary conducting phases such as superconducting contacts to exploit exotic interactions and excitations. The characterization revealed additional challenges, such as variate interfaces between different crystal structures, polycrystalline regions, or intricate and inconspicuous compositional gradients. From these, GSu devices showed multiple materials to approach the superconducting phase, from traditional Al, to Al-based heterostructures, to first-time reported superconducting phases like PtSiGe. [45, 47] The common feature towards the excellent transport properties is the flawless quantum well with total mismatch compensation with elastic strain, a sharp interface between the semiconductor and the superconductor, and high-quality isolating protective top layers. Remarkably, these devices and their characterisation have overall supposed huge improvements to the state of the art. Nonetheless, there is still room for improvement, such as the need for better control over the annealing process of the PtSiGe layer to increase the uniformity of its grain size and composition.

Furthermore, the presented research has showcased diverse device morphologies and platforms that target major challenges differently. MZM with proximitised NWs, or spin qubits with NWs instead of QWs, are paradigms of major advances in the field. [54] Indeed, the objectives set for this chapter have been effectively addressed, showcasing a wide range of cases that are of interest from both a (S)TEM perspective and from a quantum properties standpoint. These cases have been useful for identifying the characterisation needs of modern devices, which

have been faced during the stepped examination of manifold characteristics and potential features to exploit in every device. In fact, the acquired wisdom constitutes a pillar for the knowledge generated in the next two chapters. Their main research unknowns precisely arise and are identified during the analysis of the devices in the current chapter, as indicated throughout the text.

3.4 Contributions

The research work displayed in the present chapter is collaborative and combines crucial inputs from many researchers. Thus, from the work presented, the author of the thesis dissertation, Marc Botifoll Moral, contributed with the following:

- Autonomously performed the complete electron microscopy imaging and spectroscopy of the data, or supervised the acquisition from another microscopist from Group of Advanced Electron Nanoscopy, or was present when the data acquisition took place by a member of Group of Advanced Electron Nanoscopy.
- Autonomously performed the thorough data analysis of the electron microscopy data, or supervised its processing from another microscopist from Group of Advanced Electron Nanoscopy.
- Actively participated in the scientific discussion on the correlation between the quantum and transport properties observed by the collaborators and the structural characteristics retrieved during the electron microscopy characterisation workflow.

The remaining work was distributed as follows:

- The team in the Institute of Science and Technology in Austria (IST Austria), led by Prof. Giorgos Katsaros, together with their collaborators, conceived the idea, theoretical basis, the growth and its optimisation, and the transport and quantum measurements of the devices in operation, behind the GSQ, GSu_01 and GSu_02, and InAs @ Al NW. [26, 45, 54]
- In the context of the GSQ devices, Dr. Daniel Chrastina designed and computed the FEM simulations from device GSQ_4, from which the whole collaborating group behind the GSQ devices discussed and compared the theoretical results with the outputs from GPA and XRD.
- The team in the QuTech, Technical University of Delft, led by Prof. Giordano Scapucci, together with their collaborators, conceived the idea, theoretical basis, the growth and its optimisation, and the transport and quantum measurements of the devices in operation, behind the SSQ, and the GSu_03-GSu_10 batches. [37, 38, 47]

- The team in the École Polytechnique Fédérale de Lausanne, led by Prof. Anna Fontcuberta i Morral, together with their collaborators, conceived the idea, theoretical basis, the growth and its optimisation, of the presented Ge NWs.

References

[1] C. H. Bennett and G. Brassard. "Quantum cryptography: Public key distribution and coin tossing". In: *Theor. Comput. Sci.* 560 (2014), pp. 7–11. DOI: 10.1016/j.tcs.2014.05.025.

[2] L. K. Grover. "A fast quantum mechanical algorithm for database search". In: *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96*. Philadelphia, Pennsylvania, United States, 1996, pp. 212–219. DOI: 10.1145/237814.237866.

[3] N. Wiebe, D. Braun, and S. Lloyd. "Quantum Algorithm for Data Fitting". In: *Phys. Rev. Lett.* 109.5 (2012), p. 050505. DOI: 10.1103/PhysRevLett.109.050505.

[4] N. Moll et al. "Quantum optimization using variational algorithms on near-term quantum devices". In: *Quantum Sci. Technol.* 3.3 (2018), p. 030503. DOI: 10.1088/2058-9565/aab822.

[5] D. Venturelli and A. Kondratyev. "Reverse quantum annealing approach to portfolio optimization problems". In: *Quantum Mach. Intell.* 1 (2019), pp. 17–30.

[6] Frank Arute et al. "Quantum supremacy using a programmable superconducting processor". In: *Nature* 574.7779 (Oct. 2019), pp. 505–510. ISSN: 1476-4687. DOI: 10.1038/s41586-019-1666-5. URL: https://doi.org/10.1038/s41586-019-1666-5.

[7] Giordano Scappucci et al. "The germanium quantum information route". In: *Nature Reviews Materials* 6.10 (Oct. 2021), pp. 926–943. ISSN: 2058-8437. DOI: 10.1038/s41578-020-00262-z. URL: https://doi.org/10.1038/s41578-020-00262-z.

[8] J. R. Petta et al. "Coherent Manipulation of Coupled Electron Spins in Semiconductor Quantum Dots". In: *Science* 309.5744 (2005), pp. 2180–2184. DOI: 10.1126/science.1116955.

[9] C. D. Bruzewicz et al. "Trapped-ion quantum computing: Progress and challenges". In: *Appl. Phys. Rev.* 6.2 (2019), p. 021314. DOI: 10.1063/1.5088164.

[10] H.-L. Huang et al. "Superconducting quantum computing: a review". In: *Sci. China Inf. Sci.* 63.8 (2020), p. 180501. DOI: 10.1007/s11432-020-2881-9.

[11] F. A. Zwanenburg et al. "Silicon quantum electronics". In: *Rev. Mod. Phys.* 85.3 (2013), pp. 961–1019. DOI: 10.1103/RevModPhys.85.961.

[12] R. Maurand et al. "A CMOS silicon spin qubit". In: *Nat. Commun.* 7.1 (2016), p. 13575. DOI: 10.1038/ncomms13575.

[13] L. R. Schreiber and H. Bluhm. "Silicon comes back". In: *Nat. Nanotechnol.* 9.12 (2014), pp. 966–968. DOI: 10.1038/nnano.2014.249.

[14] M. Veldhorst et al. "An addressable quantum dot qubit with fault-tolerant control-fidelity". In: *Nat. Nanotechnol.* 9.12 (2014), pp. 981–985. DOI: 10.1038/nnano.2014.216.

[15] E. Kawakami et al. "Electrical control of a long-lived spin qubit in a Si/SiGe quantum dot". In: *Nat. Nanotechnol.* 9.9 (2014), pp. 666–670. DOI: 10.1038/nnano.2014.153.

[16] J. J. Pla et al. "A single-atom electron spin qubit in silicon". In: *Nature* 489.7417 (2012), pp. 541–545. DOI: 10.1038/nature11449.

[17] R. Pillarisetty. "Academic and industry research progress in germanium nanodevices". In: *Nature* 479.7373 (2011), pp. 324–328. DOI: 10.1038/nature10678.

[18] H. Watzinger et al. "A germanium hole spin qubit". In: *Nat. Commun.* 9.1 (2018), p. 3902. DOI: 10.1038/s41467-018-06418-4.

[19] N. W. Hendrickx et al. "Fast two-qubit logic with holes in germanium". In: *Nature* 577.7791 (2020), pp. 487–491. DOI: 10.1038/s41586-019-1919-3.

[20] D. Paul. "The progress towards terahertz quantum cascade lasers on silicon substrates". In: *Laser Photonics Rev.* 4 (2010), pp. 610–632.

[21] J. W. Matthews and A. E. Blakeslee. "Defects in epitaxial multilayers: III. Preparation of almost perfect multilayers". In: *J. Cryst. Growth* 32 (1976), pp. 265–273.

[22] Siavash Nikravesh, Donghyeon Ryu, and Yu-Lin Shen. "Instabilities of Thin Films on a Compliant Substrate: Direct Numerical Simulations from Surface Wrinkling to Global Buckling". In: *Scientific Reports* 10.1 (Mar. 2020), p. 5728. ISSN: 2045-2322. DOI: 10.1038/s41598-020-62600-z. URL: https://doi.org/10.1038/s41598-020-62600-z.

[23] Rahul Panat, K. Jimmy Hsia, and David G. Cahill. "Evolution of surface waviness in thin films via volume and surface diffusion". In: *Journal of Applied Physics* 97.1 (Dec. 2004), p. 013521. ISSN: 0021-8979. DOI: 10.1063/1.1827920. eprint: https://pubs.aip.org/aip/jap/article-pdf/doi/10.1063/1.1827920/14679343/013521_1_online.pdf. URL: https://doi.org/10.1063/1.1827920.

[24] J. P. Dismukes, L. Ekstrom, and R. J. Paff. "Lattice Parameter and Density in Germanium-Silicon Alloys1". In: *The Journal of Physical Chemistry* 68.10 (1964), pp. 3021–3027. DOI: 10.1021/j100792a049. eprint: https://doi.org/10.1021/j100792a049. URL: https://doi.org/10.1021/j100792a049.

[25] Cengiz S. Ozkan, William D. Nix, and Huajian Gao. "Anisotropic Behaviour of Surface Roughening in Lattice Mismatched Heteroepitaxial Thin Films". In: *MRS Online Proceedings Library* 436.1 (Dec. 1996), pp. 487–492. ISSN: 1946-4274. DOI: 10.1557/PROC-436-487. URL: https://doi.org/10.1557/PROC-436-487.

[26] Daniel Jirovec et al. "A singlet-triplet hole spin qubit in planar Ge". In: *Nature Materials* 20 (2021), pp. 1106–1112. DOI: /10.1038/s41563-021-01022-2. eprint: https://doi.org/10.1038/s41563-021-01022-2. URL: https://doi.org/10.1038/s41563-021-01022-2.

[27] M.J. Hÿtch, E. Snoeck, and R. Kilaas. "Quantitative measurement of displacement and strain fields from HREM micrographs". In: *Ultramicroscopy* 74.3 (1998), pp. 131–146. ISSN: 0304-3991. DOI: https://doi.org/10.1016/S0304-3991(98)00035-7. URL: https://www.sciencedirect.com/science/article/pii/S0304399198000357.

[28] M. J. Hÿtch, J.-L. Putaux, and J.-M. Pénisson. "Measurement of the displacement field of dislocations to 0.03Å by electron microscopy". In: *Nature* 423 (2003), pp. 270–273.

[29] H.G. Weller et al. "A Tensorial Approach to Computational Continuum Mechanics Using Object Orientated Techniques". In: *Computers in Physics* 12 (Nov. 1998), pp. 620–631. DOI: 10.1063/1.168744.

[30] R. P. Graham et al. "THE ISOTOPIC CONSTITUTION OF GERMANIUM". In: *Canadian Journal of Chemistry* 29.1 (1951), pp. 89–102. DOI: 10.1139/v51-011. eprint: https://doi.org/10.1139/v51-011. URL: https://doi.org/10.1139/v51-011.

[31] Oussama Moutanabbir et al. "Nuclear Spin-Depleted, Isotopically Enriched 70Ge/28Si70Ge Quantum Wells". In: *Advanced Materials* 36.8 (Dec. 2023). ISSN: 1521-4095. DOI: 10.1002/adma.202305703. URL: http://dx.doi.org/10.1002/adma.202305703.

[32] V. A. Shah et al. "Reverse graded relaxed buffers for high Ge content SiGe virtual substrates". In: *Applied Physics Letters* 93.19 (Nov. 2008), p. 192103. ISSN: 0003-6951. DOI: 10.1063/1.3023068. eprint: https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/1.3023068/13163580/192103_1_online.pdf. URL: https://doi.org/10.1063/1.3023068.

[33] J. J. Pla et al. "High-fidelity readout and control of a nuclear spin qubit in silicon". In: *Nature* 496.7445 (2013), pp. 334–338. DOI: 10.1038/nature12011.

[34] N. Samkharadze et al. "Strong spin-photon coupling in silicon". In: *Science* 359.6380 (2018), pp. 1123–1127. DOI: 10.1126/science.aar4054. eprint: https://www.science.org/doi/pdf/10.1126/science.aar4054. URL: https://www.science.org/doi/abs/10.1126/science.aar4054.

[35] Stephan G. J. Philips et al. "Universal control of a six-qubit quantum processor in silicon". In: *Nature* 609.7929 (Sept. 2022), pp. 919–924. ISSN: 1476-4687. DOI: 10.1038/s41586-022-05117-x. URL: https://doi.org/10.1038/s41586-022-05117-x.

[36] Guoji Zheng et al. "Rapid gate-based spin read-out in silicon using an on-chip resonator". In: *Nature Nanotechnology* 14.8 (Aug. 2019), pp. 742–746. ISSN: 1748-3395. DOI: 10.1038/s41565-019-0488-9. URL: https://doi.org/10.1038/s41565-019-0488-9.

[37] Brian Paquelet Wuetz et al. "Reducing charge noise in quantum dots by using thin silicon quantum wells". In: *Nature Communications* 14.1 (Mar. 2023), p. 1385. ISSN: 2041-1723. DOI: 10.1038/s41467-023-36951-w. URL: https://doi.org/10.1038/s41467-023-36951-w.

[38] Davide Degli Esposti et al. "Low disorder and high valley splitting in silicon". In: *npj Quantum Information* 10.1 (Mar. 2024), p. 32. ISSN: 2056-6387. DOI: 10.1038/s41534-024-00826-9. URL: https://doi.org/10.1038/s41534-024-00826-9.

[39] Morten Kjaergaard et al. "Superconducting Qubits: Current State of Play". In: *Annual Review of Condensed Matter Physics* 11.Volume 11, 2020 (2020), pp. 369–395. ISSN: 1947-5462. DOI: https://doi.org/10.1146/annurev-conmatphys-031119-050605. URL: https://www.annualreviews.org/content/journals/10.1146/annurev-conmatphys-031119-050605.

[40] R. M. Lutchyn et al. "Majorana zero modes in superconductor–semiconductor heterostructures". In: *Nature Reviews Materials* 3.5 (May 2018), pp. 52–68. ISSN: 2058-8437. DOI: 10.1038/s41578-018-0003-1. URL: https://doi.org/10.1038/s41578-018-0003-1.

[41] Chetan Nayak et al. "Non-Abelian anyons and topological quantum computation". In: *Rev. Mod. Phys.* 80 (3 Sept. 2008), pp. 1083–1159. DOI: 10.1103/RevModPhys.80.1083. URL: https://link.aps.org/doi/10.1103/RevModPhys.80.1083.

[42] J. Q. You, J. S. Tsai, and Franco Nori. "Scalable Quantum Computing with Josephson Charge Qubits". In: *Phys. Rev. Lett.* 89 (19 Oct. 2002), p. 197902. DOI: 10.1103/PhysRevLett.89.197902. URL: https://link.aps.org/doi/10.1103/PhysRevLett.89.197902.

[43] Lukas Grünhaupt et al. "Granular aluminium as a superconducting material for high-impedance quantum circuits". In: *Nature Materials* 18.8 (Aug. 2019), pp. 816–819. ISSN: 1476-4660. DOI: 10.1038/s41563-019-0350-3. URL: https://doi.org/10.1038/s41563-019-0350-3.

[44] Amir Sammak et al. "Shallow and Undoped Germanium Quantum Wells: A Playground for Spin and Hybrid Quantum Technology". In: *Advanced Functional Materials* 29.14 (2019), p. 1807613. DOI: https://doi.org/10.1002/adfm.201807613. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adfm.201807613. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/adfm.201807613.

[45] Kushagra Aggarwal et al. "Enhancement of proximity-induced superconductivity in a planar Ge hole gas". In: *Phys. Rev. Res.* 3 (2 Apr. 2021), p. L022005. DOI: 10.1103/PhysRevResearch.3.L022005. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.3.L022005.

[46] Ch.J Raub. "Superconductivity of the platinum metals and their alloys". In: *Materials Design* 5.3 (1984), pp. 129–136. ISSN: 0261-3069. DOI: https://doi.org/10.1016/0261-3069(84)90081-5. URL: https://www.sciencedirect.com/science/article/pii/0261306984900815.

[47] Alberto Tosato et al. "Hard superconducting gap in germanium". In: *Communications Materials* 4 (2023), p. 23. DOI: /10.1038/s43246-023-00351-w. eprint: https://doi.org/10.1038/s43246-023-00351-w. URL: https://doi.org/10.1038/s43246-023-00351-w.

[48] D. Laroche et al. "Scattering mechanisms in shallow undoped Si/SiGe quantum wells". In: *AIP Advances* 5.10 (Oct. 2015), p. 107106. ISSN: 2158-3226. DOI: 10.1063/1.4933026. eprint: https://pubs.aip.org/aip/adv/article-pdf/doi/10.1063/1.4933026/12984215/107106_1_online.pdf. URL: https://doi.org/10.1063/1.4933026.

[49] A Yu Kitaev. "Unpaired Majorana fermions in quantum wires". In: *Physics-Uspekhi* 44.10S (Oct. 2001), p. 131. DOI: 10.1070/1063-7869/44/10S/S29. URL: https://dx.doi.org/10.1070/1063-7869/44/10S/S29.

[50] S. M. Albrecht et al. "Exponential protection of zero modes in Majorana islands". In: *Nature* 531.7593 (Mar. 2016), pp. 206–209. ISSN: 1476-4687. DOI: 10.1038/nature17162. URL: https://doi.org/10.1038/nature17162.

[51] David Aasen et al. "Milestones Toward Majorana-Based Quantum Computing". In: *Phys. Rev. X* 6 (3 Aug. 2016), p. 031016. DOI: 10.1103/PhysRevX.6.031016. URL: https://link.aps.org/doi/10.1103/PhysRevX.6.031016.

[52] M. T. Deng et al. "Majorana bound state in a coupled quantum-dot hybrid-nanowire system". In: *Science* 354.6319 (2016), pp. 1557–1562. DOI: 10.1126/science.aaf3961. eprint: https://www.science.org/doi/pdf/10.1126/science.aaf3961. URL: https://www.science.org/doi/abs/10.1126/science.aaf3961.

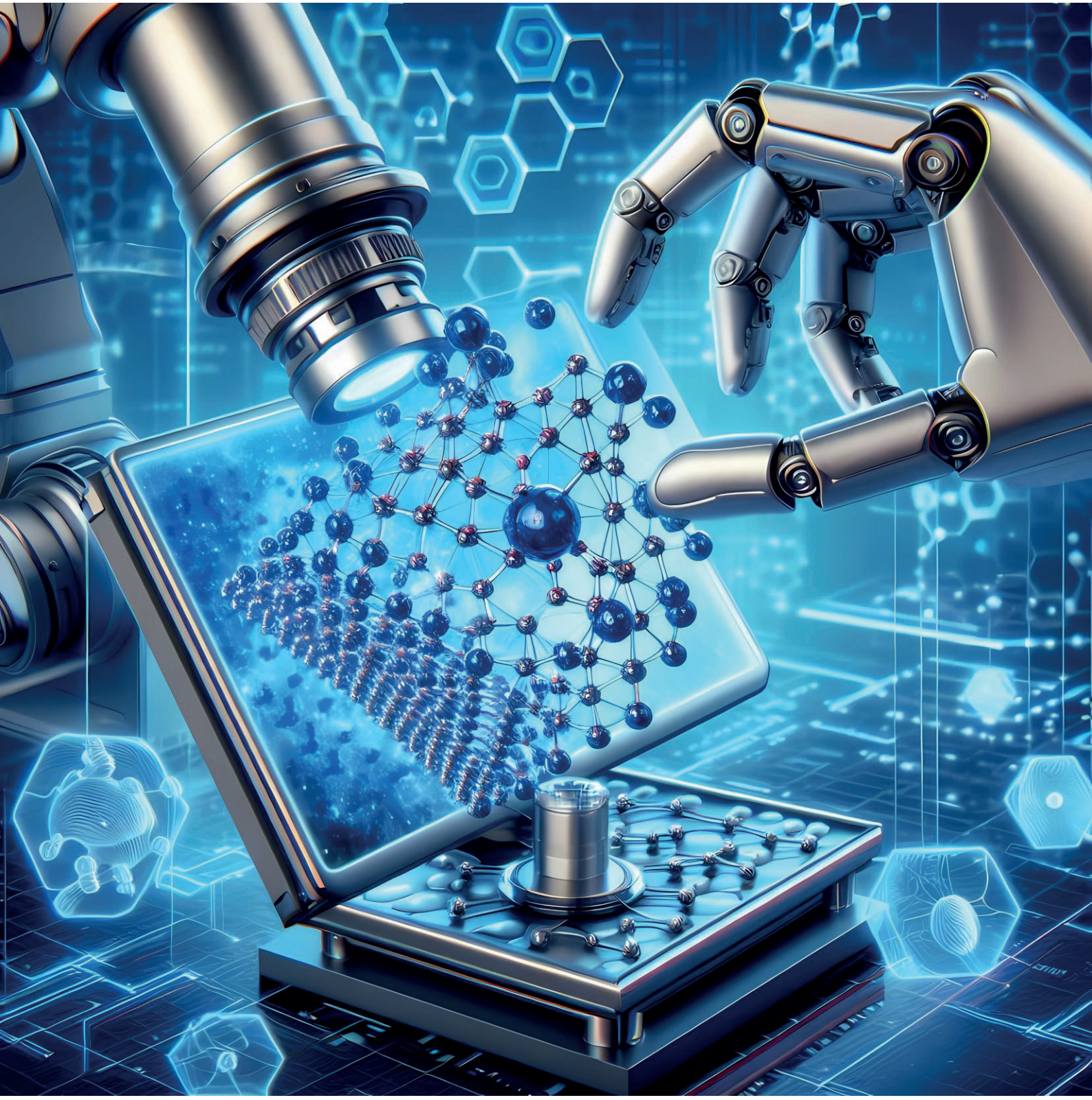
[53] Sankar Das Sarma, Michael Freedman, and Chetan Nayak. "Topological quantum computation". In: *Physics Today* 59.7 (July 2006), pp. 32–38. ISSN: 0031-9228. DOI: 10.1063/1.2337825. eprint: https://pubs.aip.org/physicstoday/article-pdf/59/7/32/16660151/32_1_online.pdf. URL: https://doi.org/10.1063/1.2337825.

[54] Marco Valentini et al. "Majorana-like Coulomb spectroscopy in the absence of zero-bias peaks". In: *Nature* 612 (2022), pp. 442–447. DOI: 10.1038/s41586-022-05382-w. eprint: https://doi.org/10.1038/s41586-022-05382-w. URL: https://doi.org/10.1038/s41586-022-05382-w.

[55] S Bernal et al. "The interpretation of HREM images of supported metal catalysts using image simulation: profile view images". In: *Ultramicroscopy* 72.3 (1998), pp. 135–164. ISSN: 0304-3991. DOI: https://doi.org/10.1016/S0304-3991(98)00009-6. URL: https://www.sciencedirect.com/science/article/pii/S0304399198000096.

[56] Santhanu Panikar Ramanandan et al. "Control of Ge island coalescence for the formation of nanowires on silicon". In: *Nanoscale Horiz.* (2024), pp. -. DOI: 10.1039/D3NH00573A. URL: http://dx.doi.org/10.1039/D3NH00573A.

[57] Martin Friedl et al. "Template-Assisted Scalable Nanowire Networks". In: *Nano Letters* 18.4 (2018). PMID: 29579392, pp. 2666–2671. DOI: 10.1021/acs.nanolett.8b00554. eprint: https://doi.org/10.1021/acs.nanolett.8b00554. URL: https://doi.org/10.1021/acs.nanolett.8b00554.



CHAPTER 4: END-TO-END ARTIFICIAL INTELLIGENCE WORKFLOW FOR THE AUTONOMOUS DISCOVERY OF PHYSICAL KNOWLEDGE FROM (S)TEM DATA

Chapter 4

End-to-end Artificial Intelligence Workflow for the Autonomous Discovery of Physical Knowledge from (S)TEM Data

The application of artificial intelligence has significantly enhanced the analysis of (Scanning) Transmission Electron Microscopy related data in materials science. This technological advancement is revolutionising the field by boosting throughput through automation and unlocking higher precision and fidelity levels. Nevertheless, the state-of-the-art methods exploiting it and their proofs are case-specific and lack generalisation capabilities. The present chapter, which constitutes the central research contribution of this thesis, reverts this blockade with a groundbreaking end-to-end and general data analysis workflow. The presentation of the workflow within the chapter is divided into two distinctive parts based on their particular although closely linked goals. The first part showcases the automation of the thorough structural characterisation applicable to a broad range of device and sample configurations built upon state-of-the-art material platforms in an intercorrelated way. The second part exploits this extracted information to create representative models of the device with unmatched fidelity. Moreover, it includes intuitive proofs of the knowledge that can be retrieved, with unprecedented precision, from experimental devices.

4.1 Motivation

The exploration and development of novel materials, as well as the optimization of existing ones for various applications, constitute a multifaceted process involving necessity identification, literature review, material proposal, device engineering,

characterization, and application testing. This iterative cycle, driven by multidisciplinarity and collaborative efforts, forms the backbone of scientific progress. However, it is often hindered by its time-consuming and expensive nature, particularly when atomic scale precision is required for understanding the functionality of materials and heterostructured devices. [1, 2, 3, 4]

As broadly discussed in the introduction (Chapter 1) and in Chapter 4, in an era marked by profound digital transformations, semiconductor heterostructures within a chip have emerged as crucial and widespread assets, driving major industrial value chains. They support advancements in both novel sectors (such as automated vehicles, cloud computing, Internet of Things, space exploration, supercomputing, and quantum technologies) and traditional ones (including computing and communications, industrial automation, entertainment, and healthcare). Miniaturization, now reaching the nanoscale and approaching the atomic limit, stands out as a primary driver of progress. This trend enhances device capabilities, lowers costs, and reduces energy consumption.

The shift towards new concepts and extreme miniaturization marks a paradigm change in device design, where "every atom matters." In contrast to traditional electronic devices, the diminishing sizes introduce significant influences, such as fluctuations in dopant concentration, interdiffusion of elements at interfaces, and local strain fields at the nanoscale, profoundly impacting device function and performance. Quantum devices, exemplifying the extreme case, can see these factors not only degrade but entirely disable their functionality. But not only for semiconductor-based devices. Ubiquitously, for any material in which miniaturisation is key towards property improvement, the precise characterisation of every single atom will be key in understanding its properties when embedded in a full device (i.e., ferroelectrics and their interatomic distances, catalytic particles and their shift into single-atom catalysts, batteries and the atom-wise *in-situ* analysis of their degradation mechanisms, and others). [5, 6, 7]

While substantial strides have been made in automating Scanning Transmission Electron Microscopy (STEM), especially in industrial settings for metrology and process characterization, a substantial challenge persists in extracting meaningful physical insights from the vast amount of data (raw images and spectra) generated during experimental analysis. This challenge hinders a deep understanding of embedded material heterostructures at the atomic level in devices. Conventional fab and lab metrology tools fall short in providing a comprehensive analysis of these intricate architectures, leading to a lack of sufficient statistical sampling for understanding performance variability among individual devices. Consequently, the demand for advanced analysis and physical modelling continues to rise, surpassing current capabilities.

In the present chapter, which constitutes the central contribution of the thesis, we address the challenges inherent in the traditional (Scanning) Transmission Electron Microscopy ((S)TEM)-related material exploration process and introduce a paradigm shift facilitated by recent breakthroughs in data analysis methodologies. [8, 9, 10] Thus, the emergence of Machine Learning (ML), Deep Learning (DL), Computer Vision (CV), and Artificial Intelligence (AI) has revolutionized the field of materials science, particularly in the context of microscopy data analysis. These advancements have enabled the attainment of analytical levels of accuracy, precision, and noise tolerance previously considered unachievable in (S)TEM-related analyses. [11, 12, 13]

The early stages of AI-driven methodologies associated with (S)TEM data analysis are characterised by the utilization of relatively straightforward unsupervised unmixing algorithms to decompose hyperspectral signals. They have evolved into the adoption of advanced models such as convolutional neural networks, autoencoders, or reinforcement learning. [4, 14, 15, 16, 17, 18] These advanced models are employed to unveil and learn features from images and high-dimensionality signals. Notably, these advancements have paved the way for the first crucial steps towards the automation of both experiments and data analysis. [19, 20, 21, 22, 23] This innovation strikes a balance, harnessing the high throughput of knowledge made accessible through automated data analysis facilitated by AI-based routines. [24, 25, 26, 27, 28]

As aimed by the central goal detailed in section 1.2, we explore the integration of these advanced techniques into a comprehensive and automated characterization workflow, aiming to overcome the traditionally slow and tedious aspects of material exploration. Not only does our solution provide a new paradigm of automation in STEM characterisation but also an easy way to further generate new knowledge which would otherwise require fine-tuned and laborious manual sample-dependant measurements. The proposed workflow starts by automating the data analysis process, traditionally considered a bottleneck, turning it into a solution for rapid and reproducible knowledge retrieval. The present chapter outlines a step-by-step approach inspired by the logical progression of human microscopists, beginning with low-magnification segmentation to gauge device morphology and culminating in phase and orientation-sensitive Geometrical Phase Analysis (GPA), for detailed structural insights on local atomic displacements, strain and defects.[29, 30]. Importantly, the workflow extends beyond mere data analysis and incorporates the automated generation of representative 3D Atomic (3DAMs) and Finite Element (FEMs) Models, utilizing the experimental data collected. Notably, these models comprehensively capture all the experimental information obtained through the preceding automated steps. Consequently, the automated workflow can simulate a device that precisely aligns with the initially designed, engineered, and grown device, a concept referred to as an experimental simulation, as these models are constructed entirely based on experimental data and parameters. For example, the

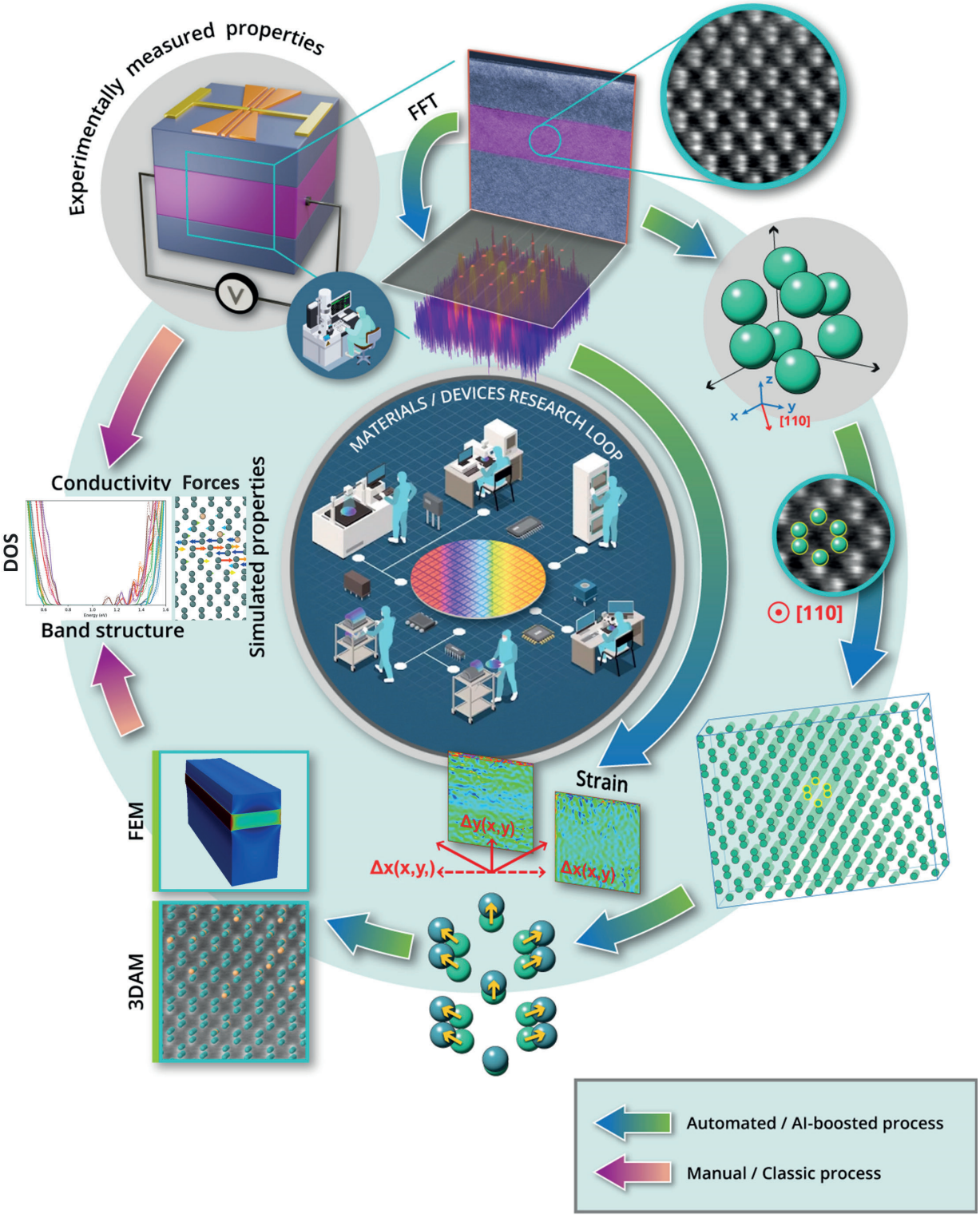


FIGURE 4.1: Scheme of the proposed workflow: Initially, we segment the image, breaking it down into distinct regions. Subsequently, we characterise the crystallography within each segmented region. This characterization allows us to map strain and finer structural details, such as interfaces or defects. The compiled information is then transferred to create Finite Element and 3D Atomic Models (FEM/3DAM). These models, in turn, serve as the basis for computing functional properties like the Density of States (DOS) or atomic forces and strain relaxations. This sequential process ensures a comprehensive analysis and understanding of the material's functional characteristics and behaviour.

finite elements of the FEMs are derived from the contours outlined by the segmentation of low-magnification images, while the atomic positions in the 3DAMs can be determined by the displacements identified through GPA. This capability allows for the exploration and answering of various questions, such as the impact of lamella thinning on strain release or the simulation of real defects in actual devices and their influence on their final performance.

The chapter demonstrates the flexibility and adaptability of the proposed workflow, emphasizing its modular nature. This modularity allows seamless integration of cutting-edge research and open-source tools from diverse sources, positioning the workflow at the forefront of AI-driven analysis of electron microscopy datasets. The potential applications of this automated approach extend to diverse fields, including energy and environmental research, classical communications, quantum technologies, mechanical engineering, and fundamental chemical research.

In the next section, we describe the backbone of the workflow and its central steps. However, the body of the chapter divides the workflow into two main parts, where we delve into the intricate details that make every step possible. Firstly, we showcase the automation of the data analysis. It provides a comprehensive structural characterisation route applicable to the, so far, widest spectrum of devices, samples, materials and STEM scenarios in the literature. Its capability to generalise and robustness makes it an exception to the current path of AI in electron microscopy to provide case-specific and tailored solutions to particular systems and inquiries. On the other hand, the second part of the chapter focuses on the generation of the models representing the target structures. Although this part is also by construction generally applicable to any system, it is physically limited by the assumption of translational invariance. Thus, multiple device morphologies can be modelled by this approach. Indeed, we validate the use of these models by performing theoretical simulations of systems that match the devices experimentally measured and that are eventually exploited in their functional context. Intriguingly, these simulations unlock a level of unprecedented sensitivity in the experimental structural properties and the functional ones. Overall, we discuss the broader implications of this AI-driven methodology and its potential to reshape the landscape of materials exploration and characterisation. [11, 31]

4.2 Workflow description

The workflow connects the data that is obtained directly from the electron microscope and the properties of the imaged heterostructure or device, in an intuitive and automated manner. The focus is on single images (numerically, single 2D matrices) or stacks of them (a number n $\neq$ 1 of 2D matrices), and in a second term, spatially correlated stacks of spectra, known as spectrum or spectral images (numerically, 3D

matrices). This work concentrates on automating image analysis rather than spectroscopy, although compositional information is incorporated when necessary to enhance fidelity.

The choice between a single micrograph and a stack of images depends on the device's or heterostructure's size, imaged features, and microscope sampling. A single micrograph is suitable when it captures both the morphology and atomically resolved structural information. Conversely, a stack of images is used for larger devices or structures, requiring a progressive decrease in magnification to transition from the morphological overview to atomic and structural details. Both situations are handled differently, with the latter requiring an additional processing step to link information gathered at each magnification.

The workflow consists of nine main modules:

4.2.1 Segmentation

Initially, our goal is to reveal the morphology of the device, encompassing its size, identification of key regions, spatial distribution, interfaces, and more. These regions typically represent different materials or may encode diverse crystal phases or orientations. Essentially, any statistically distinctive feature in the acquired signal can be set as a segmented feature. [32, 33] The workflow has been successfully demonstrated for both TEM and STEM data, where different physical phenomena generate the acquired intensity, such as phase versus composition. Despite these differences, both signals exhibit sufficient variability for proper data segmentation.

The segmentation model relies on either contour detection, suitable for STEM micrographs, or a trained convolutional neural network, which performs well in both STEM and TEM. This choice is determined by a binary hyperparameter defined at the workflow's outset. Interestingly, these hyperparameters, along with another mentioned shortly after, are the only two modifiable ones in the entire workflow. They allow users to slightly adjust the process based on the type of device or structure under analysis. A detailed explanation can be found in section 4.I.1.

4.2.2 Fast Fourier Transform peak finding

After segmentation, automated structural characterization begins by transforming real space images into reciprocal space patterns using Fast Fourier Transform (FFT).[34, 35, 36] Two methods have been used for regulating the access to Fourier space during image processing. The first involves masking the image with every segmented region and applying the Fast Fourier Transform (FFT) algorithm, preserving crystallographic information from the full segmented region. Alternatively, the region can be accessed by finding the largest square fitting within its masked pixels and extracting its FFT. This choice serves as a binary hyperparameter, being one of the two modifiable options in the algorithmics, with implications in how

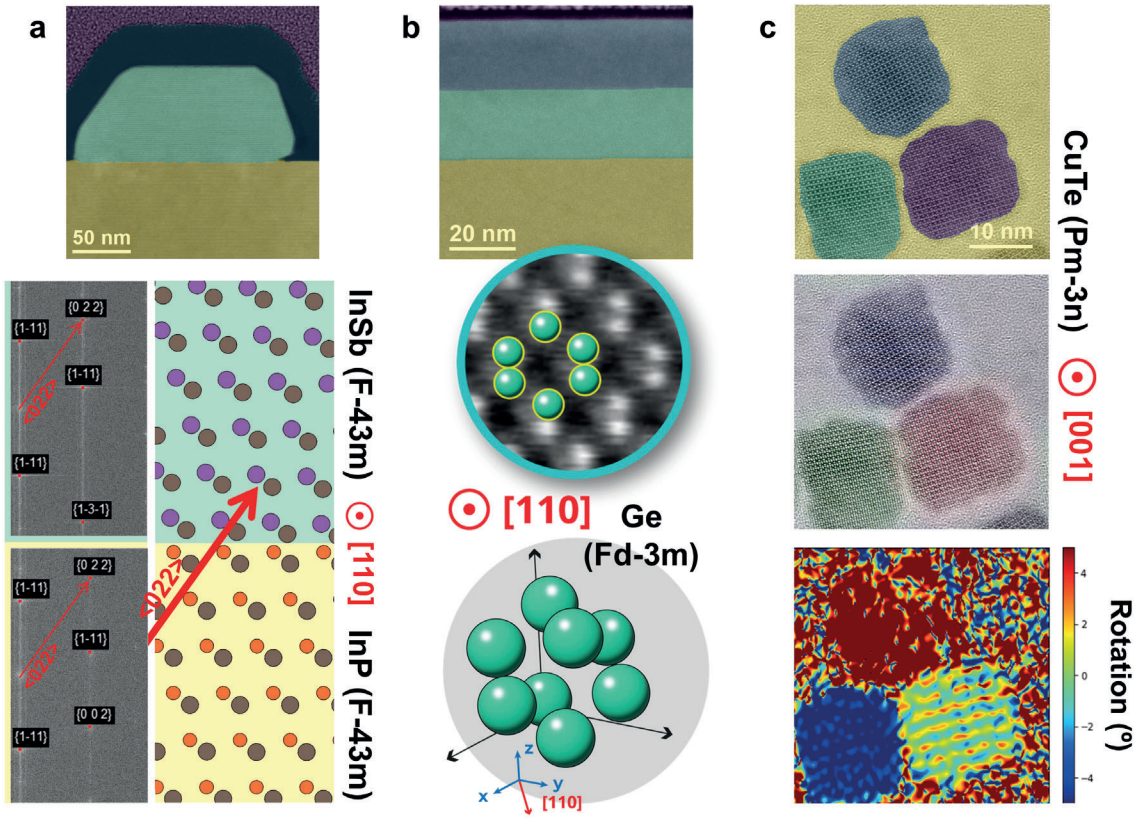


FIGURE 4.2: Outputs from the automated workflow for the comprehensive structural characterisation of devices. a) Scanning Transmission Electron Microscopy (STEM) image of an InSb nanowire cross-section on an InP substrate oriented along [011] axis. The contour-based segmentation accurately separates the four distinguishable regions of the heterostructure within the lamella. The automated phase retrieval indexes the Fast Fourier Transform (FFT) per region and finds the spatial orientation of both constituting phases. [37] b) STEM micrograph of a $Si_{0.3}Ge_{0.7}/Ge/Si_{0.3}Ge_{0.7}$ quantum well (device GSQ_4 from Chapter 3). [38] As in a), the segmentation and phase identification conclude the device is entirely formed by Ge with diamond structure, along [011]. Not surprisingly, the stoichiometry of $Si_{0.3}Ge_{0.7}$ from the buffer layers is not retrieved, as this information is not self-contained in the FFT. The next steps show how spectroscopy complements the workflow in these challenging scenarios. c) High-resolution parallel beam TEM micrograph of three CuTe square particles. The SAM segmentation shines in particle-like formats and highlights how the workflow is general and applicable to multiple device morphologies to retrieve the strain and the Bragg filtering. [39]

representative and precise the FFT from the selected region is. The goal in either case is to find the coordinates of the FFT frequency peaks representing crystallographic planes, as an encoding of the information of the imaged crystallographic phases. Crucially, since the workflow is designed with universality in mind, it necessitates a peak-finding method that can address all possible cases, including the detection of amorphous materials, single crystals, and polycrystalline materials. The proposed peak-finding model represents, to the best of our knowledge, the most accurate and robust autonomous FFT peak-finding method available in the literature. This model relies on the successive application and evaluation of up to three distinct methods:

ML-based 2D Gaussian fitting, a trained encoder-decoder convolutional neural network, and a 1D profile scanning. While not flawless, it achieves the optimal balance between true positives and false positives and demonstrates adaptability across a wide range of materials and corresponding image types. [32, 40] A detailed explanation can be found in section 4.I.2.

4.2.3 Phase identification

By disclosing the pixel coordinates of spectral frequencies corresponding to imaged crystallographic planes, the option to compare them with a database of candidate unit cell models is unlocked. This process involves ranking each potential crystal phase and assigning a score based on how well the experimental "diffraction" pattern aligns with their theoretical kinematical diffraction pattern. Each crystal phase possesses a distinctive kinematical diffraction pattern when aligned through a specific zone axis (i.e., crystal direction parallel to the electron beam), computable if the unit cell of the crystal phase is provided (i.e., the primitive cell with its space group). [41, 42] In summary, the distances and angles between experimentally identified spots are compared, checking if they align with a crystal phase from the database (physical modelling involved). This yields a stack of scores per unit cell, ranking the likelihood of each checked phase in every segmented key region. Consequently, the first valuable output includes, for each region in the device, the crystal phases that best explain the observed crystallographic planes, their zone index and spatial orientation. Figure 4.1 illustrates the central role of the indexing process within the whole workflow, while figure 4.2 showcases the success in generalising the phase identification task and how it enables further outcomes, such as Bragg filtering. A detailed explanation can be found in sections 4.I.5, 4.I.3, 4.I.7 and 4.I.8.

4.2.4 Low-to-high magnification correlation

At this point, the fourth step might shine on its own merit, even though it is actually performed before step 1. Although it is not required in the single-image scenario, this step becomes crucial when dealing with a stack of images as it aligns high-magnification micrographs with low-magnification ones. It applies to an image stack, as it circumvents the insufficiency of a single image to capture both morphological and structural details, especially in large devices or varied focal conditions. The process puts together images with varying experimental parameters like magnification, focus, and orientation, which are adjusted during consecutive image acquisition. The core idea involves sorting and matching these images by magnification, allowing the segmentation of low-magnification images to reveal device morphology. Next, the automated reciprocal space analysis is performed on higher-magnification images with structural details. This enables mapping phases from

atomically resolved images into the lowest magnification image, providing structural information in fields of view without atomic resolution. Interestingly, the process is optimised to ensure pixel size-aware maximization of cross-correlation between high and low-magnification images, emulating the benefits of 4D-STEM acquisition without handling large datasets. A detailed explanation can be found in section 4.I.4.

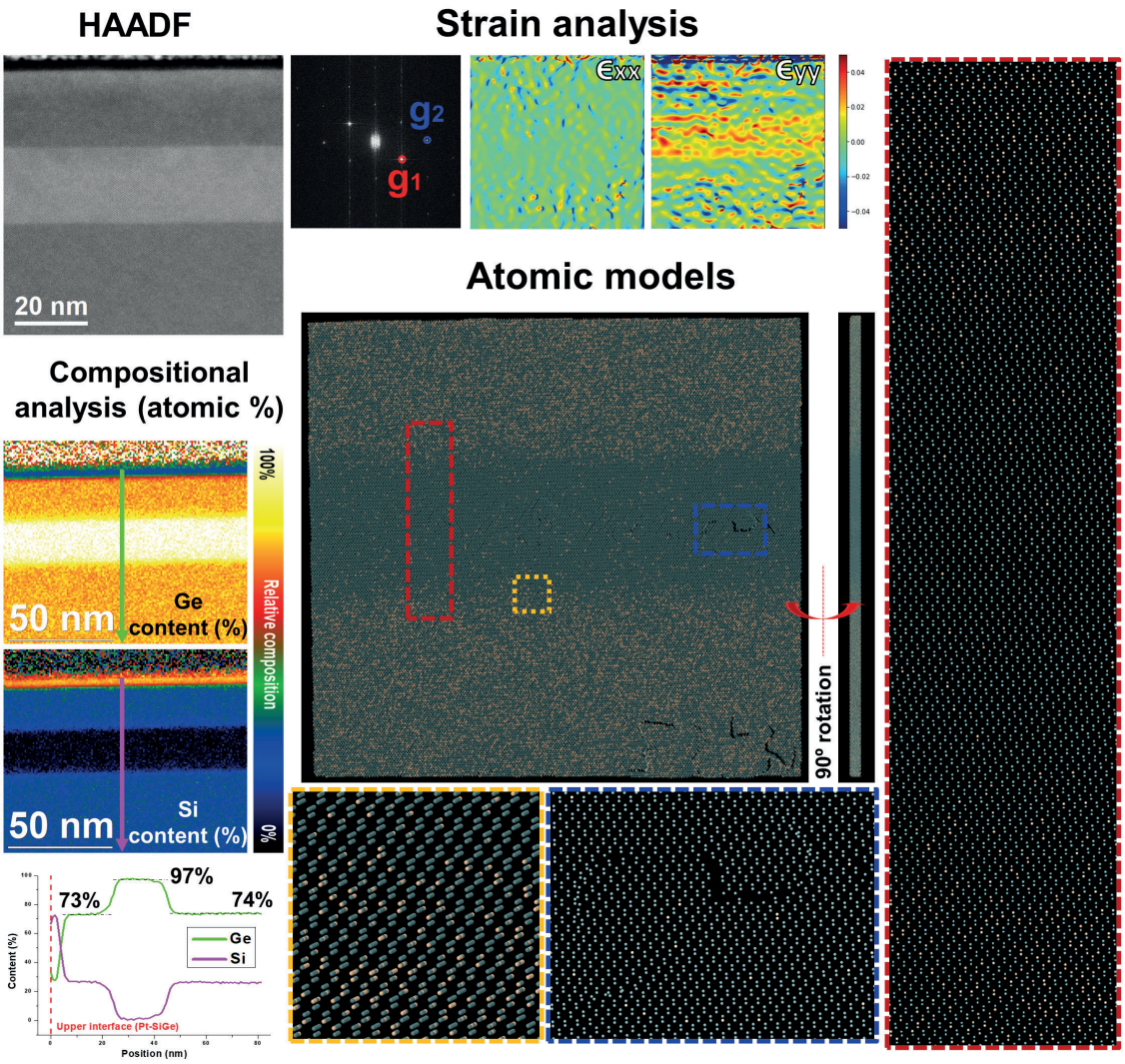


FIGURE 4.3: Automated generation of atomic models from the convergent knowledge from the analytical segment of the workflow. Representatively, the process is exemplified with the $Si_{0.3}Ge_{0.7}/Ge/Si_{0.3}Ge_{0.7}$ quantum well (device GSQ_4 from Chapter 3). [38] Its creation demands a collaborative effort between imaging and spectroscopy. (Top left) First, the analysis of the Scanning Transmission Electron Micrograph (STEM) micrograph renders the phase identification (i.e., Ge, diamond structure) and (top centre) the strain analysis. (Left) Secondly, the quantitative Electron Energy Loss Spectroscopy (EELS) (Ge and Si quantitative maps and profile on the left, showcasing the $Si_{0.3}Ge_{0.7}/Ge/Si_{0.3}Ge_{0.7}$ layered structure) locally defines stoichiometry as experimentally retrieved. (Centre) The combined knowledge allows to create a representative 3D atomic model which illustrates the device and allows further theoretical insights into a model created from experimental data. The thickness of the model is limited to 1.3 nm as its translational invariance is captured as long as a single unit cell is capture along this direction.

4.2.5 Strain analysis

Back in the single-image scenario, having a complete description of reciprocal space is powerful. Combining FFT indexing with knowledge of crystallographic planes, morphology, and physical proximity reveals a global vision of material arrangement. This information facilitates mapping epitaxial relations, identifying heterojunctions or single crystalline blocks. Correlating the mapped reciprocal space through Geometrical Phase Analysis (GPA) in an automated fashion is possible, leveraging parameters obtained from real and reciprocal space information. The GPA process involves selecting a crystalline segmented region as a reference and identifying optimal planes as g vectors based on the relative orientation of all the indexed crystallographic planes of the image. [29, 30, 41, 42] Crystallographic frequencies from other regions are compared to chosen g vectors, and a mask is drawn to fit neighbouring spots with appropriate spatial resolution. GPA reveals spatial dependencies of the elastic strain tensor components, addressing an until now unattended gap in the workflow: interfaces. Beware that these outputs refer to relative strain, not absolute, as extensively differentiated in Chapters 2 and 3. Not surprisingly, this will have strong implications and become key for the overall workflow, particularly in the last stages.

Assessing the relative orientation of crystals and their correlation is valuable for visually mapping crystal phases and planar defects, particularly with the automated Bragg filtering process added to the workflow. GPA excels in highlighting minor structural details within different interfaces of the device, going beyond the capabilities of previous phase identification steps. It detects defects, dislocations, and alterations from the expected perfect crystal lattice, offering detailed insights not easily discernible by the naked eye. However, interpreting GPA output remains a manual task for a trained materials scientist.

The workflow, which initially makes no assumptions about the device heterostructure, is versatile and applicable to various morphologies, crystal systems, Bragg lattices, space groups, and settings. It provides general segmentation and complete structural information (i.e., crystal phases, FFT indexing, Bragg filtering, strain) for devices or structures. Moving forward, the workflow assumes translational invariance (i.e., projecting 3D information into 2D micrographs), which is reasonable for many cases, especially when imaging lamellae extracted from 1D (e.g., nanowire quantum networks) or 2D (e.g., 2D electron gases or 2D hole gases based on quantum wells) heterostructures. This assumption allows for extending the workflow with extra steps to build representative models of the device, enabling the extraction and simulation of functional properties straightforwardly. Figures 4.2 and 4.3 illustrate the capability of the workflow to retrieve the strain tensor components and rotations. Importantly, this knowledge will be fundamental for the creation of the 3D models we aimed for in the central objective proposed in section 1.2, as depicted in figure 4.3. A detailed explanation can be found in section 4.I.6.

4.2.6 Finite Element Model (FEM) building

The first type of models generated by the workflow will be Finite Element Models (FEMs). FEMs describe complex physical systems using discrete geometric elements. This enables numerical calculations that would otherwise be impractical when considering the entire global system simultaneously. [43, 44, 45] The workflow's prior segmentation of the device facilitates the creation of FEMs, representing the morphology as contour vectors in a boundary elements model encoded into a Graphic Data Stream (.gds) file. This file contains contour information, materials, and identified phases. Additionally, a crystallography calculator determines the crystal planes aligned with the Cartesian axes based on the former automated FFT indexing. This information proves valuable for incorporating crystal phase-specific or orientation-dependent properties in the simulation, such as anisotropic strain tensors or conductivities.

The workflow-based model generation outperforms traditional approaches to FEM work. It is automated, fast, and does not require manual input for system properties or tedious building of morphology with simple polygons. Moreover, it is more accurate and representative of the actual heterostructure or device being simulated. However, the introduction of simulation properties (e.g., mesh size, material properties database, boundary conditions) still requires manual setup in simulation software based on extracted knowledge. This is illustrated in Figure 4.1, and the FEM results, presented in Figure 4.4, demonstrate the type of results achievable. A detailed explanation can be found in section 4.II.2.

4.2.7 3D Atomic Model (3DAM) building

Compared to FEMs, 3DAMs provide a more precise description of the actual device and its structural properties. [46, 47] While this precision allows for higher theoretical levels in simulations that extract functional properties, it comes at a greater cost in computing time and resources. [41, 42, 48, 49, 50, 51] The fundamental concept here is to generate an atomic model, a file containing the three spatial coordinates of all atoms constituting the entire device or a specific region of interest. These positions accurately represent the actual atom positions within the experimental device. Perfect crystals of identified phases are constructed, aligning with the found conventional unit cell and device shape (including faceting). A detailed explanation can be found in section 4.II.3.

4.2.8 Strain transfer to atomic model

The initial 3DAM volume, populated with atoms based on crystal phase symmetries, is subsequently correlated with GPA analysis, particularly focusing on displacement maps. The supercell volume is populated with a reference virtual unit cell, modified to match experimental distances indicated by two g-vectors used as a reference in

GPA. As a result, GPA-discovered relative displacements are directly transferred to the atomic model. An iterative algorithm displaces every atom from the reference based on displacement maps, resulting in an atomic model reflecting features highlighted by GPA as shifts from the perfect crystal lattice. Figure 4.3 evinces that combining the GPA strain fields, with spectroscopic information obtained by Electron Energy Loss Spectroscopy (EELS), unlocks the creation of the representative atomic models we were initially aiming for (i.e., research goals, section 1.2). The exact way this knowledge intercommunication is achieved is explained throughout the present chapter. A detailed explanation is provided throughout section 4.II.3.

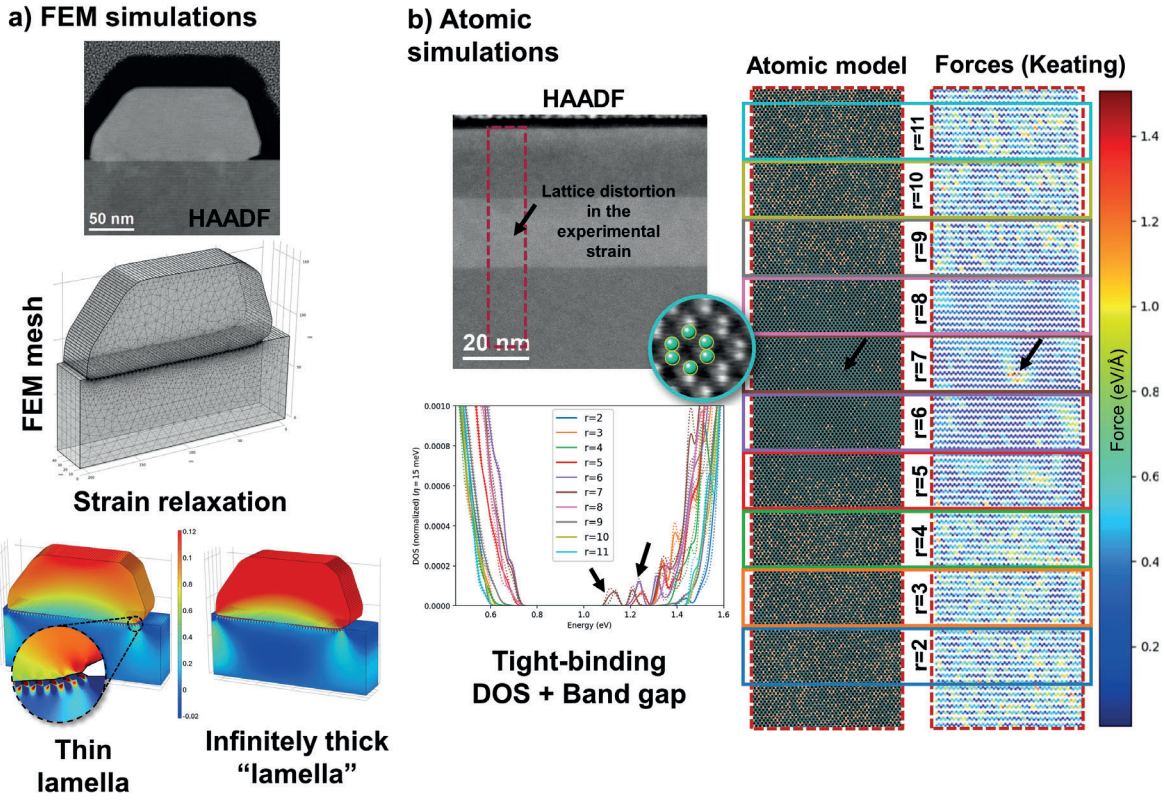


FIGURE 4.4: Finite Element Model (FEM) and atomic simulations out of the workflow's outputs. a) FEM simulations of an InSb nanowire grown on InP. The segmentation allows us to extract its morphology, and we make a 3D geometrical mesh to compute the strain relaxations even in the presence of fine features like dislocations (zoom-in). b) Atomic simulations from the 3D Atomic Model (3DAM) of the Ge quantum well. The complementary simulations of forces (local peak in the forces) through the Keating model, and electronic structure through tight binding (presence of mid-gap states), revealed an invisible lattice distortion artefact caused by scanning noise in the electron micrograph acquisition, highlighting the potential of the workflow to detect subtle structural features.

4.2.9 Simulation of functional properties

Once the atomic model accounts for elastic displacements, it becomes an input for simulating functional properties with precision achievable only when considering the discrete atomic nature. Simulation paradigms exploiting the atomistic nature, such as molecular dynamics or *ab initio* calculations, can relax structural strain and

compare it with FEM simulations or experiments. This enables the computation of functional properties like conductivity in complex devices, bandgap, dispersion relations, or any property relevant to understanding intricate devices like quantum wells and hybrid quantum NWs. Key examples, along with additional configurations adding extra value and insight into the workflow's capabilities, validate its effectiveness. [37, 38, 52, 53, 54] Figure 4.4 underscores the possibilities behind the simulations that can be done after retrieving the 3D models. Figure 4.4,a showcases the ease to tune and adjust FEM simulations to our requirements. For instance, to see the effect of the lamella thinning in the experimental strain, as evaluated in Chapter 3. On the other hand, Figure 4.4,b highlights the achievable atomic simulations, which include forces calculations but also *ab initio* simulations through models like tight-binding. Intriguingly, we are able to unveil preliminary fascinating effects in the electronic properties caused by subtle experimental structural features captured by the previous steps of the workflow. Indeed, this structural-functional correlation closes the loop proposed in the research goals section 1.2, while awaiting for additional experimental output from Chapter 5. A detailed explanation can be found in section 4.II.5.

In summary, the workflow, which assumes translational invariance in the studied heterostructures, is versatile and applicable to various devices. Thorough proofs of it on real experimental devices and samples can be found in section 4.II.4. We concludes by demonstrating the simulation of these previously mentioned functional properties using FEMs and 3DAMs, providing insights into complex heterostructures and devices such as the Ge QW serving as a 2D Hole Gas (2DHG) device, previously used to demonstrate a singlet-triplet hole spin qubit in planar Ge, [38] (see GSQ_4 in Chapter 3, and figure 4.4,b) or the InSb nanowire cross-sections extracted from a hybrid semiconductor-superconductor nanowire quantum network [37] (fig. 4.2 and 4.4,a), among others.

This succession of steps closes the workflow and makes it a solid unity in which the sum of parts is stronger than their standalone contribution. However, the importance and value of each module as an independent unit must not be disregarded. The details, validity, key aspects, strengths and weaknesses of every module are comprehensively described in the text below, with the twofold chapter division in analysis and modelling as justified above.

Part I

Comprehensive and General Automation of the Structural Analysis of (Scanning) Transmission Electron Microscopy Data

Part I of Chapter 4 comprises all the steps or modules of the workflow involved in the thorough structural characterisation of devices. This comprehensive knowledge of our device is obtained after 1) Segmentation, 2) FFT peak finding, 3) Phase Identification and FFT indexing, 4) Low-to-high magnification correlation, and 5) Strain analysis. The provided level of detail should be enough to help in the readability of the code once open and publicly available, and even encourage the custom reproduction of the workflow, if wanted.

4.I.1 Segmentation

The initial step of the workflow demands the development of a segmentation model capable of discerning between various regions within the samples. Establishing such a model is essential for subsequent algorithms to independently analyze each component of the device. Given that distinctions among different areas in EM images can be based on signal intensity and position (elaborated in more detail in the methodology), clustering algorithms were employed to construct this segmentation model. Among the myriad options available for such algorithms, ML algorithms stand out as highly effective. Specifically, ML algorithms can be categorized into supervised and unsupervised types. Supervised ML algorithms utilize labelled data to autonomously learn regression and classification processes, while unsupervised ML algorithms classify data by defining clusters based on shared similarities. Given the complexity of acquiring large labelled datasets for training models, unsupervised ML algorithms emerged as the preferred choice in this case.

Within this project, various unsupervised ML algorithms were tested, including K-means clustering [55], Agglomerative clustering [56], and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [57]. An alternative approach to constructing this algorithm involves detecting the edges of each region instead of identifying distinct regions. To explore this avenue, a classical border-finding algorithm was also tested. [58]

4.I.1.1 Methods

4.I.1.1.1 Data treatment

The Python programming language was utilized to code all the algorithms required for both models and to conduct statistical analyses on the final results. We selected Python due to its open-access nature and the abundance of packages that facilitate the implementation of Machine Learning (ML) algorithms (see Methodology). [59]

Python also offers ease of working with TEM images. Upon acquisition, these images are initially in a proprietary .dm3 format, which can be read using the open-access hyperspy Python package, enabling their conversion into a matrix. [60, 61] We worked with 2048x2048 pixel images, for that, the resulting matrix, denoted as M, has a size of 2048x2048. The initial step before commencing data treatment was to reduce M to a 256x256 matrix, denoted as RM (reduced matrix). This downsampling is crucial for two reasons: it diminishes intensity variance in each region of the device, facilitating segmentation, and significantly reduces the time and memory requirements of the algorithms.

Unsupervised ML algorithms aim to segregate n-dimensional data into different clusters using the Euclidean metric. In our case, we dealt with 3D data: two spatial dimensions and the intensity of each pixel as the third dimension. To ensure compatibility with the data in an appropriate format, we had to transform the RM matrix into a 3x65536 (256^2) matrix using the following conversion:

$$m_{ij} \rightarrow [i, j, m_{ij}] \tag{4.I.1}$$

where m_{ij} is the i, j value of the RM matrix. We designate this new matrix as DM (data matrix). Using DM as input, the unsupervised ML algorithm divides all the data into n different clusters, resulting in a new 256x256 matrix LM (labeled matrix). The $i, j-th$ value of this matrix represents the cluster number assigned to the $i, j-th$ pixel of the reduced image. Additionally, depending on the specific algorithm used, certain extra hyperparameters must be introduced, as explained below.

4.I.1.1.2 Tested algorithms

We tested three distinct unsupervised ML algorithms, employing the Scikit-learn package. [62] As indicated in the Methodology chapter, Scikit-learn is a Python module specialized in machine learning, allowing us to apply these algorithms without the need to code from scratch. We also conducted tests with border-finding algorithms, specifically utilizing the Canny edge detection algorithm implemented in the OpenCV package. OpenCV is a module specialized in computer vision and ML algorithms applied to images. [63]

The first of these algorithms is K-means clustering. The K-means algorithm clusters data by grouping samples into n clusters of equal variances, minimizing a criterion known as inertia or within-cluster sum-of-squares:

$$\sum_{i=0}^{n} \sum_{j=0}^{n'} \left\| x_j^i - \mu_i \right\|^2 \tag{4.I.2}$$

where x_j^i is $j-th$ point of the $i-th$ cluster and μ_i is the centroid position of $i-th$ cluster. These centroids are initialized in random positions, and at each step, they are moved to minimize 4.I.2. Moreover, the clusters are redefined throughout the process based on the proximity of the points to different centroids. [64] Once the algorithm finds the minimum, it returns the final clusters using the LM matrix. For this, the algorithm requires manually entering the number of clusters as a hyperparameter. [65]

The second algorithm is Agglomerative clustering. This algorithm belongs to the family of hierarchical clustering algorithms, which build nested clusters by successively merging or splitting them. In particular, Agglomerative clustering employs a bottom-up approach where each data point starts in its own cluster, and clusters are successively merged. This process continues until 4.I.2 is minimized, similar to the K-means algorithm. [64] However, in some cases, the results may be similar but not identical because the segmentation process is different. This algorithm also necessitates manually entering the number of clusters as a hyperparameter. [66]

The third unsupervised ML algorithm is DBSCAN. This algorithm views clusters as areas of high data density separated by areas of low density. A cluster consists of core data points (close to each other) and non-core data points (close to a core data point but located in areas of lower density or cluster edges). A core point is defined as a point in the dataset with n other points within a distance of r, which are considered neighbours of the core point. Therefore, a cluster is built by recursively taking a core point, finding all of its neighbours that are core points, and repeating the process. Points that are neighbours of core points but not core points themselves are also included in the cluster as non-core points. [64] Unlike the previous two algorithms, DBSCAN does not require entering the number of clusters as a hyperparameter because it is defined dynamically during the process. Instead, manual input of n and r values is necessary. [67]

Edge-finding algorithms leverage derivative calculations to identify abrupt changes in pixel intensity, which are labelled as edges. In the case of the Canny algorithm, the gradient in each pixel is calculated and used to identify edge pixels. As the gradient direction is perpendicular to the edges, the algorithm detects pixels maximizing the gradient modulus 4.I.3 in the gradient direction as edges. Additionally, this algorithm allows for adaptation to the samples using the $maxVal$ and $minVal$ hyperparameters: if 4.I.3 > $maxVal$, the pixel could be considered as an edge, and if 4.I.3 < $minVal$, the pixel is not considered as an edge. [68] In our case, we focused on the $maxVal$ hyperparameter because it allowed us to threshold pixels with lower intensity variance that could be incorrectly detected as an edge.

$$\|\nabla p\| = \sqrt{p_x^2 + p_x^3} \tag{4.I.3}$$

4.I.1.1.3 Metrics

To assess the model's performance, we employed five distinct samples, incorporating multiple images per sample in certain cases to encompass diverse regions, magnifications, and orientations. This deliberate variability enables us to obtain an initial assessment of the models under general conditions. The five samples fall into two primary categories: Ge QWs for 2D hole gas devices (fig. 4.I.1, a and b [69, 70]) and III-V semiconductor nanowires (fig. 4.I.1, c, d, and e [2, 37]). The choice of samples was motivated by their inherent ease of distinguishing different parts within the devices or nanostructures. This characteristic facilitates the establishment of a metric for evaluating the accuracy of our model, as further detailed in the subsequent section.

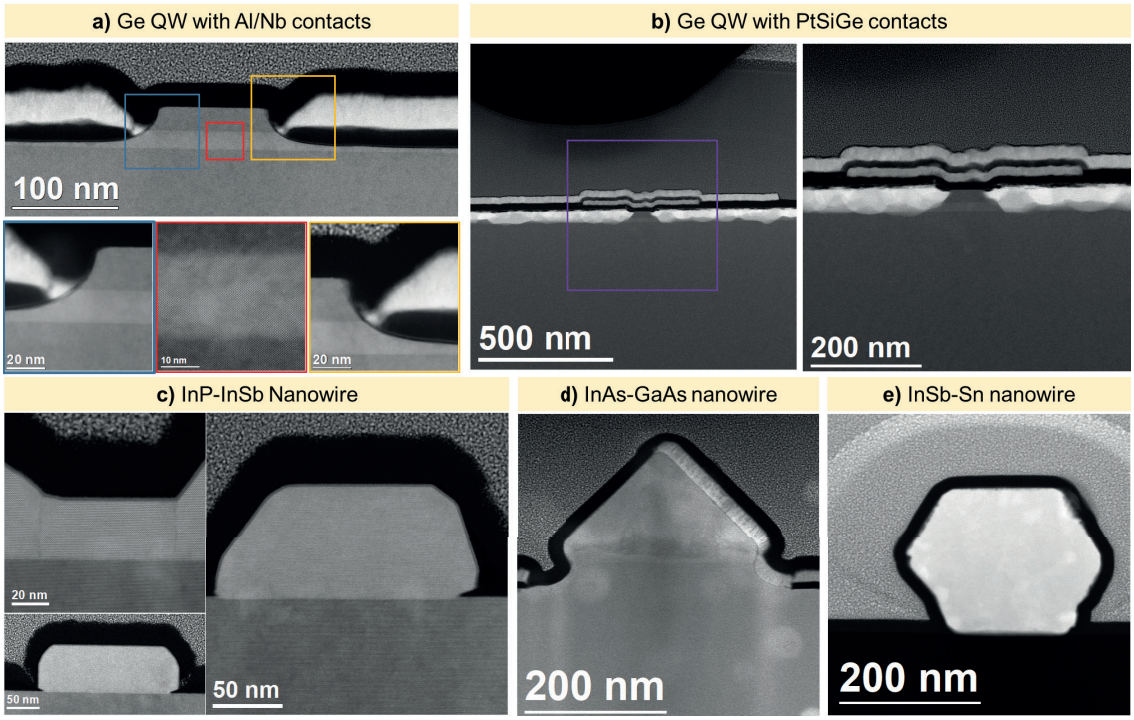


FIGURE 4.I.1: High-Angle Annular Dark-Field Scanning Transmission Electron Microscopy (HAADF–STEM) images utilized for model testing. a) Germanium Quantum Well (QW) with Al and Nb contacts for 2D hole gas devices, corresponding to device GSu_02 from Chapter 3. [69] Ge QW interfaces were oriented along the [011] zone axis. Cross-section images are presented in low (top) and high (bottom) magnification conditions. High-magnification images offer atomic resolution for the detailed examination of various areas of interest. b) Ge QW with PtSiGe contacts, corresponding to device GSu_10 from Chapter 3 [70]. A very high magnified image is presented for a comprehensive overview of the device (left), and a less magnified image (right). c) and d) present, respectively, InP on InSb, and InAs on GaAs nanowires. [37] e) Presents a vapour-liquid-solid nanowire alloying Sn and InSb. [2]

The micrographs evaluated in the present section were acquired using a FEI TITAN microscope in HAADF–STEM imaging mode (see Methodology). Operating under

constant thickness conditions ensured that brightness solely depended on the composition of each sample region (i.e., Z^2, see Methodology). As highlighted previously, this condition is pivotal for creating a segmentation model capable of permitting the ultimate goal of distinguishing different crystal phases within a sample.

To establish our metric, we initially performed manual segmentation on all images used in this study (fig. 4.I.1). This manual segmentation involved applying different intensity and position thresholds customized for each image. The manually segmented images were then used as the ground truth. Subsequently, to calculate accuracy, we systematically paired each region in the ground truth images with a corresponding region in the images segmented by a clustering algorithm. This pairing process was executed with the constraint of not repeating regions in the assigned pairs and aimed at maximizing pixel matches between the respective paired regions. All pixels matched by the paired regions were deemed as well-segmented pixels. Accuracy was computed by calculating the ratio of well-segmented pixels to the total number of pixels in the image.

4.I.1.2 Results

In this section, we present and discuss the results of the clustering algorithms' performance on the devices outlined in the methodology from the previous paragraphs (fig. 4.I.1). We divided the study into three stages, each with a higher degree of complexity and generality than the previous one. In the first stage, we tested the accuracy and time requirements of all the algorithms on the same sample. We selected the InP-InSb nanowire grown along the [112] direction and cut transversally (fig. 4.I.1, c, bottom left) for these tests, denoted as nanowire 1 for simplicity. This sample was chosen as a starting point due to its four well-differentiated regions, facilitating both visual and numerical evaluation of the results. For each algorithm, we tuned hyperparameters to maximize accuracy. In the case of K-means and Agglomerative clustering, we achieved the maximum accuracy with four clusters (nc=4), corresponding to the four different regions of the sample. For the DBSCAN algorithm, we fixed the neighbour's distance at r=0.1 (where 1 is the length of the image side) and maximized accuracy with n=100. Note that, in this algorithm, the clusters consider data density, proportional to the ratio between n and r. Therefore, fixing one hyperparameter and maximizing the other is sufficient. Furthermore, fixing n and maximizing r yields the same result. Finally, for the Canny edge detection algorithm, we achieved maximum accuracy with $maxVal$=50. In table 4.I.1, we present the performance results in accuracy and time for each algorithm. According to this metric, accuracy between 90-100% is considered good segmentation, 0-80% is considered bad segmentation, and percentages in between represent intermediate results. It's crucial to note that accuracy below 80% implies that more than one-fifth of the image is poorly segmented, indicating high noise levels in each region and/or entire undetected regions in the sample studied.

As shown in table 4.I.1, all the algorithms exhibit high accuracy, with Canny edge detection outperforming the others. This superior performance is attributed to the fact that the Canny algorithm detects edges rather than regions. Consequently, unlike unsupervised ML algorithms, it only introduces noise at the edges of regions and not in the core. The segmentation outputs for this first stage are presented in figure **??**, a). Additionally, we evaluated the time required to segment one image for each algorithm. We observed that all the algorithms perform segmentation in a reasonable time, except for Agglomerative clustering. This algorithm necessitates over 13 minutes to segment a single image, more than 200 times longer than the other algorithms. This renders the Agglomerative clustering impractical for scaling the model for larger data flows or images. Consequently, we excluded this algorithm from subsequent stages of the study, despite its high accuracy.

Algorithm	**Accuracy (%)**	**Time (s)**
K-means clustering	93.3	0.4
Agglomerative clustering	96.8	784
DBSCAN	95.6	3.2
Canny edge detection	97.8	1.4

TABLE 4.I.1: Performance of Clustering Algorithms in InP-InSb Nanowires. The accuracy measures the percentage of well-segmented pixels, as per the metric outlined in the methodology. Additionally, the table presents the time required for the image segmentation for each algorithm.

In the second stage, we tested the remaining algorithms on every NW using the same hyperparameters as the previous stage. The objective of this phase is to evaluate how well the algorithms can generalize to different samples with similar sizes and morphologies. It is important to note that the analyzed areas containing the NWs exhibit a consistent pattern: they consist of four distinct regions. An overview of these samples is presented in figure 4.I.2, b. To simplify the notation of the samples, we number them as follows:

1. InP-InSb NW grown along [112] direction and cut transversally (fig. 4.I.1, c, bottom left). This is the NW used in the previous stage.
2. InP-InSb NW grown along [112] direction and cut longitudinally (fig. 4.I.1, c, top left).
3. InP-InSb NW grown along [011] direction (fig. 4.I.1, c, right).
4. InAs-GaAs NW (fig. 4.I.1, d).
5. InP-Sn NW (fig. 4.I.1, e).

This part of the study provides insight into how well the algorithms generalize across different samples with similar characteristics. All nanowires maintain precisely four distinct regions, contributing to the overall understanding of algorithm performance across variations in morphology and sample composition.

Nanowire	K-means clustering	DBSCAN	Canny edge detection
1	93.3	95.6	97.8
2	95.4	66.2	66.9
3	96.0	67.8	96.7
4	55.3	64.8	97.3
5	71.4	87.8	89.4
Mean	**82.3**	**76.4**	**89.6**

TABLE 4.I.2: Accuracy of clustering algorithms in the nanowire devices. Accuracy of well-segmented pixels, in %, of each algorithm applied to each nanowire. The mean accuracy for each algorithm is also presented.

Table 4.I.2 showcases that some nanowires for each algorithm were not well segmented. In the case of K-means, NWs 4 and 5 did not yield good results. This occurred because both images have a relatively small region compared to the others. The algorithm merged these small regions with a neighbouring region and split the largest one into two parts. This significantly reduced accuracy and serves as a clear example of a bad segmentation. Regarding DBSCAN, NWs 2, 3, and 4 were incorrectly segmented. In all three nanowires, the algorithm merged different regions into one. This algorithm struggles to generalize the process for different samples, as finding the optimal value of the hyperparameter n for each would be required. Finally, in the case of the Canny algorithm, only NW 2 did not achieve high accuracy. This occurred because the algorithm could not detect the edge between two specific regions and merged these two regions into one. Another important observation is that nanowire 5 obtained an accuracy of less than 90% in all the algorithms. This is due to the fact that all the algorithms detected the substrate and the covering layer of the nanowire as the same region, since both exhibit very similar intensities.

The algorithm that produced the best results in this stage was the Canny edge detection algorithm. It achieved almost 90% accuracy on average, with only one sample not achieving high accuracy. Moreover, changing the value of the *maxVal* hyperparameter from 50 to 25 in the case of NW 2 increased accuracy in this sample to 91.2%. One possible reason for the different performance of the Canny algorithm in this nanowire compared to the others could be the magnification of the nanostructure image. The size of this NW image is 100 nm, while all the other images are larger than 200 nm. As this algorithm yielded better results than the others in both stages, we only considered this one in the next stage. The segmentation outputs of the Canny algorithm in the second stage, with the highest accuracy, are presented in figure 4.I.2, b.

In the third stage, we tested the Canny algorithm on all the devices and nanostructures in this study. The primary goal of this stage was to analyze how well this algorithm generalizes across more diverse samples, and for this, the Ge QW images were a perfect choice. They represent a different type of device and also encompass

a wide range of magnifications from 40 nm to 1000 nm image size. An overview of these devices is presented in figure 4.I.2, c. To simplify the notation of the devices, we number them as follows:

1. Ge QW with Al/Nb contacts (fig. 4.I.1, a, top). GSu_02 device from Chapter 3. We also denote the high magnification images of this device as 1l for the left part (fig. 4.I.1, a, bottom left), 1c for the central part (fig. 4.I.1, a, down center), and 1r for the right part (fig. 4.I.1, a, bottom right). [69]
2. Ge QW with PtSiGe contacts (fig. 4.I.1, b). GSu_10 device from Chapter 3. We denote 2a for the low magnification image (fig. 4.I.1, b, left) and 2b for the more magnified image (fig. 4.I.1, b, right). [70]

Ge QW	**Canny edge detection**
1	93.4
1l	59.8
1c	53.7
1r	55.7
2a	92.3
2b	96.4
Mean	**71.6**

TABLE 4.I.3: Accuracy in well-segmented pixels % of Canny edge detection algorithm applied to each Ge quantum well device. The mean accuracy for each algorithm is also presented.

As seen in table 4.I.3, the devices that did not achieve good accuracy were 1l, 1c, and 1r, corresponding to the highest magnification images: 76 nm, 38 nm, and 76 nm, respectively. We observed the same result as in stage 2, where at the highest magnification, the algorithm did not properly detect all edges. However, by adapting the *maxVal* hyperparameter to these devices, similar to what was done with NW 2, we obtained good segmentation outputs. For devices 1l and 1r, changing the *maxVal* from 50 to 20 increased accuracy to 94.2% and 82.3%, respectively. In the case of 1r, we did not achieve 90% accuracy because the model did not detect the upper region due to the great variability of intensities it presents. For 1c, changing the *maxVal* from 50 to 10 increased accuracy to 96.6%. With these adjustments, the algorithm achieved an average accuracy of 92.4%. All segmentation outputs with the highest accuracy are presented in figure 4.I.2, c. We observed that the value of the *maxVal* hyperparameter that maximizes the accuracy of the segmentation is inversely proportional to the device magnification. Thus, the Canny algorithm offers the possibility of creating an autonomous segmentation model by correlating these two hyperparameters.

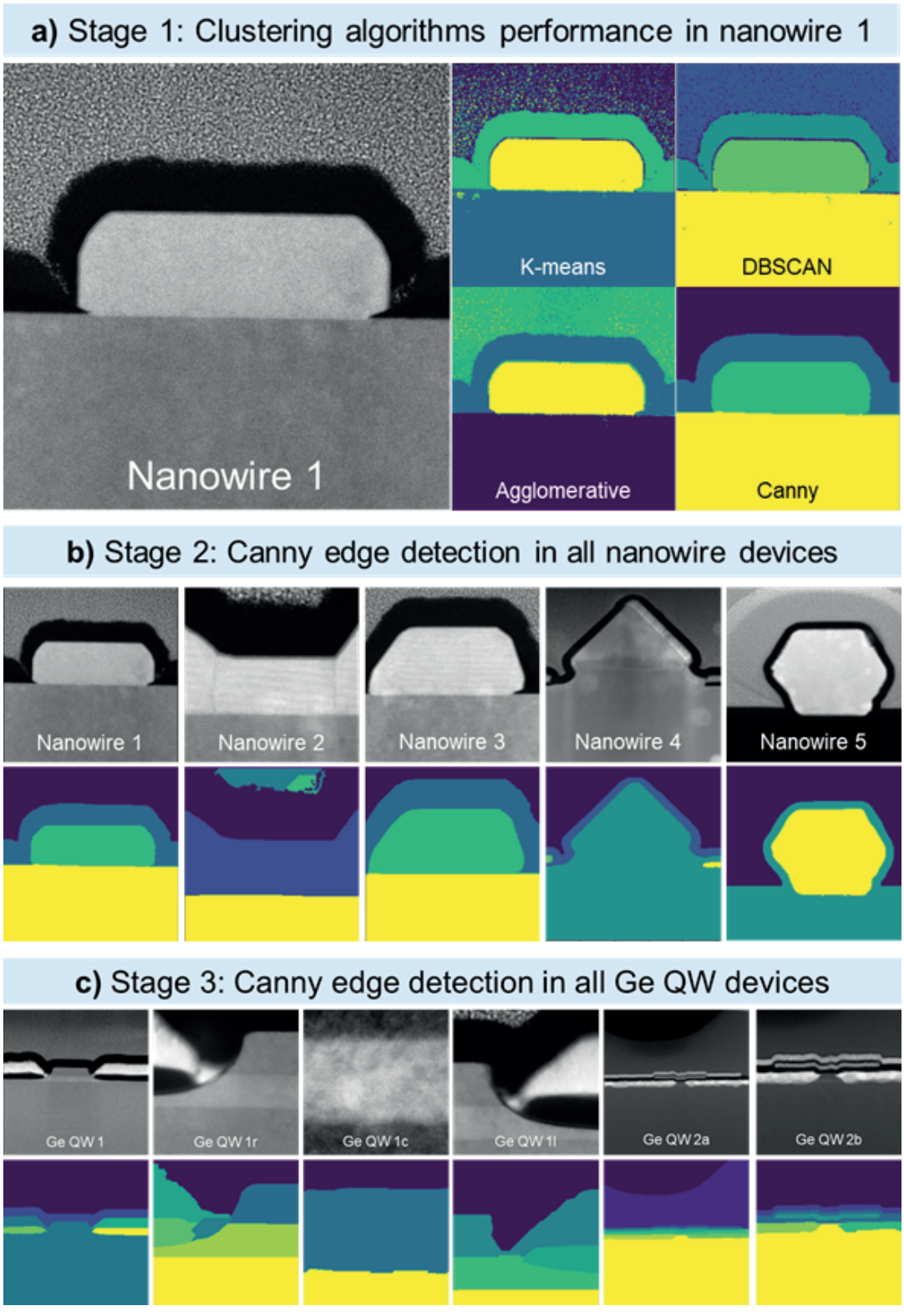


FIGURE 4.I.2: a) Stage 1: Results of all the clustering algorithms applied to nanowire 1. The HAADF-STEM image of the nanowire (left) and the output segmented images of each algorithm (right) are exposed. In segmented images, each pixel with the same colour is labelled as the same cluster. b) Stage 2: Results of the Canny algorithm applied in the nanowires. For each image pair, the HAADF-STEM nanowire image (top) and the respective segmented image (bottom) are shown. c) Stage 3: Results of the Canny algorithm applied in the Ge QW devices with the same disposition than b). The segmented images correspond to those that obtained the highest accuracy.

4.I.1.3 Final segmentation model proposed

After the evaluation of the different algorithms proposed in the previous section, the metrics obtained, and their generalisation capabilities, we ended up with the final model we propose for correctly segmenting electron micrographs. We were initially limited by the absence of labelled data to directly deploy a robust supervised model. Therefore, an aspect of primordial importance is the trade-off between resources deployed into the generalisation of the model, its final performance, and the avoidance of mechanical Turk-like approaches. Thus, we wanted to create a model which performs well in the demanded task, but it is not purely based on the brute force training of a neural network. This way, we discarded the approach of manually labelling data and the posterior supervised approach due to the absence of knowledge gained in the labelling process, the lack of high magnification micrographs which base the interest in segmentation models, and the possibility of consisting of a future refined routine from the obtained model. Instead, we pushed for algorithms which intrinsically are not as good in generalising, despite probably ending up with a less general segmentation model. However, its manual tailoring allowed us to gain insight into the data we worked with, as well as try multiple approaches in a short time, providing valuable insight into the unapparent challenges of a robust segmentation. In fact, the segmentation is oriented to low-magnification images (i.e., Fields Of View (FOVs) displaying the full device and its morphology). That is because atomically-resolved features displayed in higher magnifications alter the intensities histogram diluting the features highlighting the morphology. Thus, if we work with a single image, it is, of course, always segmented and processed following the full workflow. In this scenario, an intermediate magnification providing a clear overview and atomic resolution constitutes the optimum input for the workflow. However, if a dataset with different magnifications is provided, the segmentation is only done in micrographs presenting FOV larger than 150 nm. In these cases, the segmentation of the high-magnification micrographs is done by template-matching correlation with the segmented low-magnification image, as explained in the template matching section a few paragraphs later (section 4.I.4)

The creation and assessment of the presented custom segmentation model was conducted alongside the evaluation of a state-of-the-art segmentation neural network, the Segment Anything Model (SAM). [71] SAM is fed out of the largest segmentation dataset to date, named SA-1B, comprising over 1 billion masks across 11 million licensed and privacy-respecting images. This model demonstrates impressive zero-shot performance (i.e., categorising data not used for training), often surpassing previous results obtained with fully supervised learning, as we show for electron micrographs.

We obtained the same metrics we got from Canny edge detection but with this CNN originally designed for general digital images but not specifically trained for electron microscopy images. Keep in mind that supervised segmentation models eventually

learn convolutional kernels which individually detect edges. Their strength is their ability to construct multiple complementary edge detection kernels, which typically boost the performance of the supervised solution (fig. 4.I.3 and 4.I.4). Therefore, a possible future SAM fine-tuning for our purpose would be based only on (S)TEM micrographs. Despite the added complexity, we will see that SAM is not unbeatable and our custom Canny edge outperforms it in several scenarios.

Interestingly, both Canny edge detection and SAM models correctly segmented images, yet both struggled with particularly challenging edges. This observation is noteworthy because general digital images present sharp edges (usually one or two pixels), whereas electron micrographs often display interfaces imaged through a few pixels with an intensity gradient. As a result, both Canny Edge and SAM identify interfaces as independent segments or as significant separations between segments, requiring post-processing to refine the output and unite regions without borders.

These differences are visible in the additional examples provided in figures 4.I.3 and 4.I.4, which compare the performance of the two segmentation models. As stated, the main differences arise when dealing with HRTEM pictures, where Canny edge totally misinterprets the segmentation. In STEM, the results are similar although SAM wins in detecting finer features like the Al shells of the hexagonal InAs nanowires in figure 4.I.4), or the twinned domains in the longitudinal cuts of the InSb-InP nanowires in figure 4.I.3. Although it is a small sample, we present the comparative metrics between Canny Edge and SAM in some of the samples and devices we have presented so far (tables 4.I.4 and 4.I.5). Note that any outlier might heavily affect the average in a small sample of samples like the one presented. Consequently, a case-by-case evaluation will be much more informative than figuring out the global picture with just the average. As said, figures 4.I.3 and 4.I.4 are self-explanatory and showcase the main strengths of each, where the major discordance is the HRTEM contrast and the smoothing of edges. Remarkably, pixelated pixels will not be a major issue since edge smoothing is always performed when building our posterior Finite Element Models (FEMs).

Nanowires	**Canny edge detection**	**SAM**
1	97.8	95.1
2	66.9	95.1
3	96.7	97
4	97.3	93.6
5	89.4	95.2
Mean	**89.6**	**95.2**

TABLE 4.I.4: Comparison of the accuracy, in well-segmented pixels %, of Canny edge detection and Segment Anything Model (SAM) algorithms applied to the reviewed nanowires. The mean accuracy for each algorithm is also presented.

Ge QW	Canny edge detection	SAM
1	93.4	94.0
1l	59.8	74.3
1c	53.7	94.1
1r	55.7	76.1
2a	92.3	92.6
2b	96.4	90.0
Mean	**71.6**	**86.9**

TABLE 4.I.5: Comparison of the accuracy, in well-segmented pixels %, of Canny edge detection and Segment Anything Model (SAM) algorithms applied to the reviewed Ge quantum well devices. The mean accuracy for each algorithm is also presented.

The tables show that SAM produces a higher score on more occasions than Canny edge, but not as much as expected from the abysmal architecture complexity difference between both. As a result, the final proposed model combines the Canny edge detection and the SAM algorithms. [71, 72] In fact, of the only two hyperparameters we can manually define in the whole proposed workflow, the choice between Canny Edge or SAM is one of them. Then, we can manually choose which one we use for the segmentation depending on the type of micrograph we work with. Generally speaking, the joint atomic mass and thickness contrast outputted by regular STEM provides a noteworthy difference in the intensities of the regions that should be segmented as different regions. Generally, then, if the atomic mass of neighbouring regions is different, the segmentation by contour detection is accurate (i.e., outputs the segments as a trained human would define them). It is important to mention that Canny edge and other edge detection algorithms (Sobel, Robert, Prewitt filters, thresholding, Marr-Hildreth) are powerful in general digital imaging, because their edges are typically drawn in just a few pixels (i.e., the transition in the intensities happens in one or two pixels). [73, 74, 75, 76]

This is not the case in electron micrographs, where the transitions can happen very gradually even if the interfaces are atomically sharp. As a result, a downscaling step is necessary to make this gradual transition be converted into single-pixel transitions that can be captured by the Canny edge detector. This downscaling is typically done by a factor of 8, for instance, reducing a 2k image to a 256-pixel image. This results in properly separating the edges of the image and its regions, but with an aliased segmentation due to the downsampling. This issue is not ideal but it is not problematic either. In the first steps of the workflow, the segmentation is only used to choose from where to extract the FFTs, for which a pixelated contour will not be detrimental. However, in the FEM and atomic model generation, the contours are important. Thus, in the generation of these models, we perform an additional step of smoothing the contours down as an antialiasing measure. It is based on reducing the total number of vertices defining the contour, which reduces the corners in the contour drawing.

These intensity differences are not as pronounced or evident in regular parallel beam TEM images. Consequently, our Canny edge automated model fails to perform effectively in this operation mode due to two main factors. First, a strong presence of transmitted but un-diffracted electrons affects the edge detection process. Second, the intensity distribution in these images differs significantly from that in STEM images. [9] Specifically, our Canny edge detection algorithm is manually optimized for STEM, particularly for HAADF imaging. However, it is not expected to perform well on regular High-Resolution TEM (HRTEM). Therefore, in this scenario, SAM is the recommended choice. Examples of the compared segmentation can be seen in figures 4.I.3 and 4.I.4. In general terms, SAM is more robust due to the enormous training stream employed, which despite not being domain-specific, not surprisingly generalises well with electron microscopy data. Its approach resembles the effort available so far in the bibliography in TEM segmentation, which mainly focuses on particle/cell counting and statistics extraction [77, 78, 79]. It accurately segments STEM data as well, although does not typically vary from Canny edge results. Its strong point though, is that does not require downsampling. Therefore, if the regular Canny edge loses too many contour details after the smoothed downsampling, it is recommended to try SAM in STEM. For instance, SAM wins in detecting finer features like the Al shells of the hexagonal InAs nanowires in figure 4.I.4), or the twinned domains in the longitudinal cuts of the InSb-InP nanowires in figure 4.I.3. The counterpart, though, is that it is significantly slower than Canny Edge. The computing time is sample-dependent, as the edge convergence needed to refine SAM is not always needed or equally expensive (i.e., more undefined contours, more time to converge them). Moreover, as it does not downsample, for big images the process becomes slower. As a result, the Canny edge detection can be calculated in about 10 seconds, while the SAM can be up to 100 times slower in complex segmented morphologies. Grossly, SAM is preferred for HRTEM, while for STEM, Canny Edge is the first choice unless finer morphological features are required.

Interestingly, the ultimate model for general segmentation could be merging both segmentation models into a single unique one. That is, to fine-tune SAM with the domain-specific labelled data as extracted by our current joint model with a manual correction of these labels to address particularly subtle features of interest if needed. [80] This would allow the incorporation in the trained model of the manual fine-tuning of the STEM-optimised Canny edge detection. In addition, the computationally expensive refinement of the wrongly categorised edges of SAM would also be integrated, which is the artefact caused by the difference in pixel thickness of the edges of general digital images and electron micrographs.

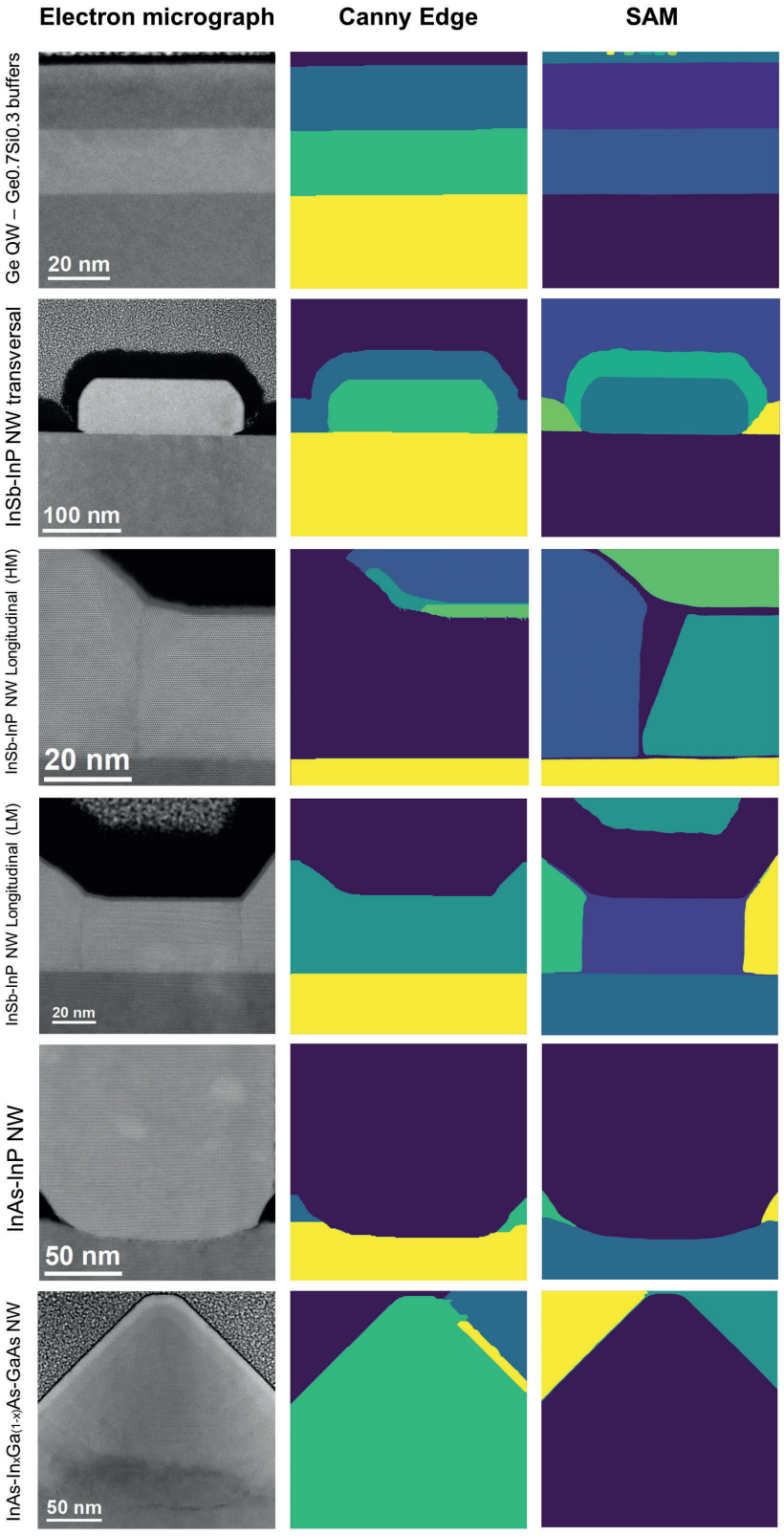


FIGURE 4.I.3: Comparative examples of the segmentation performed with Canny Edge vs SAM (Part 1). The examples presented here are scanning transmission electron microscopy high-angle annular dark field (STEM-HAADF) micrographs, for which our Canny edge detector is tailored. The results are comparable with both algorithms, although SAM can find subtle features like the twinned domains in the longitudinal cut of the InSb-InP nanowires, at the expense of more computing time. This is not necessarily a particular advantage as the posterior phase identification will notice the twins either way, but just a detail to keep in mind when choosing the segmentation approach. In addition, SAM smoothes down the contours by default, although pixelation is still present and needs to be smoothed further in both scenarios.

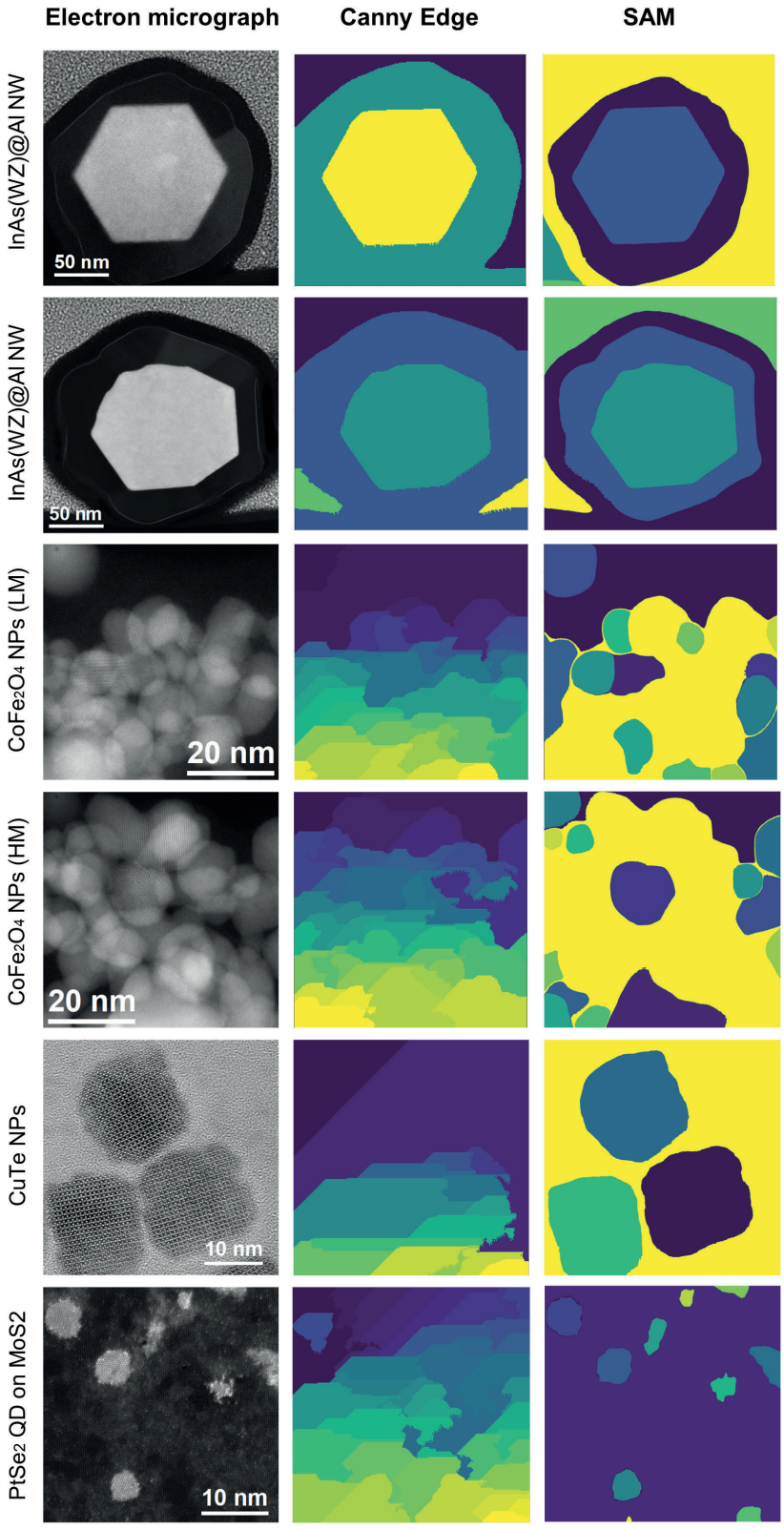


FIGURE 4.I.4: Comparative examples of the segmentation performed with Canny Edge vs SAM (Part 2). The examples presented here are scanning transmission electron microscopy high-angle annular dark field (STEM-HAADF) micrographs (two first rows) and parallel beam high-resolution (HR)-TEM (last four rows). It highlights the performance differences of the Canny Edge algorithm, working fine in STEM and completely confusing the HR-TEM intensities (specifically in the micrographs of the nanoparticles). In the latter case, SAM is always the best alternative.

4.I.2 Peak finding in the Fourier spectra

In our pursuit of developing solutions for a general segmentation routine, we encountered challenges when trying to create a peak-finding model applicable to any FFT derived from electron micrographs. Rather than relying solely on manual labelling of as many FFTs as possible to train a model, we aimed to gain deeper insights into the data's nature and requirements. As a result, we explored three different approaches. Surprisingly, none of these approaches stood out as the best individually, but their combination resulted in our proposed robust solution. In the following subsections, we describe each approach in detail and discuss how they can be effectively combined.

4.I.2.1 Machine Learning-based 2D Gaussian fitting

The core of this approach is the traditional 2D Gaussian fitting. It is a well-known method based on fitting Gaussian curves into the input image, to try to find all the patterns that fit to a Gaussian distribution, and not the ones that do not. Given the likelihood of the local distribution fitting to a Gaussian distribution, the coordinate (if the input is an image, the row and column of the pixel) expressing the maxima of the Gaussian curve fitted is outputted. This process, though, depends on a key hyperparameter which is the density of spots expected per area. It is tuned to address scenarios in which spots are expected closer to one another, and oppositely, cases where the spots to be found are relatively far away from their neighbours.

Complementarily, we can threshold the image to adjust the fitting even more to our needs by only fitting the region where the threshold is met. As a result, this turns the algorithm into a twofold hyperparameter algorithm. This is the starting point of the routine we have developed, which requires the definition of these two hyperparameters to output a result: peak density and threshold. [40, 81]

The goal was to automatically define the two hyperparameters with no supervision towards automating the algorithm to ensure an accurate peak finding with absolutely no human intervention. In other words, to optimise the hyperparameters given the information extractable from the micrograph. We correlated the two hyperparameters with the micrograph metadata to seek a tendency which we can model. Interestingly, we achieved this by correlating the optimal hyperparameters found manually with the pixel size of the image. This is justified because the pixel size defines the spacing of the spots in the reciprocal space when computing the FFT, or effectively, its areal density. For instance, a large pixel size refers to spots well spaced along the whole FFT, while a small pixel size typically means that all the frequencies corresponding to crystal planes are grouped in the very low-frequency circular domain of the FFT. This idea is captured in figure 4.I.5,c, in which we manually optimised the 2D Gaussian fitting to adjust its density of fitted spots based on the actual density of frequencies displayed. In this way, we gathered micrographs of

different natures and pixel sizes. We then manually optimised its two hyperparameters paying special attention to the density of spots, and finally, we trained a simple machine-learning model with it. As we did use a relatively small sample size, we used a simple regression model to avoid overfitting, as the expected tendency should be *a priori* simple enough not to demand complex models like neural networks. We used 42 image conditions (i.e., different pixel sizes, FOVs, number of pixels) to fit a polynomial curve of seventh degree that captures the nature of this correlation in the domain of FOVs used. This simple tendency is afterwards manually adjusted considering its performance with varying numbers of pixels of the analysed image, which is ultimately equivalent to the direct correlation with the pixel size.

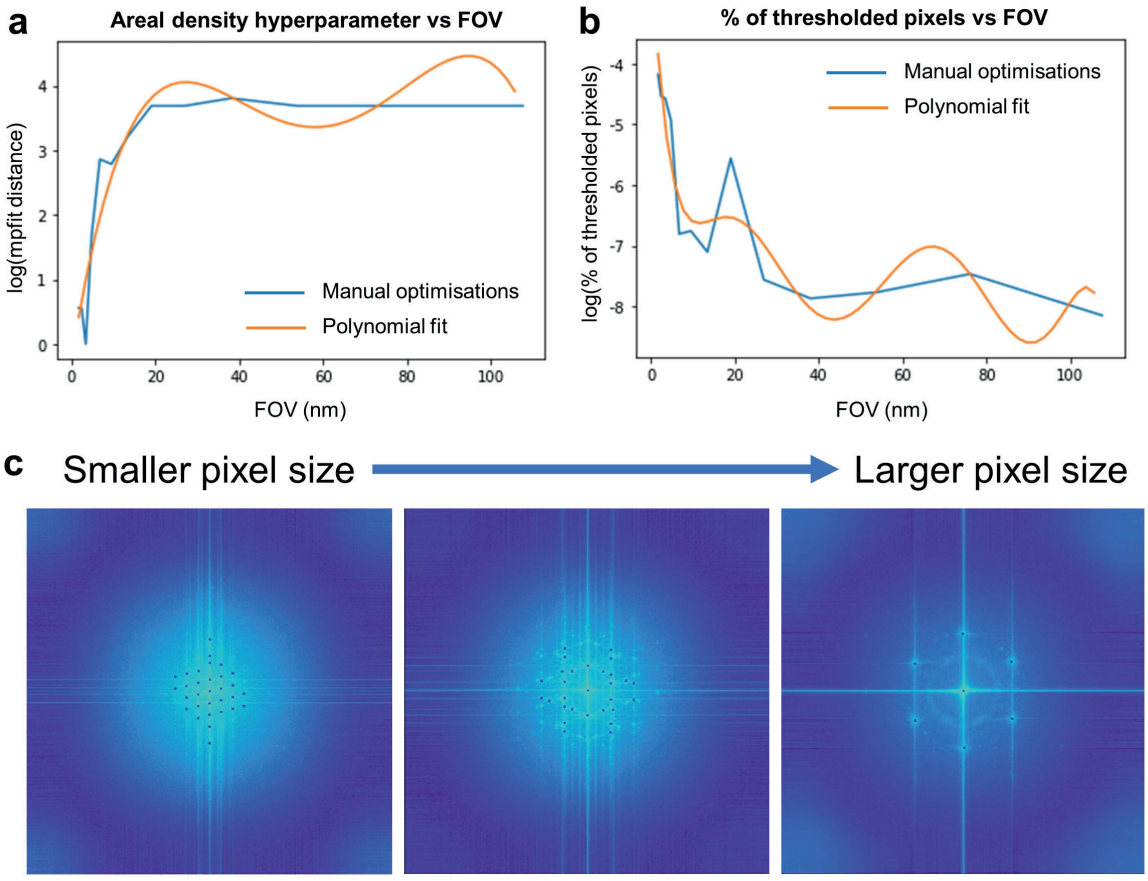


FIGURE 4.I.5: The two polynomial regressions correlating the two hyperparameters of the 2D Gaussian fitting, a) the areal density as a function of the Field Of View (FOV), and b) the percentage of pixels after thresholding as a function of FOV. The blue curves represent the manual optimisations grouped and averaged per FOV, and the orange line draws their (seventh-degree) polynomial fitting. c) Three Fast Fourier Transforms (FFT) representative of different expected areal densities of peaks (i.e., different pixel sizes) and how the 2D Gaussian fitting algorithm handles them. If the density is uniform, the autonomous tuning of the hyperparameters can output an almost perfect result (right and left FFT), while a heterostructure represented by a heterogeneous peak placing will have meaningful peaks kept (central FFT).

To prepare the regression, we keep the distance or areal density parameter unmodified, but we change the way of expressing the threshold. Thus, as the pixels of the

FFT holding the crystal planes frequencies are a fraction of the most intense ones, we transform the threshold parameter into the percentage of pixels, as the fitting parameter (i.e., the same threshold in images of different numbers of pixels would lead to a different percentage of pixels considered). As a result, we train a model in which our independent variable is the FOV read from the micrograph metadata, and the output is the areal density and the percentage of pixels to keep. To do so, we group the manual optimisations based on the FOV and fit the polynomial curve to it. This way, we end up with two polynomial regressions correlating the areal density as a function of the FOV, which can be seen in the figure 4.I.5,a, and the percentage of pixels after thresholding as a function of FOV in 4.I.5,b. Then, when an image is inputted, the FOV enters as a variable and the model outputs the two hyperparameters to be used in the 2D Gaussian fitting (i.e., the percentage is converted into real thresholding intensity).

Figure 4.I.5, c, showcases the main strong and weak points of this approach. Its robustness when facing a uniform expected density can be seen in the right and left FFTs when properly adjusting the hyperparameters. However, challenges arise when dealing with heterostructures or double diffractions that can break the homogeneous symmetry of well-placed spots, typical of perfect crystals. In these cases, as the density is not homogeneous, the algorithm will either skip contiguous spots if the density is low, or add noise if a higher density is selected instead. These characteristics of the method need to be taken into account and will be important when discussing the global model proposed in the section 4.I.2.4.

4.I.2.2 Encoder-decoder convolutional neural network

To create a peak detector in FFT spectra, we employed a CNN with a U-net architecture (encoder-decoder). [21, 32] U-net, specifically designed for image segmentation tasks, plays a crucial role in improving the precision of peak detection within the frequency domain of electron microscopy images. This choice is justified in the Methodology chapter. For training, we utilized over 5000 simulated spectra. These simulations rely on the kinematic diffraction theory to generate 2D diffraction patterns corresponding to the given model crystal phase. [9] The underlying philosophy behind the training process and the desired results is illustrated in figure 4.I.6, where the CNN is trained to identify the positions of the frequency maxima. Let's explore the subsequent pipeline and its results.

To ensure the network could learn and generalize complex patterns associated with peak features, our training dataset encompasses distinctive crystal phases oriented along different directions. We focus on six distinct crystalline phases: Ge (diamond cubic), GaAs (wurtzite), InP (zinc-blende), InSb (zinc-blende), Si (diamond cubic) and ZnO (wurtzite). These specific crystal phases were selected due to their comparable but still differentiable atomic geometries, providing a comprehensive representation of structural variations. The dataset encompasses crystal orientations along

crystallographic axes (i.e., [001], [100], [011], [110] and [111]) allowing the model to learn and generalize peak features across diverse directions within the crystal lattice. Furthermore, the dataset incorporates three different pixel sizes: 0.015 nm/pixel, 0.03 nm/pixel, and 0.05 nm/pixel. We work then with 90 different base FFT types. The partial equivalence between zone axes and phases combined with the variation in pixel sizes push towards the robustness of the trained model, accommodating variability of representative frequency peaks commonly encountered in experimental FFTs.

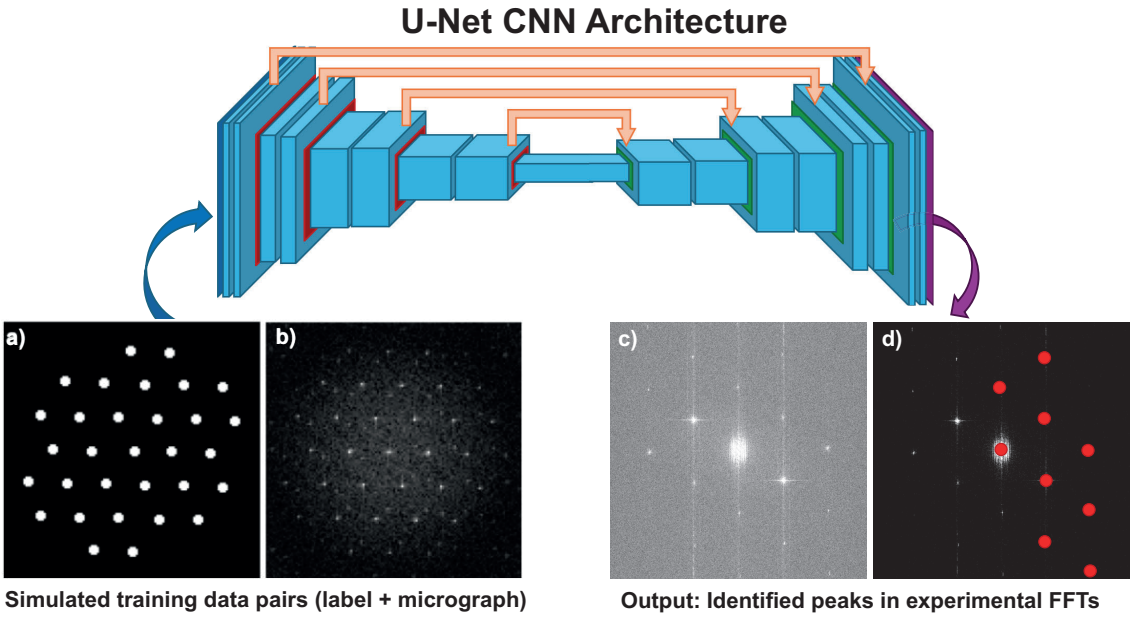


FIGURE 4.I.6: Scheme on the peak finding model training. Given the encoder-decoder architecture of the U-Net network, we input a training set with pairs of kinematically simulated b) Fast Fourier Transforms (FFTs) and their respective a) labels (multipixel peaks). [32] After its training with 5000 data pairs, we obtain a model capable of isolating the frequency peaks from the FFT (c). (d) Postprocessing unlocks retrieving the very centre of the peaks corresponding to crystallographic planes.

This combination of parameters and conditions can be viewed as a pathway for spreading intensity peaks across a background. However, these peaks follow characteristic and varied patterns dictated by the specified zone axes. Notably, each simulated element serves as a background with null intensity, featuring pixel-wise squared intensity peaks (fig. 4.I.7, a). To enhance the resemblance to experimental data and improve subsequent training performance of the simulated FFT spectra, we transform these resulting binary matrices into realistic peak intensity distributions. To achieve this, we introduce cumulative noise layers using the following stepwise approach (fig. 4.I.7, top, a)-e), and table 4.I.6):

1. **Drift Transformation:** Initially, a drift transformation is applied to the theoretical FFT spectrum (fig. 4.I.7, a). Although it is not strictly speaking noise per se but a systematic effect, we fit it within noise as it modifies the ideal

output. This drift layer displaces frequencies in a horizontal distance, followed by a displacement slope s, and a rotation at an angle θ. These values are centred at 0 to enable drift in any direction. Subsequently, the inverse FFT (IFFT) is computed to return to the real space, resulting in an atomic image with drift. The benefit of this layer is twofold. First, we add variability to the input data to more extensively sample the infinite range of experimental possibilities within FFTs. Moreover, we ensure the NN does not learn perfect symmetric Fourier spectra. Thus, the idea of having a general and case-flexible peak-finding model would clash with its learning of fixed symmetrical patterns. Although the approach could be potentially useful for other purposes, it would be detrimental in this case. The drift breaks the perfect symmetry by unevenly displacing certain intensity peaks along a displacement vector, telling the model that symmetry is an important constraint to learn but not the only one.

2. **Gaussian Atom Transformation:** The punctual atoms in the drifted image are transformed into bidimensional atoms using a Gaussian filter. The size of these atoms is regulated by the σ value of the Gaussian function.

3. **Noise Filters on Gaussian Atoms:** Various noise filters are applied to the Gaussian atoms in this step. These include scanning noise modulated by Gaussian distributions with their respective hyperparameters (σ_h and σ_v) and Poisson noise modulated by its intensity λ hyperparameter (fig. 4.I.7, b-d). Essentially, they modify the until now perfect Gaussian intensity distribution of the atoms into a more realistic representation of pixel-wise intensities.

4. **Noisy FFT Spectrum Calculation:** Finally, the FFT of the noisy atom image is computed to obtain the simulated FFT spectrum with added noise (fig. 4.I.7, e).

Variables	μ	σ'	**Interval (99.9%)**
s **(adimensional)**	0	0.06	(-0.02, 0.02)
θ **(º)**	0	0.91	(-2, 2)
σ **(pixels)**	1.5	0.15	(1, 2)
σ_h **(pixels)**	0.8	0.06	(0.6, 1)
σ_v **(pixels)**	0.6	0.06	(0.4, 0.8)
λ **(intensity, a.u.)**	30	6.06	(10, 50)

TABLE 4.I.6: Range of values when adding the noise layers in the synthetic data preparation for the Convolutional Neural Network model for peak finding. Every value follows a normal distribution N(μ, σ'), and the interval embraces 99.9% of the distribution.

By incorporating these noise layers, we aim to emulate the complexities present in experimental data, thereby enhancing the robustness and generalization capabilities of the trained neural network model. This approach ensures that the simulated data

closely aligns with real-world scenarios, contributing to the model's ability to handle diverse and noisy experimental FFT images.

To train the U-NET neural network, we utilize the AtomAI Python package. It provides a convenient solution to use U-Net with its prebuilt "atomai.models.Segmentator" class. We employ the cross-entropy loss function, as the standard for categorical problems such as the one occupying us. Consequently, the last layer is a softmax converging to either 1 or 0, meaning having or not a peak in that pixel, respectively. The neural network undergoes training for 500 epochs with a batch size of 32 images, and a learning rate of 0.001. Additionally, we enable the stochastic weight averaging parameter to enhance the method's accuracy.

To enhance the training accuracy of these models and improve the detection of the peaks, we adjust the point labels corresponding to frequency locations to small circles. This adjustment instructs the model that pixels near the frequencies should also be classified as peaks. Therefore, the pixels contributing positively to the training loss and consequently, its accuracy, are increased. In other words, the single-pixel label puts the whole pressure of a positive identification on just a single pixel. If the image is small, around 128 or 64 pixels per side, one pixel is a distinguishable unit in the whole matrix. However, if the image is sampled according to the typical rates of microscopy (i.e., 1K to 4K), a single pixel is a negligible unit. Therefore, the impact on the loss of correctly classifying a single pixel as a peak, 1, is negligible when compared to correctly classifying the rest as background. Interstingly, we can revert this trend with larger labels indicating that a peak is a distribution rather than just one single pixel. As a result, the weight of a correctly identified peak in the total loss is not negligible anymore. The bottom panels of figure 4.I.7 compare the training curves, both the loss and the precision, when training with punctual and circular labels. At first glance, their comparable final performance metrics blindly mean that both are outstanding models. The trick lies in whether these models excel at their intended tasks. Surprisingly, the single-pixel labelling engineers an exceptional model for identifying null backgrounds, even though that was not the original intention. Equally intriguing is the perfectly smooth loss curve observed in that model, which raises suspicions of unexpected and unintended behaviour. On the other hand, the oscillating pattern observed during training with circular labels suggests that the model adapts to the data it encounters epoch after epoch. The definitive confirmation occurs during the validation process, where only measurable peaks are obtained in the model originating from the broad labels (fig. 4.I.8).

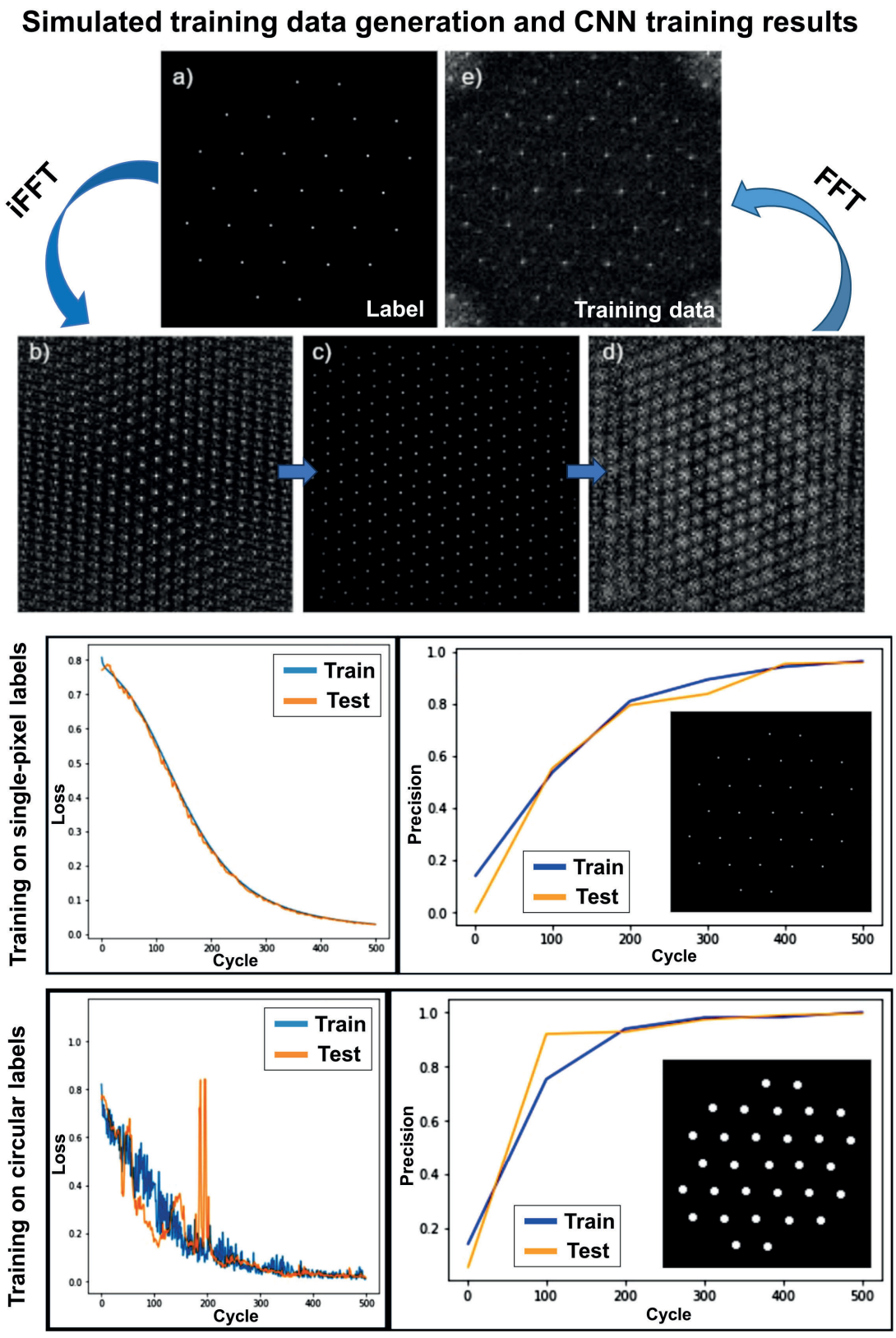


FIGURE 4.I.7: Generation of training data for the supervised peak finding model based on U-Net, and its training results. Top half: The (a) original kinematic diffraction pattern is (b) inversely Fast Fourier Transformed (iFFT) to (c) locate the position of the atoms. The noise filters are applied to generate (d) realistic atoms, whose (e) FFT is used to train the U-Net. Bottom half: Training of the network with 5000 simulated FFTs. The training using the circular labels makes the loss oscillate much more than the one with single-pixel labels. However, the achieved precision, both in the test and training set is superior to the circular one, making it our preferred model to tackle experimental FFTs.

As seen in figure 4.I.8, it is noteworthy that the implemented neural network is not directly tailored for peak detection in FFTs. Instead, its primary function lies in denoising the images and enhancing the quality and interpretability of signals in the frequency domain. This denoising approach significantly facilitates the subsequent detection of peaks through a post-processing step. Thus, the post-processing consists of the detection of local maxima, as detailed in the next section.

4.I.2.3 1D profile scanning

The primary challenge in detecting peaks in denoised FFTs lies in the presence of straight lines of noise surrounding the most intense frequencies generated during the scanning process in the acquisition phase. The 2D peaks detector performed poorly in these spectra as it identified numerous noisy points within these noisy straight lines. To address this issue, we proposed the development of an algorithm utilizing a 1D peak detector.

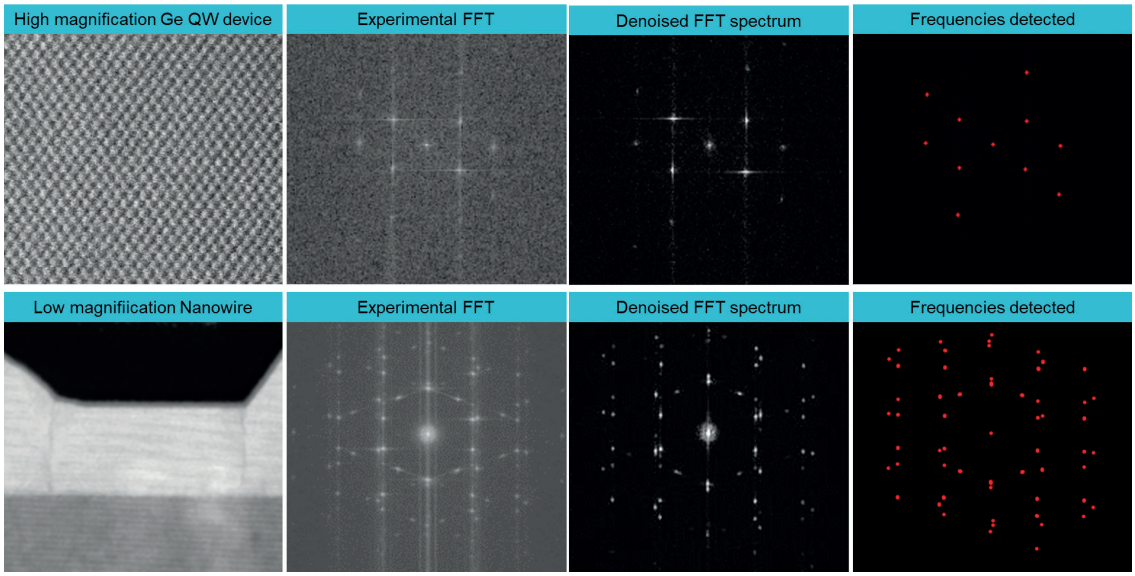


FIGURE 4.I.8: Convolutional Neural Network (CNN), U-Net-based peak finding model performance in FFT spectra. Atomic resolution region of a Germanium quantum well (Ge QW) heterostructure for 2D hole gas (2DHG) devices (top) and an InP-InSb nanowire (bottom) are used to test the peak finding model in high and low magnification conditions and in monocrystalline and polycrystalline materials. From left to right, the different steps to find the peaks in the FFT are listed: compute the FFT spectra, denoise these spectra using the peak finding model, and detect the final frequencies using the 1D scanning post-processing algorithm.

This algorithm operates by scanning the spectra both vertically and horizontally to detect the various peaks within the image. By retaining only the peaks detected in both scans, we mitigate the impact of noise. For the 1D peak detection, we employed a method provided in the SciPy package. [82] This method enables the detection of

peaks in a 1D array while allowing adjustment of hyperparameters such as minimum height, minimum distance between peaks, or minimum peak thickness. [82]

Addressing this problem in a one-dimensional manner yields notable benefits. When scanning the spectrum horizontally, the vertical noise lines are overshadowed by the central frequency that generates them, resulting in the algorithm primarily detecting this frequency. Similarly, when conducting a vertical scan, the horizontal noise lines are obscured by their central frequency. Ultimately, by solely considering the frequencies detected in both scans, we effectively filter out noise stemming from these artefacts. In figure 4.I.8, we present the outcomes of this post-processing algorithm for both high and low-magnified samples.

4.I.2.4 Global combined peak finding model

After developing the three described peak-finding solutions, we realised that none of them outperformed the others consistently in all scenarios. Whether dealing with amorphous regions, perfect crystals, polycrystalline segments, or equivalent FFTs obtained for twinned domains, no single solution stood out as universally superior. Surprisingly, the task of distinguishing between an FFT corresponding to an amorphous region versus one from an actual crystal proved more challenging than foreseen. In other words, when peaks are detected in an FFT, can we confidently assert that these are genuinely identified spots, or could they be mere noise occurring in an amorphous region where peaks should not exist?

The solution to this issue emerged through the combination of the three methods within our proposed final model. To the best of our knowledge, this joint peak-finding model stands as the most robust FFT peak-finding approach documented in the literature, effectively eliminating the need for human intervention. The model integrates the machine learning-based 2D Gaussian fitting (referred to as 2DGML), the encoder-decoder convolutional neural network trained on kinematical diffraction patterns (referred to as CNN), and the 1D Scanner (referred to as 1DScan). These components work sequentially to optimise performance based on the specific characteristics of the analyzed device and region.

The central idea is to first unveil whether the FFT corresponds to an amorphous or crystalline region, and if crystalline, optimise the output minimising the likelihood of adding false positives because of noise. Figure 4.I.9 shows the algorithm, its steps and crosschecks, to get to the final output of the identified spots in the FFT.

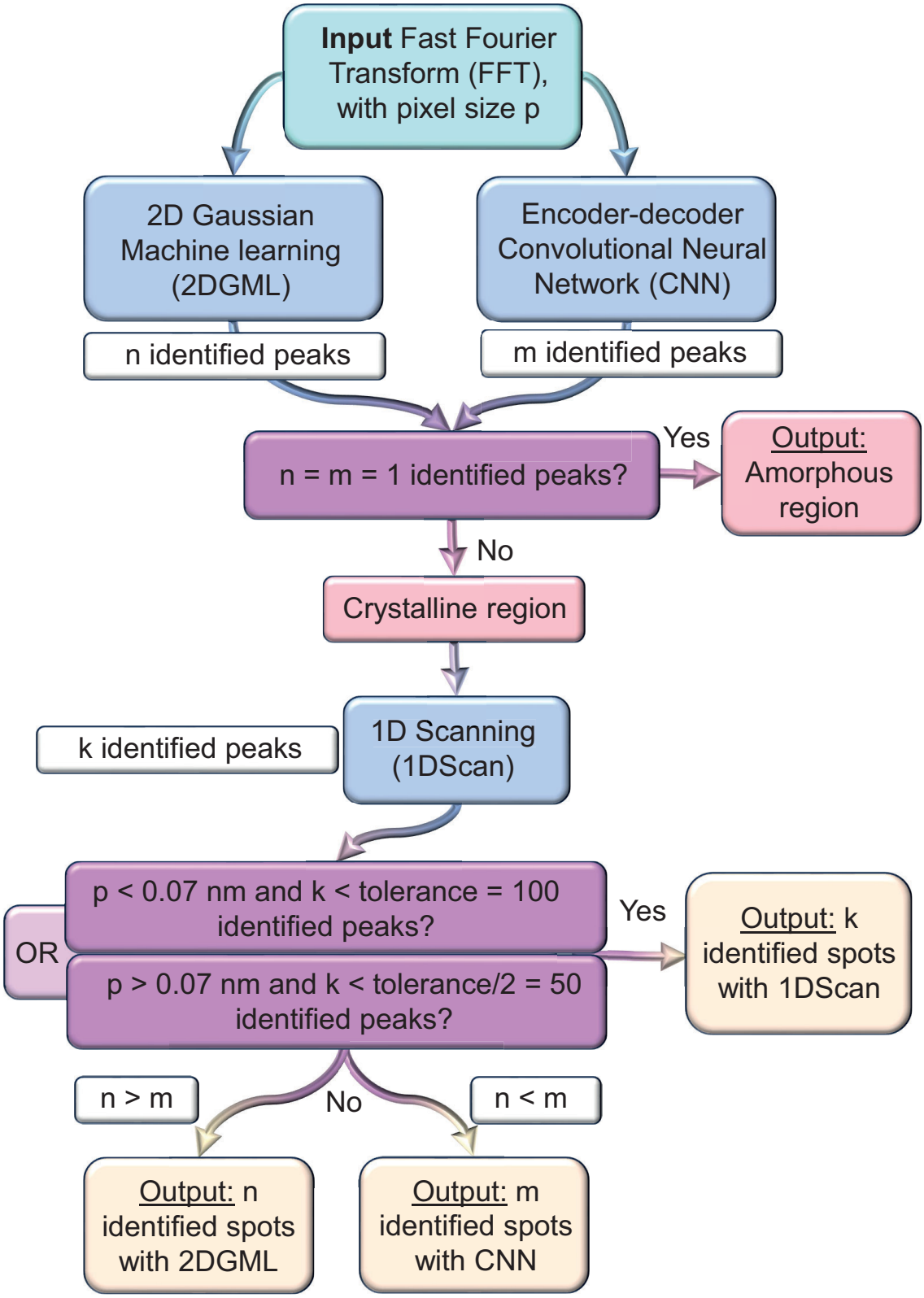


FIGURE 4.I.9: Flowchart of the global peak finding model, displaying the sequential application of the three methods: the machine learning-based 2D Gaussian fitting (2DGML), the encoder-decoder convolutional neural network (CNN) and the 1D Scanner models. We unveil whether the region is amorphous or crystalline, and if crystalline, we optimise the output based on the situations in which each model performs better.

Both 2DGML and CNN are accurate in detecting amorphous FFTs when they are actually amorphous. Thus, if only the central spot corresponding to the transmitted beam is present and no diffracted crystallographic planes are visible, only a single central spot is returned. On the other hand, in this situation, the 1DScan method typically explodes outputting an enormous number of random spots. This is, the whole background is filled with identified but unwanted spots. That is because its criteria to decide whether a signal fragment is a peak or not depends on the overall levels of signal, and if no real peaks appear, then the tolerance for peak accepting is reduced and the noise is understood as real peaks.

Therefore, if the two methods agree on just detecting the single central spot, the algorithm considers there are no actual spots to be detected. Consequently, the region must be amorphous. Oppositely, if this condition is not met, then the 1DScan is computed, as for crystalline FFTs it is the most accurate and flexible method of the three. Nevertheless, we impose a security check to avoid the unwanted situation of 1DScan still exploding and finding thousands of false positives. Although it is a minority, this issue can still occur, especially in small pixel sizes with low resolution, where just a handful of spots are localised in the centre of the FFT and the rest of the high-frequency domain perfectly resembles an amorphous FFT. This security check is based on only accepting 1DScan output if the total number of identified spots is below a defined tolerance, which is pixel size-dependent. For bigger pixel sizes (i.e., larger than 0.07 nm), as fewer spots are expected to be found, the tolerance is reduced (i.e., less than 50 spots). On the other hand, for smaller pixel sizes (i.e., smaller than 0.07 nm), where a higher density of spots can populate the FFT, the maximum number of spots accepted increases (i.e., less than 100 spots). These tolerance values are set because it is extremely rare to find FFTs in which a human would count up to 100 meaningful crystal plane frequencies. As a result, if fewer spots are found, the output of the global model is that of the 1DSCan.

Contrarily, if more spots are found, the most likely scenario is that the 1DScan algorithm has exploded in unwanted false positives, for which we must accept the output of either of the two remaining methods. In this case, we will simply accept the algorithm that outputted more spots in the first step of the flowchart. That is because both are robust to noise, but often not sensitive enough to clear spots. That is why generally speaking, in 2DGML and CNN, more identified spots imply a better result. From these two, 2DGML generally outputs more spots, whereas CNN, due to the way it is trained, gives more confidence to every detected spot at the expense of not being as sensitive. Examples of the combined global peak finding in action with a wide variety of FFTs are present in figures 4.I.10, 4.I.11 and 4.I.12.

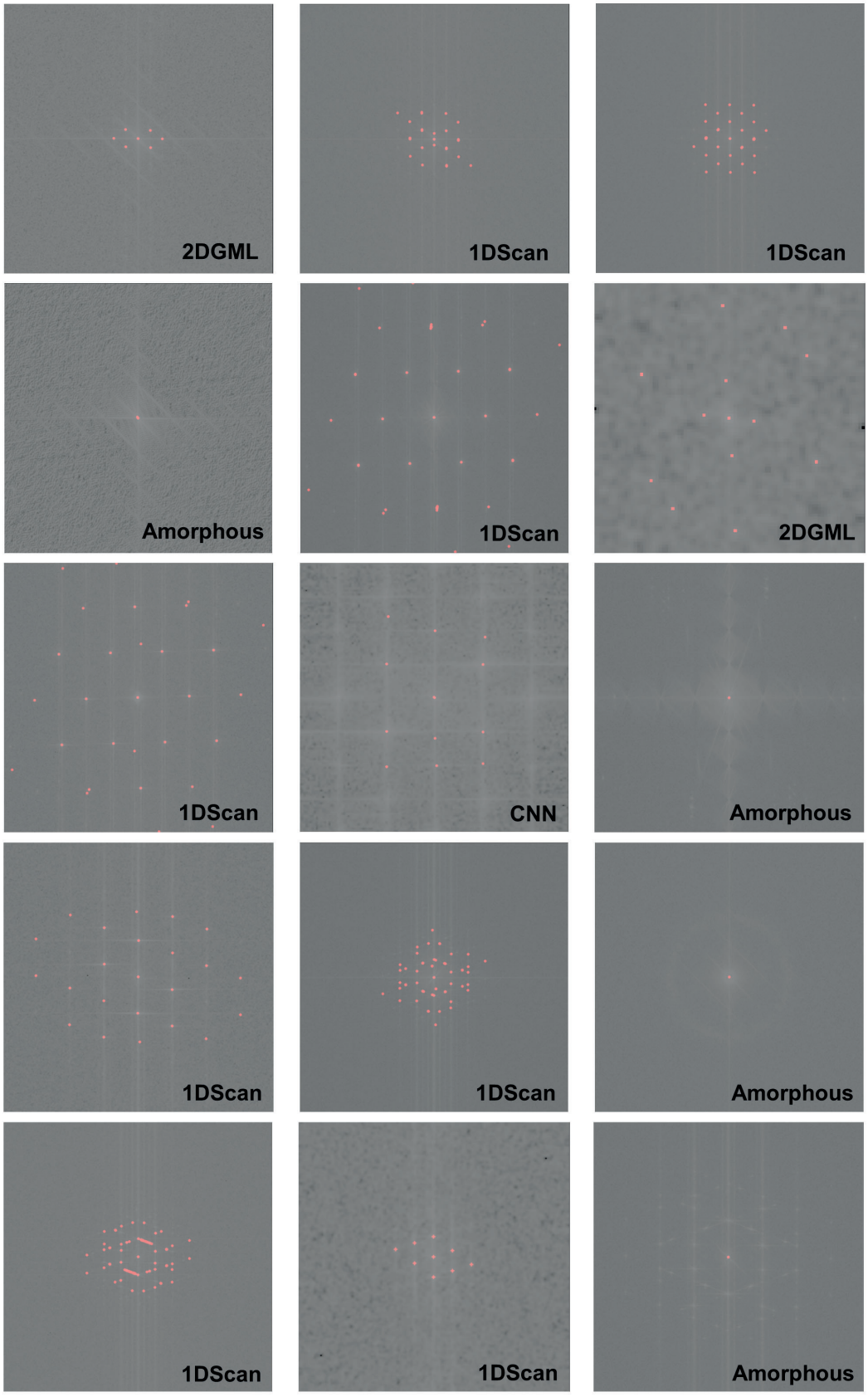


FIGURE 4.I.10: Part 1 of examples of the performance of the global peak identification model in a wide variety of Fast Fourier Transforms (FFT), extracted from different electron micrographs of different total size, pixel size, field of view, segmentation-based masking, as well as techniques (both parallel beam Transmission Electron Microscopy (TEM) and Scanning TEM (STEM).

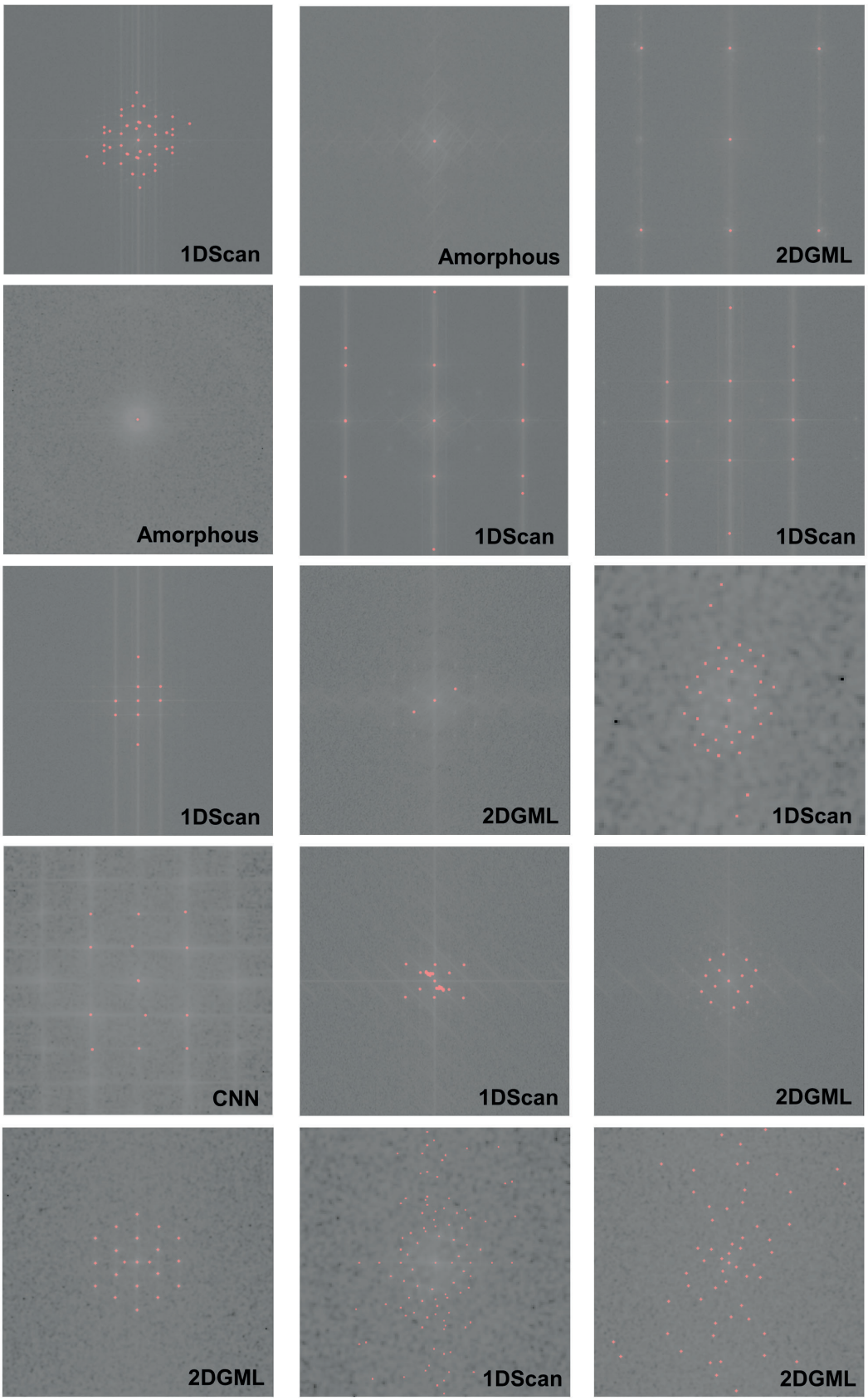


FIGURE 4.I.11: Part 2 of examples of the performance of the global peak identification model in a wide variety of Fast Fourier Transforms (FFT), extracted from different electron micrographs of different total size, pixel size, field of view, segmentation-based masking, as well as techniques (both parallel beam Transmission Electron Microscopy (TEM) and Scanning TEM (STEM).

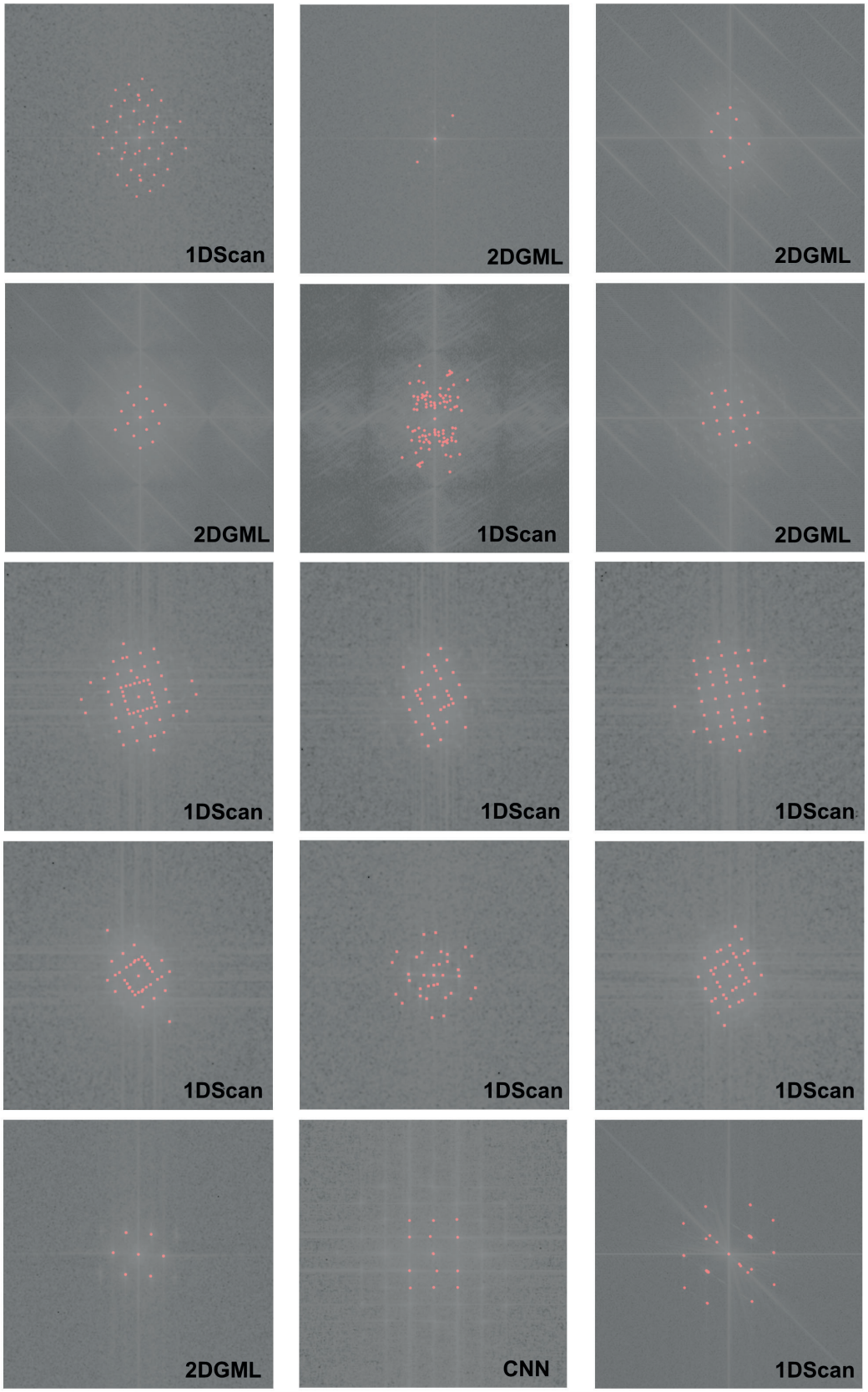


FIGURE 4.I.12: Part 3 of examples of the performance of the global peak identification model in a wide variety of Fast Fourier Transforms (FFTs), extracted from different electron micrographs of different total size, pixel size, field of view, segmentation-based masking, as well as techniques (both parallel beam Transmission Electron Microscopy (TEM) and Scanning TEM (STEM).

4.I.2.5 Peak finding performance metrics

We have manually labelled up to 1000 FFTs of different samples, crop sizes, and FOVs to obtain a metric for the performance of each of the three methods and the overall model. The manual labelling was conducted using a custom interface that enabled fast and precise labelling of the spots by automatically locating the maximum intensity within a specified window range.

For each of these methods, we calculate the predictions for the 1000 FFTs and compare them with the labelled ones. We classify the different peaks into three classes:

1. True Positives (TP): Peaks present in both images, the labelled and predicted ones. These are the correctly detected peaks.
2. False Positives (FP): Peaks present in the prediction but not in the labelled data. These are the incorrectly detected peaks.
3. False Negatives (FN): Peaks present in the labeled data but not in the prediction. These are the undetected peaks.

We utilize precision (equation 4.I.4), recall (equation 4.I.5), and F1 (equation 4.I.6) score as metrics. Precision indicates how many of the detected peaks are correctly identified, recall signifies the completeness, indicating how many of the actual peaks are detected, and F1 score is the harmonic mean of these two values.

$$Precision = \frac{TP}{TP + FP} \tag{4.I.4}$$

$$Recall = \frac{TP}{TP + FN} \tag{4.I.5}$$

$$F1 = 2\frac{Precision \times Recall}{Precision + Recall} \tag{4.I.6}$$

The metrics we get for each peak-finding approach are gathered in the tables A.1, A.3, A.5 and A.7 for the full labelled dataset, and tables A.2, A.4, A.6 and A.8 for the elements which provide a precision higher than 10% (complete tables in appendix A.1). Tables A.2, A.4, A.6 and A.8 highlight the significant increase in the three metrics when removing just the worst performing 10%, which is mostly attributed to abnormal detections like single peak (i.e., (000) reflection) detection in crystalline samples or noise explosions in amorphous regions. Examples of the computation of these metrics in FFTs that are representative of a certain crystallinity scenario can be seen in figures 4.I.14 and 4.I.15.

To fully explain the tabulated values, let us recap. *Recall* informs us about how many spots are found out of the total that should be found. On the other hand, precision describes how many spots are actual spots from all the found ones. In

other words, if the method adds noise. Therefore, detecting all spots in unessential for the recall metric. Interestingly, F1 averages both metrics giving a global sense of the performance of the algorithm.

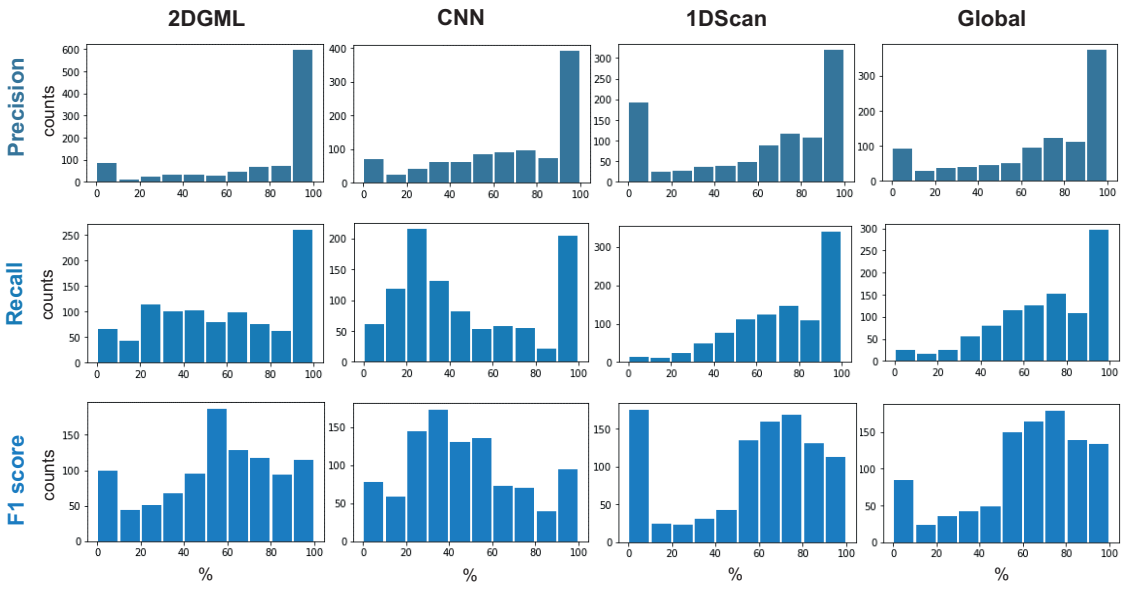


FIGURE 4.I.13: Histogram collection displaying, for every peak finding method, the distribution of the precision, recall and F1 metrics.

With these ideas in mind, we can agree that 2DGML is the most precise method, while the 1DScan has the best recall, and the global method has the best F1 score. Accordingly, we can actively rely on 2DGML if the absence of noise is a must. However, the global method will be the choice if we want the majority of points to be detected, even at the expense of extra noise. Importantly, CNN stands out in precision and 1DScan is the source of the high recall of the global method. This is important to notice because this combination of two high-precision methods and a high-recall one creates the robust global one we use in our workflow. Although the values in the tables 4.I.7, 4.I.8 and 4.I.9 and histograms in figure 4.I.13 help in understanding this trend, they do not encompass the full picture behind the major benefit of mixing the methods. For instance, one of the major achievements of the joint global peak finding model is the robustness and confidence in distinguishing between a crystalline and an amorphous model (figure 4.I.15, left column). However, this information is diluted within the full labelled dataset, which includes manifold sample types. Furthermore, the second major achievement of the global method is the capability of prioritizing high recall when crystallinity is found, also keeping a notable precision.

Essentially, as displayed in figure 4.I.9, the 1DScan (i.e., high-recall method) is the priority output unless a peak outbreak is detected (i.e., do not exceed the maximum number of allowed peaks per pixel size). This scenario is met in figure 4.I.15, in the middle column, where the CNN is the global output despite the higher recall of

1DScan that is discarded because of adding too much noise. If this scenario is met, either 2DGML or CNN is selected based on the one that generates more spots (as both stand out in precision, more spots will generally score better).

Precision	2DGML	CNN	1DScan	Global
Counts	1000	1000	1000	1000
Mean	78.33	69.03	62.33	69.78
St. dev.	32.63	31.98	37.06	32.5
Min.	0	0	0	0
25%	64.12	46.67	33.33	51.29
50%	100	76.92	73.91	78.52
75%	100	100	100	100
Max.	100	100	100	100

TABLE 4.I.7: Precision comparison over all the peak finding methods.

Recall	2DGML	CNN	1DScan	Global
Counts	1000	1000	1000	1000
Mean	59.13	48.55	73.87	70.89
St. dev.	31.61	32.19	24.42	25.84
Min.	0	0	0	0
25%	33.33	21.88	56.42	53.85
50%	57.89	36.84	76.47	73.68
75%	92.45	73.48	100	100
Max.	100	100	100	100

TABLE 4.I.8: Recall comparison over all the peak finding methods.

F1	2DGML	CNN	1DScan	Global
Counts	1000	1000	1000	1000
Mean	55.16	46.42	56.25	61.87
St. dev.	28.38	26.85	31.55	27.51
Min.	0	0	0	0
25%	37.5	28.57	38.78	50
50%	58.2	42.86	64.71	67.86
75%	76	63.64	78.95	81.25
Max.	100	100	100	100

TABLE 4.I.9: F1 score comparison over all the peak finding methods.

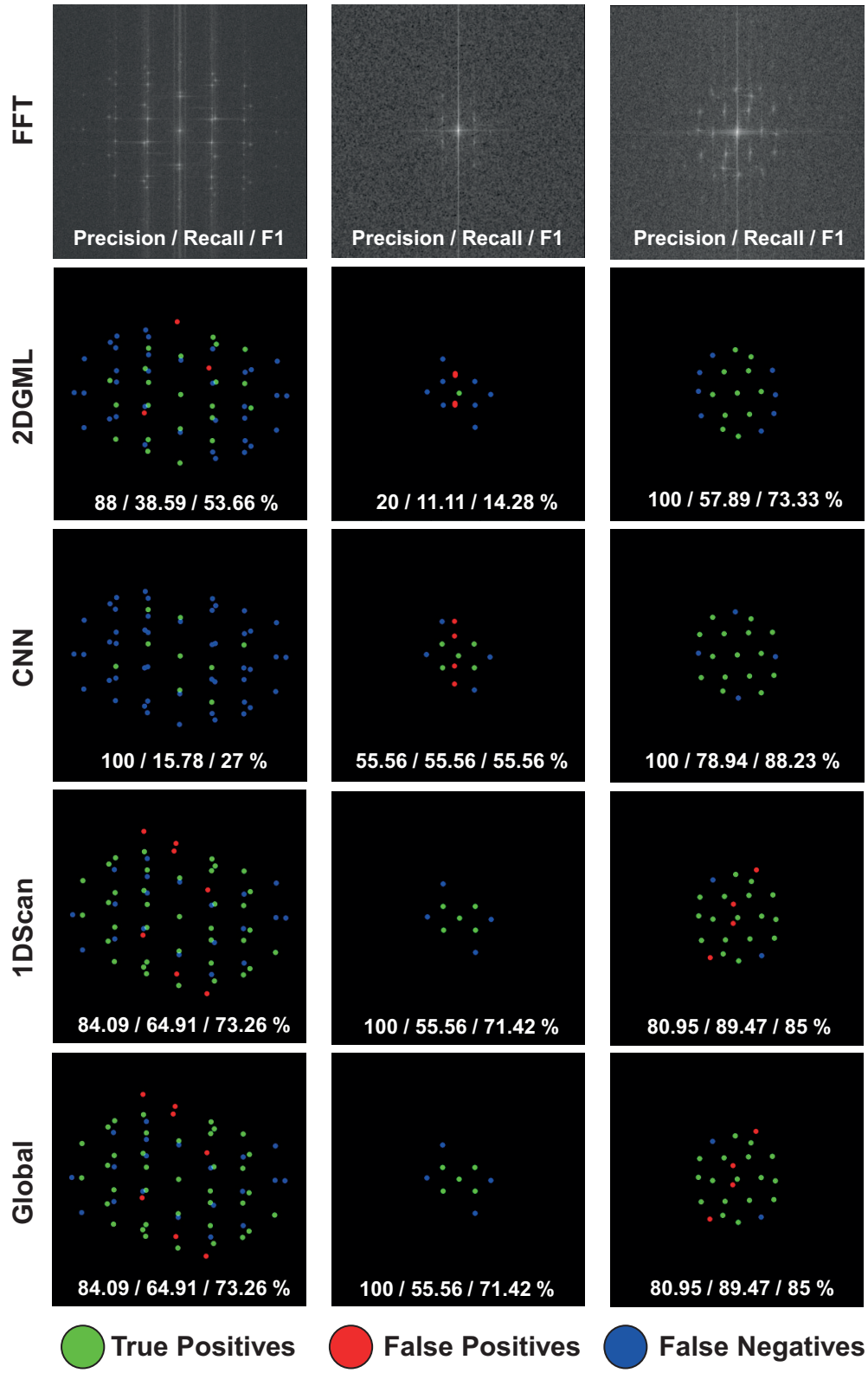


FIGURE 4.I.14: Part 1 of representative examples on the assessment of the three used performance metrics of all the peak finding methods on the labelled dataset. The colour scheme of the spots indicates true positives in green (i.e., correctly predicted spots), false negatives in blue (i.e., actual peaks not identified) and false positives in red (i.e., identified peaks that are not actual peaks but noise).

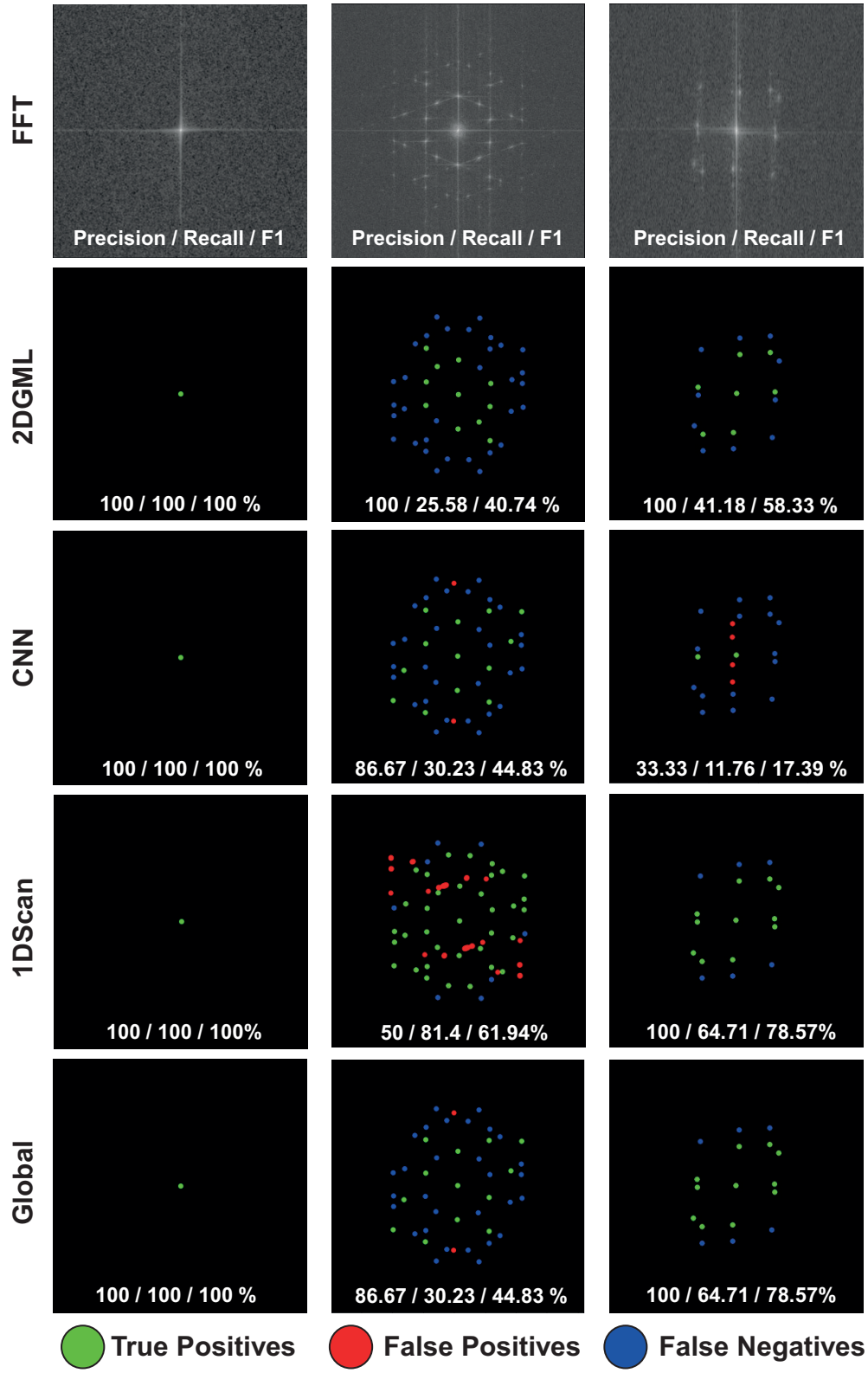


FIGURE 4.I.15: Part 2 of representative examples on the assessment of the three used performance metrics of all the peak finding methods on the labelled dataset. The colour scheme of the spots indicates true positives in green (i.e., correctly predicted spots), false negatives in blue (i.e., actual peaks not identified) and false positives in red (i.e., identified peaks that are not actual peaks but noise).

The high recall of 1DScan enables the detection of close by spots corresponding to polycrystalline samples (figure 4.I.14 left column, and 4.I.15 middle and right columns). This is relevant because of the nature of the phase identification algorithm. A high recall is preferred over a high precision because if multiple spots from multiple phases are found, then all the corresponding crystal phases have a chance of being identified. In fact, despite finding multiple phases, the workflow is built to work with the best-fitting one, which generally points to the most abundant one (i.e., the broadest area over a region of interest). As a result, finding additional phases is only relevant if checking more than one phase is intended. For that, noise is not as harmful as could be with a less robust phase identification step. Similarly, suppose the sample is single crystalline and noise is added. In that case, a major match with significantly higher goodness of fit will appear for the main phase that groups its corresponding spots. Additionally, secondary negligible phases will follow this result pointing to spare spots not reproducibly interconnected along a common zone axis, yet with inaccurate goodness of fit. In conclusion, the take-home message is the global robustness of the whole peak finding plus phase identification method, reached by building them individually to mutually reinforce the strengths of the other and counteract their weaknesses.

4.I.3 Segmentation - Peak finding interconnection

The overall robustness of our workflow stems from its design as a cohesive whole, surpassing the mere sum of its individual modules. To achieve this synergy, we meticulously address communication between these modules, leveraging their complementary strengths to compensate for each other's weaknesses. In that sense, the interplay between segmentation and peak finding plays a pivotal role in shaping our phase-finding module. This cross-communication occurs through two distinct mechanisms, both of which being controlled by the only two eligible hyperparameters within the entire process, alongside the choice of the segmentation algorithm.

The tunability of our approach determines from where the FFT is extracted in the image, and therefore which crystallographic information is accessible. We can either segment the image and find the biggest square which we can crop inside each region, or mask the image to only receive signal from the pixels where the mask is applied.

The "crop" setting identifies the largest square within each masked region using dynamic programming. Subsequently, we compute the FFT of this cropped square to extract crystallographic information from the corresponding image region. Assuming that our segmented region is crystalline, this approach achieves two critical goals: First, by maximising the size of the crop, we enhance the intensity of the crystalline frequencies. Larger crops encompass more relevant signal associated with each crystal plane. Second, during phase identification, transforming the coordinates of identified peaks into distances and angles becomes more precise. The higher

the number of pixels within the crop (i.e., a higher sampling rate), the finer the accuracy of the measurement. This precision is advantageous in terms of signal-to-noise ratio. If the original sampling of the crystal planes exceeds the Nyquist limit, the FFT unambiguously depicts the crystalline frequencies, increasing the likelihood of proper identification by the global peak finding algorithm.

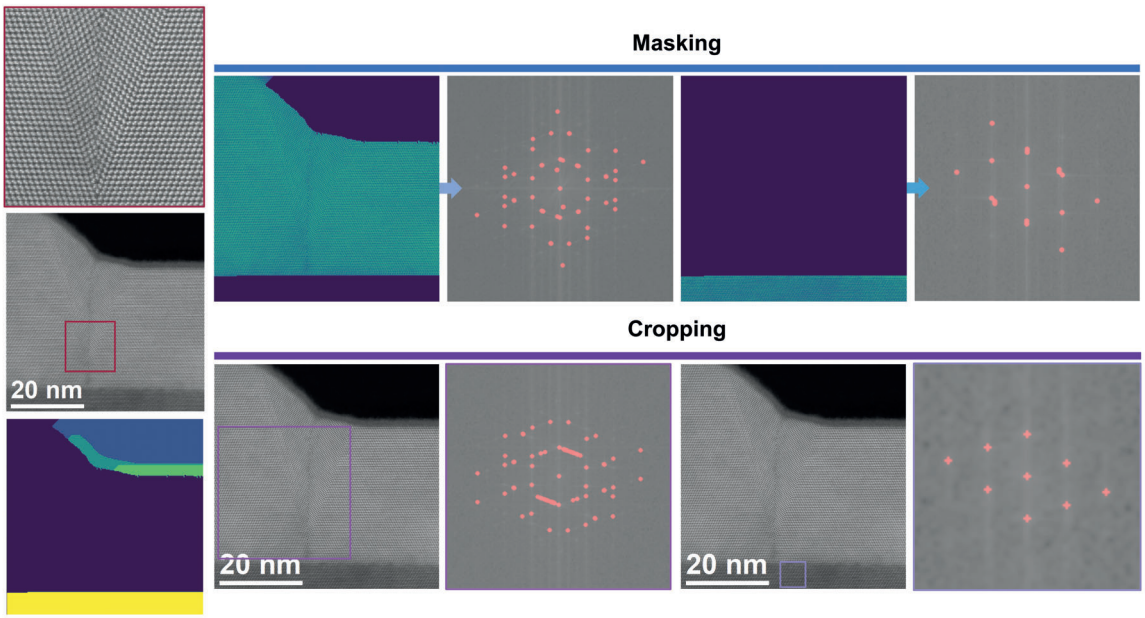


FIGURE 4.I.16: Comparison of how the masking and cropping may affect the comprehensive characterisation of the analysed devices. In the longitudinal cut of an InSb on InP nanowire with multi-twinned domains, the masking of the segments (above) when extracting the Fast Fourier Transform (FFT) is the preferred choice to capture different crystalline orientations within the same segment. On the other hand, the cropping of the segments (below) only extracts the FFT within the crops. The crop found in the InP substrate highlights how a reduced FOV might be detrimental due to the noticeable pixelation in the FFT and consequent loss in the precision of the calibration. [37]

However, the counterpart of this approach has two distinct aspects. First, from a physical perspective, the crystallographic signal representing the segment originates exclusively from the pixels within the cropped square. In other words, if the segmented region exhibits heterogeneity in terms of crystallinity (such as facets or polycrystals), the retrievable crystallographic information is limited to the cropped square. Information contained in the remaining masked pixels becomes inaccessible (see Figure 4.I.16). Secondly, the other aspect considers the possibility of cropping regions that are insufficiently large to yield meaningful and interpretable crystallographic information. If the largest square within a segmented region is small, only a few counts contribute to the crystal plane frequencies. Consequently, these frequencies appear as diffuse distributions in the final FFT, complicating their identification. Moreover, even if identified, the small number of pixels results in low measurement precision. Each pixel represents a significant jump in distance and angle measurements (fig. 4.I.17). In such cases, despite accurate peak identification, the subsequent

phase identification begins from a non-optimal starting point, potentially leading to incorrect phase assignments.

To rigorously address the significance of a small crop, we established a correlation between the minimum usable FOV and a given pixel size. We computed an experimental regression that describes the minimum workable FOV for a specific pixel size. Below this critical pixel size, we refrain from performing peak identification and subsequent phase finding routines due to the lack of expected statistical significance in the data. To derive this regression, we analysed a dataset from which peak finding metrics were extracted. We progressively cropped the dataset and reduced the FOV to identify the ranges at which we lost the typical distribution associated with a peak representing a crystallographic plane. The pixel size served as the independent variable in our analysis. Only if the FOV of the crop exceeded the value predicted by the regression did we proceed with peak identification for that segmented region. In essence, a sufficient number of pixels is necessary to construct the distribution sought by the peak finding algorithm. Given a specific pixel size, this requirement is met only beyond a certain FOV, which we empirically compute and retrieve.

The other setting, namely the "mask" setting, considers the entire segmented region obtained during the segmentation step. In this approach, we extract the FFT of the original image after applying a binary mask to it. The pixels within the segmented region retain their original values, while those outside the mask are set to zero. Consequently, the FFT signal arises solely from the segmented region, preserving the pixel size of the original image due to the FFTs adherence to the original sampling conditions. As a result, we obtain a comprehensive representation of the crystallinity across the entire segmented region, regardless of its size and the nature of the biggest crop within it.

If the segmented region contains a polycrystal-like structure, all crystallographic planes within that region will be represented in the FFT. Therefore, this setting is advantageous when such a nature is expected in the analysed sample. Moreover, by maintaining the original pixel size, we define precision based on the original image rather than the crop size, as in the previous setting. However, there is a trade-off: the intensity of the planes within the FFT will generally be lower than if locally cropped. This is due to the smaller ratio between pixels contributing to a specific frequency and the total number of pixels. Consequently, if subtle features are expected and are represented by only a few crystallographic planes, masking may obscure them, whereas cropping would highlight them within a smaller FFT or region of interest (fig. 4.I.16).

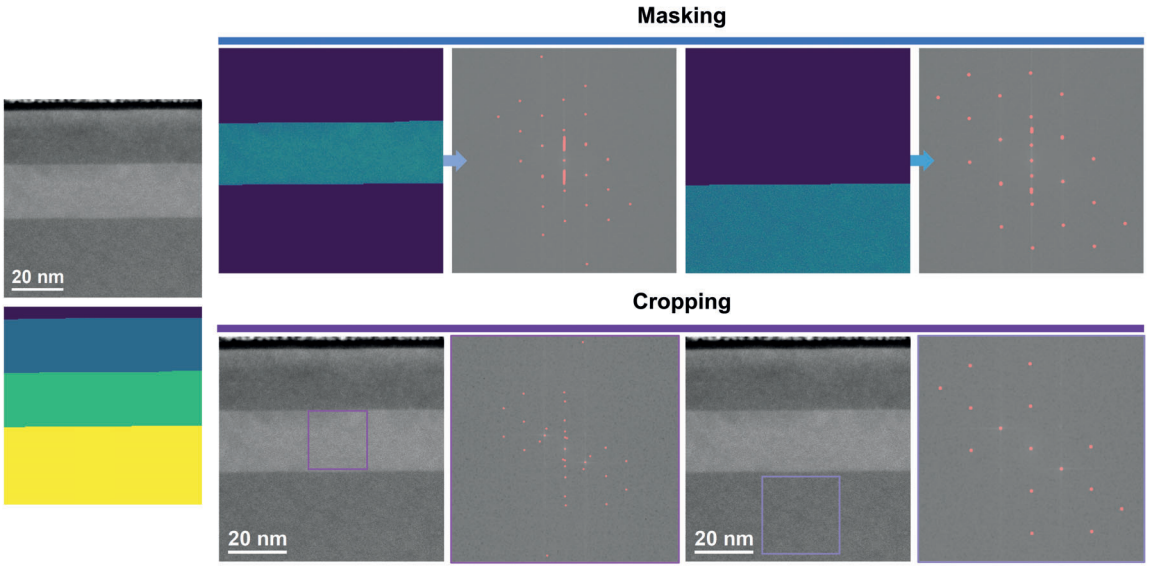


FIGURE 4.I.17: Comparison of how the masking and cropping may affect the comprehensive characterisation of the analysed devices. In this 2D Ge quantum well, GSQ_4 (Chapter 3), both masking and cropping deliver an equivalent result due to the homogeneity of the crystals within each segment found. [38]

Two examples of how the two settings substantially affect the extraction of the characteristic FFT of each segmented region can be found in figures 4.I.16 and 4.I.17. Figure 4.I.16 displays the importance of masking the segments when expecting a polycrystalline device. In this case, where we have a [110] InSb with three twins generating up to four different in-(image)-plane orientations (see the red square inset in figure 4.I.16), the masking fully captures the whole spatial extension of each orientation domain in the global FFT. On the other hand, the cropping just captures a reduced spatial extension of some of the orientations due to the positioning of the square to crop. The attentive reader will notice that in this particular sample morphology, the crop conveniently coincides with where the twinned domains sit. However, if the twins were placed in a different spatial position, they would not have been captured either in the crop or its FFT. The safe alternative then is the "mask" setting, which will always consider the whole region despite the position of the crystalline domains within the segment. In fact, figure 4.I.16 exemplifies how in the cropping of the InP substrate, whose size approaches the before mentioned minimum FOV, the pixelation of the FFT starts becoming evident. As mentioned, if bordering FOVs are met, the pixelation might affect the posterior phase identification step.

On the other hand, figure 4.I.17 showcases an instance much easier to deal with. It displays a 2D Ge quantum well resulting in three main segmented regions. Interestingly, each region is homogeneous in its crystallinity, for which we would not skip

details when cropping within each region. The choice in these scenarios would then be based on which setting gives us the best signal-to-noise vs pixelation balance.

4.I.4 Low-to-high magnification correlation

As indicated in the main text, the workflow behaves slightly differently when we input a single image or when we input a complete dataset of electron micrographs. When a single image is inputted, we expect that it presents the whole device morphology together with its atomically resolved structural features. If this is the case, we can segment and process the FFT with the micrograph, allowing us to continue the workflow finding the crystallographic phases, the strain and the 3D models.

On the other hand, if the sample is large for the maximum sampling the microscope can reach, this condition is not met. Thus, the overview of the sample structure will be sampled with a frequency below the Nyquist frequency of the atomic planes. In other words, to get atomic resolution we can only picture a portion of the whole device. This is the case of some of the devices explored in Chapter 3. For instance, GSu_02 and GSu_10 are huge devices extending along the micrometre range. Thus, their overview is contained in a huge FOV incapable of displaying atomic resolution at regular sampling ranges. For example, an atomically resolved overview of GSu_10, with a FOV of 1197 nm, would demand an image of around 20000 pixels per side (fig. 4.I.19 and 4.I.21). GSu_02, and a variation of the InAs@Al NW, also in Chapter 3, propose a similar scenario but presenting different FOV requirements. [3, 69, 70] Therefore, they are restricted to separate the overview information on the device morphology and its disposition, with the local identification of the underlying crystallography.

Our proposed solution to circumvent this apparent technical limitation is working with a dataset of micrographs instead of just a single one. The adequate datasets must be wisely acquired in no different way than the usual mental process the microscopist follows: start with low magnification images that highlight the morphology of the devices, and progressively zoom in decreasing the FOV to image the local structural details, now with atomic resolution. As a result, we end up with a dataset in which the low-magnification images (i.e., large pixel sizes) with no atomic resolution represent the global morphology of the device, and the high-magnification ones (i.e., small pixel sizes) depict just a region of the device from which the crystallographic information is accessible. Therefore, the determination is to use the low-magnification images to segment them and unveil the morphology of the device, and parallelly use the high-magnification ones to locally find the crystal phases populating the device. This idea forms the foundation for extending the end-to-end workflow proposed in this chapter to larger devices, even when high-resolution information is not always accessible. Specifically, it involves creating a comprehensive model of the large working device to gain deeper insights into both short-range and

long-range effects on its performance. However, there are numerous challenges associated with this approach, and achieving this goal falls outside the scope of the thesis.

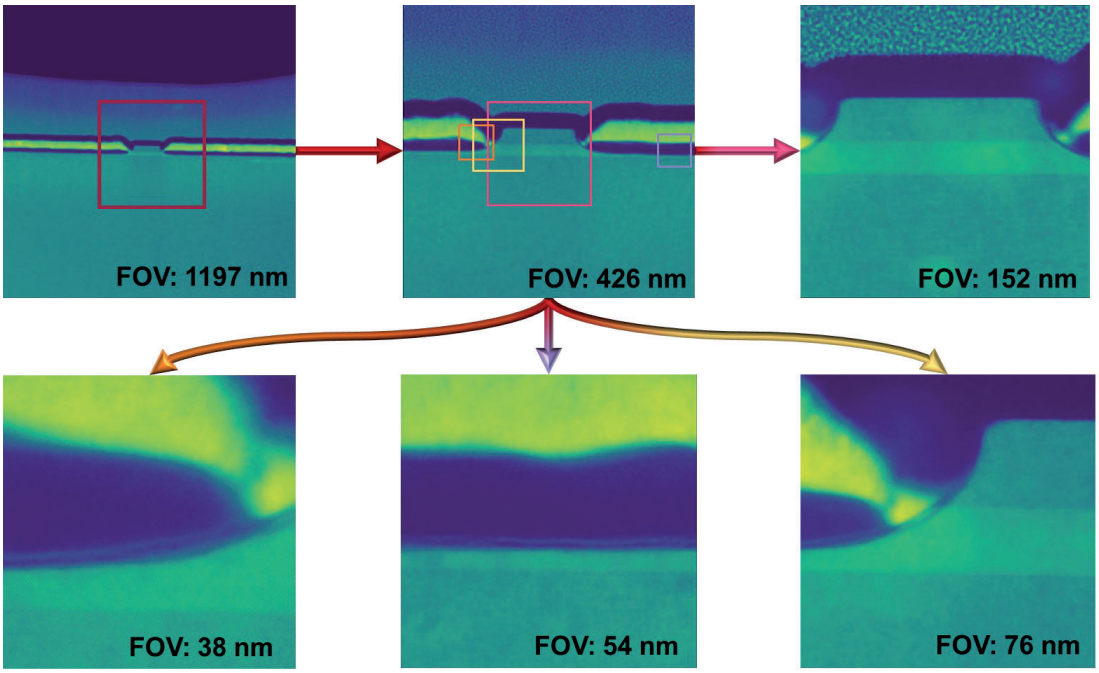


FIGURE 4.I.18: Multi-scale template matching chain of micrographs in a Ge quantum well with Al-Nb metallic contacts (GSu_02, Chapter 3) [69]. The coloured boxes highlight the position where the consecutive image connected by the arrow of the same colour lies according to the matched template. This case study highlights how a key image like the top-central one can host multiple template allocations.

This is done in several steps, the first of which is sorting the micrographs by pixel size to map the magnifications captured during the acquisition. The pixel size is the sorting variable as it correlates the micrographs with the presence of atomic resolution. This information is key in the posterior phase finding step as meaningful outputs are only expected in atomic resolution micrographs.

The next step is the key of the process, which is the global multi-scale template matching routine to generate template-query pairs. Thus, every image except the largest pixel size one is the template found to match a region within another query image with a larger pixel size. In other words, every image except the smallest pixel size ones are the query of another image with a smaller pixel size that acts as the template. Therefore, as the micrographs are sorted by pixel size, the smallest pixel-size images are set as candidates to become templates of a query image belonging to the group of images sorted with a higher pixel size. We optimised the correlation of the pixel sizes among them by allowing not less than three sorted jumps in the

magnification knob. That is, keeping the same number of pixels, to reduce the magnification three times to ensure the FOV of the template is well-placed within the query.

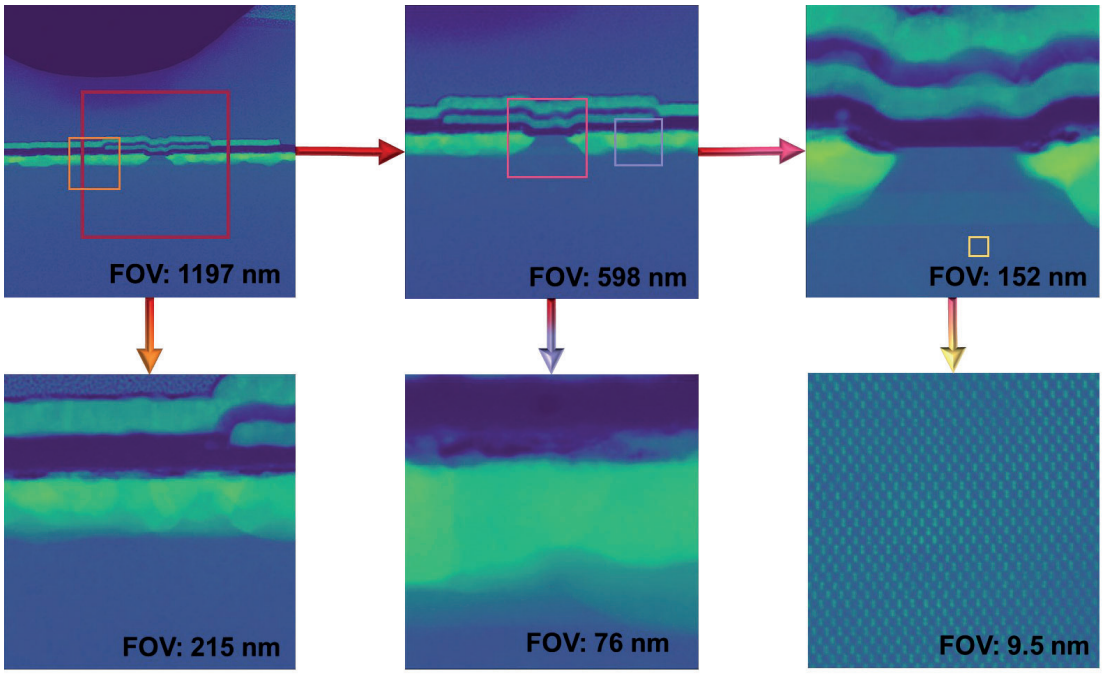


FIGURE 4.I.19: Multi-scale template matching chain of micrographs in a Ge quantum well with PtSiGe metallic contacts [70]. The coloured boxes highlight the position where the consecutive image connected by the arrow of the same colour lies according to the matched template. This case study highlights a chain of template-query pairs which exhibits fewer ramifications than what is observed in figure 4.I.18. The chain of micrograph pairs outputted by the process links the full device overview of a huge field of view, 1197 nm, with its crystallography, revealed by a micrograph of just 9.5 nm imaged.

Throughout the entire dataset, this process involves pairing every set of images that meet the imposed three-pixel sizes difference. For each pair, a metric is computed based on their pairing. At this stage, the multi-scale template matching algorithm plays a crucial role role in selecting the best template-query pairing across the dataset. Firstly, the multi-scale template matching accounts for the calibration difference between the two micrographs being paired. It achieves this by downsizing the smaller pixel size image (the template) to match the pixel size of the query. Intuitively, this downscaling corresponds to the ratio of the pixel sizes. Importantly, this effective calibration adjustment ensures that if the query region aligns with the template, they should exhibit a pixel-wise correspondence, as the cropped area within the query must have the same number of pixels as the downscaled template.

Afterwards, our first task is to identify the optimal template-query pair, followed by determining the coordinates of the most suitable region within the query that aligns with the template. These two steps are performed simultaneously by leveraging the template-matching algorithm and the associated metric. The

algorithm relies on cross-correlation, which assesses the alignment of template pixels across the query image. A heatmap is generated, with its peak indicating the highest degree of alignment between the template and the query. In addition, the metric $R(x,y)$ is normalised for every template-query pair (i.e., $T(x',y') - I(x+x',y+y')$), ranging from 0 (no alignment) to 1 (perfect alignment). This normalization enables straightforward comparison between pairs by seeking the global maximum (as expressed in equation (4.I.7)). [63, 83] As a result, we obtain the best template-query pairs and the corresponding coordinates within the query where the template exhibits the strongest match. The core implementation of this algorithm is practically achieved using the OpenCV Python module, specifically the cv2.matchTemplate(query_image, downscaled_template, cv2.TM_CCORR_NORMED) function. [63, 83]

$$R(x,y) = \frac{\sum_{x',y'}(T(x',y') \cdot I(x+x',y+y'))}{\sqrt{\sum_{x',y'} T(x',y')^2 \cdot \sum_{x',y'} I(x+x',y+y')^2}} \tag{4.I.7}$$

A schematic representation of the results obtained from this algorithm can be observed in figures 4.I.18 and 4.I.19. In these figures, we see a collection of query images with coloured squares and arrows, indicating where the templates align best within the FOV. Additionally, the template-query pairs are linked sequentially when the dataset is sufficiently large, as depicted by the different arrows connecting the micrographs in figures 4.I.18 and 4.I.19. The main advantage of this process lies in our ability to identify atomic resolution regions within a micrograph - even in cases where no crystallographic planes are visibly discernible (i.e., achieving atomic resolution within a few µm FOVs). Similarly, 4D-STEM allows us to access the crystallography in large FOVs, for which we can somehow emulate it without the counterpart of dealing with massive datasets. In the end, this allows us to get maps similar to those shown in Figures 4.I.20, 4.I.21 and 4.I.22, where all micrographs from a dataset are superimposed to explicitly reveal the conditions that optimise global template-query matching. It is worth noting that this multi-scale matching approach facilitates accurate segmentation of high-magnification images by extrapolating and refining the segmentation performed on their lower-magnification counterparts. This capability enables us to navigate the progressive gradients present in atomically-resolved interfaces and overcome the diffuse segmentations typically encountered in high-magnification micrographs (as detailed in section 4.I.1).

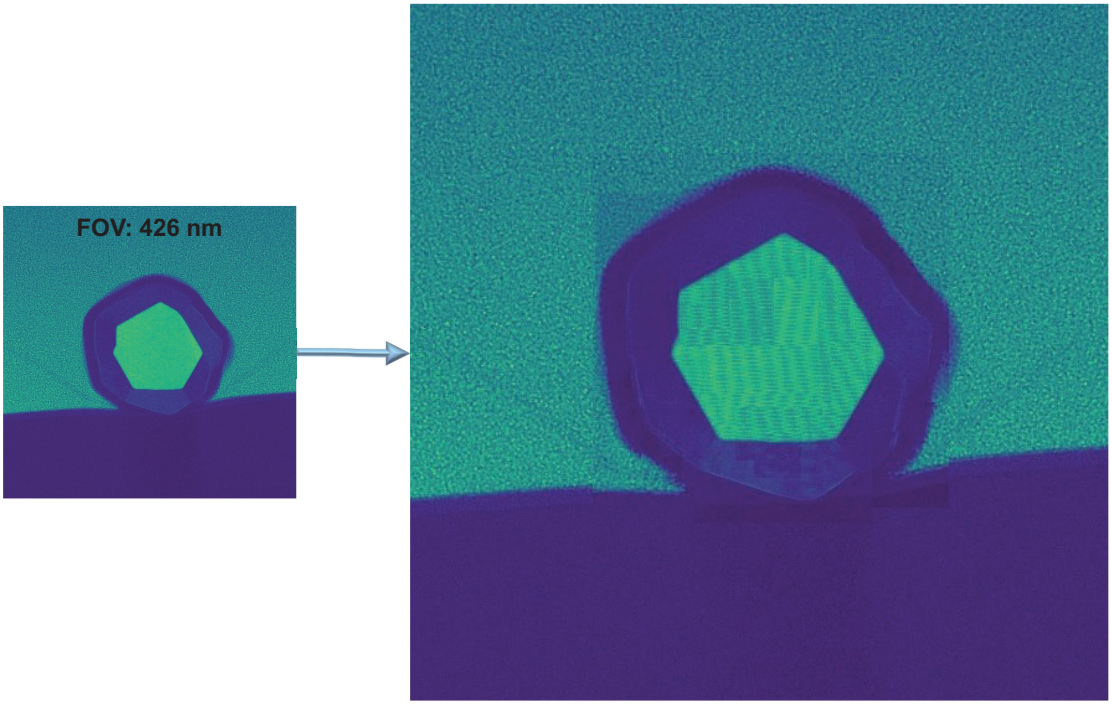


FIGURE 4.I.20: Example of the full dataset collage obtained after superimposing all the template-query pairs of an InAs nanowire with an Al shell (variation of the device in Chapter 3, but same study). [3] The original micrograph at left shows the absence of atomic resolution at a FOV of 426 nm, while the patched image highlights the atomic resolution insets within the nanowire.

Understanding the strengths and limitations of this algorithm for performing the envisioned template-query pair correlation is crucial. As an efficient and robust method, it works well when dealing with a rigid acquisition scheme, which is the case in most (S)TEM scenarios. However, the algorithm currently does not handle rotations or zone axis tilts gracefully. While, in-(image)-plane rotations could be incorporated alongside multi-scale considerations as an additional variable inducing these rotations, more aggressive tilts would still pose challenges. Therefore, although the existing multi-scale template matching approach is robust and suitable for the scope of our present study, a more general solution would require computationally expensive algorithms. Notable examples include feature-finding methods like Scale-Invariant Feature Transform (SIFT)[84], Speeded-Up Robust Feature (SURF)[85], Orientated FAST and Robust BRIEF (ORB)[86], among others[87].

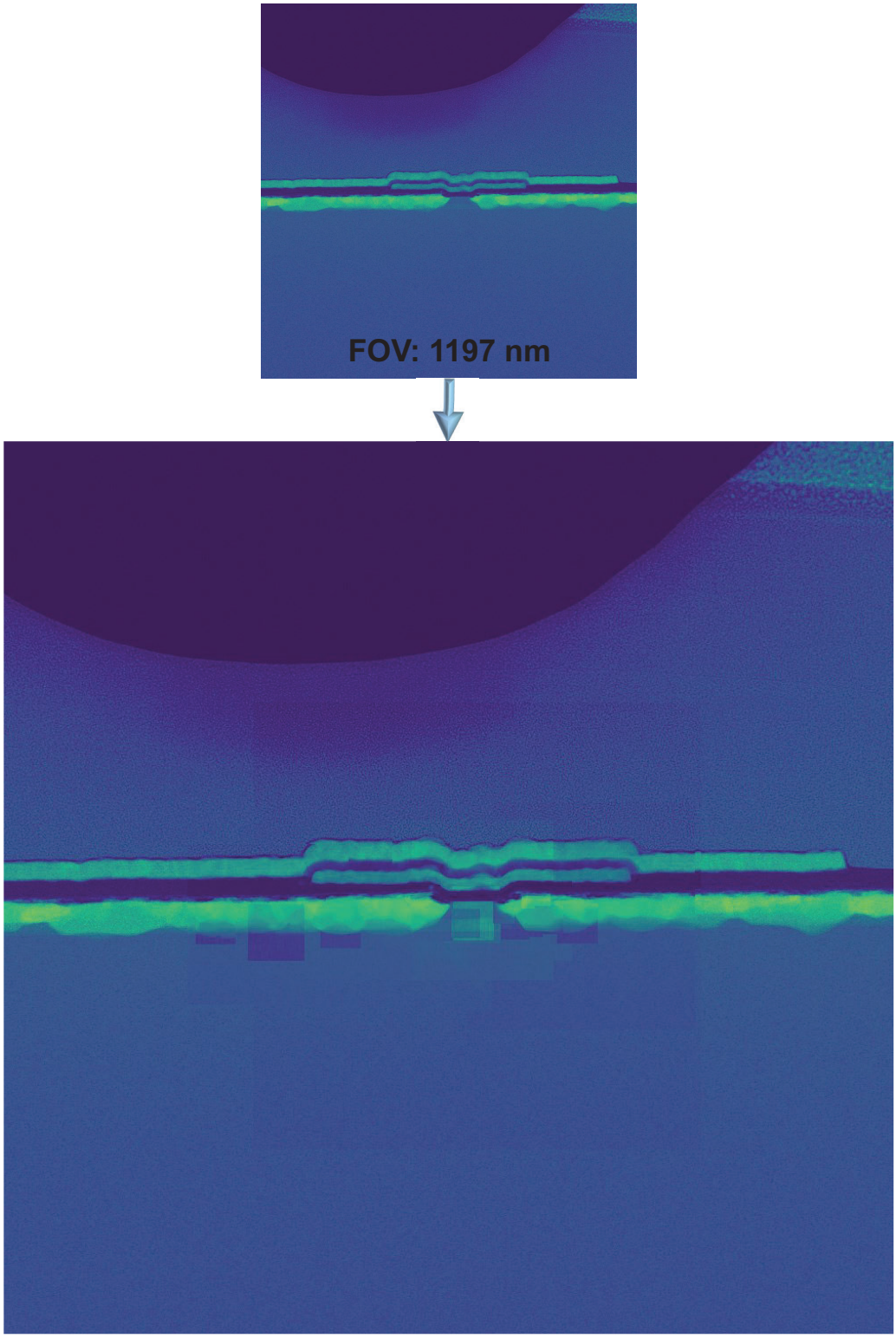


FIGURE 4.I.21: Example of the full dataset collage obtained after superimposing all the template-query pairs of a Ge quantum well with PtSiGe metallic contacts (Chapter 3). [70] The original micrograph shows the absence of atomic resolution at an ultrawide FOV of 1197 nm, while the patched image highlights how higher-magnification insets are allocated within the overview, allowing us to access atomically resolved information at a large FOV.

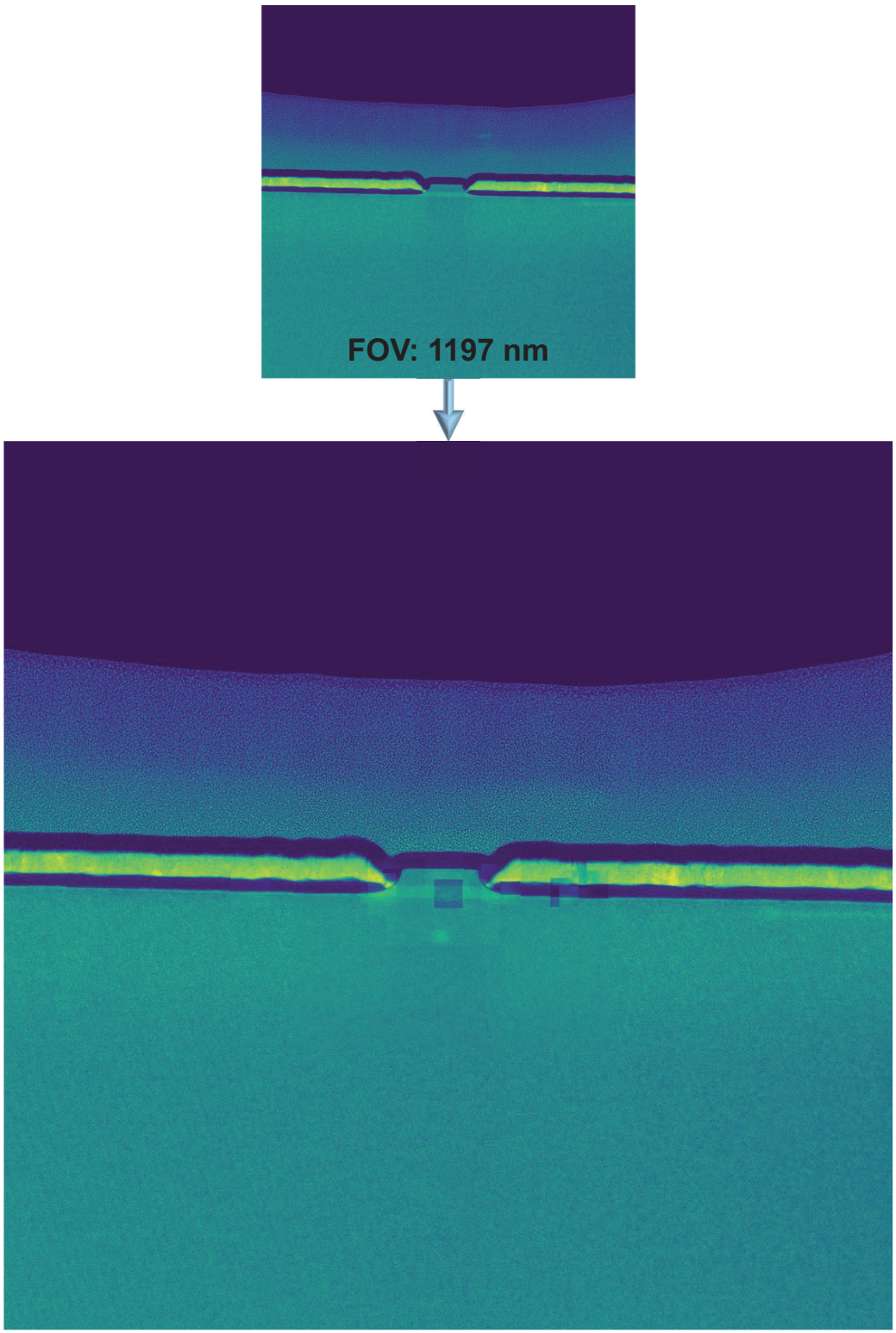


FIGURE 4.I.22: Example of the full dataset collage obtained after superimposing all the template-query pairs of a Ge quantum well with Al-Nb metallic contacts (Chapter 3). [69] The original micrograph shows the absence of atomic resolution at an ultrawide FOV of 1197 nm, while the patched image highlights how higher magnification insets are sparsely allocated within the overview, allowing us to access atomically resolved information at a large FOV.

4.I.5 Phase identification

The former segmentation and peak finding modules are modelled in a way that combines robustness for the phase identification, much like like a symbiotic relationship. Specifically, within each segmented region, we compute its masked or cropped FFT and identify its peaks. From these peaks, we extract the crystallographic phase that better matches our database of candidate phases. Of course, the limitations of the method coincide with the same one would face if doing the FFT-based model comparison manually. If the crystallographic similarity raises coincident diffraction patterns, for instance, due to equal interplanar spacings of different materials, this methodology will be incomplete. Moreover, fine compositional details will not be captured in reciprocal space, for which its unveiling will require spectroscopy, as we extensively comment in Part II. Yet, in order to address a wide range of scenarios that can benefit from this approach, we suggest a procedure guided by multiple factors.

The identified peaks, represented by FFT pixel coordinates, are then transformed into real space distances and angles between peaks. Interestingly, we only use the right hemisphere of the FFT since the FFT exhibits centrosymmetry. Our ultimate objective is to establish a link between these peaks and a zone index. The zone index corresponds to a vector normal to the mathematical plane parallel to the image plane (i.e., a real space vector). Additionally, we aim to associate the peaks with Miller indices of the crystallographic planes, which represent them more accurately (i.e., a reciprocal space vector). To define a mathematically heterogeneous plane in space, we require two vectors that belong to the plane. Consequently, given a single FFT, the process involves accounting for all spot pairs and linking each of them to a zone axis and an indexation. Furthermore, we seek to identify the best zone axis or axes that explain all the observed peaks. In essence, this entails correlating the spot pairs with the results obtained for each peak.

The process of obtaining the best candidate phase, zone axis and indexing based on a calibrated spot pair follows the pseudocode outlined in Algorithm 4.I.1 and is visually represented in Figure 4.I.23. The underlying idea is to define a tolerance window in reciprocal space. Within this window, we accept that a pair of reflections from a candidate theoretical diffraction pattern matches the experimental spot pair. The tolerance encompasses both distances and angular domain. Specifically, we allow for a deviation in the interplanar distance of the experimental reflections compared to the theoretical ones. Additionally, we consider a deviation between the angle formed by the two experimental spots and the angle between the coinciding theoretical reflections. To ensure accuracy and to restrict matches to realistic crystal phases, we set the tolerance at 5% (i.e., a range of 5-10% is reasonable). If these conditions are met, the spots are indexed according to the proposed reflections. In Figure 4.I.23, the green circular sectors drawn atop the experimental spot pair indicate the reciprocal space area where the candidate theoretical reflection pairs must

reside to be accepted. Finally, the zone axis is computed by taking the cross product of the two reciprocal space vectors represented by the Miller indices of the identified planes.

Algorithm 4.I.1 Crystal phase, zone axis and indexation finding for every spot pair

for all spot pairs sp_l $(dist_{exp,1,l}, dist_{exp,2,l}, angle_{exp,l})$ **do**
  for all candidate crystal phases ph_i **do**
    $dist_{theo,i,hkl_{j,k}}, angle_{theo,i,j-k} \leftarrow ph_i$
    if $dist_{exp,1,l} + tolerance \cdot dist_{exp,1,l} \geq dist_{theo,i,hkl_j} \geq dist_{exp,1,l} - tolerance \cdot dist_{exp,1,l}$ **then**
      if $dist_{exp,2,l} + tolerance \cdot dist_{exp,2,l} \geq dist_{theo,i,hkl_k} \geq dist_{exp,2,l} - tolerance \cdot dist_{exp,2,l}$ **then**
        if $angle_{exp,l} + tolerance \cdot angle_{exp,l} \geq angle_{theo,i,j-k} \geq angle_{exp,l} - tolerance \cdot angle_{exp,l}$ **then**

$$\sqrt{\left(\frac{dist_{exp,1,l} - dist_{theo,i,hkl_j}}{dist_{theo,i,hkl_j}}\right)^2 + \left(\frac{dist_{exp,2,l} - dist_{theo,i,hkl_k}}{dist_{theo,i,hkl_k}}\right)^2 + \left(\frac{angle_{exp,l} - angle_{theo,i,j-k}}{angle_{theo,i,j-k}}\right)^2} \leftarrow Score\ metric_{\,l,i} \quad (4.I.8)$$

          $\overrightarrow{v_{hkl_j}} \times \overrightarrow{v_{hkl_k}} \leftarrow [uvw_l]$
          Store $hkl_{j,l,i}, hkl_{k,l,i}, uvw_{l,i}$, Score metric$_{\,l,i}$
        end if
      end if
    end if
  end for
  Ascendant sort of $(hkl_{j,l,i}, hkl_{k,l,i}, uvw_{l,i})$ by Score metric$_{\,l,i}$ criteria
  $(hkl_{j,l,i}, hkl_{k,l,i}, uvw_{l,i})_{min} \leftarrow$ min(Score metric)
  return $(hkl_{j,l,i}, hkl_{k,l,i}, uvw_{l,i})_{min}$ (i.e., zone axis and crystal phase) of sp_l
end for

In this evaluation, we consider every model cell proposed as a candidate. Consequently, a single-spot pair might match multiple candidate crystal phases, zone axes, and even different zone-coincident configurations. To assess the goodness of fit, we introduce a metric for our proposed theoretical reflection pair given a crystal phase and zone axis in conjunction with the experimental spot pair (see Equation 4.I.8). Specifically, the best candidate model for each spot pair minimises the Euclidean distance with the experimental spot pair.

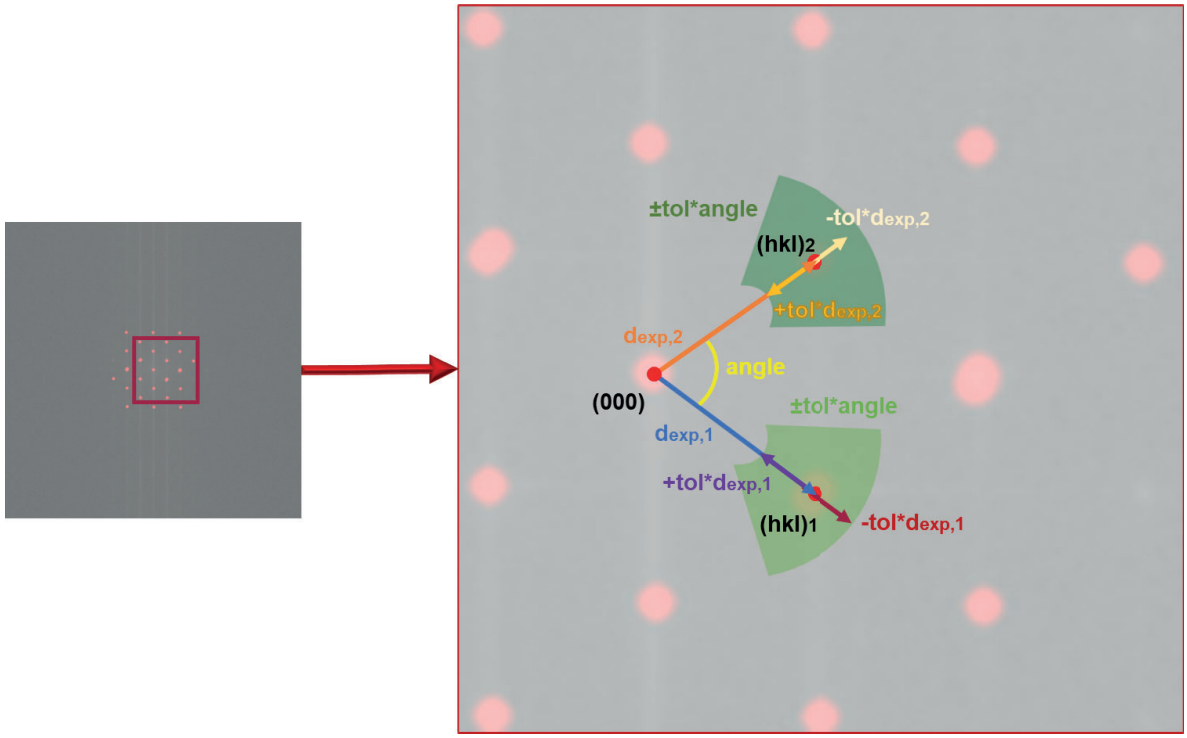


FIGURE 4.I.23: Fast Fourier Transform of a [110] Ge quantum well, with an accurate identification of its peaks. The inset identifies the quantities (i.e., distances and angles) that are computed to define the reciprocal space area (i.e., the green circular sectors) within which the candidate reflection pairs of a given crystal phase are accepted as a valid matching model.

Ultimately, we end up with $\frac{n(n-1)}{2}$ spot pairs, where n represents the number of identified peaks in the FFT. Consequently, there are $\frac{n(n-1)}{2}$ potential crystal phases and orientations. However, this outcome might not be conclusive, as these multiple outputs need to be interconnected. The second part of the process relies on the correlation between spot pairs. The pseudo-code for this process can be found in Algorithm 4.I.2. The approach involves grouping the spot pairs into families, each representing an individual phase identified in the image. Initially, we identify spot pairs with coinciding phases and zone axes, as they are potential candidates for forming one of these spot families associated with a single crystal. The family of spots initially emerges with the first checked spot pair. Afterward, the goal is to add spots to the family if they belong to the same crystal. The condition for accepting a new spot is evaluated by considering spot pairs between the existing spots in the family and the candidates to join it. The new spot must have previously formed a pair with at least two spots already present in the family, sharing the same zone axis and crystal phase. Although this condition may seem restrictive, it ensures robustness. Only peaks resulting from shared phase and axis among multiple pairs are added. Occasionally, isolated spots might arise due to insufficient tolerance, especially when high-index (high-frequency) reflections are abundant. By imposing this constraint,

we prevent incorrect assignments from misleading the final identified phase. Essentially, we prioritise creating an additional family for outliers rather than adding them to an existing family. Subsequently, we assess the goodness of fit or reliability of each family (see Algorithm 4.I.2).

Algorithm 4.I.2 Recursive global crystal phases and zone axes finding within a single Fast Fourier Transform

[ ] ← spots_families ▷ Empty list to allocate families of crystal phases/spots
function CORRELATE SPOT PAIRS($spots, spot_pairs, spots_families$)
  if all sp_m in spots_families **then**
    return spots_families
  end if
  for all spots (identified) s_k **do**
    for all spot pairs sp_m $(CrystalPhase_m, ZoneAxis_m)$ AND sp_n $(CrystalPhase_n, ZoneAxis_n)$ **do**
      if s_k in sp_m AND s_k in sp_n **then**
        $s_{i,l}$, $s_k \leftarrow sp_m$
        $s_{j,l}$, $s_k \leftarrow sp_n$
        if ($CrystalPhase_m = CrystalPhase_l$ AND $ZoneAxis_m = ZoneAxis_l$) AND ($CrystalPhase_n = CrystalPhase_l$ AND $ZoneAxis_n = ZoneAxis_l$) **then**
          $s_{k,l}$ (in family$_l$ in spots_families) ← s_k
        else
          $s_{k,p\neq l}$ (in family$_{p\neq l}$ in spots_families) ← s_k ▷ Find p in next iteration
        end if
      end if
    end for
  end for
  function CORRELATE SPOT PAIRS($spots, spot_pairs, spots_families$)
  end function
  for family$_l$ in spots_families **do**
    $n = \frac{m(m-1)}{2}$ spot pairs ← m spots in family$_l$

$$ScoreMetric_l = \frac{1}{n^2}\sum_{i=1}^{n} sp_i \tag{4.I.9}$$

    Store ScoreMetric$_l$
  end for
  Ascendant sort of family$_l$ in spots_families by ScoreMetric$_l$ criteria
  $family_{l,min}$ ← min(ScoreMetric$_l$)
  return $family_{l,min}$ (i.e., most likely zone axis and crystal phase)

This assignation of spots into families is accomplished by recursively removing spots and adding them to the families until no spots are left unassigned. Ideally, a single crystal would exhibit spot pairs with the same phase and zone axis. In such a scenario, assigning them to a family would create just one family, and all spots would be progressively added due to their correlation with each other. However, interestingly, if one of the previous spot pairs does not match the same zone axis (for example, assigning high Miller indexes from another candidate phase or zone axis to a reflection pair with small interplanar distances), two families would be created. The main family would include all spots except one, corresponding to the actual phase. Additionally, there would be another family containing the outlier phase found in a single spot pair. Despite this restrictive condition, even when a single spot pair is not well identified, all the spots in the FFT would still be linked to the same phase, as they would be correctly associated with the rest of spots.

Eventually, we compute a score for each family. This metric is based on the goodness of fit of each spot pair within the family and the total number of spots it contains. Our restrictive condition for adding spots to a family ensures that the more spots a family contains, the higher the likelihood of achieving a true positive in the phase identification step. Therefore, we can sort the identified families of correlated spots and phases from most likely (i.e., smallest value and primary output) to least likely (i.e., largest value) based on this score. The score rewards more detections and can be represented by Equation 4.I.9.

Importantly, the restrictive constraint of twofold spot pair confirmation goes beyond merely adding robustness to the algorithm. Alongside the identification of single crystals, it enables the detection of polycrystalline domains and structural deviations visible in the FFT. Specifically, if a defect or structural element produces additional frequency peaks in the FFT, the algorithm identifies them as independent families and separates the corresponding phases. Interestingly, even when a reflection is shared between multiple crystals (as in some twinned structures), the algorithm recognises that the spot representing this reflection is shared between two structures with the same zone axis but different in-plane orientations.

While the methodology offers significant benefits and impressive generalization capabilities, it is essential to recognise its context scope. No single approach can universally address all materials science challenges. Instead, the methodology excels within specific contexts and applications. Let us list some thoughts about it, which not only define better its current state but also facilitate and set perspectives on what can be achieved in the future:

- As extensively detailed, the phase identification model cannot leave the full-of-possibilities but incomplete realm of model testing. Therefore, we are limited to testing already reported crystal phases. Nowadays, there are thousands of identified phases and most of the state-of-the-art materials science implies

known phases rather than phase discovery. However, with the advent of the AI-powered discovery of materials, identifying and fully characterising unreported crystal phases might become increasingly important. [88] This obligatorily demands electron diffraction. Thus, 3D electron diffraction, either by exhaustive tomography or by specific rotation along zone axes of interest will be required. AI has already facilitated the creation of solutions that start targeting this concern in an automated way, although they are just the beginning of the eventual solution which could be perfectly merged with the present workflow to deliver an even broader landscape of crystal phase identification. [89]

- In this context, we fully leverage the strengths of our methodology. We have rigorously validated its robustness within the known-phase constraint, a critical and widely applicable aspect of ongoing research. Throughout the thesis, this validation is evident and supported by manifold proofs that we have gathered. However, many more have been tested although not explicitly included for the sake of cohesion and coherence with the materials, samples, and devices studied. As explained in the present section, our approach involves the comprehensive encoding of all space groups. However, the true extent of its capabilities remains to be determined through further diverse investigations and evaluations. In that regard, we envision continuous improvement over the years to come.

- However, we need to account for this strength with precaution. Importantly, this applies if only one material is found along the longitudinal direction of the electron beam. That is, an in-depth stack of materials will produce complex double diffraction patterns that will surely pose a challenge to the methodology. For instance, a positive peak detection will then identify the double diffracted peaks from the (000) overall peak. If so, the algorithm should then find a second positive phase matching candidate and experiment. This is programmatically equivalent to two or more phases coinciding in the image plane. Nevertheless, the Bragg filtering would reveal a spatial coincidence with the rest of the longitudinally stacked crystal phases. On the other hand, if these additional spots do not lead to the second identified phase, the supplementary double-diffracted spots will not unlock the phase identification and would otherwise cause confusion and distort the guessed phases. 3D electron diffraction would facilitate this task. Additionally, focal series could also timely answer this concern. In fact, this approach would probably be an easier integration with the current workflow, as it could lead to a sliced phase identification in which the filling of the 3D information would follow the axis defined by the focal change.

  The take-home message is that the algorithm is conceived to take this possibility into account, although not thoroughly. For instance, an additional validation could be done by selecting every spot as the new centre of a shifted

FFT. If any, this would reveal crystal phases underneath the main identified ones. Nonetheless, this would not come computationally free, as its cost would square the already factorial cost of the spot pair checks, as described in Algorithms 4.I.1 and 4.I.2.

- Similarly, the translational invariance can be broken by in-depth defects of the same crystal phase. If these defects modify the resulting FFT, mainly two scenarios can occur. On the one hand, if the structural features alter the FFT in a way that still a zone axis and a 3D orientation could be mapped by a trained scientist, then the workflow should be prepared to handle it. On the other hand, if the structural variations distort the FFT but do not provide extra information to extract a secondary phase, the information would inevitably be lost with the current version of the workflow. In fact, it could even distort and damage the identification of the primary phase. A possible solution would lie in the inclusion of physical models stimulating the response in the presence of any possible defect along the electron beam path. A more realistic approach, though, entails training a supervised model to correlate FFT or real space (when convenient, as part of the decision tree model) modulations with in-depth defects. In any case, it seems an extremely complex endeavour only addressable through the careful effort of combined physics modelling and AI/CV.

- Exotic pseudo-crystalline structures like quasicrystals, molecular crystals, serpentine materials, or non-centrosymetrical phases, would pose an interesting challenge to the algorithm. Importantly, the crystallographic information encoded in the algorithm ignores anything outside the conventional 3D space groups. [90, 91] Therefore, an additional encoding would be required to capture these nuances. In fact, a paradigm change could even be needed as encoding unconventional crystallographic structures of non-periodic cells is not trivial. Despite its fascinating challenge, its interest is not as extended as the previous raised concerns. As a result, we would prioritise widening the scope of conventional materials in the future stages.

  Larger structures and potentially more complex diffractograms like proteins, viruses, molecular crystals, and in general biological structures, would be out of the current reach of the workflow. The full universality we seek would inevitably imply adapting to these cases. Fortunately, cryo-electron microscopy has already been tackling this problem since its conception in the decade of 1980. [92, 93] Thus, extremely complex routines based on (electron beam) phase determination for structure discovery were implemented as the basis of modern biological structure characterisation. [94, 95] Indeed, this revolution was enabled by a computational evolution, similar to the current emergence of AI. In any case, these complex routines for peak-wise analysis and direct blind structure determination can serve as a guide of the equivalent method for the

present materials science context. In fact, the physics would not differ much, and a collaborative merge would be the wisest option to optimally address the first raised limitation.

- As introduced before, the most obvious limitation lies in the compositional analysis and equivalence of FFTs. These scenarios limit the operability in the same ways one would struggle with manually and requiring spectroscopy. Dynamical effects and subtle distinctive information between phases corresponding to individual peaks, as well as ambiguous FFT making the identification impossible with a single image would not be considered. The template matching approach would be a first step to merge multiple rotations and converge into a global most common crystal phase per region (section 4.I.4). However, further physical modelling is required to fully account for these effects.

- The polarity is not considered directly in the workflow. Thus, once we set the in-(image)-plane orientation, we need to define the polarity through a manual intervention if polarity is expected. Practically, if the orientation of the dumbbells does not correspond to the experimental observations, a 180° in-(image)-plane rotation must be manually set. This is discussed later in Part II, especially when building the atomic models of III-V semiconductors (section 4.II.3).

  An additional polarity validation could be retrieved from real space, if statistically possible. On the other hand, if the polarity is foreseen in the FFT, for instance, in cases where the difference in dumbbell intensities is especially strong, the full analysis could be kept in the reciprocal space. Situations in which the planes of the light atom of the dumbbell are clearly distinguishable from the heavy ones should underscore visible modulations in the FFT in case stacking faults, twins or anti-phase boundaries are observed. Again, specifically trained models would be required for such cases to converge into the correct polarity. [96, 97]

- Interstitial atoms, vacancies, substitutional atoms, and similar are not distinguishable in the reciprocal space. Although they are indirectly considered in the next section about the strain calculations (section 4.I.6), we still require a tailored real-space solution to identify them if statistically possible. In fact, we foresee this complexity in the Introduction when we refer to the current case-specific literature as the perfect complement to a more general method. Therefore, the proposed workflow could settle a general basis upon which to add modules specifically tackling these singular effects, and avoid being misled by confusing effects such as static atomic displacement. [98, 99] However, this is a delicate aspect that might lie close to the boundaries of the intrinsic limitations of TEM. At least the boundaries of regular TEM imaging unless diving into the realm of tomography to resolve in-depth features.

Naturally, these limitations will persist throughout the following sections. Some will likely be exacerbated, at the very least, by the crucial constraint of translational invariance exploited in Part II. These limitations are discussed in detail throughout the chapter, underscoring the workflow's current capabilities while revealing its potential for future growth.

4.I.6 Strain analysis: Geometrical Phase Analysis automation

To obtain local information about the strain that complements the crystallography revealed in the previous steps of the workflow, we employed Geometrical Phase Analysis (GPA). The understanding and applicability of GPA are extensively documented in the literature, and a detailed explanation is unnecessary for this text. [29, 30]. Nonetheless, the fundamental steps of GPA are outlined in the following subsections, along with explanations of how they are automated and integrated into the overall workflow.

4.I.6.1 Definition of the reference region

Generally, the choice of the reference region which offsets the chosen g vectors is assigned to a fully relaxed area of the device. This decision launches a calculation whose output is more interpretable as directly transforms the computed geometrical phase and the dilatation into the absence of strain in the chosen region. In any case, overall, there are no incorrect choices in picking up a region of the image as a reference region, but some choices make the results easier to interpret than others.

Therefore, as the reference region is an arbitrary choice, we opt for the segmented region that is more likely to exhibit a fully relaxed state. Specifically, we focus on the crystalline segmented region closer to the bottom of the image. Typically, this region corresponds to the substrate on which heterojunctions grow. If the device is imaged with a different orientation (for example, deliberately rotating interfaces from the scanning directions), it is advisable to rotate the image to ensure that the substrate occupies most of the lower portion of the image. Importantly, due to the way the 3D atomic model is built (based on virtual unit cells), the choice of the reference region does not significantly impact the final models (see section 4.II.3.2 for details on virtual cell creation). Furthermore, while some routines for highlighting points of interest can be automated (such as identifying local areas of maximum variability in the GPA maps), the interpretation of the results ultimately rests with the scientist. These discussions will be of interest especially when simulating functional properties out of the built 3D models (see section 4.II.5).

Regarding the size and positioning of this area of reference, the algorithm computes the largest possible square that can fit within the selected segmented region. This

process mirrors the same procedure used in the "crop" setting of the segmentation and peak finding interlinking (see section 4.I.3). Subsequently, a uniform reduction of 20% is applied to the square's side to prevent it from capturing any phase shifts from neighbouring regions.

4.I.6.2 Selection of the optimal g vectors pair and mask resolution

The g vectors represent reciprocal space vectors that point to crystallographic planes (i.e., peaks in the FFT). These vectors are used for the geometrical phase calculation. The selection of the two g vectors is a crucial step in the GPA automation. It leverages all the knowledge acquired from the preceding steps to identify the optimal pair of vectors.

After defining the reference region, the indexed crystallographic planes within this region are directly compared to the indexed planes found in other regions using the phase identification algorithm. Initially, the planes corresponding to the best-fitting crystallographic phase in the reference region are used to calculate the average interplanar distance. This average distance defines a central ring of distances around which the search for the g vectors begins. The starting point for this search aligns with the location where most of the planes are found, pinpointing the first Laue zones. It's important to note that only planes that have been correctly indexed and assigned to a specific phase contribute to this average distance.

Generally, the best planes to act as g vectors belong to the zero or first-order Laue zones. Interestingly, these planes are the most intense and representative crystallographic planes, resulting in a smoother representation of the phase calculation. The algorithm takes this into account by checking the planes with an interplanar distance higher than the average. If two or more meet this criterion, the search for the best pair of g vectors begins. However, if only a single plane is found below the average (i.e., closer to the (000) reflection), the process is equivalent and selects candidate spots within a ring of progressively increasing size in reciprocal space.

If multiple spots are present below the mask, these potential candidates for g vectors are evaluated based on their scores when matching the best-fitting phase and their relation with the rest of the indexed planes from other best-fitting phases within each crystalline segmented region. For all the candidate planes, we search for planes from other phases that may represent epitaxy with the reference ones. Specifically, if a candidate plane is near (i.e., exhibits similar interplanar distances and small angular misalignment) another plane from a different phase, it is likely that they are representing epitaxial growth. In addition, we strengthen this hypothesis by checking if the space groups and the indexation of both neighbouring planes coincide. Although this scenario may seem specific, it is recurrent in heterostructures within the semiconductors industry. Therefore, we select the spot pair (i.e., pair of planes,

pair of g vectors) that minimises the joint difference with its closest spot from another phase. If two or more planes from the reference phase have a clear epitaxial neighbour, we choose the pair of spots that minimise the differences in distance and angle (fig. 4.I.24, a). However, if only one or no spots present an obvious neighbouring spot, the chosen spot pair adheres to the proximity criteria and still minimises the differences, even if they are not obvious to the eye. In such cases, heteroepitaxy is not identified due to the lack of proximity in the selected spots (fig. 4.I.24, b). This evaluation occurs during the search for the diameter of the mask used for GPA.

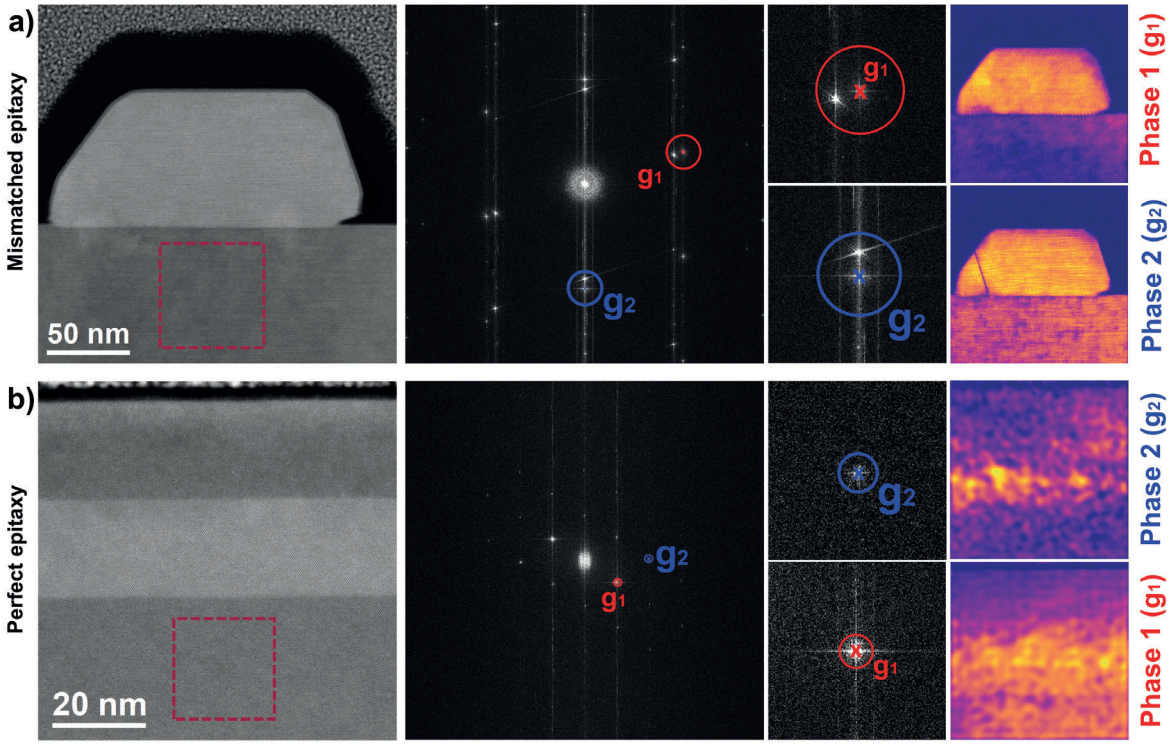


FIGURE 4.I.24: Two representative cases of how the automated Geometrical Phase Analysis (GPA) handles the selection of the g vectors and the radius of the mask based on the relative orientation of the materials and their epitaxy. a) Highlights the epitaxial growth of an InSb nanowire on InP substrate, whose lattice mismatch of 10% is detected by the routine and opens a mask of 2.3 nm to fully capture both spots in the GPA. [37] b) Similarly, the Ge quantum well sandwiched by two buffer layers of $Ge_{0.7}Si_{0.3}$ showcases the case of a perfect epitaxy where no mismatch is visible (GSQ_4). [38] The perfect epitaxy of this scenario is programmatically equivalent to not finding any epitaxy at all (i.e., spots from single phases far away from each other). Although physically the situations are opposite, the workflow just finds a single spot and limits the mask to 4 nm. The distinction between perfect epitaxy or no epitaxy will come from posterior steps, basically the phase identification, the Bragg filtering, and the results of the GPA strain maps.

The mask size is computed following the same logic as the choice of the g vectors. Taking advantage of the knowledge extracted from this process, we assess whether our g vectors fall within a logical threshold representing lattice mismatch. The mask size is typically defined in nanometers and corresponds to a circumference with a radius centered at the FFT ((000) reflection). Consequently, the radius of just a few

reciprocal space pixels corresponds to an extremely low frequency and a large distance of a handful of nanometers. This mask size defines the resolution of the calculation (see Methodology). A larger mask (i.e., smaller radius in nm) can consider a higher lattice mismatch, resolving finer features, but it is also more sensitive to noise. Interestingly, when centring this mask on the reflections indicated by the chosen g vectors, it represents a possible mismatch that depends on the interplanar distance of the g vector, denoted as d, and the resolution of the GPA calculation, denoted as m. The possible mismatch lies within the interval $[100\frac{-d}{m+d}, 100\frac{d}{m-d}]$ (in %). For convenience, we set a maximum mismatch of m = 2 nm in mask resolution. Considering typical interplanar distances of d = 0.2 and 0.4 nm, this opens intervals like [-9, 11]% and [-17, 25]%, respectively. These intervals strike a balance between capturing almost any physically meaningful real mismatch between two materials and avoiding distant reflections that are not representative of the epitaxy. If this condition is met, the mask is defined as a distance equivalent to the furthest epitaxial spot found, plus a 20% of this distance. This ensures that the ring captures the full picture of its geometrical phase (fig. 4.I.24, a).

If no sign of epitaxy is found within the m = 2 nm mask resolution, the algorithm assumes that no neighbouring spot should be considered around it. As a consequence, only the information around the reference g vector needs to be taken into account. To ensure this, a default resolution of m = 4 nm is set. This value captures the information around a single spot independently of its spread, striking a balance between resolution and noise (fig. 4.I.24, b).

4.I.6.3 Computation of the rotation angle of the Cartesian axes

To ensure that the x component of the dilatation maps is parallel to the interface, and consequently the y component is perpendicular to it, we compute the angle of the interface to the x axis of the image (i.e., the horizontal edge). To do this, we use the segmentation of the region defined as a reference. From its contours, we draw the profile of the segment and search for those contours that do not belong to an image edge while maintaining the same vertical position. Due to pixelation, a contour is defined as a vector from one pixel to another. If the starting and ending coordinates of the contour share the vertical component, then this vector is a fragment of a real horizontal interface.

Aliasing in the segmentation can potentially lead to abrupt changes in the calculation of this angle when small contour fragments are considered. If only a single contour is used, the angle will always be either 0º or 90º. However, if groups of contours are employed, this effect is mitigated, and intermediate angles can be obtained. Consequently, we perform calculations using an increasing number of contours, computing the angle to the horizontal edge of the image for each contour fragment. The angle should be more accurate as it considers multiple contours. However, if there is noise in the segmentation or if the interface is not perfectly straight, the angle

computed in the largest segment might not be the most representative. Therefore, we retrieve the average of the computed angles from all possible contour fragments. Interestingly, these contours originate from the central contour of the interface and extend in both directions, first separately and finally jointly, to fully map all possible tangents to the interface.

4.I.6.4 GPA base code

The base code for performing GPA is adapted from STEM_CELL. [41, 42] It is integrated into the Python-based code through a custom plugin that maintains its core functionalities. In other words, the Python-based workflow automates the extraction of the parameters required for the original code. Subsequently, the base code is called to produce displacement maps, the four components of the strain tensor, and rotation and shear maps. Prior to applying the GPA automation, the image undergoes processing with a Wiener filter to enhance the smoothness of its outputs.

As a minor variation from the original code, we made slight adjustments to the use of the refinement reference area to suit the conditions of the global workflow better. The original code checks the sub-pixel refinement for the FFT of the reference area and selects these positions as the exact coordinates of the g vectors. This approach works reliably when the total number of pixels in the reference region allows for precision in defining the sub-pixel routine that marks the peaks' maximum. However, if the number of pixels in the reference region is small, this routine could inadvertently shift the maxima due to the low sampling rate of the reference region. Although this is not a common issue, it can occur when the reference is a small region of interest with just a few pixels. To address this, instead of directly finding new maxima from the crop, we compute the gradient shift of the geometrical phase within the crop. We then correct the original g vector using this gradient shift.

4.I.6.4.1 Proofs of the automated GPA

In this subsection, we present several representative examples of the output obtained from the fully automated GPA routine. These examples, which cover a wide variety of devices and sample morphologies, demonstrate that the GPA automation routine is reproducible both as an integral part of the workflow and as an independent module for strain analysis. Notably, the examples highlight how the automated g vector selection and adaptive resolution effectively reveal the rotation of domains, defects, and dislocations.

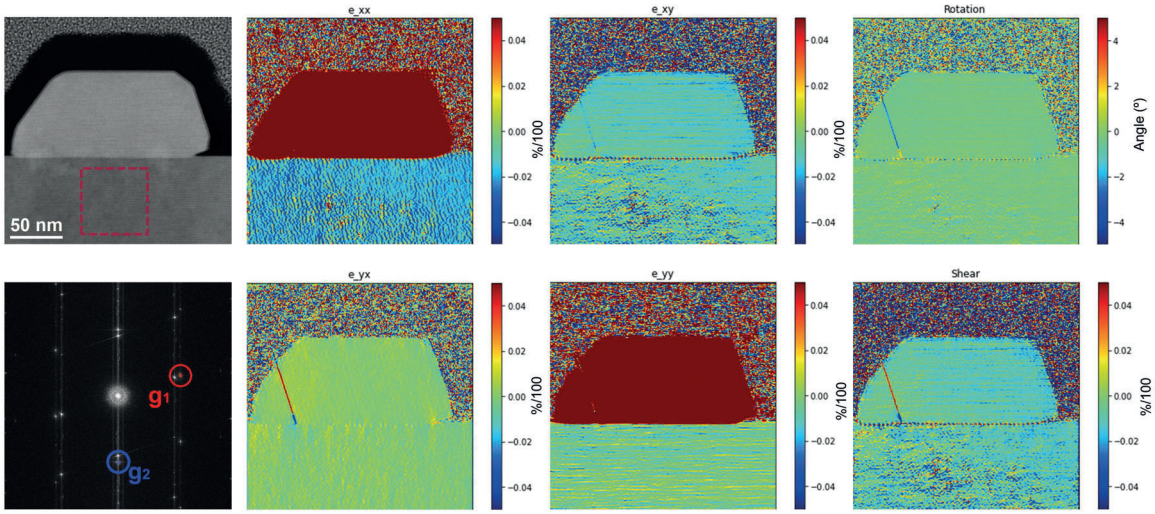


FIGURE 4.I.25: Geometrical Phase Analysis (GPA) automation, example 1. Highly mismatched, 10.3%, InSb nanowire on InP (transversal view, [110]). The automated GPA identifies the peaks corresponding to heteroepitaxy and adjusts the mask to include them in the analysis. This process highlights the array of equally spaced misfit dislocations within the interface and the stacking fault traversing the nanowire. It is the output of the phase analysis detailed in figure 4.I.24, a.

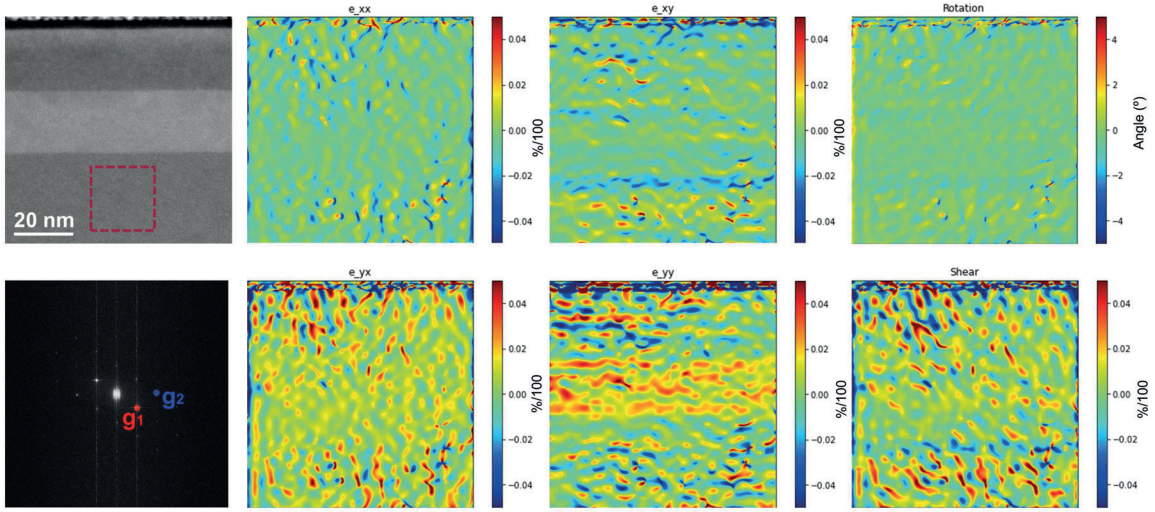


FIGURE 4.I.26: Geometrical Phase Analysis (GPA) automation, example 2. Ge quantum well sandwiched by two buffer layers of $Ge_{0.7}Si_{0.3}$, [110]. The automated GPA finds no neighbouring spots indicating mismatched epitaxy, as the epitaxy is perfect in the strained quantum well. In this case, the routine defines a 4 nm mask that trades off the resolution of the maps against their noise. It is the output of the phase analysis detailed in figure 4.I.24, b.

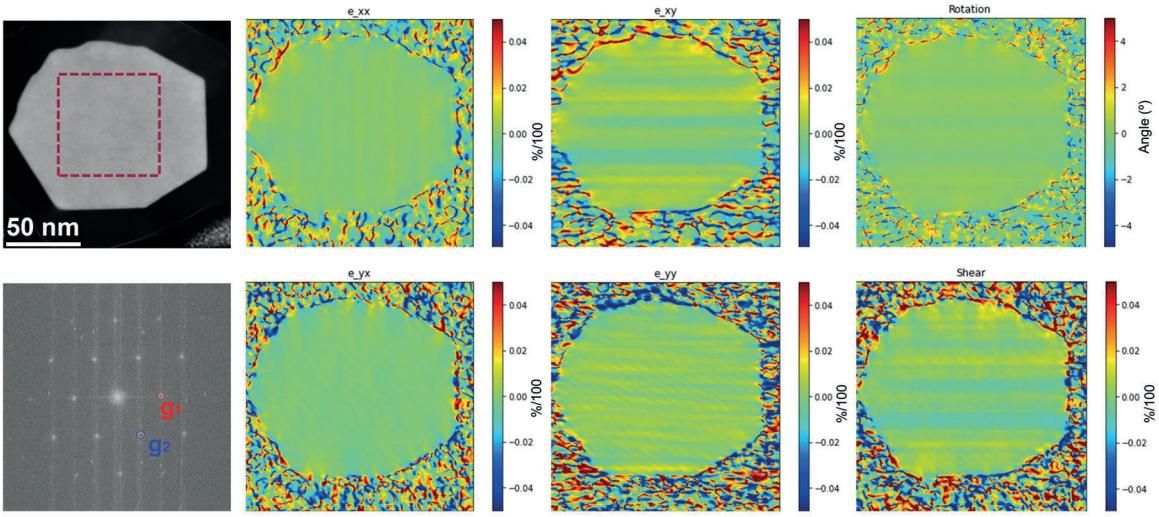


FIGURE 4.I.27: Geometrical Phase Analysis (GPA) automation, example 3. InAs (wurtzite, [0001]) nanowire surrounded by polycrystalline Al. In this case, no epitaxy is found: the nanowire is the main found phase, and no polycrystalline Al is perfectly in axis, for which no signs of heteroepitaxy are found. For that, the routine sets the 4 nm which smoothes the map down, showcasing there is no distortion in the nanowire structure.

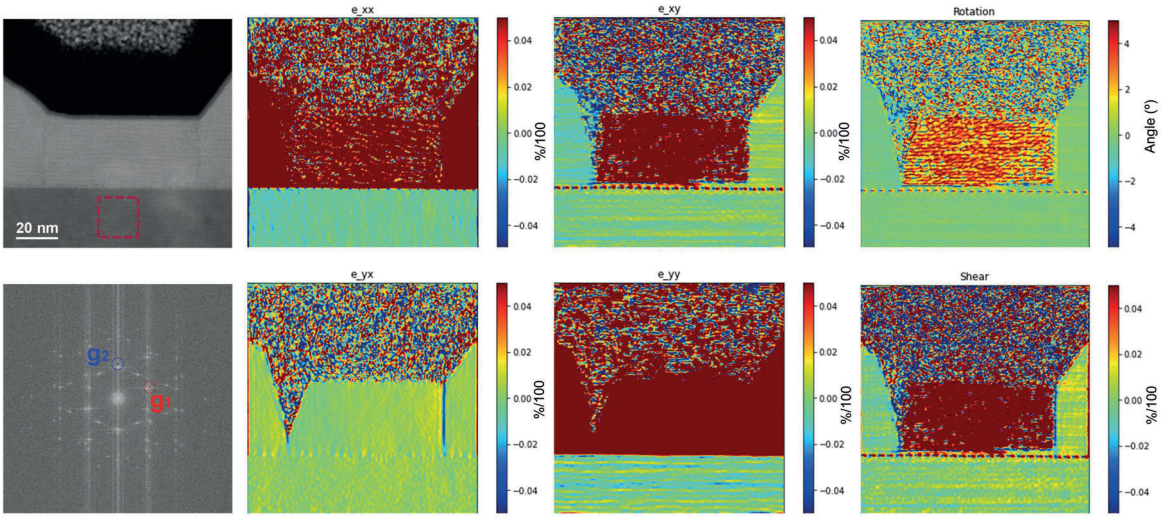


FIGURE 4.I.28: Geometrical Phase Analysis (GPA) automation, example 4. Highly mismatched, 10.3%, InSb nanowire on InP (longitudinal view, [110]). The system exhibits multiple twins that rotate the nanowire differently concerning the substrate. The substrate defines the reference for this GPA analysis. Notably, the rotation map reveals up to 4 different orientations of the nanowire. This effect is discernible due to the presence of multiple twins. The automated routine identifies mismatched domains that align with the substrate orientation. It then opens the mask to fit these domains accurately. Consequently, the remaining domains are coloured based on their relative rotation concerning the substrate.

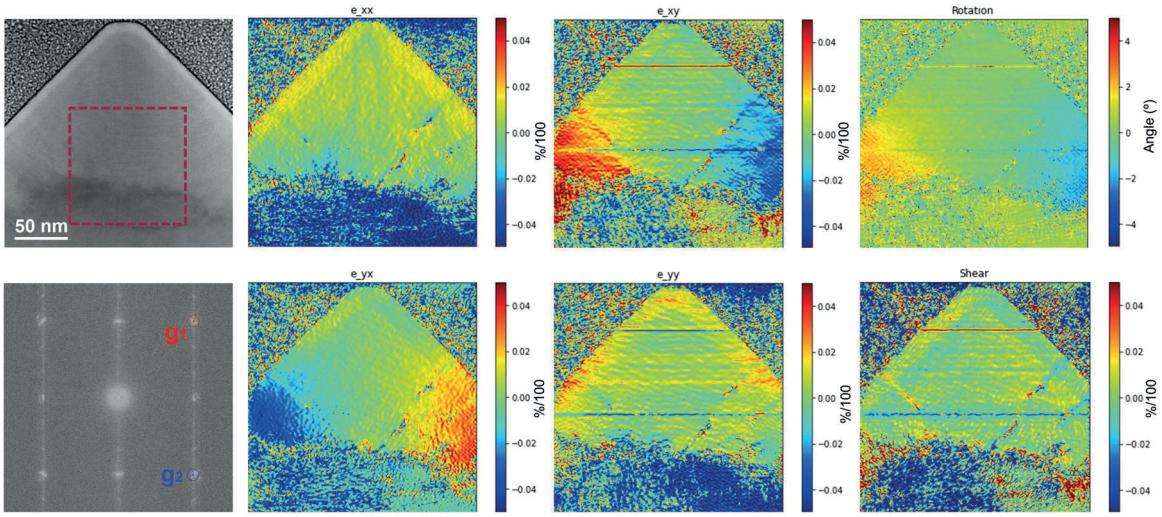


FIGURE 4.I.29: Geometrical Phase Analysis (GPA) automation, example 5. InAs – $Ga_xIn_{1-x}As$ (compositional gradient) on GaAs substrate, [001]. The small contrast differences due to the compositional gradient make the segmentation routine fail to distinguish the regions. As a result, only a main phase is found despite the Fast Fourier Transform (FFT) presenting separated spots indicating heteroepitaxy. Therefore, the 4 nm mask only partially captures the information of the substrate, for which its representation is noisy and heterogeneous. Despite the noisy substrate, the nanowire body is well-represented, highlighting a symmetric rotation of the lattice at the edges of the nanowire, as well as horizontal scanning artefacts.

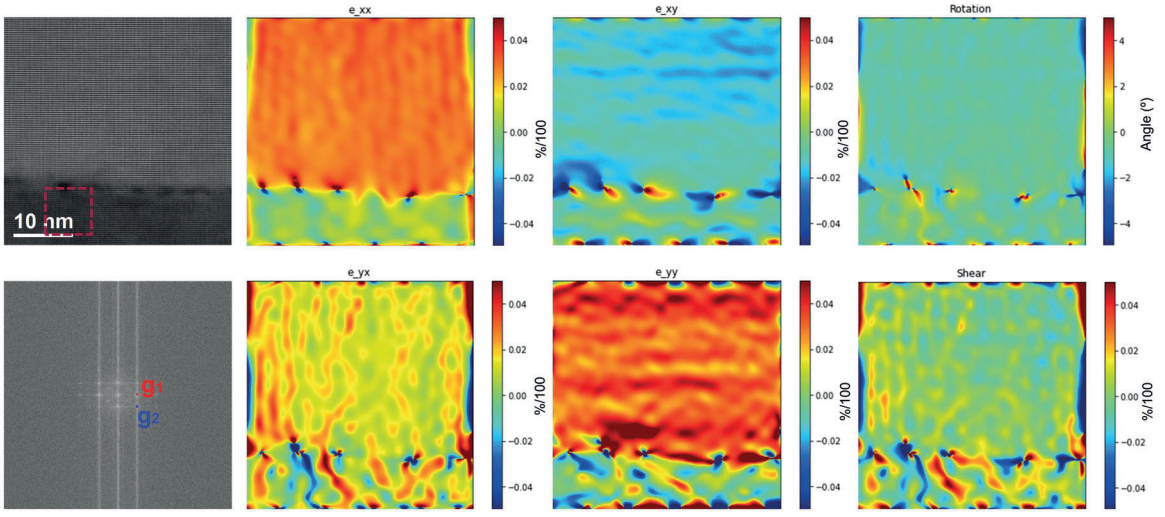


FIGURE 4.I.30: Geometrical Phase Analysis (GPA) automation, example 6. Mismatched InP on InAs nanowire, [112]. The automated GPA finds close neighbouring spots and a heteroepitaxy presenting a lower mismatch, 3.15%, than the InSb-InP (10.3%). A mask is then opened to capture both planes, which due to the high magnification of the micrograph, appear smaller than the previous study cases. Nevertheless, it is enough to smoothly map the strain components and highlight the dislocations.

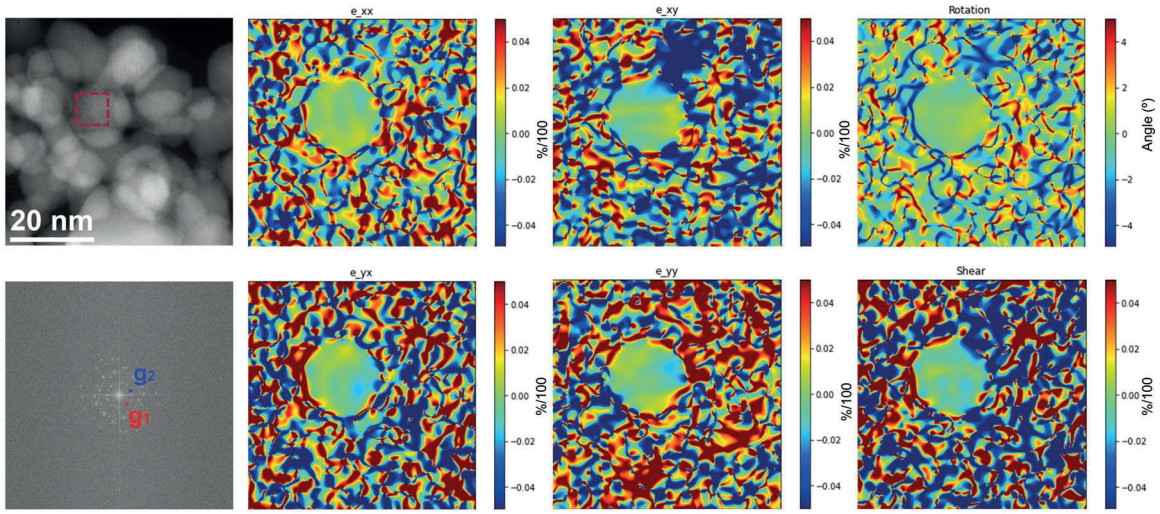


FIGURE 4.I.31: Geometrical Phase Analysis (GPA) automation, example 7. Cubic (Fd-3m) $CoFe_2O_4$ nanoparticles dispersion. In this effectively polycrystalline sample, multiple particle orientations (i.e., zone axes) are present. The GPA analysis selects one of these orientations. Specifically, the chosen particle has a [110] orientation. Notably, the resulting maps for these particles appear homogeneous, indicating that no strain significantly affects them. Interestingly enough, the algorithm selects g vectors that avoid surrounding spots from other particle orientations. It is important to note that the remaining areas of the map contain noise, as the selected spots only provide information within the local domain of the particle.

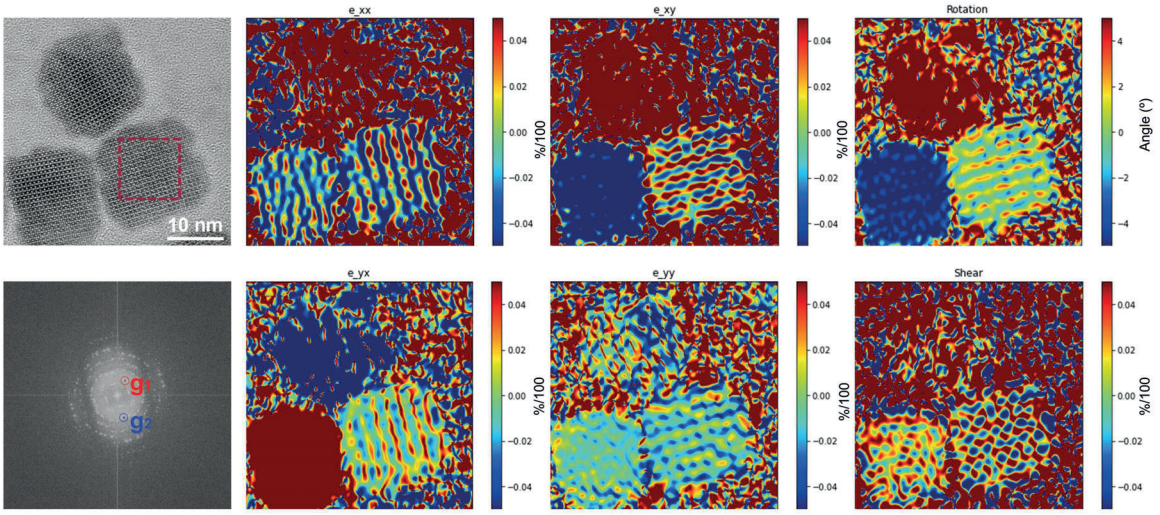


FIGURE 4.I.32: Geometrical Phase Analysis automation, example 8. Cubic (Pm-3n) CuTe nanoparticle dispersion. This interesting case features three particles, all oriented along the [001] axis but varying in-plane rotations (-4° and 11° from the reference orientation). The algorithm treats them as if they were epitaxially growing, given the close proximity of planes in the [001] zone axis. As a result, a mask is opened to embrace the spots from the two particles that deviate from the reference particle. The nuances of in-plane rotation mentioned are accurately represented in the rotation map.

4.I.7 Bragg filtering automation

Bragg filtering allows us to map crystallographic planes in the original micrograph. Thanks to all the knowledge acquired by the workflow in the previous steps, Bragg filtering can be easily automated to intuitively represent this knowledge.

Firstly, the process involves sorting the crystallographic phases that have been identified for each segmented region based on the accuracy of their fit with the candidate unit cells. Then, for each segmented region, we obtain a sorted list of crystallographic phases along with their respective indexed crystallographic planes. Using this information, we can place a circular binary mask of a specified radius on top of each spot assigned to a particular phase. By inverting the masked FFT, we can map these crystallographic planes in real space. If the size of the masked reciprocal space is appropriate, this process should precisely locate the origin of these crystallographic planes within the micrograph.

An appropriately sized binary mask will capture frequency information broadly enough to avoid missing regions with slight variations in the mapped interplanar distance, while also avoiding noise and satellite frequencies. To determine the mask size, we leverage the information used to compute the GPA mask and adapt it for this new purpose. If the GPA automation routine detects a mismatched heteroepitaxy or an equivalent situation in Fourier space, the Bragg filtering reduces this mask to a fraction of the GPA mask. This adjustment prevents interference from neighbouring spots that were retained for GPA. If no neighbouring spot is found, the GPA remains unchanged, capturing a single plane across its entire extension. Generally speaking, a reliable Bragg mask is approximately 5-10 times smaller than a GPA mask. Once computed, the same mask is used for all crystallographic planes and phases. It is important to note that only features represented in reciprocal space will have their equivalent mapped back to real space.

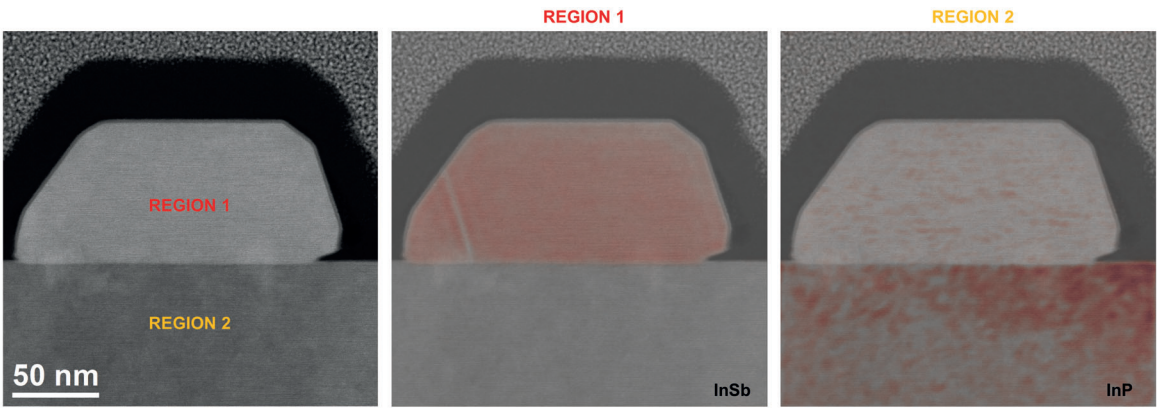


FIGURE 4.I.33: Automated Bragg filtering of a highly mismatched InSb nanowire on InP (transversal view, [110]). [37] Each region is identified with just one crystallographic phase, and mapped into the real space. Interestingly, the Bragg filtering of the InSb nanowire highlights the stacking fault as a less intense fringe because some of the crystallographic planes do not belong to it.

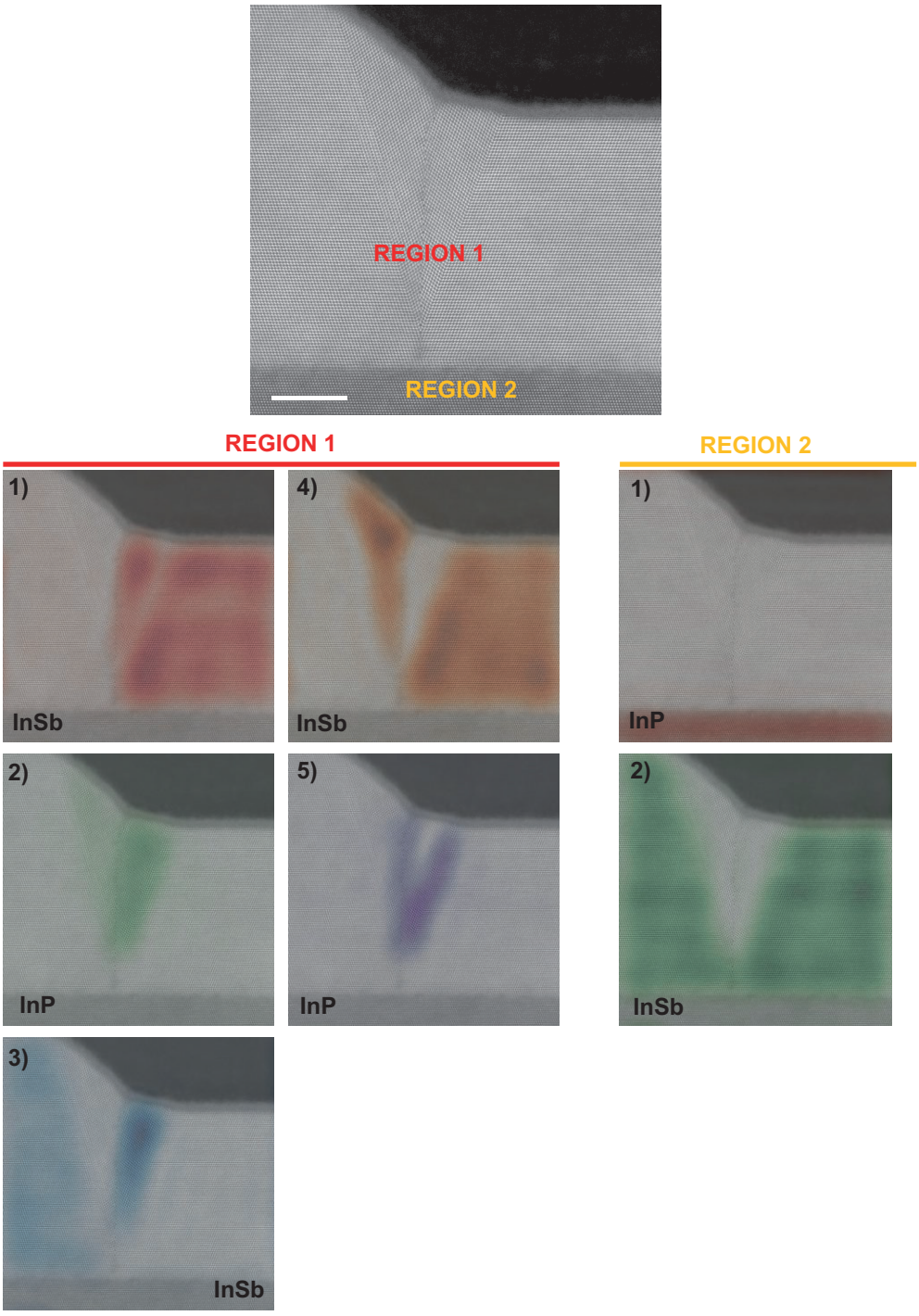


FIGURE 4.I.34: Automated Bragg filtering of a highly mismatched InSb nanowire on InP (longitudinal view, [110]). [37] In this case, every region outputs more than one phase. The number in each Bragg map sorts the phase from more likely to less (i.e., the smaller the number the better). Region 1 presents three twin boundaries generating four different orientations of the InSb. Importantly, the automated Bragg filtering showcases four regions orienting the phases differently and sharing crystallographic planes (as numbers two and five filter the same spots). Wrongly, numbers 2 and 5 confuse the InSb phase with InP, most likely due to the consideration of extra peaks not perfectly fitting the maximum of the peaks. The same happens with the InP substrate, which is correctly identified and located. Nonetheless, it adds an InSb region which Bragg-filters planes from the actual nanowire. Again, this is because of misplacing maxima or duplicating them due to non-Gaussian distributions of the peaks. The scale bar is 10 nm.

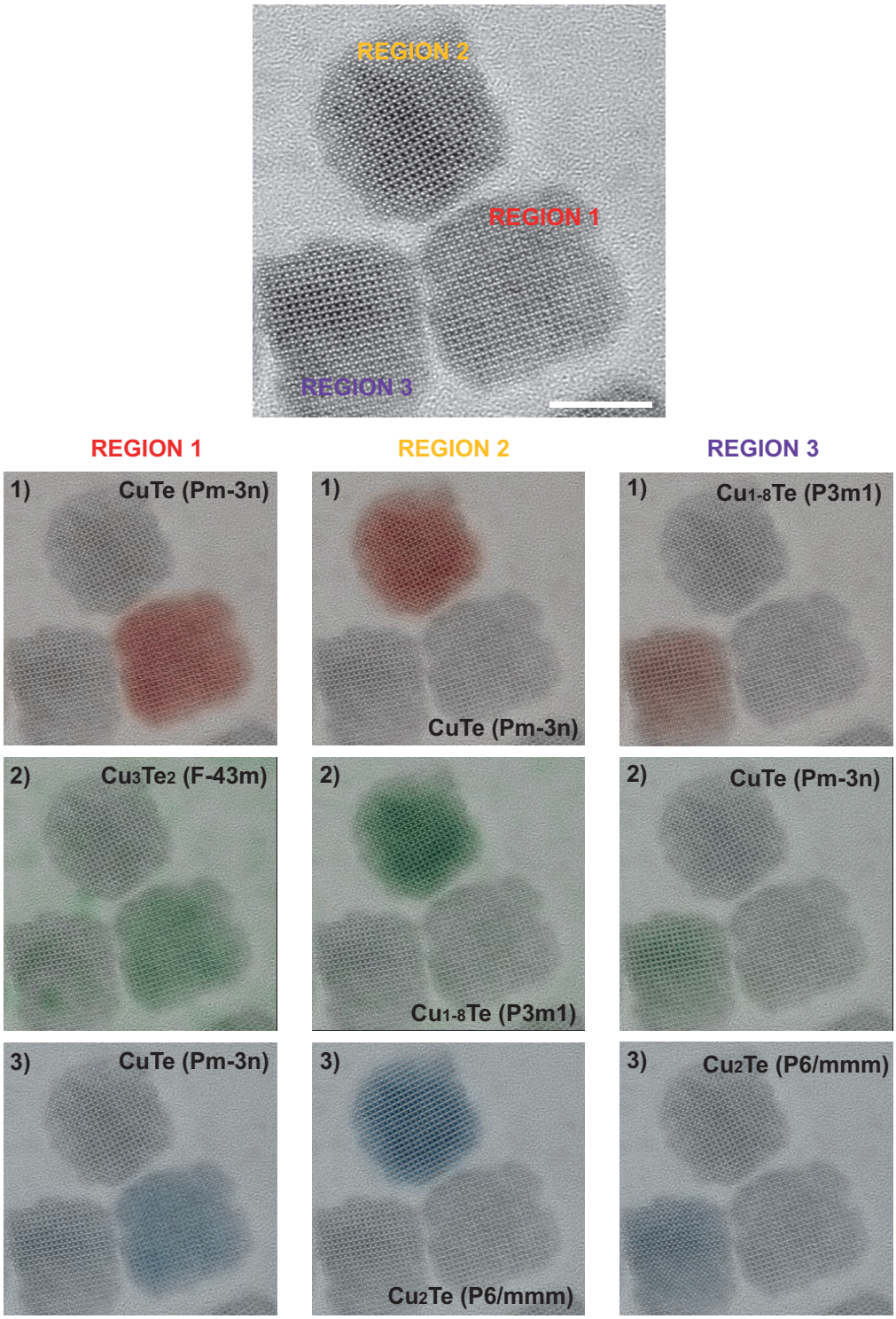


FIGURE 4.I.35: Automated Bragg filtering of three cubic (Pm-3n) CuTe nanoparticles oriented along the [001] axis. [39] In the three regions coinciding with the three particles, three phases are sorted from better to worse goodness of fit with different model unit cells. The overall score leans the balance in favour of the cubic Pm-3n as the most likely model explaining the experimental data. In this case, the region-wise Bragg maps outputting the same intensity are mapping the same crystallographic planes but assigned to further possible candidate phases that do not fit as well as the very first model unit cell. This is especially perceptible in region 2. Generally, if the automated routine finds a better fit and more representative planes for a given phase, the Bragg map will become more intense. The scale bar is 10 nm.

Figures 4.I.33, 4.I.34 and 4.I.35 showcase how this knowledge can be useful in particular device configurations, especially in heterostructures and polycrystalline samples (or their equivalents, such as particle dispersions). These examples illustrate how Bragg mapping can reveal the distribution of multiple crystal phases in different locations or orientations within a single segmented region, providing a realistic representation of the device. In cases where multiple phases are identified within a single segmented region, it is essential to carefully evaluate the relative location of the plane's peaks contributing to the Bragg-filtered image. Misplaced detected peaks, such as duplicated peaks arising from elongated tails from scanning artefacts, can erroneously shift the identified phase (fig. 4.I.34). Alternatively, the same set of indexed planes may lead to identifying another crystallographic phase with lower probability, indicating significantly worse fitting than the most likely choice (fig. 4.I.35). In any case, if doubts remain unresolved, complementary compositional information, such as stoichiometry, can help tip the balance towards the correct phase.

4.I.8 Automated phase identification proofs

This section showcases results from paradigmatic cases, spanning from segmentation to the phase identification. Each selected example highlights a characteristic crystallographic particularity, emphasising the general nature of the workflow presented thus far (with the mentioned limitations). Indeed, this workflow has undergone extensive testing across multiple device types, configurations, crystal systems, materials, and imaging conditions. The results presented below represent only a concise selection, chosen for the sake of brevity.

Remarkably, these results are obtained entirely automatically. The only manual intervention required is in selecting the segmentation algorithm (such as contour detection vs SAM, as discussed in section 4.I.1) and choosing the segmentation-peak finding approach (either crop or mask, as described in section 4.I.3). Optionally, one may limit the candidate crystal phases in the database based on prior knowledge (for example, hints about composition). Comparing with fewer model cells optimises computing time and increases the chances of correctly identifying the phase.

- (Figure 4.I.36) STEM-HAADF micrograph of a longitudinal cut of an InSb nanowire on InP substrate. [37] Characteristically, the InSb nanowire presents three twins at the kink, sharing a common origin and displaying up to three different in-(image)-plane rotation variations of the InSb. Both InSb and InP, in the zinc-blende structure, serve as candidate cells. The algorithm excels at identifying the three rotations of InSb within the same zone axis [011]. The evidence lies in the region-wise selection of the crystal plane frequencies during Bragg filtering. This process effectively separates the rotated domains, thereby highlighting the locations of twin boundaries. In this specific case,

the "mask" segmentation-peak finding approach is essential for detecting this effect, as it occurs within an independent region of the segmentation result.

- (Figure 4.I.37) HR-TEM micrograph of three square CuTe nanoparticles, with a "mask" intercommunication setting. [39] The most remarkable aspect is the flawless identification of the correct crystal phase: CuTe, Pm-3n space group, specifically at the [001] zone axis. The presence of the CuTe Pm-3n phase, highlighted in red within the white square, unequivocally confirms its identification without any secondary and less likely candidates. The significance of this crystal phase lies in its original discussion during the crystallographic characterisation, which prompted a thorough investigation. [100] The CuTe Pm-3n phase was directly extracted from this study and compared with other possible CuTe phases across different crystal systems (including cubic, orthorhombic, hexagonal and tetragonal). The workflow successfully identified the same phase and zone axis for all three particles, revealing their particle-wise in-(image)-plane rotation (as observed in the GPA analysis, as shown in Figure 4.I.32). The Bragg filter further reinforces these findings by independently colouring the particles based on their particle-wise crystal planes (as illustrated in Figure 4.I.35).

- (Figure 4.I.38) Same HR-TEM micrograph of three square CuTe nanoparticles, as in figure 4.I.37, with a "crop" intercommunication setting instead. [39] In this case, we revisit the same candidate phases and scientific question as in the previous scenario (see Figure 4.I.37). However, our focus shifts to comparing the results obtained using the "crop" intercommunication setting with those from the "mask" setting (as shown in Figure 4.I.37). While the "crop" approach improves peak finding, it introduces a trade-off: the accuracy of eventual phase identification. Although it is easily detectable by comparing the peaks indexation (i.e., red peaks vs green peaks), two out of the three particles exhibit false positives. The explanation lies in the loss of sampling precision due to cropping a fragment of the original image. While the real space pixel size remains consistent during cropping, reducing the total number of pixels alters the reciprocal space sampling, resulting in larger steps between reciprocal space pixels. In other words, the precision of peak finding and its translation into real-space interplanar distances diminishes due to fewer pixels representing the same reciprocal distances. Consequently, there is less agreement between the theoretical candidate and the experimental data because the cropping factor now undersamples the experimental distances. Smaller crops further exacerbate the loss of precision. However, the outcome is not critical; the second most likely phase still exhibits more detected peaks, confirming it as the correct crystal phase. In this context, SAM plays a crucial role in providing a representative segmentation of HRTEM-imaged nanoparticles, leading to positive results.

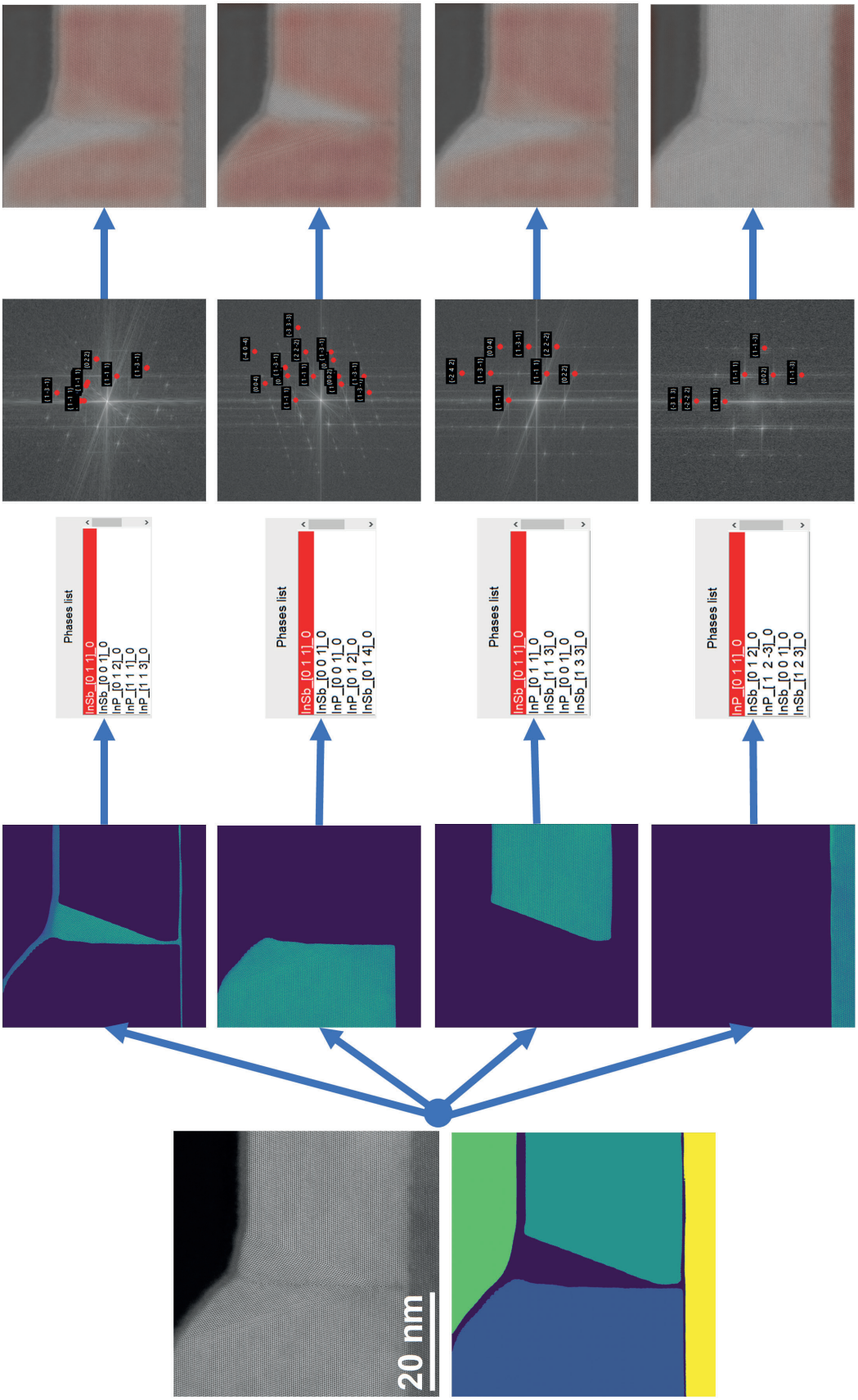


FIGURE 4.I.36: Phase identification proof in a longitudinal cut of a twinned InSb nanowire on InP, imaged by Scanning Transmission Electron Microscopy (STEM). [37]

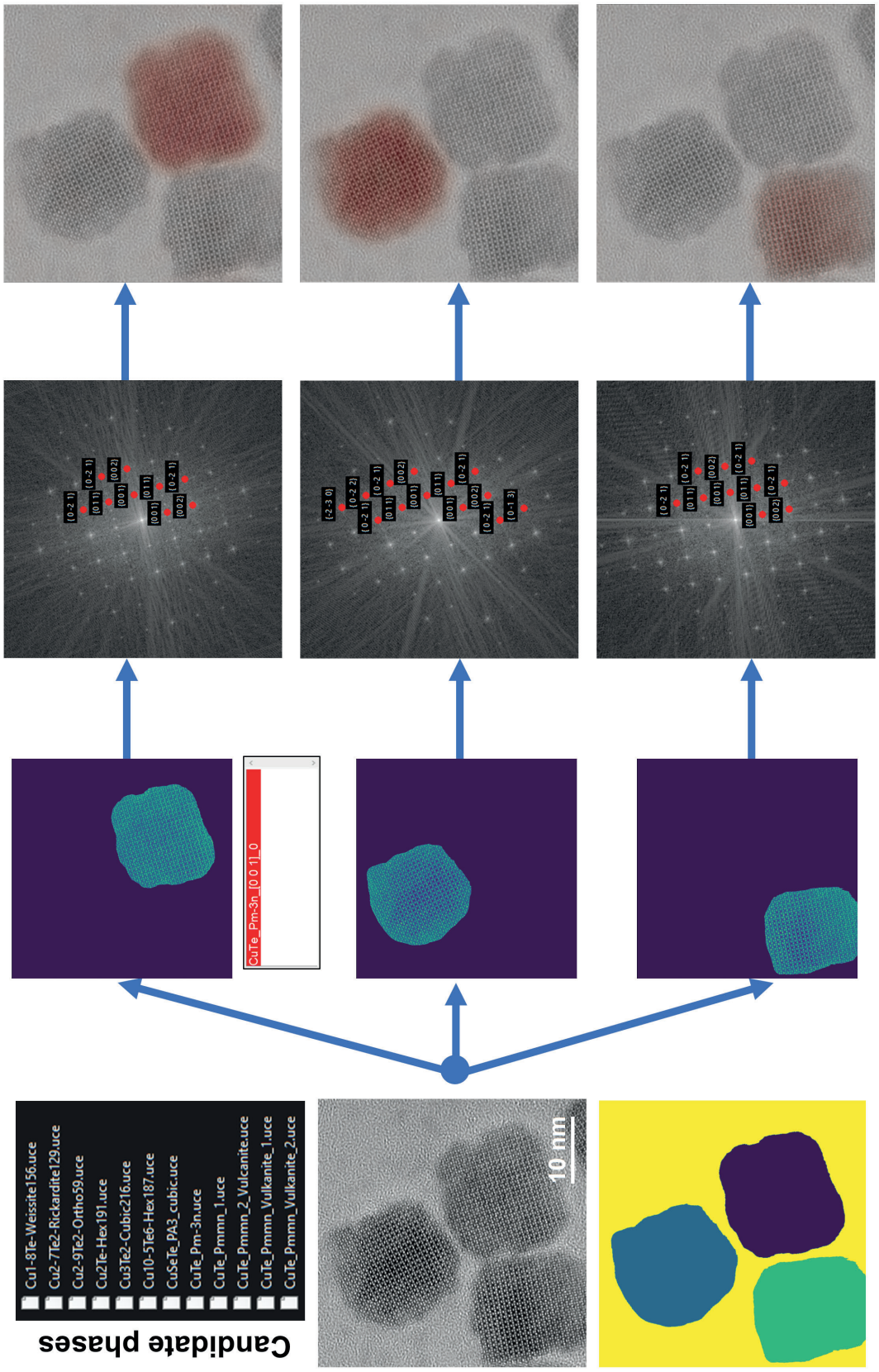


FIGURE 4.I.37: Phase identification proof in a parallel beam High-Resolution Transmission Electron Microscopy (HRTEM) image of three CuTe nanoparticles, with the "mask" segmentation-peak finding intercommunication. [37]

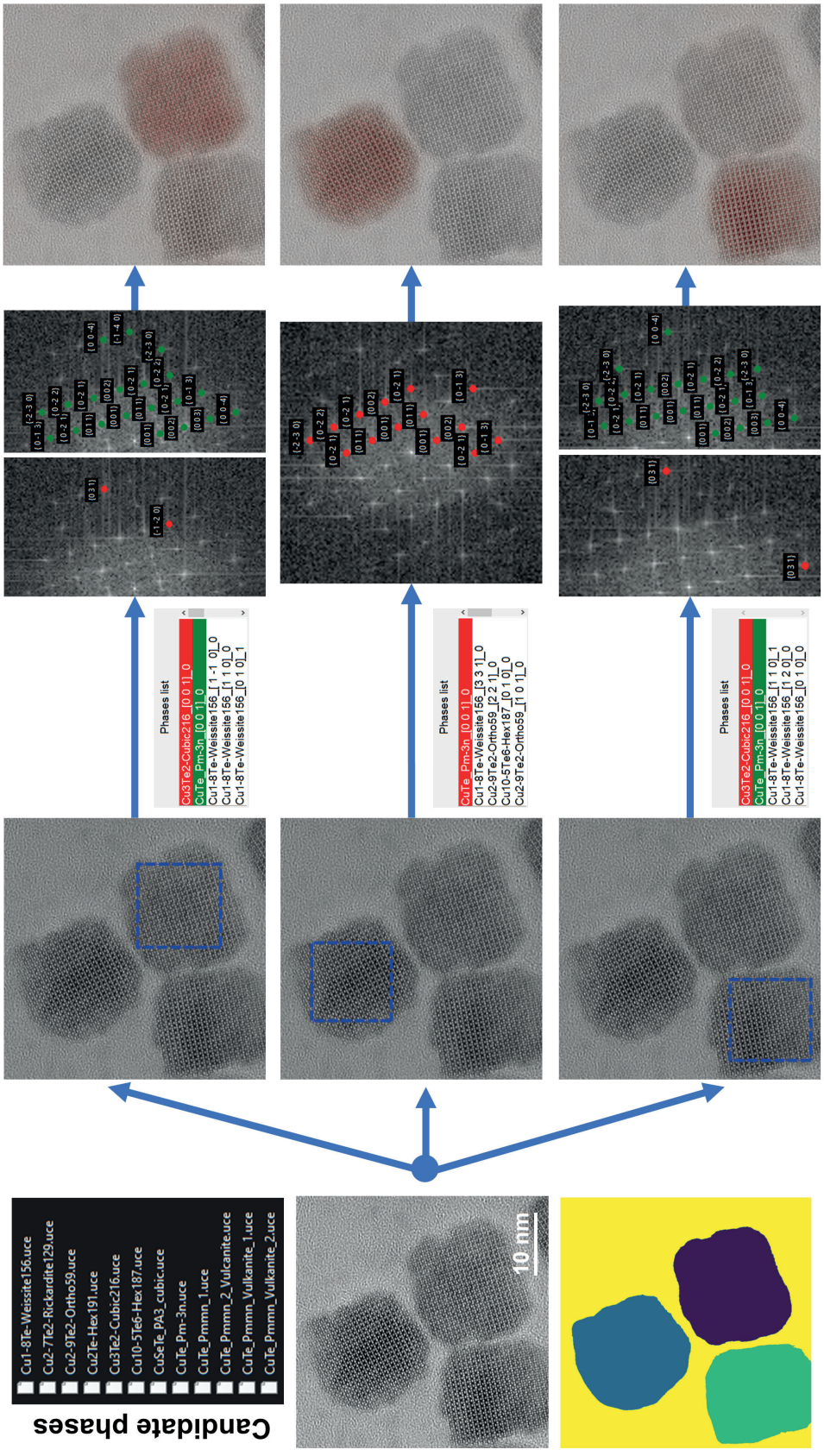


FIGURE 4.I.38: Phase identification proof in a parallel beam Transmission Electron Microscopy (TEM) image of three CuTe nanoparticles, with the "crop" segmentation-peak finding intercommunication. [37]

- (Figure 4.I.39) STEM-HAADF micrograph of a transversal cut of an InSb nanowire on InP substrate. [37] Characteristically, the InSb nanowire presents a diagonal stacking fault. Two candidate cells, InSb and InP (both zinc-blende structures), are considered. The phases are distinctively found correctly for both "crop" and "mask" settings, with minor differences in the peak finding, but equivalent indexation. Interestingly, the Bragg filtering accentuates the presence of the stacking fault by shadowing the iFFT intensity in the coinciding region.

- (Figure 4.I.40) STEM-HAADF micrograph of a transversal cut of a Ge QW, $Si_{0.3}Ge_{0.7}/Ge/Si_{0.3}Ge_{0.7}$ from device GSQ_4 from Chapter 3. Based on prior knowledge of the device, the candidate crystal phases are the diamond structure of both Si and Ge. Not surprisingly, the phase identification outputs Ge [011] as the most likely phase in the three regions. Thus, the difference in interplanar distance between Ge and $Si_{0.3}Ge_{0.7}$ is too minimal to be detected, at least in the displayed resolution. The truth scenario is that the Ge QW is strained to match the lattice of the buffer $Si_{0.3}Ge_{0.7}$, for which not even this difference is expected, as remarked in the Bragg filtering. As a result, as $Si_{0.3}Ge_{0.7}$ is closer to Ge than Si in cell parameters (see Chapter 3), our output is directly Ge. The compositional information needs further refinement to output the exact characterisation. This process is explained in the paragraphs below.

- (Figure 4.I.41) The same philosophy as in figure 4.I.40 for the Ge QW (buffer-QW-buffer of $Si_{0.3}Ge_{0.7}/Ge/Si_{0.3}Ge_{0.7}$), but a STEM-HAADF of a Si QW ($Si_{0.7}Ge_{0.3}/Si/Si_{0.7}Ge_{0.2}$), corresponding to device SSQ_3 from Chapter 3. [54] The candidate phases include diamond cubic Si and Ge. The situation is similar to the previous case, where the buffer $Si_{0.7}Ge_{0.3}$ strains the Si layer to match its cell parameter. An interesting scenario arises when the most likely output is Ge instead of the correct Si (since Si is the closest to $Si_{0.7}Ge_{0.3}$). The Si [011] zone axis also falls within the accepted tolerance, but it is identified as the second most likely crystal phase. The segmentation and peak finding perform nearly flawlessly in this case. Additionally, the "mask" setting ensures optimal sampling of the reciprocal space based on the initial raw data. However, there exists a slight calibration shift in the image that affects the true output. The algorithm functions as intended, but the issue lies in the original pixel size assigned to the image. Therefore, calibration is a sensitive parameter, especially in STEM, emphasising the need for accurate calibration to achieve true positives and stay within the allowed tolerance described in section 4.I.5.

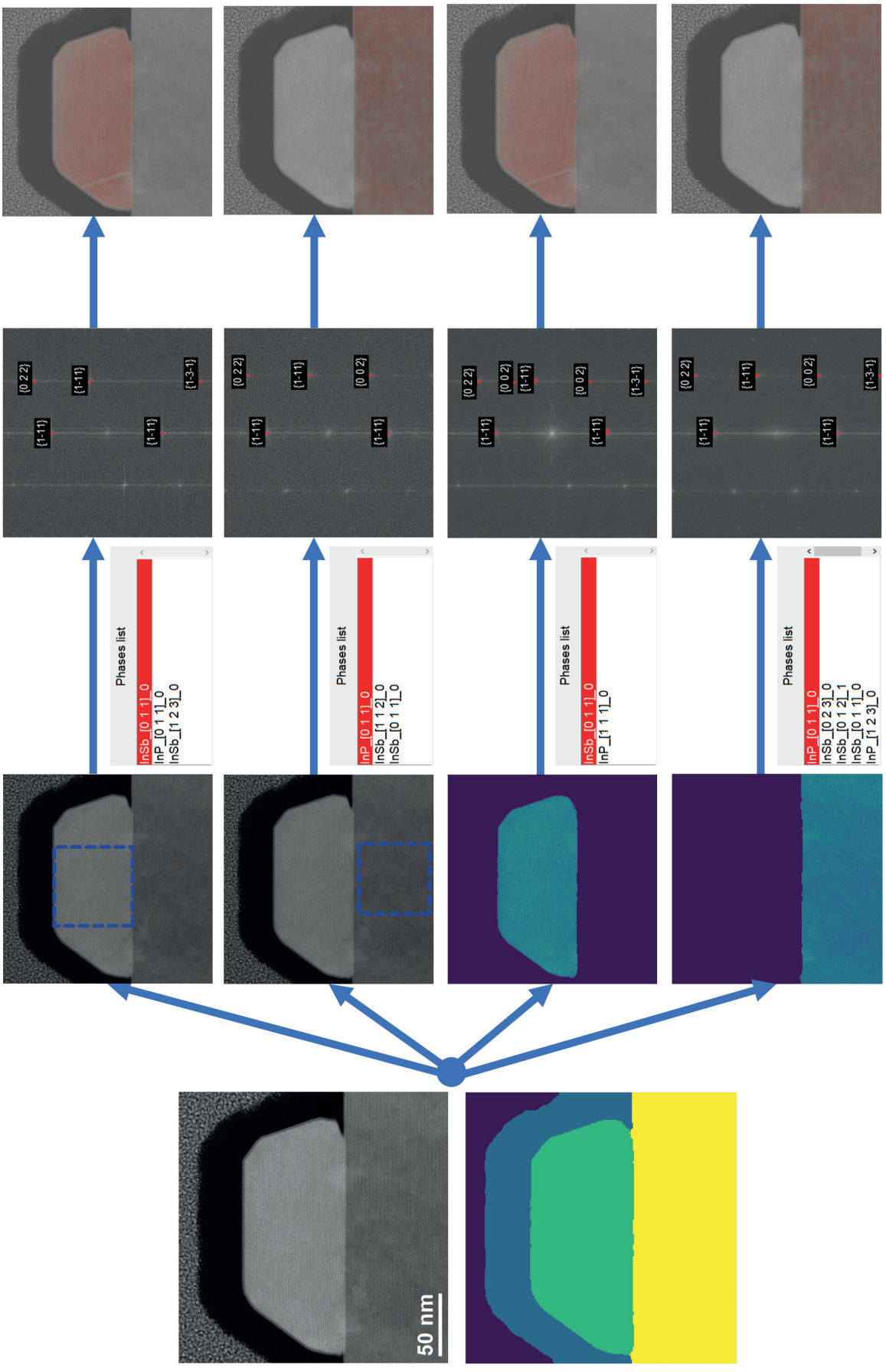


FIGURE 4.I.39: Phase identification proof in a transversal cut of InSb nanowire, presenting a stacking fault, on InP, imaged by Scanning Transmission Electron Microscopy (STEM). [37]

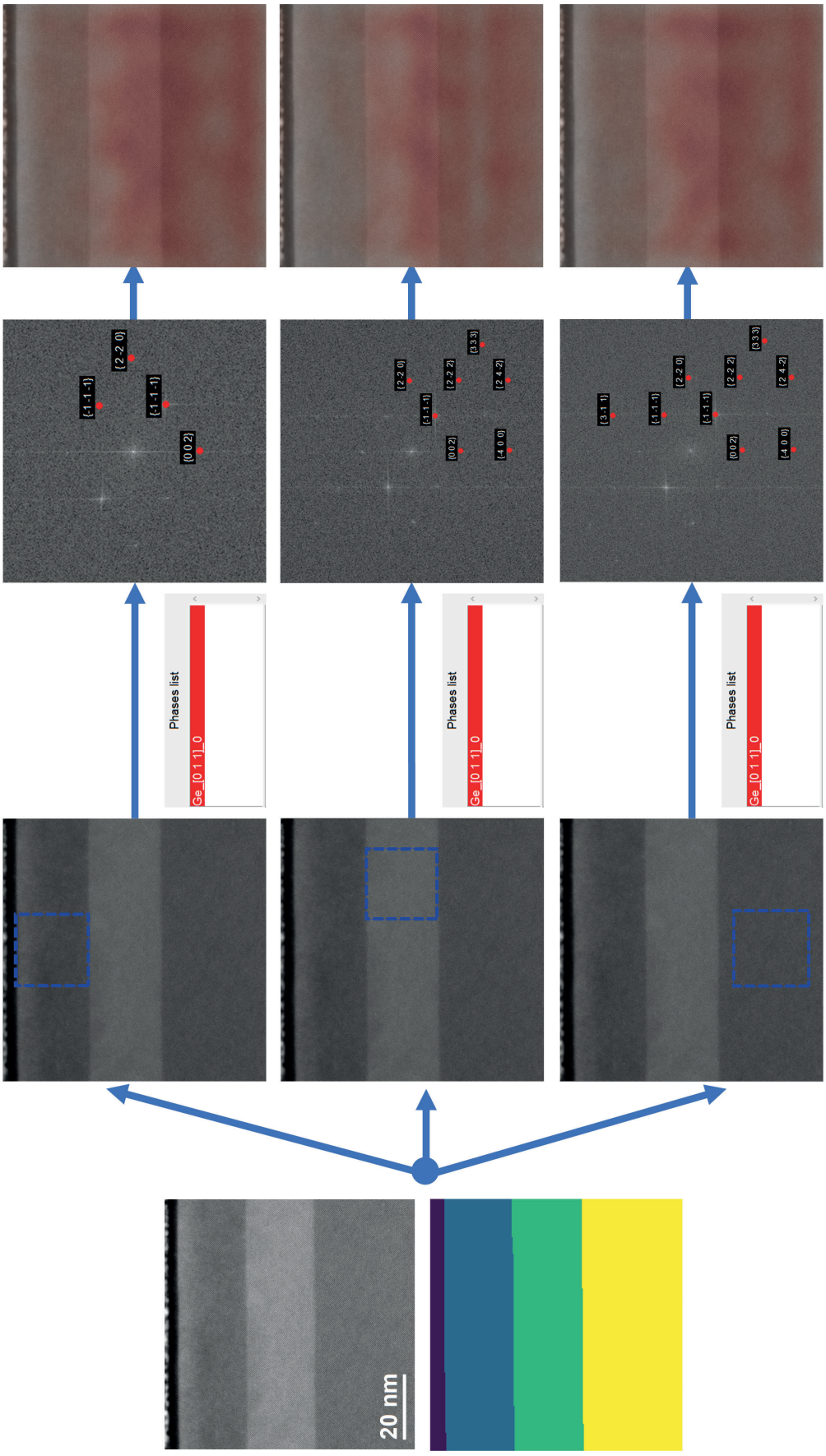


FIGURE 4.I.40: Phase identification proof in a $Si_{0.3}Ge_{0.7}/Ge/Si_{0.3}Ge_{0.7}$ quantum well imaged by Scanning Transmission Electron Microscopy (STEM). [38]

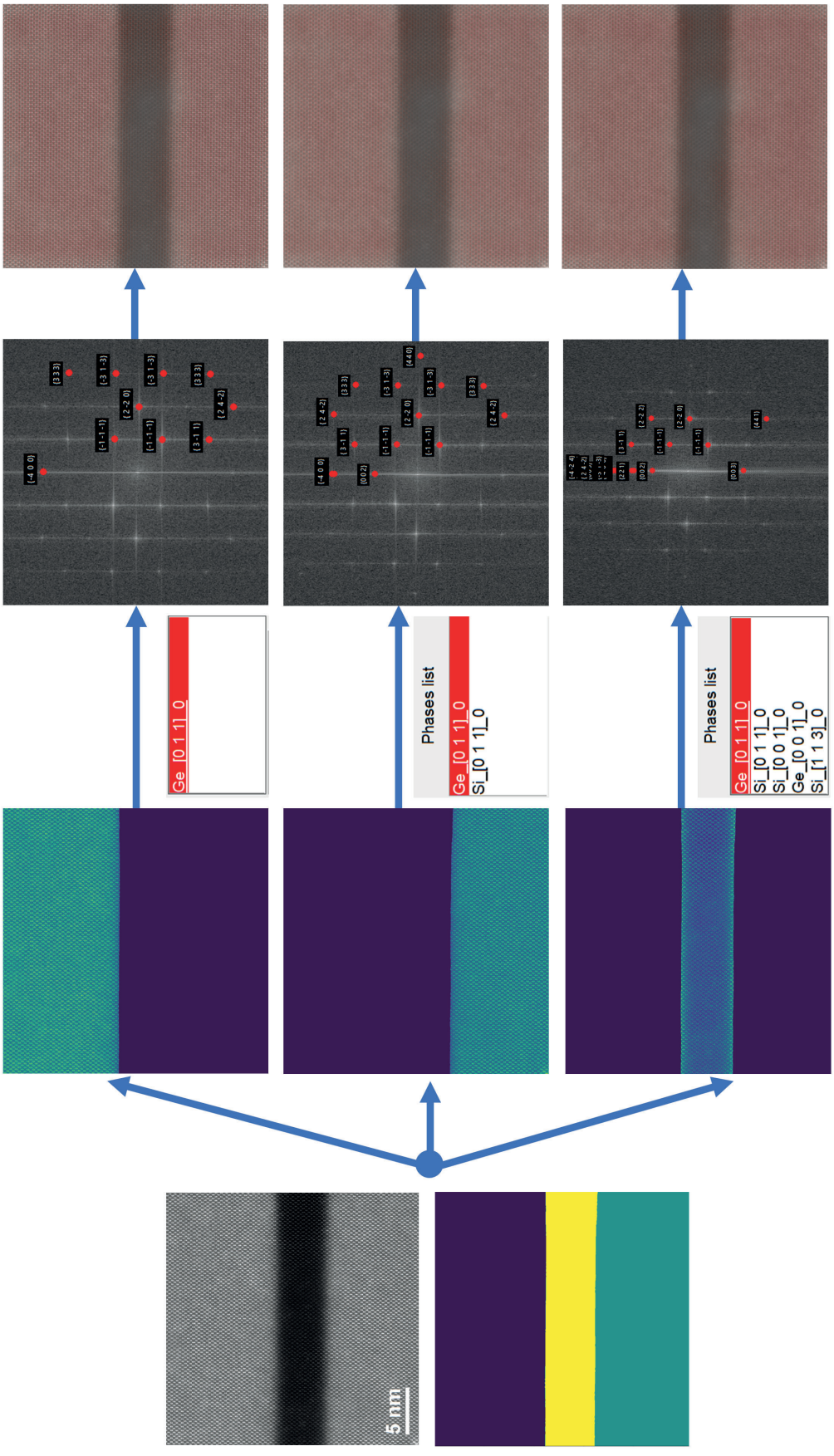


FIGURE 4.I.41: Phase identification proof in a $Si_{0.7}Ge_{0.3}/Si/Si_{0.7}Ge_{0.2}$ quantum well imaged by Scanning Transmission Electron Microscopy (STEM). [38]

- (Figure 4.I.42) Two STEM-HAADF micrographs of a dispersion of $CoFe_2O_4$ nanoparticles. [101] In this scenario, the dispersion contains only a few particles perfectly exposing a clear zone axis. Specifically, there are two particles in the first micrograph and one in the second. The algorithm successfully detects the zone axis associated with these particles, whose phase corresponds to the spinel $CoFe_2O_4$ crystal phase (space group Fd-3m, number 227). This example highlights the workflow's ability to determine the spatial orientation of crystal phases, whether it involves a single identified pair of crystal planes or multiple frequencies correlating with a single zone axis. However, the algorithm does not detect any other zone axes due to the absence of found peaks: For instance, in the 4^{th} row, where precision and recall are both zero, the algorithm cannot identify an already nonexistent zone axis. If noisy and random peaks are mistakenly selected, the algorithm either repeats the same behaviour or associates them with high-order, unphysical zone axes.

- (Figure 4.I.43) Same CuTe cubic system as in figures 4.I.37 and 4.I.38, also imaged by HRTEM. [39] The segmentation process plays a crucial role, particularly in how the choice of the "crop" FFT setting impacts our understanding of the sample. In this context, one of the segments groups two particles. Consequently, a single crop may only capture information from one particle. Figure 4.I.43 illustrates this scenario, where the crop characterises only the right particle. Interestingly, the Bragg filtering also identifies the two particle groups as segmented. As a result, we can conclude that the in-(image)-plane relative orientation of both particles is identical. In summary, by combining various sources of information, we effectively compensate for their individual weaknesses and achieve a comprehensive characterisation of the micrograph.

- (Figure 4.I.44) Phase identification proof in a STEM-HAADF dataset of an InSb nanowire on InP. The system is similar to the one appearing in Figure 4.I.36. [37] The difference between both cases is that the current NW is too big to unlock atomic resolution in an overview image. As a result, the multi-image approach explained in section 4.I.4 is necessary. With the four-image dataset (overview + three atomically resolved micrographs of a portion of the device), we do the following. We first correlate the high-resolution images with their overview and its segmentation, as indicated by the spatial origin of the high-magnification images from the overview (coloured boxes). Next, the phase identification process, equivalent to the single image processing, is done and the most likely phases are extracted. Then, we chose to output the first and, if any, the second most likely crystal phase found. Next, these phases are mapped in the low magnification overview based on the centre of the crops from where the atomic resolution information is extracted, as pinpointed by the red circles.

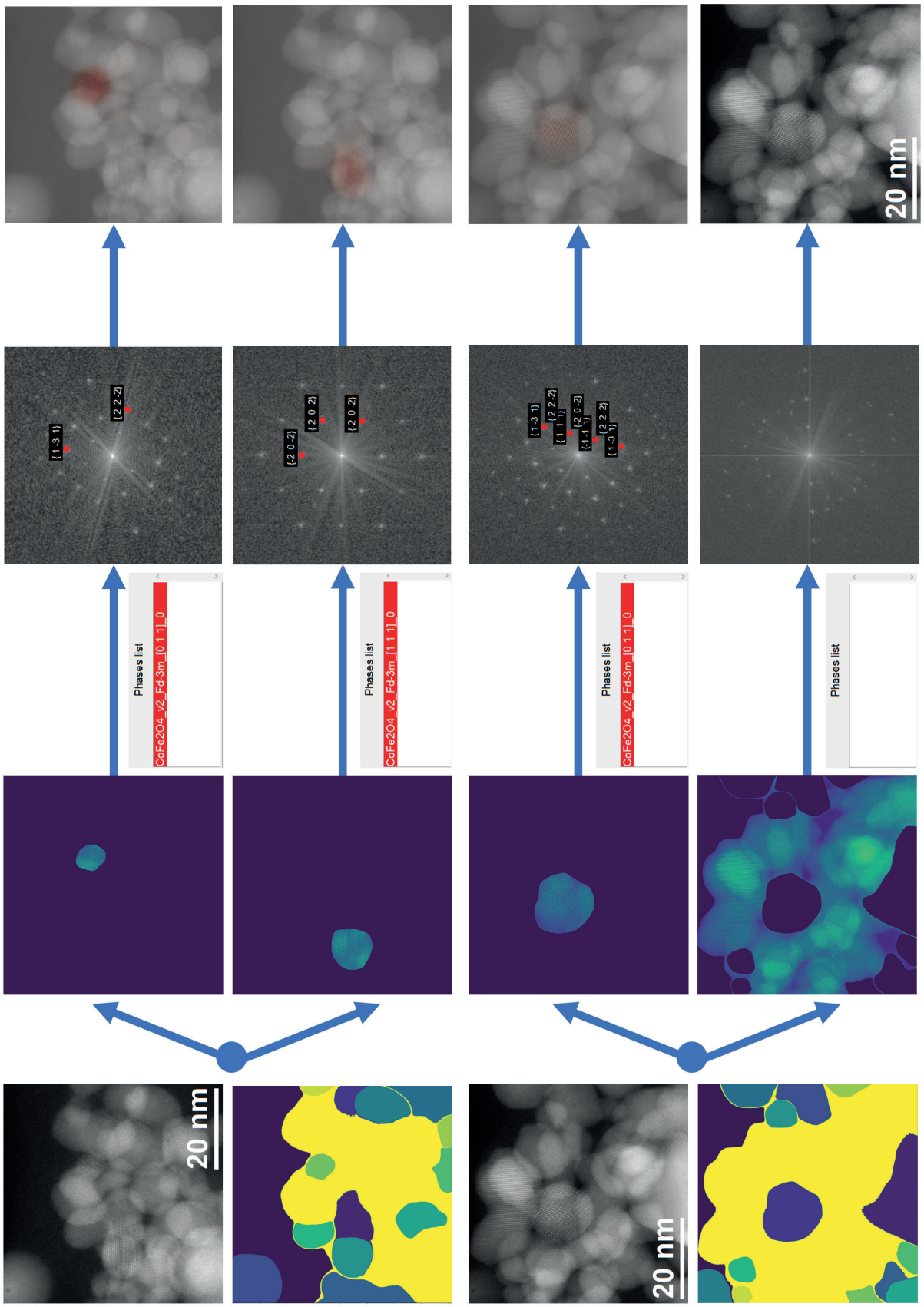


FIGURE 4.I.42: Phase identification proof in two Scanning Transmission Electron Microscopy (STEM) images of a dispersion of $CoFe_2O_4$ nanoparticles. [101]

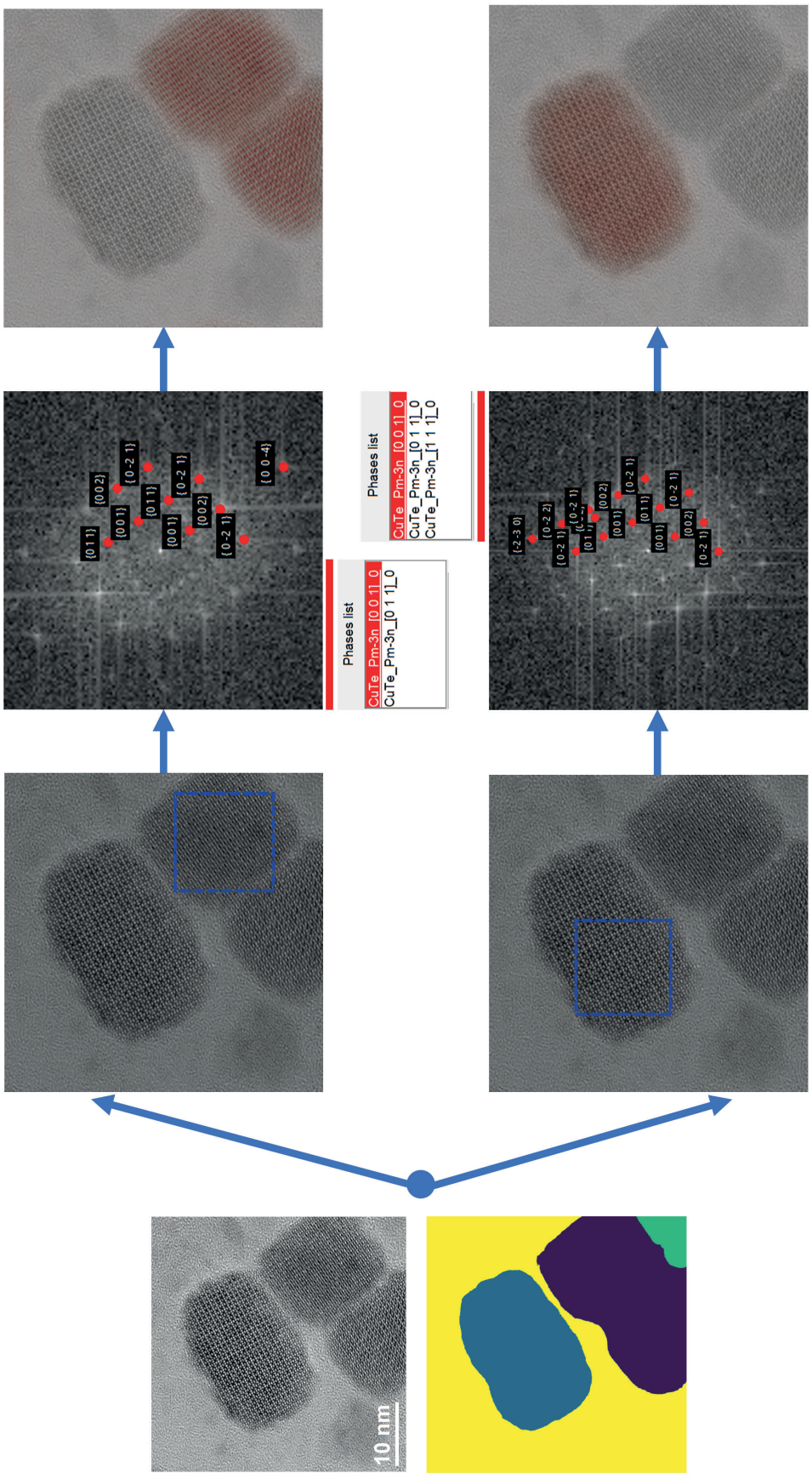


FIGURE 4.I.43: Phase identification proof in a parallel beam Transmission Electron Microscopy (TEM) image of three CuTe nanoparticles, with the "crop" segmentation-peak finding intercommunication and its compensation based on Bragg filtering. [37]

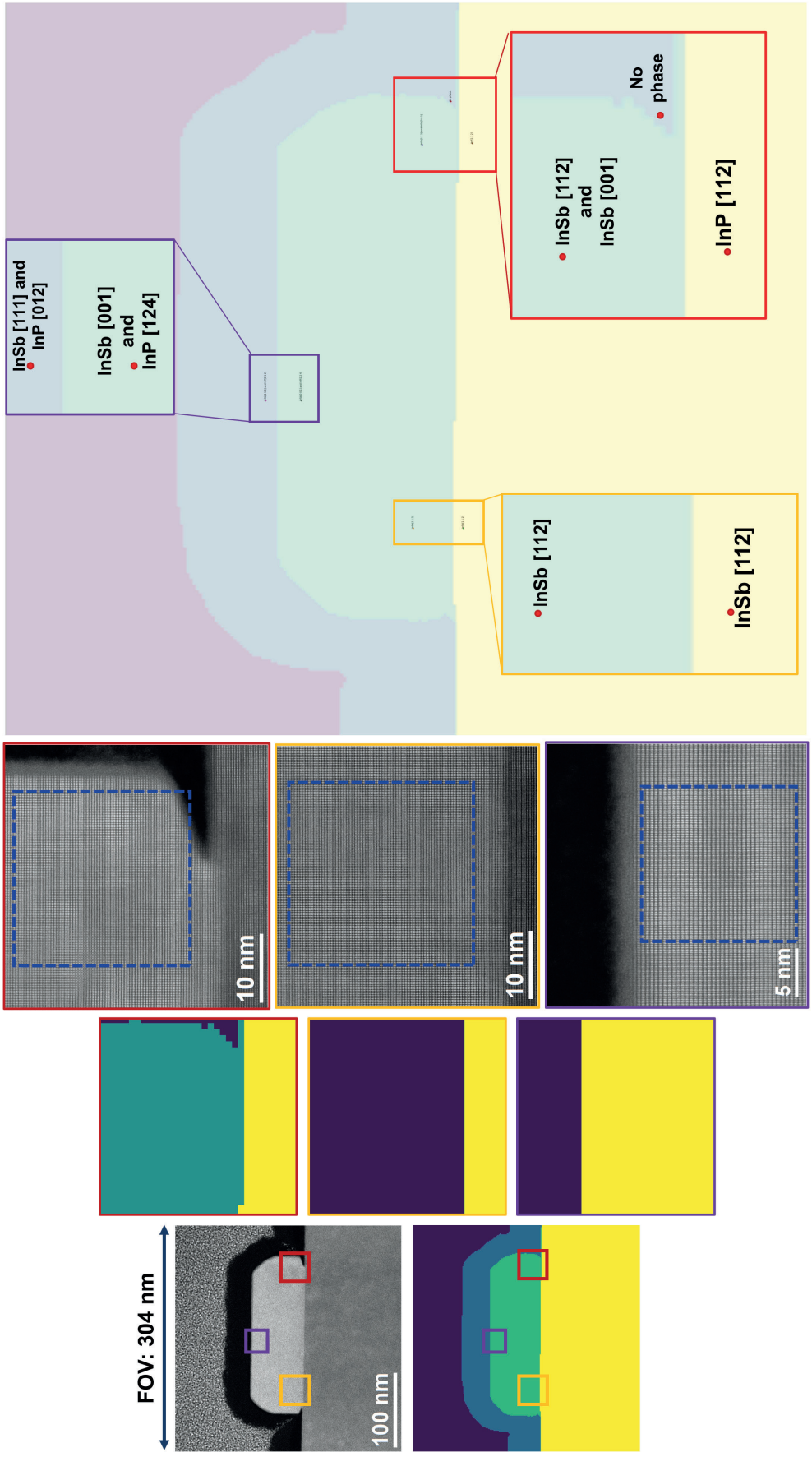


FIGURE 4.I.44: Phase identification proof of a full Scanning Transmission Electron Microscopy (STEM) dataset through the low to high magnification correlation: transversal cut of an InSb nanowire on InP substrate. [37]

The most common phase found in each segmented region should be interpreted as the resulting identified phase for that segment. Therefore, InSb [112] is the most likely candidate for the NW. In the substrate, there is a choice between InSb [112] and InP [112]. To determine the correct phase, one could identify the image where the largest area of the substrate can be cropped. This criterion improves the accuracy of positive identifications. In this specific case, the micrograph of the NW edge (indicated by the red margin) suggests that the phase is InP as visualized along the[112] zone axis. A similar situation occurs with the shell, where the absence of a crystalline phase ("No phase") competes with InSb [111], which is raised due to imprecise peak finding. In summary, this example highlights that a reliable identification based on the most common output criteria is favoured when there is a large number of images in the dataset. This principle applies to the next examples where selecting a crystal phase and orientation per segment is straightforward.

- (Figures 4.I.45 (Part 1) and 4.I.46 (Part 2)) The full dataset now corresponds to the huge Ge QW with PtSiGe contacts, namely GSu_10, whose overview's FOV is larger than a micron. [70] The idea of the multiscale template matching chain is reviewed in figures 4.I.19 and 4.I.21, from section 4.I.4, and its interpretable results are gathered in the double figure 4.I.45 (part 1) + 4.I.46 (part 2). The first part (shown in Figure 4.I.45) demonstrates the multiscale template matching chain combined with the segmentation of representative low magnification micrographs form the dataset. These low magnification micrographs serve as both the templates for the overview on the left and the queries for the high magnification micrographs on the right. Interestingly, the segmentation of the higher magnification images does not originate from those images themselves but rather from the lower magnfication ones. The contours in the lower magnification images are sharper, making them more suitable for Canny edge detection (as discussed in Section 4.I.1).

  This information chain is exploited in part 2 (fig. 4.I.46), where the phase identification on every dataset image and its segments are processed. The result is the identified phases mapped along the overview. Thus, we achieve the direct correlation of atomic and crystallographic information with a single STEM micrograph displaying a FOV larger than a micron. The inset in Figure 4.I.46 displays the output phases spread along the main region of interest in the device. This dataset comprises up to 34 images. In addition, their segmentation, based on mid/low magnification images, increases the number of output sources. The resulting identified phases are mapped throughout the inset, and the position of each coloured circle indicates their origin. Importantly, the large number of outputs in this case (identified phases per single segment) greatly reinforces the consensus on the most likely phase for each segment. For example, the best output for the Ge QW and surrounding buffer layers consistently

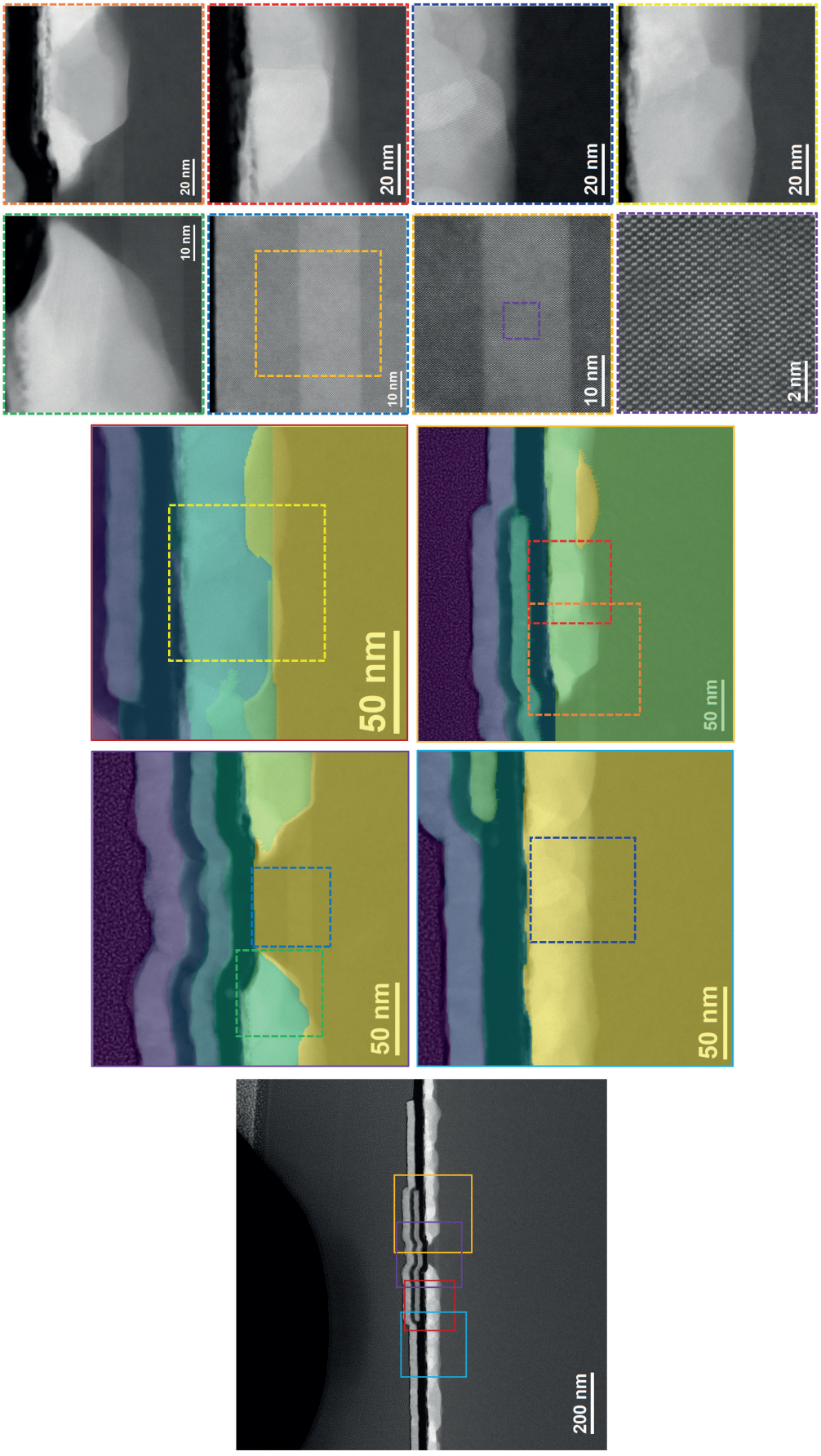


FIGURE 4.I.45: Phase identification proof of a full Scanning Transmission Electron Microscopy (STEM) dataset through the low to high magnification correlation: Ge quantum well with PtSiGe contacts. Focus on the multiscale template matching chain. [70]

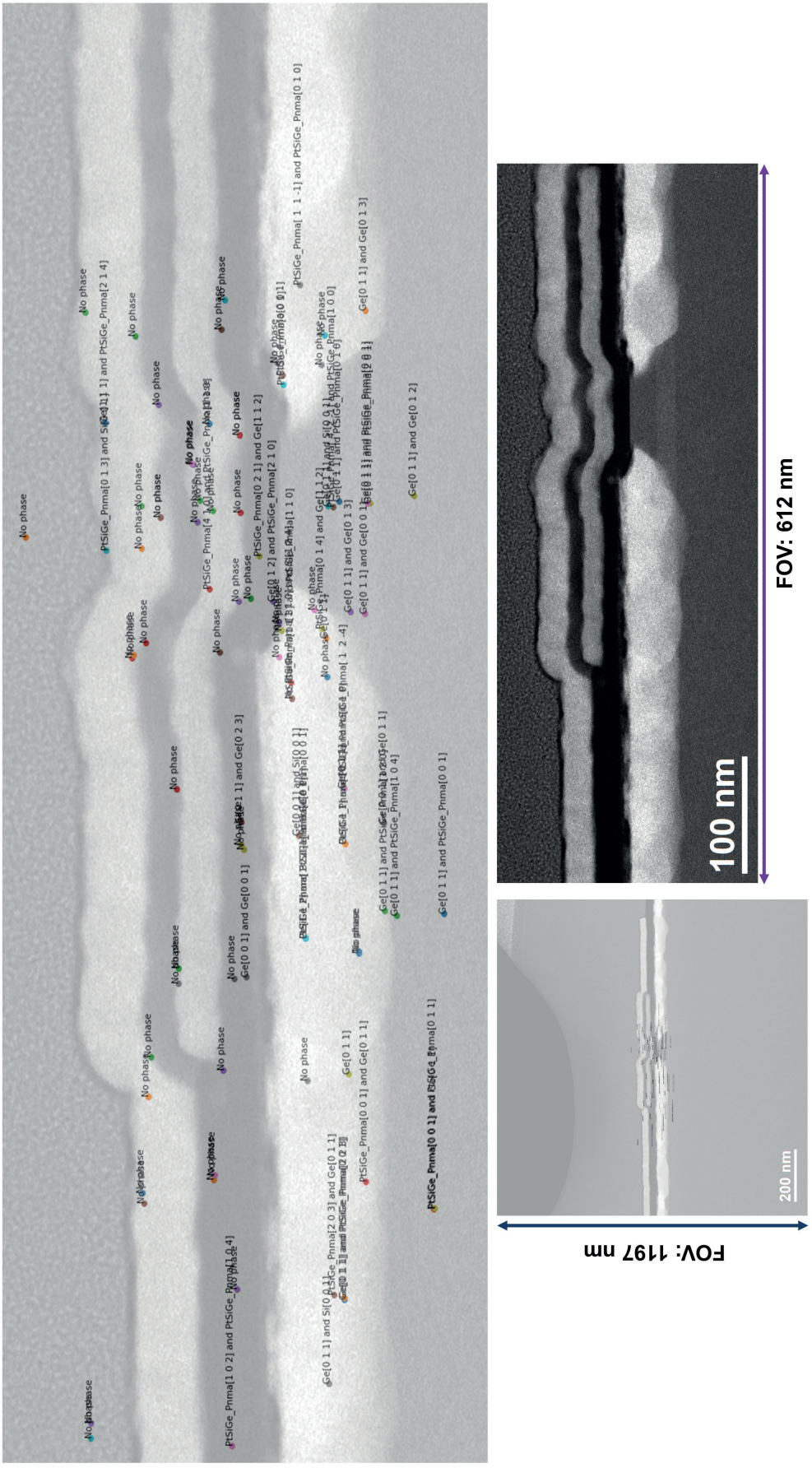


FIGURE 4.I.46: Phase identification proof of a full Scanning Transmission Electron Microscopy (STEM) dataset through the low to high magnification correlation: Ge quantum well with PtSiGe contacts. Focus on the identified crystal phases. [70]

corresponds to Ge as visualized along its[011] zone axis, which unifies the consensus. Remarkably, the polycrystalline nature of PtSiGe contacts is supported by the manifold zone axes of PtSiGe (Pnma) obtained. Despite this, consensus is still found within individual grains. The same pattern occurs with the top metallic gates, which are not crystalline. Therefore, they mostly present a "No phase" label, indicating either an amorphous or a misoriented phase.

- (Figures 4.I.47 (Part 1) and 4.I.48 (Part 2)) Same phase identification proof in a 8-images STEM-HAADF dataset, but in a core-shell Ge/Si nanowire on Si substrate, as presented in the last section of Chapter 3. [102] Part 1 (fig. 4.I.47) displays the template matching chain of a device in the limit of accessing atomic resolution at its overview FOV. Interestingly, the segmentation occurs directly in the overview, and it is applied to the higher magnification images. It cannot distinguish the thin (2-3 nm) Si shell from the protective layer, for which it will not be found.

  Complementing, part 2 (fig. 4.I.48) showcases the phases identified for every image/segment information source. The consensus regarding the resulting crystal phases per region is clear: Si [011] is the main option for the substrate, while Ge [011] is even more prevalent. The imaging session specifically targeted the top and bottom interfaces of the device, which explains the concentration of outputs in those areas. However, a remarkable absence in this characterisation is the unidentified Si shell, mainly due to the limitations of the contour-based segmentation method.

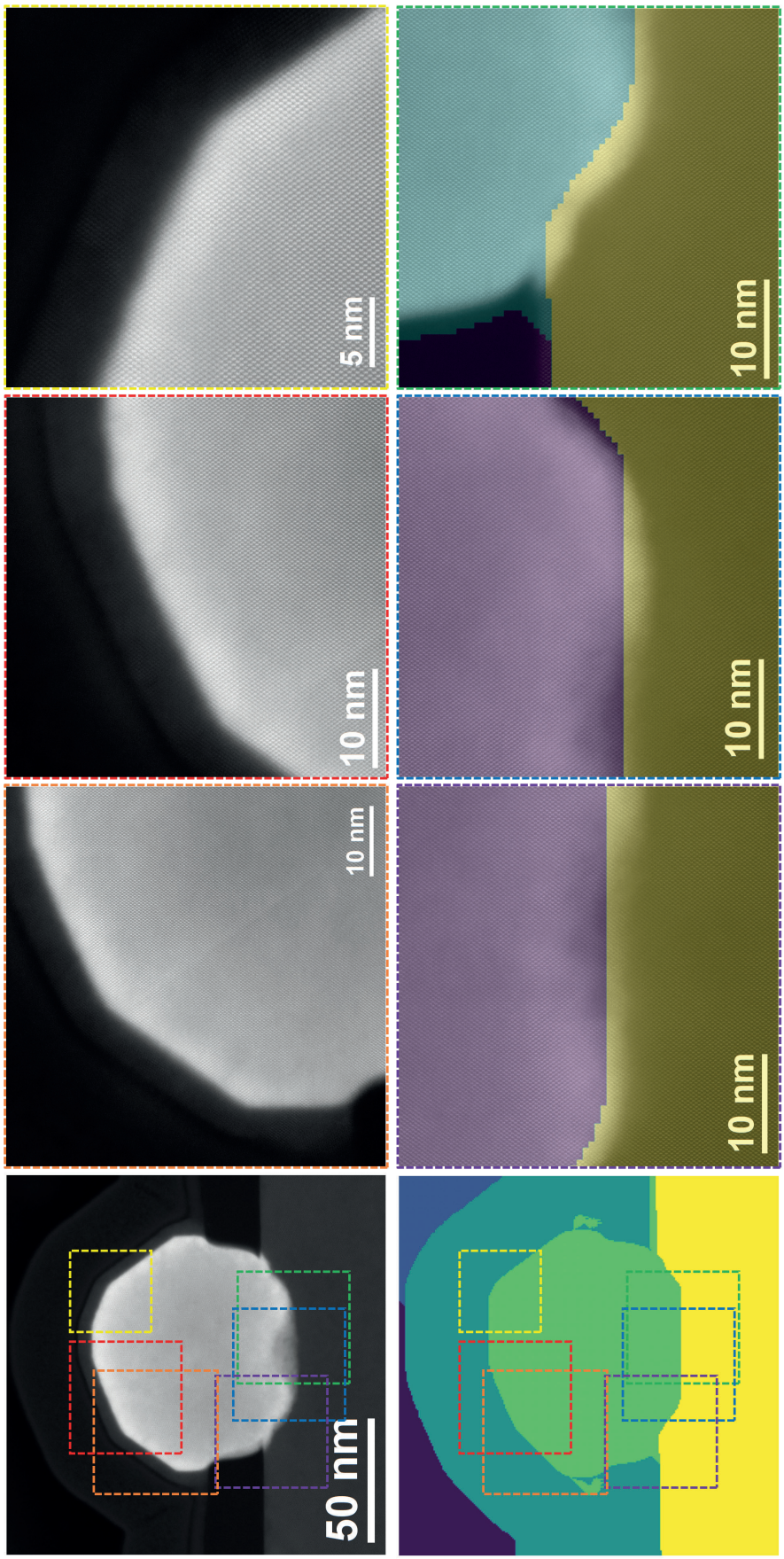


FIGURE 4.I.47: Phase identification proof of a full Scanning Transmission Electron Microscopy (STEM) dataset through the low to high magnification correlation: Ge nanowire with Si shell grown on a Si substrate well. Focus on the multiscale template matching chain. [102]

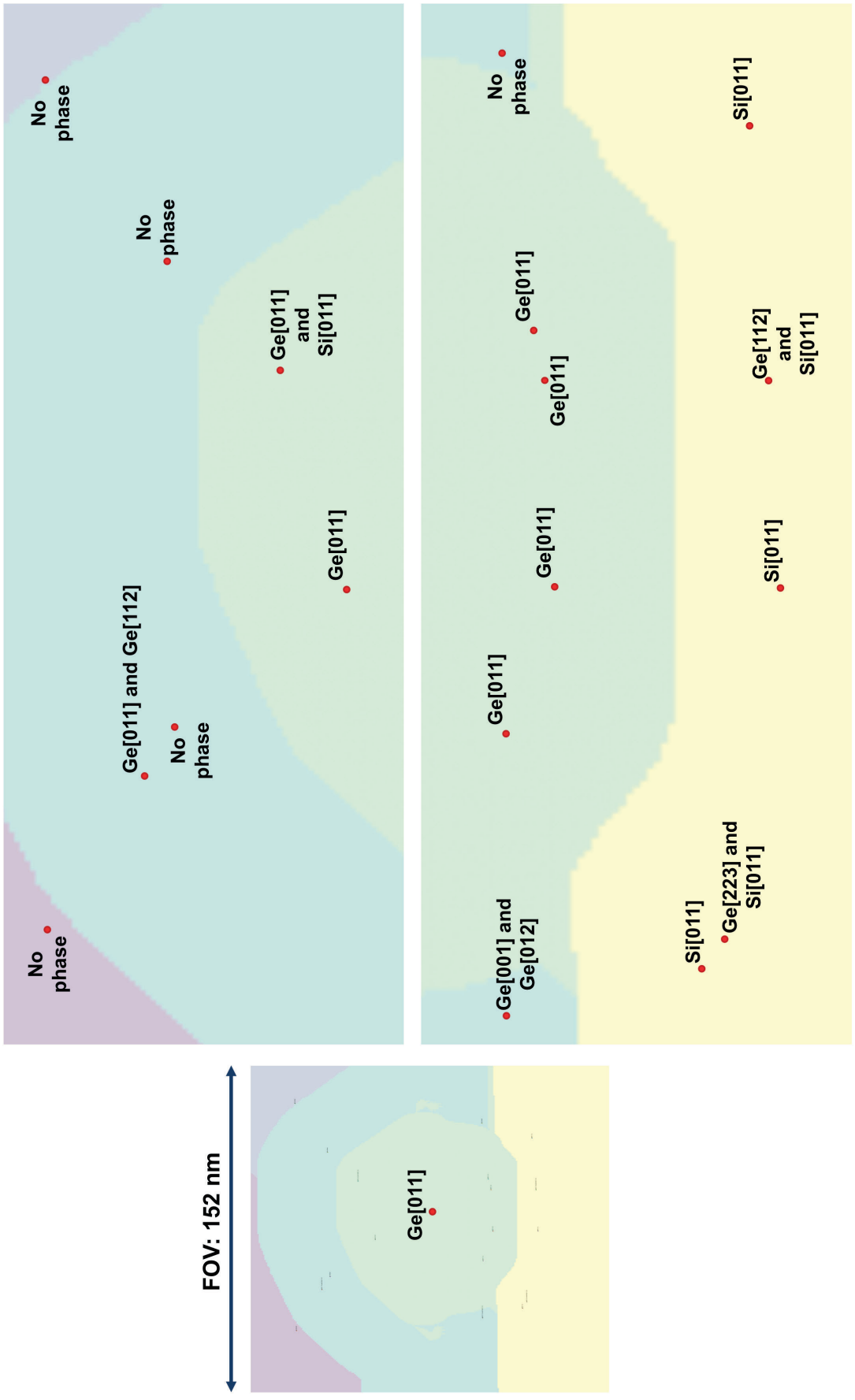


FIGURE 4.I.48: Phase identification proof of a full Scanning Transmission Electron Microscopy (STEM) dataset through the low to high magnification correlation: Ge nanowire with Si shell grown on a Si substrate well. Focus on the identified crystal phases. [102]

- (Figures 4.I.49 (Part 1) and 4.I.50 (Part 2)) Phase identification proof in a STEM-HAADF dataset of an InAs NW with a polycrystalline Al shell. 16 micrographs compose the dataset, which corresponds to the same InAs@Al NW presented in Chapter 3. [3] Moreover, it is the refined version of the device introduced in the template matching collage in figure 4.I.20. The approach with this device resembles what was analysed in figures 4.I.45 and 4.I.46 for device GSu_10. In GSu_10, its polycrystalline nature was confirmed by the segmentation but also by the phase-finding results. Contrarily, the present device's segmentation unifies the whole polycrystalline Al layer. Consequently, a more detailed evaluation of the workflow's output is required, and the criteria of the most common phase per segment is not enough.

  Part 1 (fig. 4.I.49): The multiscale template matching chain uses two low magnification images as segmentation sources. However, both images merge the polycrystalline shell into a single segment. Consequently, multiple aluminium zone axes are expected in the shell, not indicating a discrepancy but rather uncaptured crystal boundaries. Part 2 (fig. 4.I.50): This section highlights a hexagonal, wurtzite, InAs [001] = [0001] core with no significant discrepancies, despite a couple of outliers. However, the algorithm behaves more erratically in the Al shell and does not provide a consistent and accurate answer. It correctly characterises the twofold [011] Al in the right kink of the hexagon, but dances discordantly everywhere else in the system. [103]

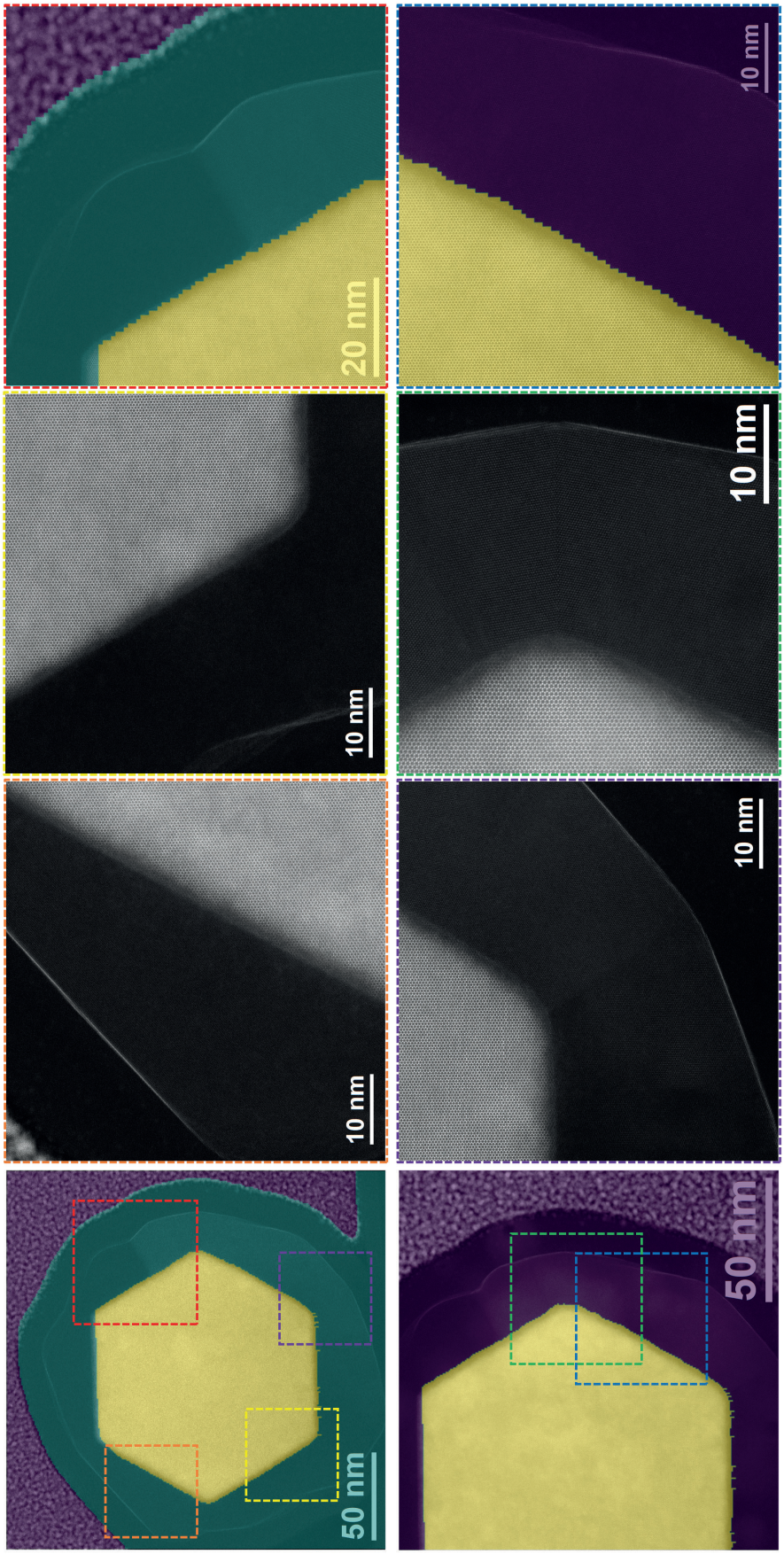


FIGURE 4.I.49: Phase identification proof of a full Scanning Transmission Electron Microscopy (STEM) dataset through the low to high magnification correlation: InAs hexagonal nanowire with polycrystalline Al shell. Focus on the multiscale template matching chain. [3]



FIGURE 4.I.50: Phase identification proof of a full Scanning Transmission Electron Microscopy (STEM) dataset through the low to high magnification correlation: InAs hexagonal nanowire with polycrystalline Al shell. Focus on the identified crystal phases. [3]

Part II

Automated Creation of Experimentally Realistic 3D Models from the Comprehensive Structural Characterisation of Devices

Part II of Chapter 4 comprises all the steps or modules of the workflow involved in the creation of 3D models from the gathered knowledge in Part I. The creation and leveraging of the models encompasses 1) Finite Element Model, 2) 3D atomic model building, 3) Strain transfer to atomic model, and 4) Simulation of functional properties. The provided level of detail should be enough to help in the readability of the code once open and publicly available, and even encourage the custom reproduction of the last segment of the workflow, if wanted.

4.II.1 Starting point: Knowledge from Part I

The workflow segment described in Part I outputs the thorough structural characterisation of the device/sample of interest. We extensively provided manifold proofs of its general and reproducible character. Excitingly, it is then suitable to analyse most device types through regular (S)TEM modes. We even hypothesise its excellent extrapolated performance in complementary detected signals, such iDPC, although we have not tried gathering experimental proofs yet.

In order to continue the workflow, we need to keep in mind that the following key information pieces can be accessed from Part I:

- The segmentation outputs a 2D matrix of integers in which each integer corresponds to an independent region (section 4.I.1). The presence of pixels as the data units unavoidably entails the aliasing of the interface between segments, which will be important for the continuation of the workflow.
- The precise identification of the intensity peaks in the FFT of the atomic resolution image (section 4.I.2). The proposed state-of-the-art global peak finding model, which attained a precision of 69.78%, recall of 70.89%, and F1 score of 61.87%, labels the regions as either amorphous or crystalline. We add to this important distinction the capacity to later characterise these peaks and the crystallographic phases they represent.
- The phase identification of the crystallographic phases within every crystalline segmented region (sections 4.I.5, 4.I.3 and 4.I.8). In other words, complete knowledge of the spatial orientation of these crystalline phases, which is characterised by their zone axis and the indexation of the FFT peaks based on their family of planes. Interestingly, this crucial quantitative information is visually displayed through an interactive graphical interface that allows the intuitive

linking between each crystallographic phase and its identified peaks (crops from the interface in section 4.I.8).

Segmentation, peak finding and phase identification blend to constitute a robust phase structural characterisation that, for instance, is resistant to false positives in the peak finding algorithm. Altogether, this integrated information forms the foundation of the subsequent steps in the workflow.

- The correlation between query-template micrograph pairs and their phase identification if working with a dataset instead of a single image (sections 4.I.4 and 4.I.8). As detailed in section 4.I.4, the intricacies of proceeding with the workflow using a dataset exceeded the scope of this thesis, despite having similar research objectives (section 1.2). As a result, this knowledge is kept to unlock accessing atomic and crystallographic information at unprecedented and *a priori* impossible FOVs with regular (S)TEM.
- The strain calculation based on the awareness of the contact between the segmented regions (section 4.I.6). In the context of heteroepitaxies, the relative in-(image)-plane rotations of separated segmented regions play a crucial role in determining the mask and resolution of GPA. Specifically, these rotations influence the quality of strain mapping. We can capture strain effects in the relative strain GPA maps by directly examining the relation between segments' crystallinity and strain. Importantly, combining knowledge of segmentation and phase identification (such as the bulk relaxed state of the phases) with information about relative strain allows for the direct and immediate mapping of absolute strain. It is worth noting that the computed strain inherits the same limitations as its foundation, GPA, where the quality of the strain mapping depends on the quality of the input image. The relevance of these observations becomes evident in the next steps of the workflow, particularly during the creation of the 3D Atomic Models (3DAMs).
- The Bragg filters qualitatively map the crystal phases and their significative planes (section 4.I.7). Although in its current state, it is purely qualitative, the visual help that Bragg filtering is proved to be crucial in complementing the output from the previous steps. Nevertheless, we do not transform the filtered images into quantitative information to be used in the next steps.

These information sources are exploited in the second part of the workflow to build the 3D models constituting the ultimate end of our central objective (see research goals, section 1.2). However, we will not be able to claim the same general character here, as we will be limited, as indicated before, by systems where translational invariance is a physically compelling argument. Let us review the path towards these models, how they combine the knowledge from part I and their implications.

4.II.2 Finite Element Modelling automation

The process of generating a Finite Element Model (FEM) primarily relies on the segmentation step. In this context, each region identified during segmentation becomes an independent material block within the FEM simulation. The segmentation process allows us to extract and define the contour of each segmented region, represented either as a list of vertices or, in finite element terminology, as a list of contours or boundary elements. We aim to create a Boundary Elements Model (BEM) that encloses a specific region. From this BEM, we have two options: we can either retain the boundary elements themselves or mesh the interior of the region to create a complete Finite Element Model. Again, it is important to note that throughout this process, we assume translational invariance within our system.

To generate the boundary elements, we begin by converting the pixelated segmented image into a coordinate map. Each pixel corresponds to a quadrilateral with four corners, and we assign 2D coordinates to these corners. Next, we define each contour fragment using a starting and an ending coordinate, which together form a vector. For convenience, we ensure that the vectors trace an anticlockwise closed circuit. Following the right-hand rule, this circuit unequivocally defines the material enclosed within the contour. Typically, this material corresponds to the most likely crystallographic phase identified by the workflow within the given segmented region.

The nature of images and their fundamental building blocks, the pixels, dictates that the magnitude of vectors always equals the pixel size. In other words, after segmentation, we end up with numerous vectors that outline the segmentation at a pixel level. However, this process includes noise introduced by the segmentation algorithm, which might fail to accurately reproduce the real device morphology. The issue becomes especially problematic in the segmentation based on contour detection. The consecutive downsampling and upsampling steps emphasise the aliasing effect. In this case, what starts as a single-pixel artefact after downsampling ends up being represented by several pixels following the final upsampling (i.e., a multiple of the upsampling factor) (see section 4.I.1).

To overcome this limitation, we smooth down the curve generated by the original contour vectors. This approach is beneficial not only for the SAM segmentation but also specifically useful for the contour-based segmentation. The proposed smoothing method involves skipping some coordinates and uniting adjacent edges into a single longer vector. As with any smoothing process, there's a trade-off: the more we smooth (or equivalently, the more pixels we skip), the more uniform the shapes become, but finer features are diluted. The optimal number of pixels to skip depends on the total image size. For images in the thousands of pixels per side range, skipping around 30 pixels strikes a balance between smoothing and preserving features. For smaller images in the hundreds of pixels per side range, 10-15 skipped

pixels achieve a similar trade-off. Importantly, regardless of this choice, triple junctions, where three or more regions meet, are always preserved. The contour vector formation and smoothing process are illustrated in Figure 4.II.1. The grid in the zoomed-in sections represents the pixels, and it demonstrates how the segmentation is drawn through them. Each pixel side that belongs to an interface contains two antiparallel vectors corresponding to the two regions it borders. In contour-based segmentations, the benefits of skipping pixels during contour generation are even more evident, as it reduces the noise introduced by the segmentation process. Ultimately, this approach allows us to create a physically smooth curve, even at the triple junctions, in the final .gds file containing the vertex coordinates of the regions, enabling accurate and representative FEM simulations.

As just mentioned, we leverage the advantages of .gds files to encode vertex information. This choice is justified because Graphic Data System (gds) files are a standard in integrated circuit design. They inherently include features like layered design and independent segment masking (i.e., also opening the possibility of automatically analysing planar views and inputting them as integrated circuit masks). [104, 105] We can conveniently access the regions to include in simulations by using .gds files. For instance, we can compare calculations globally in a heterostructure and separately with each material as the full system. Importantly, this approach introduces complexity by enabling easy manual tuning and rapid hypothesis testing. In addition, several FEM software packages officially support the .gds format as a valid input format containing the vertex coordinates of a surface intended to become a simulated translational invariant volume. [106, 107]

4.II.2.1 In-(image)-plane alignment of the experimental-model rotation

Beyond the contours and vertexes lists that form the foundation of FEM simulations, we can leverage the crystallographic computations obtained during phase identification to enhance our simulations. Specifically, when dealing with simulation parameters exhibiting anisotropic behaviour, it becomes advantageous to project these quantities along the orientation of the materials within the simulation system. For instance, nextnano enables the direct assignment of a crystallographic plane parallel to the cartesian x-axis, while COMSOL would require a manual projection. [106, 107]

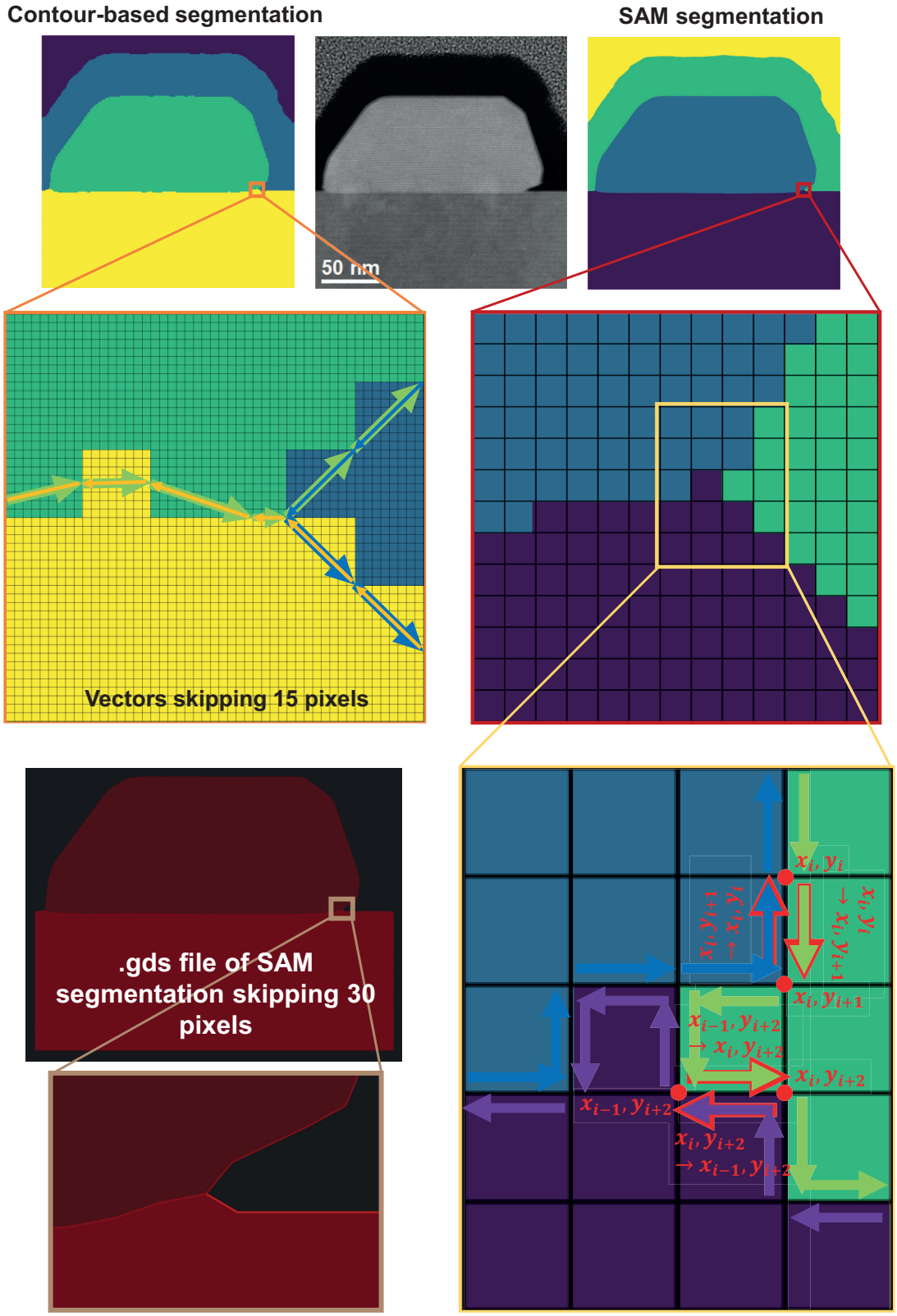


FIGURE 4.II.1: Process of creating and smoothing the contour vectors down out of a segmented micrograph to build its representative boundary element model. The segmented image is read on the basis of its pixel coordinates, and the vectors are built based on this. Every region encoding a material is closed by a list of vectors anticlockwise looping through its edge. To avoid the noise generated by the segmentation, we skip vertexes coordinates to smooth the curves down. Eventually, we encode this information into a Graphic Data System (.gds) file readable by finite elements simulation software.

Therefore, we convert the crystallographic planes indexed by the phase identification into a crystallographic plane oriented along the horizontal axis of the image. By quantifying the angle between any indexed plane and this virtual axis, we can determine the crystallographic plane that aligns with the eventual abscissa axis of the FEM. This calculation is described in figure 4.II.2. The reference plane we use for this purpose belongs to the $[uvw]$ zone axis and is selected based on having the least multiplicity. Since plane indexation is family-aware, we choose this plane as a reference. In addition, an extra plane helps us in corroborating the found orientation. Our equations system yields a two-vector basis, representing the intersection of the cone formed by all the vectors at a given angle with the reference plane and the plane defined by the zone axis (figure 4.II.2). The use of the second vector allows us to distinguish which of the two vectors we need to select to orient our system. It is worth noting that this approach parallels the in-plane orientation of the atomic model, as detailed in section 4.II.3.

This is encoded in the following equations, which using the concepts of metric and reciprocal metric tensor, allow us to compute the reciprocal space vectors (and their Miller indexes) fulfilling these conditions, $g^*_{x,1}$ and $g^*_{x,2}$.

In general terms, these conditions are satisfied by a vector with Miller indices $(h_x k_x l_x)$. Later, due to construction, the twofold separation of the indices automatically emerges. These three indices represent our unknowns. First, the plane must lie in the zone defined by the zone axis we have identified $[uvw]$. Therefore, the zone equation must be satisfied (equation 4.II.10).

$$h_x u + k_x v + l_x w = 0 \tag{4.II.10}$$

Secondly, we need to ensure that the new vector h^*_x forms a given angle θ_x with the reference vector g^*. Since the angle between two vectors is directly related to their dot product and the cosine of the angle, we utilize the dot product encoded with the reciprocal metric tensor (since we are now dealing with reciprocal space vectors) (equation 4.II.11).

$$|g^*_i| = \sqrt{g^*_i g^*_{i,j} g^*_j} \tag{4.II.11}$$

Where $g^*_{i,j}$ represents the reciprocal metric tensor of the unit cell we are working with. It is important to note that i and j do not correspond to specific elements; rather, they express the orientation of the vector during vectorial operations. Specifically, i represents a horizontal vector, while j represents a vertical one.

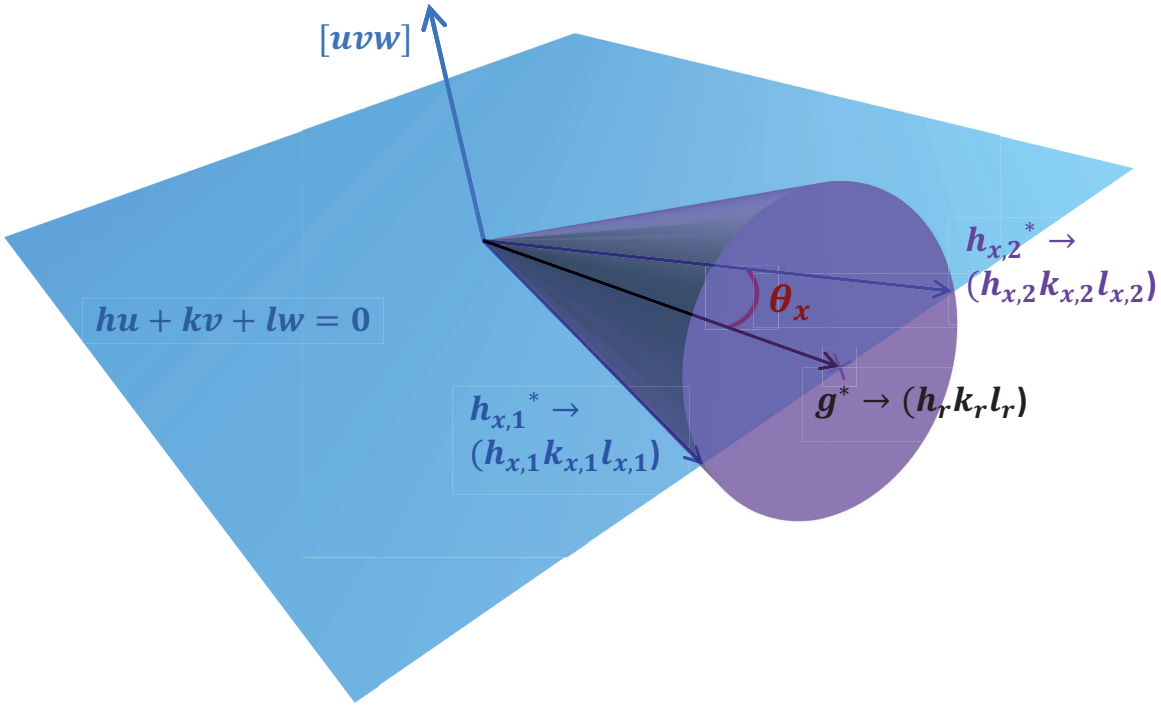


FIGURE 4.II.2: In light blue, the representation of the $hu + kv + lw = 0$ plane or zone defined by the zone axis $[u, v, w]$. In purple, the cone of all the possible vectors forming an angle θ_x with the reference (reciprocal space) vector g^*, which belongs to the former plane, with Miller indices $(h_r k_r l_r)$. The vectors we aim to find to orient our atomic or finite element models are the intersections of these two objects, $h^*_{x,1}$ (with Miller indices $(h_{x,1}k_{x,1}l_{x,1})$ and $h^*_{x,2}$ (with Miller indices $(h_{x,2}k_{x,2}l_{x,2})$.

With this convenient way to express vectors dot products and modules, we define the angle between two vectors like:

$$cos\theta_x = \frac{g^*_i g^*_{i,j} h^*_j}{\sqrt{g^*_i g^*_{i,j} g^*_j}\sqrt{h^*_i g^*_{i,j} h^*_j}} \tag{4.II.12}$$

And by expressing the dot product of the two vectors in the following convenient way only dependent on our three unknowns $(h_x k_x l_x)$:

$$g^*_i g^*_{i,j} h^*_j = g^*_i g^*_{i,j} h^*_j = [A, B, C]\begin{bmatrix} h_x \\ k_x \\ l_x \end{bmatrix} = Ah_x + Bk_x + Cl_x \tag{4.II.13}$$

Then, by combining equations 4.II.12 and 4.II.13, we end up with the second condition to be fulfilled:

$$cos\theta_x = \frac{Ah_x + Bk_x + Cl_x}{\sqrt{g^*_i g^*_{i,j} g^*_j}\sqrt{h^*_i g^*_{i,j} h^*_j}} \tag{4.II.14}$$

Finally, we can make a statement with the module of the vectors to constrain the system.

$$|h_i^*| = N\,|g_i^*| \tag{4.II.15}$$

For which, combining equations 4.II.10, 4.II.14 and 4.II.15, we have an equation system with three equations to solve for the three unknown Miller indices $(h_x k_x l_x)$:

$$\begin{cases} h_x u + k_x v + l_x w = 0 \\ cos\theta_x = \frac{Ah_x + Bk_x + Cl_x}{\sqrt{g_i^* g_{i,j}^* g_j^*}\sqrt{h_i^* g_{i,j}^* h_j^*}} \\ |h_i^*| = N\,|g_i^*| \end{cases} \tag{4.II.16}$$

Going directly to the solutions of the system of equations is enriching to identify interesting issues intrinsic to the physical problem. h_x and k_x are calculated by:

$$k_x = \frac{u}{Bu - Av} N|g_i^*|\,cos\theta_x + \frac{Aw - Cu}{Bu - Av} l_x = R + Pl_x \tag{4.II.17}$$

$$h_x = \frac{N|g_i^*|\,cos\theta_x}{A} - \frac{BR}{A} - \left(\frac{BP + C}{A}\right) l_x = M + Ql_x \tag{4.II.18}$$

And l_x can be calculated by expanding equation 4.II.15 with the reciprocal metric tensor terms, and substituting h_x and k_x with their solutions (equations 4.II.17 and 4.II.18 respectively):

$$|h_i^*| = N\,|g_i^*| \rightarrow [h_x, k_x, l_x] \begin{pmatrix} g_0 & g_1 & g_2 \\ g_3 & g_4 & g_5 \\ g_6 & g_7 & g_8 \end{pmatrix} \begin{bmatrix} h_x \\ k_x \\ l_x \end{bmatrix} = N^2\,|g_i^*|^2 \tag{4.II.19}$$

Equation 4.II.19 makes the system non-linear, as every unknown is squared. We can group all the resulting terms from the expansion and substitutions to express equation 4.II.19 as a two-degree polynomial equation and its solution for l_x:

$$D{l_x}^2 + El_x + F = 0 \rightarrow l_x = \frac{-E \pm \sqrt{E^2 - 4DF}}{2D} \tag{4.II.20}$$

with its terms being:

$$D = Q^2 g_0 + PQ(g_1 + g_3) + Q(g_2 + g_6) + P^2 g_4 + P(g_5 + g_7) + g_8 \tag{4.II.21}$$

$$E = 2MQg_0 + RQ(g_1 + g_3) + PM(g_1 + g_3) + M(g_2 + g_6) + 2RPg_4 + R(g_5 + g_7) \tag{4.II.22}$$

$$F = M^2 g_0 + RM(g_1 + g_3) + R^2 g_4 - N^2 \left|g_i^*\right|^2 \tag{4.II.23}$$

This pair of indices $(h_{x,1}k_{x,1}l_{x,1})$ and $(h_{x,2}k_{x,2}l_{x,2})$ constitutes a solution to the proposed system of equations (4.II.16). As pointed out before, the twofold solution (represented by the two vectors marked in Figure 4.II.2) arises from purely geometric reasons. This becomes evident when the system introduces squared terms, implying the existence of more than one solution. However, it is important to note that the proposed solution for the system of equations is not the general solution for the physical system we are investigating. A closer examination of the h_x and k_x solutions (equations 4.II.17 and 4.II.18, respectively) reveals physical scenarios in which the mathematical solution is not uniquely determined. For instance, if the zone axis contains zeros (e.g., $[0, 0, w]$), or if the symmetry of the system leads to conditions such as $(Bu - Av = 0)$ or $(A = 0)$ for a given g^*, the solution becomes ambiguous. Therefore, the provided solution is not sufficiently general, both in mathematical and physical terms. To address this, while we do not explicitly present complementary solutions, we document all possible exception cases for the sake of reproducibility. These exceptional cases are appropriately handled in the implemented working code.

The exceptions are encoded in algorithm 4.II.1. Their specific solutions can be readily obtained by imposing the exceptional conditions on the Miller indices within the initial problem defined by the equations system 4.II.16. Therefore, each of these instances or deviations in Algorithm 4.II.1 requires an independent analytical solution that is invoked within the code when the designated scenario arises. These scenarios involve identifying mathematical indeterminates (i.e., locating zeros in the denominators of the general solutions while solving for each unknown). By accounting for all such cases, we develop a comprehensive algorithm that works for any crystallographic reference plane, zone axis, and crystal system.

Algorithm 4.II.1 Exceptions in the calculation of the Miller indices of crystallographic planes oriented at a certain angle from a reference reciprocal space vector

1: **if** $[u, v, w] == [u, 0, 0]$ **then**
2: Exception handling: Specific solution number 1
3: **end if**
4: **if** $[u, v, w] == [0, v, 0]$ **then**
5: Exception handling: Specific solution number 2
6: **end if**
7: **if** $[u, v, w] == [0, 0, w]$ **then**
8: Exception handling: Specific solution number 3
9: **end if**
10: **if** $[u, v, w] == [0, v \neq 0, w \neq 0]$ **then**
11: **if** $abs(v) == abs(w)$ AND $abs(B) == abs(C)$ **then**
12: **if** $abs(A) == 0$ **then**
13: Exception handling: Specific solution number 4
14: **else**
15: Exception handling: Specific solution number 5
16: **end if**
17: **else**
18: Exception handling: Specific solution number 6
19: **end if**
20: **end if**
21: **if** $[u, v, w] == [u \neq 0, 0, w \neq 0]$ **then**
22: **if** $abs(u) == abs(w)$ AND $abs(A) == abs(C)$ **then**
23: **if** $abs(B) == 0$ **then**
24: Exception handling: Specific solution number 7
25: **else**
26: Exception handling: Specific solution number 8
27: **end if**
28: **else**
29: Exception handling: Specific solution number 9
30: **end if**
31: **end if**
32: **if** else **then**
33: **if** $abs(A) == 0$ **then**
34: Exception handling: Specific solution number 10
35: **end if**
36: **if** $(A == B$ AND $abs(A) == abs(C))$ OR $Bu == Av$ **then**
37: **if** $abs(C) == 0$ **then**
38: Exception handling: Specific solution number 11
39: **end if**
40: **if** $w == 0$ **then**
41: Exception handling: Specific solution number 12
42: **else**
43: Exception handling: Specific solution number 13
44: **end if**
45: **else**
46: Exception handling: Specific solution number 14
47: **end if**
48: **end if**
49: Return $(h_{x,1}k_{x,1}l_{x,1})$ and $(h_{x,2}k_{x,2}l_{x,2})$, else =0

4.II.3 Atomic model building

All the knowledge acquired through the workflow converges into generating an atomic model that represents the experimentally analysed device. Two models of different complexity are constructed based on this compiled information. Firstly, an atomic model is generated by filling translational invariant supercells with the shape revealed by the segmentation routine. In this model, every segmented region is fully filled by the identified crystal phase, oriented along the indexed zone axis and planes. Building this model is computationally and methodologically straightforward. It follows the typical methodology found in the literature, where manual engineering is required, and inputs such as materials, zone axis, in-plane orientation, and morphology need to be manually specified. A brief literature search would reveal multiple available tools designed for this purpose. [41, 42, 108, 109, 110, 111] However, our proposed innovation lies in automating this process and incorporating experimental materials/device-specific parameters. Of course, to the detriment of the translational invariance assumption. Despite this limitation, the constructed model serves a valuable purpose: it highlights mismatches in interfaces and displays the theoretically expected (misfit) dislocation array in the heterostructure.

On the other hand, we can also build an atomic model that emulates the experimentally observed atomic arrangement. This approach leverages all the knowledge obtained from the workflow, including strain, to create a base model. Subsequently, we distort this base model to closely resemble the experimental device. To the best of our knowledge, this represents the first demonstration of a reproducible method for building representative models of both experimental and real devices. The innovative aspect lies in the automation of this process, coupled with the ability to manually intervene for finer adjustments, but above all capturing structural details which would be otherwise overlooked. This approach introduces a new paradigm in atomic model construction and the resulting insights it can provide to the scientific community.

The workflow, which is almost fully coded in Python, relies on the Atomic Simulation Environment (ASE) to complete the atomic modelling steps. [112] ASE is the main available Python API for this purpose, and some of its prebuilt functions and objects serve as valuable building blocks for achieving our desired realistic models. We use ASE's capability to read unit cells (.cif files) to convert the information about the crystal phase obtained during the phase identification process into a programming object that represents an atom. This atom is characterized by its type (corresponding to an element in the periodic table) and its position in space (specified by three spatial coordinates). After reading the cell file, we orient it along the experimentally determined zone axis, and then we replicate the unit cell to fill the entire supercell. Interestingly, we use the contour information as explained in the FEM automation (section 4.II.2), to determine whether the replicated atoms lie inside the

segmented region corresponding to the identified material and specific orientation. Specifically, an atom is considered inside a region if the closer contour vectors in all four cardinal directions are oriented anticlockwise.

4.II.3.1 ASE adaptation: general crystal surface building

When a unit cell is inputted into ASE, and a crystallographic direction (zone axis) is specified, ASE creates a crystal surface. This crystal surface exposes a facet that is perpendicular to the real space vector defined by the inputted zone axis. Our goal is to find an additional vector within the plane defined by the crystal surface. This additional vector is chosen such that it collapses all the infinite in-plane orientations into the experimentally observed one. By default, ASE orients this plane according to its standard procedure as detailed in the documentation. To achieve the desired orientation, we apply specific transformations to the default orientation provided by ASE.

After creating the crystal surface from the original base unit cell, a new basis (so-called surface basis) of vectors is generated to explain the spatial orientation and dimensions of the cell. For convenience, and unless the opposite is specified, the following vectors are defined within the Cartesian basis. Thus, independently of whether they can be better encoded in the real or reciprocal space basis, every vector has by default three components in the Cartesian basis. This is convenient because the Cartesian basis does not depend on the orientation of the cell, while the real and reciprocal cell bases do. The problem then simplifies to expressing the indexed planes in the Cartesian basis and computing their angles with the vectors that constitute the surface basis (also expressed in the Cartesian basis or any fixed vector basis). To achieve this, we convert the Miller indices (which encode reciprocal space vectors) into vectors within the general Cartesian basis. Similarly, directions representing real space vectors are converted into Cartesian coordinates.

Writing this idea in mathematical terms, if we have a unit cell of basis vectors, $\vec{a_1}, \vec{a_2}, \vec{a_3}$, where $\vec{a_1}, \vec{a_2}, \vec{a_3}$ are vectors of the form

$$\vec{a_n} = [a_i, a_j, a_k] \begin{bmatrix} \hat{e_x} \\ \hat{e_y} \\ \hat{e_z} \end{bmatrix} \tag{4.II.24}$$

where $\hat{e_x}, \hat{e_y}, \hat{e_z}$ are the unitary vectors forming the Cartesian basis.

Then, the reciprocal space vectors $\vec{b_1}, \vec{b_2}, \vec{b_3}$ are defined by:

$$\vec{b_1} = \frac{\vec{a_2} \times \vec{a_3}}{\vec{a_1} \cdot (\vec{a_2} \times \vec{a_3})} \tag{4.II.25}$$

$$\vec{b_2} = \frac{\vec{a_3} \times \vec{a_1}}{\vec{a_1} \cdot (\vec{a_2} \times \vec{a_3})} \tag{4.II.26}$$

$$\vec{b_3} = \frac{\vec{a_1} \times \vec{a_2}}{\vec{a_1} \cdot (\vec{a_2} \times \vec{a_3})} \tag{4.II.27}$$

Then, a vector written in terms of the reciprocal space $\vec{g_i^*}$, with Miller indices (h_i, k_i, l_i), is expressed in the Cartesian basis $\vec{g_i}$ by

$$\vec{g_i^*} = (h_i, k_i, l_i) \rightarrow \vec{g_i} = [h_i, k_i, l_i] \begin{bmatrix} \vec{b_1} \\ \vec{b_2} \\ \vec{b_3} \end{bmatrix} \tag{4.II.28}$$

These equations allow us to express the vectors representing crystallographic planes on a fixed basis. To get the other vectors needed, the ones belonging to the crystal surface, we must allocate these vectors on the shared basis. Then, if we create a crystal surface oriented along the crystallographic direction or zone axis $[u, v, w]$, which exposes a plane (h_s, k_s, l_s) (note that in high symmetry crystal systems these are equivalent, but not generally, for which the transformation to the actual parallel plane is computed), the crystal surface basis $\vec{c_1}, \vec{c_2}, \vec{c_3}$ is generated from the main base vectors. Any pair of linearly independent vectors from this basis can span the desired surface:

$$\vec{t_1} = k\vec{a_1} - h\vec{a_2} \ ; \ \vec{t_2} = l\vec{a_1} - h\vec{a_3} \ ; \ \vec{t_3} = l\vec{a_2} - k\vec{a_3} \tag{4.II.29}$$

This definition is limited when the (h_s, k_s, l_s) contain zeros, in which case the surface basis is formed by the original cell vectors. In any other general case, we make the vector normal to the surface as small as possible by finding the greatest common divider of the (h_s, k_s, l_s) indexes to ensure the normal vector is the smallest. This limits the solution for the main two vectors of the surface basis, which are also determined by how we want to visualise and present the surface. The details on this process can be found in the ASE documentation. In the end, what we are interested in is the final values of $\vec{c_1}, \vec{c_2}, \vec{c_3}$ and how they correlate with our initial lattice vectors. This is indicated in the next expressions.

$$[\vec{c_1}, \vec{c_2}, \vec{c_3}] \begin{bmatrix} \vec{a_1} \\ \vec{a_2} \\ \vec{a_3} \end{bmatrix} = \begin{bmatrix} \vec{a_1^*} \\ \vec{a_2^*} \\ \vec{a_3^*} \end{bmatrix} \tag{4.II.30}$$

The generated first surface vector $\vec{a_1^*}$ is by construction fixed along the x Cartesian axis. The basis is defined by

$$[\vec{a_1'}, \vec{a_2'}, \vec{a_3'}] = [\vec{a_1^*}, \vec{a_2^*}, (\vec{a_1^*} \times \vec{a_2^*})(\frac{\vec{a_3^*}(\vec{a_1^*} \times \vec{a_2^*})}{|\vec{a_1^*} \times \vec{a_2^*}|^2})] \tag{4.II.31}$$

We can access these vectors when the surface is generated, for which, computing the angle between the vector $\vec{a_1^*}$ and the vector $\vec{g_i}$ informs us about the default angle between this plane and the Cartesian x-axis of the surface. Interestingly, the third vector $\vec{a_3'}$ is parallel to the Cartesian z-axis, for which its cross product with $\vec{a_1^*}$ returns a vector also in the surface and at 90º from the x-axis (i.e., along the y-axis). These properties are clear in what is directly obtained when checking the final surface parameters:

$$\begin{pmatrix} |\vec{a_1'}| & 0 & 0 \\ \frac{\vec{a_1'} \cdot \vec{a_2'}}{|\vec{a_2'}|} & \sqrt{|\vec{a_2'}|^2 - |\frac{\vec{a_1'} \cdot \vec{a_2'}}{|\vec{a_1'}|}|^2} & 0 \\ 0 & 0 & |\vec{a_3'}| \end{pmatrix} \tag{4.II.32}$$

At this point, the calculation of the angle we are interested in is obvious. Our goal is to rotate the surface with the rotation axis being the vector normal to the image plane. The rotation must place the vector representing an indexed crystallographic plane in the in-(image)-plane direction found experimentally. The indexed plane (h_i, k_i, l_i) is found at an angle θ_{exp} to the horizontal axis of the image. From the surface construction, the angle θ_{surf} between the vector normal to (h_i, k_i, l_i), $\vec{g_i}$, and the vector parallel to the x-axis of the generated surface, $\vec{a_1^*}$, is calculated by

$$\theta_{surf} = arccos(\frac{\vec{a_1^*} \cdot \vec{g_i}}{|\vec{a_1^*}||\vec{g_i}|}) \tag{4.II.33}$$

Then, we first revert the angle found in the surface to point the vector back to the horizontal x-axis of the surface, and then we add the rotation unveiled experimentally

$$\theta_{corr} = \theta_{exp} - \theta_{surf} \tag{4.II.34}$$

The evaluation of the angles to rotate θ_{corr} is reproduced for the planes used as g vectors in the GPA analysis. As the workflow indexes the crystal planes according to the family of planes and these planes do not present their in-plane orientation collapsed, we secure the calculation by verifying the multiplicity of the plane family in the exposed zone. Therefore, we give a higher validity to the families of planes with fewer reflections represented in the zone defined by indexed zone axis $[u, v, w]$. As a result, the possibility of ambiguously pairing an angle with a wrong plane is reduced. In summary, low multiplicity planes define the angle θ_{corr} and the rest help to validate the hypothesis.

4.II.3.2 Virtual unit cell calculation

The major key point in building the strained atomic model is the construction of the base atomic model from which to apply the experimentally calculated strain fields. Only a model fulfilling certain criteria would, after applying the distortion field found experimentally, systematically represent the experimental image and corresponding device. To present these conditions, a deep understanding of GPA is needed.

In GPA, we select a region in which we fix the zero displacement. Thus, this region sets the cell parameters which act as a reference to compute relative variations in the rest of the image. Note that the strain we retrieve from GPA is not absolute but relative to the image. As a micrograph by itself is not either compositionally aware or status aware (i.e., how the materials would be arranged in perfect standalone unstressed situations), the considerations from GPA can only apply to the interplanar distances self-contained in the image. Therefore, the definition of the distances to act as the reference is essential. For this purpose, a common choice is to select a relaxed material. In such a material, the chosen region experiences no distance dilatation (i.e., average dilatation equals zero), and it is also unstrained (i.e., absolute strain equals zero). While this choice significantly helps in interpreting the results, practically, any region with a constant lattice serves as a valid GPA reference.

This idea is fully exploited in our proposed method. We just require a reference region that informs us about the distances to expect in the undilated regions (i.e., where dilatation equals zero). From this region, we compute the cell parameters that adapt perfectly to the lattice measured in the reference region. We call this custom unit cell the virtual unit cell of the given reference. As a result, building a supercell with this virtual unit cell should exactly match the lattice distances observed in the region taken as a reference. As a result, the dilatations with respect to the reference region can be transferred into the supercell built upon this reference virtual unit cell. This allows us to locally capture the relative experimental distances. The key lies in having a shared reference both in GPA and in the atomic model. Consequently, the reference information is always directly transferred to the base atomic model from which the final model is built. As a result, the choice of the reference region is not critical.

Note that GPA automation becomes more than just a throughput booster and the pure automation of a usually time-consuming process. It is key in transferring its idea and mathematical fundamentals to the building of the posterior representative model with the experimentally measured strain. In that sense, the attentive reader may have noticed that the workflow has already outputted all the necessary knowledge to retrieve absolute strain. As we get GPA, the segmentation, and the reference "bulk" crystal phases per segment, we have information on the original or relaxed status of the materials involved in the device. Therefore, the calculation of

the absolute strain is direct both in average (i.e., computing average experimental cell parameters per region), or mapped as local strain (i.e., comparing the local cell parameter from the dilatation maps with the relaxed reference one recognised in the phase identification step).

The virtual unit cell parameter is computed according to the following logic. We want to extract the cell parameters that would give rise to the indexed reflections. Particularly, to the two reflections used as g vectors in GPA. Then, a base unit cell with cell parameters $a, b, c, \alpha, \beta, \gamma$ must be converted to a virtual unit cell with cell parameters $a', b', c', \alpha', \beta', \gamma'$. As we want to avoid distorting the system symmetries, the angles are kept, for which $\alpha = \alpha', \beta = \beta', \gamma = \gamma'$. As we are working with two reflections or crystallographic planes, two reciprocal space vectors arise, which correlate with the interplanar distances found and their indexation.

$$\vec{g^*_{1,ref}} \rightarrow d_{1,ref} \rightarrow hkl_{1,ref} \tag{4.II.35}$$

$$\vec{g^*_{2,ref}} \rightarrow d_{2,ref} \rightarrow hkl_{2,ref} \tag{4.II.36}$$

Then, we use the crystal system-independent definition of the interplanar distance of a plane indexed (hkl) to correlate these three quantities.

$$\frac{1}{d^2_{hkl}} = \vec{g^*_i} g^*_{i,j} \vec{g^*_j} = [h,k,l] g^*_{i,j} \begin{bmatrix} h \\ k \\ l \end{bmatrix} \tag{4.II.37}$$

where $g^*_{i,j}$ is the reciprocal metric tensor of the considered unit cell, which for any crystal system, is calculated as the following determinant:

$$g^*_{i,j} = \frac{1}{V^2} \begin{vmatrix} b^2c^2sin^2\alpha & abc^2F(\alpha,\beta,\gamma) & ab^2cF(\gamma,\alpha,\beta) \\ abc^2F(\alpha,\beta,\gamma) & a^2c^2sin^2\beta & a^2bcF(\beta,\gamma,\alpha) \\ ab^2cF(\gamma,\alpha,\beta) & a^2bcF(\beta,\gamma,\alpha) & a^2b^2sin^2\gamma \end{vmatrix} \tag{4.II.38}$$

where

$$V^2 = a^2b^2c^2(1 - cos^2\alpha - cos^2\beta - cos^2\gamma + 2cos\alpha cos\beta cos\gamma) = a^2b^2c^2\theta \tag{4.II.39}$$

and

$$F(\alpha,\beta,\gamma) = cos\alpha cos\beta - cos\gamma \tag{4.II.40}$$

As what we retrieve experimentally are the distances of the indexed planes, and supposing we do not modify the angles between the cell vectors, we have three unknowns, a', b', c', for which we would rigorously need three equations:

$$\begin{cases} \frac{1}{d^2_{hkl_1}} = \vec{g^*_{i1}}\vec{g^*_{i,j}}\vec{g^*_{j1}} \\ \frac{1}{d^2_{hkl_2}} = \vec{g^*_{i2}}\vec{g^*_{i,j}}\vec{g^*_{j2}} \\ \frac{1}{d^2_{hkl_3}} = \vec{g^*_{i3}}\vec{g^*_{i,j}}\vec{g^*_{j3}} \end{cases} \tag{4.II.41}$$

For convenience, we work with two g vectors in GPA. It is already a challenge to automatically choose two appropriate g vectors, for which a third would likely increase the noise of the calculation (as is often the case even in regular manual GPA). For that, from the equation system 4.II.41, we would be capable of using just two equations. This would hinder the unique solution of the system. However, we can make some assumptions that simplify the system into just a two equations system, or even a single equation. The idea is to keep certain cell symmetries and the relation between cell parameters. These symmetry constraints make our solutions attainable. Although we do not implement it practically, a simple example would imply a cubic system, for which we would only have a single unknown by effectively constraining that $a' = b' = c' = a'$.

Although it is not as obvious, we can generalise these approximations for any crystal system and space group. If we want to compute the virtual cell parameters $a', b', c', \alpha', \beta', \gamma'$ from the two g vectors, we propose, combining the two first usable equations of the system 4.II.41 with 4.II.38:

$$\begin{cases} \frac{1}{d^2_{hkl_1}} = \frac{1}{V^2}[h_1, k_1, l_1] \begin{bmatrix} b'^2c'^2sin^2\alpha' & a'b'c'^2F(\alpha', \beta', \gamma') & a'b'^2c'F(\gamma', \alpha', \beta') \\ a'b'c'^2F(\alpha', \beta', \gamma') & a'^2c'^2sin^2\beta' & a'^2b'c'F(\beta', \gamma', \alpha') \\ a'b'^2c'F(\gamma', \alpha', \beta') & a'^2b'c'F(\beta', \gamma', \alpha') & a'^2b'^2sin^2\gamma' \end{bmatrix} \begin{bmatrix} h_1 \\ k_1 \\ l_1 \end{bmatrix} \\ \frac{1}{d^2_{hkl_2}} = \frac{1}{V^2}[h_2, k_2, l_2] \begin{bmatrix} b'^2c'^2sin^2\alpha' & a'b'c'^2F(\alpha', \beta', \gamma') & a'b'^2c'F(\gamma', \alpha', \beta') \\ a'b'c'^2F(\alpha', \beta', \gamma') & a'^2c'^2sin^2\beta' & a'^2b'c'F(\beta', \gamma', \alpha') \\ a'b'^2c'F(\gamma', \alpha', \beta') & a'^2b'c'F(\beta', \gamma', \alpha') & a'^2b'^2sin^2\gamma' \end{bmatrix} \begin{bmatrix} h_2 \\ k_2 \\ l_2 \end{bmatrix} \end{cases} \tag{4.II.42}$$

At this point, we can say that $a' = a'$ and $c' = c'$, but $b' = na'$, where n is an adimensional proportionality constant which, when wisely chosen, captures the symmetry we want to keep from the original unit cell. An appropriate choice comes from inspecting the ratio between the original cell parameters a, b, $n = b/a$. This allows us to introduce a third equation and to solve the equations system with a physically reasonable assumption. Later we will see that the assumptions can go a little beyond to make the calculation simpler.

Therefore, by developing the equation system 4.II.42 with equation 4.II.39 and $b' = na'$, and expressing it in terms of the unknowns a', c', we end up with:

$$\begin{cases} h_1^2 n^2 c'^2 sin^2\alpha + 2h_1 k_1 n c'^2 F(\alpha', \beta', \gamma') + 2h_1 l_1 n^2 a' c' F(\gamma', \alpha', \beta') \\ + k_1^2 c'^2 sin^2\beta + 2k_1 l_1 n a' c' F(\beta', \gamma', \alpha') + l_1^2 n^2 a'^2 sin^2\gamma = \frac{n^2 a'^2 c'^2 \theta}{d^2_{hkl_1}} \\ h_2^2 n^2 c'^2 sin^2\alpha + 2h_2 k_2 n c'^2 F(\alpha', \beta', \gamma') + 2h_2 l_2 n^2 a' c' F(\gamma', \alpha', \beta') \\ + k_2^2 c'^2 sin^2\beta + 2k_2 l_2 n a' c' F(\beta', \gamma', \alpha') + l_2^2 n^2 a'^2 sin^2\gamma = \frac{n^2 a'^2 c'^2 \theta}{d^2_{hkl_2}} \end{cases} \tag{4.II.43}$$

Although it is possible to analytically solve this system of equations, the number of terms is prohibitive. Consequently, a more realistic approach comprises a numerical solution outputting a', c', and being completed with $b' = na'$. The numerical calculation was accomplished through the Python package SymPy (sympy.solve() method). [113]

The use of this symbolic maths package was key in solving the non-linear system while avoiding an analytical solution. This approach allowed us to quickly identify a practical issue arising from the deformation of the virtual cell when considering the original unit cell and its resulting change in space group. When using three g vectors in GPA (equation4.II.41) or two (as in equation4.II.43), the computation of virtual cell parameters can lead to a shift in the crystal system and, consequently, the space group. Interestingly, ASE recognises translational replication robustly, but its crystal surface generation is less reliable if this shift is imposed. This discrepancy arises because ASE calculates directions based on symmetry constraints from the original space group. While complex, one possible solution would involve correcting these distances based on the new symmetry constraints.

Alternatively, to overcome this limitation, as briefly pointed out before, we introduce a second assumption correlating the old and new cell parameters. Thus, $a' = a'$, $b' = n_1 a'$, and $c' = n_2 a'$, where, as before, n_1 and n_2 are constants informing us on how the original cell parameters correlate to each other, $n_1 = b/a$ and $n_2 = c/a$. By physically constraining the system with these guidelines, we just require a single equation from 4.II.41 to deploy a solution for the unknown a':

$$\begin{aligned} &h_1^2 n_1^2 n_2^2 a'^4 sin^2\alpha + 2h_1 k_1 n_1 n_2^2 a'^4 F(\alpha', \beta', \gamma') + 2h_1 l_1 n_1^2 n_2 a'^4 F(\gamma', \alpha', \beta') \\ &+ k_1^2 n_2^2 a'^4 sin^2\beta + 2k_1 l_1 n_1 n_2 a'^4 F(\beta', \gamma', \alpha') + l_1^2 n_1^2 a'^4 sin^2\gamma = \frac{n_1^2 n_2^2 a'^6 \theta}{d^2_{hkl_1}} \end{aligned} \tag{4.II.44}$$

from which we can analytically extract the virtual cell parameter a':

$$a' = \frac{d_{hkl_1}}{n_1 n_2 \sqrt{\theta}} \sqrt{R}$$

$$R = h_1^2 n_1^2 n_2^2 sin^2\alpha + 2h_1 k_1 n_1 n_2^2 F(\alpha', \beta', \gamma') + 2h_1 l_1 n_1^2 n_2 F(\gamma', \alpha', \beta') + k_1^2 n_2^2 sin^2\beta + 2k_1 l_1 n_1 n_2 F(\beta', \gamma', \alpha') + l_1^2 n_1^2 sin^2\gamma \quad (4.II.45)$$

This information is used to edit a copy of the original .cif file containing the crystallographic information of the base unit cell, where we update the three cell parameters by keeping the rest unchanged. Importantly, this new .cif file representing the virtual unit cell will be used in the creation of the base atomic supercell from which the creation of the distorted model starts.

4.II.3.3 Strain transfer from GPA to atomic model

After computing the GPA and the virtual cell, we generate a supercell filled with the virtual unit cell. The virtual unit cell is already oriented along the found zone axis, and in-(image)-plane rotated to match the found indexation of the in-zone planes. The dimensions of the supercell cover the entire segmented region, including a security margin. Thus, if the applied distortion compresses the supercell, we avoid creating empty spaces. On the other extreme, an expansion of the lattice is not an issue as we eventually retain only the atoms that are within the selected segmented region. To determine whether a displaced atom lies inside the segmented region, we take advantage of the contour vectors defined for the FEM. In the translational invariant direction, the thickness would depend on the numerical needs and the possibilities and performance of our computing system. Thus, a thicker model will be slower and memory-consuming. As it represents a translational invariant pattern, we must choose wisely the situations inevitably requiring the expansion of the transversal cell. For instance, if a large mismatch arises in z, and dislocations are to be found.

After constructing the supercell, the atoms are prepared for displacement. The algorithm is based on interpolating the x and y displacement maps (i.e., dimensionally, a length) for every atomic coordinate, and moving it until the induced displacement matches these maps. In other words, the displacement maps locally inform us about how much an atom from the reference virtual cell in that position should be moved to match the experimental information. The displacement maps must be in the same unit as the atomic coordinates (i.e., angstroms for convenience). The algorithm calculates the new coordinates for all the atoms in the model in an iterative manner, ruled by a moving rate (i.e., length steps of the displacement) and by accuracy (i.e., how different the applied displacement can be when compared to the mapped one). Of course, all the atoms keep their original z coordinate, as no distortion is considered in depth, along the electron beam direction. The moving rate is set as a percentage

of the interpolated displacement missing, and about 50-60 % is sufficient to displace the atoms unless artefacts affect the GPA maps (i.e., 10 to 100 % displacement percentages proved reliable). The accuracy criterion is set to accept displacement only if they differ below the sub-picometer level from the experimental ones.

We perform a GPA calculation for each reference region, selecting their optimal g vectors, engineering the virtual unit cell, applying the distortion to its atomic model and finally filtering the atoms that are within the closed circuit of each segmented region. The separate calculations ensure a sustained relaxed starting point for every material in the device. Of course, the relaxation is in terms of the virtual unit cell. In the end, the heterojunction rises by joining the spatially filtered region according to the segmentation, already with the region-specific displacements applied. We have found that occasionally the interface delimitation of GPA differs from the one produced by the segmentation algorithm, both in morphology and in position. For instance, the contour detection segmentation and its successive downsampling-upsampling produce a misplacing approximately equal to the downsampling factor times the pixel size. As a result, to adequately capture the interface for the atomic model construction, the presence of this artefact cannot be disregarded. Interestingly, the discrepancies are difficult to avoid even with sharp segmentations because the geometrical phase calculation is tightly dependent on the mask size and chosen g vectors. This scenario is reminiscent of the atomic model we can additionally, purely based on morphology and steady relaxed supercell building. Consequently, a task to improve and focus on in future studies is definitely related to this matching, for instance, to have a better atomic representation of dislocations.

In systems where epitaxy is well-matched and lattice mismatch is minimal, a single supercell can be created using a single shared GPA calculation. The common virtual unit cell, based on this single reference, can then be locally refined to address compositional details identified spectroscopically or through phase identification (more details in sections 4.II.3.4 and 4.II.3.5). Using a single base supercell smooths the interface and enhances realism, addressing the issue mentioned in the previous paragraph. However, its limitation lies in its inability to accurately reproduce large lattice mismatches.

The central significance of this process lies not only in its automated nature but also in its criticality, as no manual procedure can replicate the capabilities enabled by this method. It is a double-edged sword, because the quality of the resulting model tightly depends on the quality of the original data. Nonetheless, it is the responsibility of the scientist or microscopist to push for high-quality data. Thus, either by directly acquiring or by post-processing it with one of the manifold available tools for microscopy data enhancement. [11, 114, 115] If the data is not good enough, we transfer to the outputted model the artefacts coming from effects such as not fully oriented samples (e.g., bent lamellae with out-of-axis regions), scan distortions, or

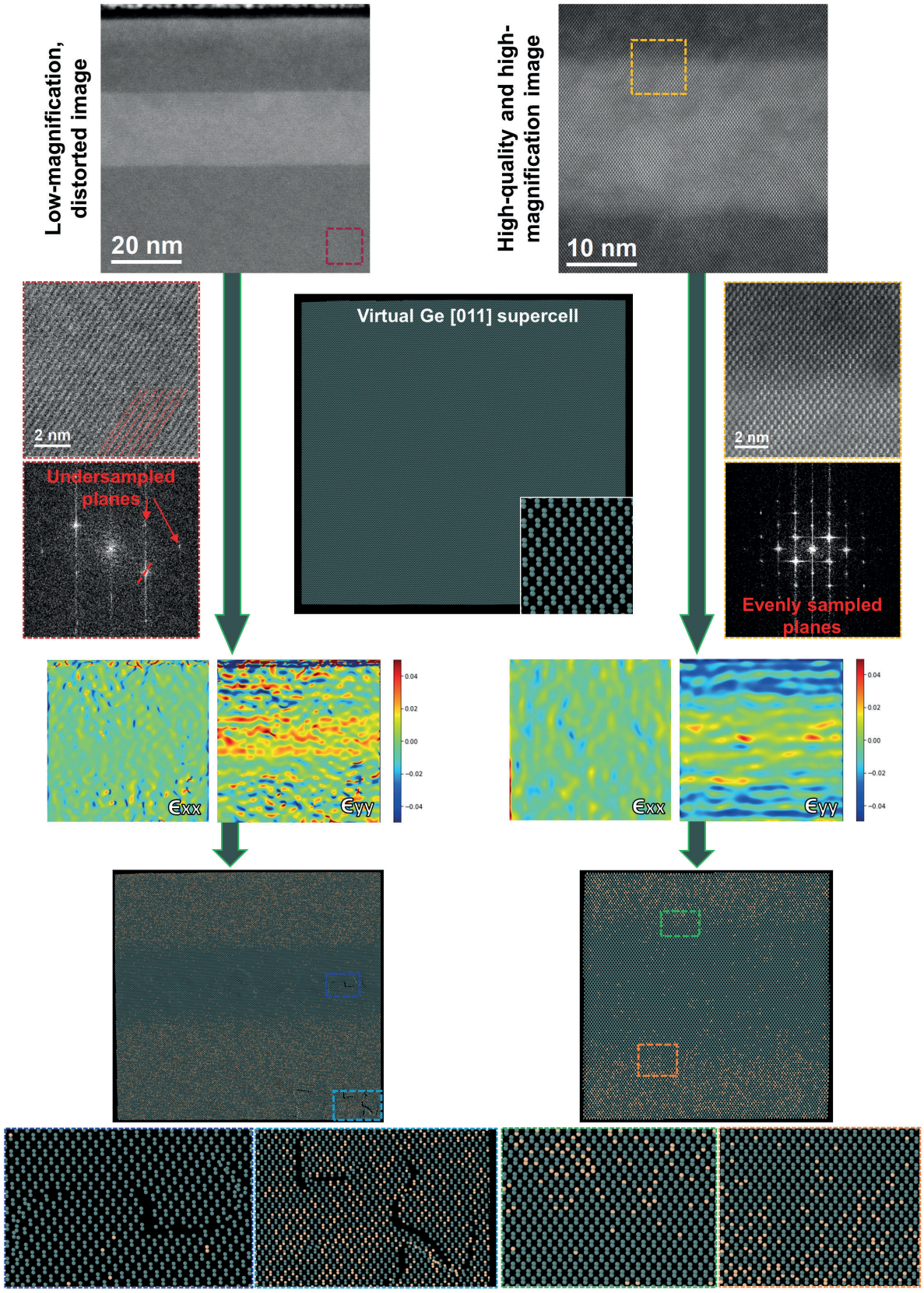


FIGURE 4.II.3: Effect of the image quality and its pivotal role to retrieve accurate and artefact-free 3D Atomic Models (3DAM). Both starting images represent the broadly tackled GSQ_4 device (chapter 3), a Ge quantum well embedded in two $Ge_{0.7}Si_{0.3}$ buffer layers. [38] The left micrograph is not as well-oriented as the right one, which is also a higher-magnification one. This is clear in the Fast Fourier Transforms (FFT), which present uneven plane intensity distributions for the top case. Their computed strain fields modify a Ge supercell built upon their respective virtual unit cell. The difference in image quality is translated into the left model full of artefacts and an accurate one for the micrograph at right.

even sample drift during the scanning process. [17, 116, 117] These effects are typically spotlighted in the GPA maps, although some might get mixed up with real structural phenomena.

Conversely, with high-quality data, genuine physical effects are accurately captured by the algorithm and reflected in the 3DAM. This precision enables the model to detect subtle features crucial for device functionality, such as minor atomic misplacements, vacancies, and dislocations. Generally, capturing any deviation from the theoretical model ideally conceived at the beginning of the research loop. Furthermore, subsequent model refinements, detailed in the upcoming sections, further enhance the alignment of the 3DAM with the experimental device to unprecedented levels. Consequently, these modifications elevate simulations to an entirely novel and unexplored realm of accuracy and detail.

These observations are crystallised in figure 4.II.3. Two micrographs representing the same $Ge_{0.7}Si_{0.3}/Ge/Ge_{0.7}Si_{0.3}$ QW are processed through the entire automated workflow to produce their representative 3DAMs. The larger FOV image at the left showcases local lattice fringes instead of atoms in a perfectly oriented zone axis, as seen in the micrograph at the right. This is visible in the left inset and its FFT, which spotlights the dominant diagonal (-11-1) planes. Importantly, its GPA is computed using these diagonal (-11-1) planes, together with the epitaxial (-220) vertical planes. The conflict arises in the image regions where the vertical planes are not correctly captured, probably due to lamellae bending or heterogeneous damage. In there, GPA is not correctly computed as no sampling can be done from the reciprocal space back to the real space strain components. As mentioned in section 4.I.6, the limitations of our strain calculation mirror those inherent to the GPA method itself. Specifically, one of the critical challenges lies in the sensitivity of the strain calculation to imperfections in the input data. Imagine this sensitivity as a double-edged sword: on one side, it allows us to extract valuable information about the material's behaviour; however, on the other side, it can magnify even minor image imperfections into nonphysical artefacts within our model. These nonphysical imperfections can significantly impact the validity of subsequent calculations based on the experimental data.

On the other hand, the right micrograph of figure 4.II.3, featuring a FOV roughly a quarter the size of the top one, displays a crystal-clear atomic resolution. Thus, most of its regions showcase the characteristic dumbbells of the Ge [011]. The computed strain fields, much smoother due to the even representation of all the in-zone planes (see the bottom inset and its FFT), are applied to the virtual supercell. The result is a model free of artefacts, resembling the artefacts-free GPA maps. This model is suitable for simulations because we can rely on its quality. Importantly, during the atom displacement process, we iteratively verify whether the atom-wise applied displacement aligns with the experimentally mapped displacements. As a result, we can identify the atoms whose displacement is imprecise after completing the

displacement cycle. This workflow serves as a means to assess the reliability of the displacements and determine whether we can or not trust our models. The process is similar to the one followed when relaxing atomic models based on calculations of atomic-wise forces (see section 4.II.5).

To conclude this section, we want to point to a serendipity based on how the distortion algorithm can help us to unveil growth mechanisms and materials dynamics. For instance, in successive frames of a live acquisition, the evolution of the displacement maps and how they rearrange the atoms could immediately be linked to the forces that explain the dynamics and growth of the imaged materials. Alternatively, the use of custom displacement maps could potentially generate any desired 3D model and complement theoretical studies with fast and versatile hypothesis testing. Therefore, one can effortlessly examine the impact of a custom strain field, such as one caused by manual indentation of a tip (e.g., AFM tip), on the digital atomic array that mirrors our experimental device. This capability enables theoretical experimentation, facilitating the optimization of the device before proceeding to more time-consuming and resource-intensive stages of the research cycle.

This concept primarily stems from the models derived from the ongoing treatment of the main InSb on InP nanowires, represented in figure 4.II.4. [37] From the whole NW, we picked an inset from the most demanding yet interesting region of the system from the materials science point of view. That is, the InSb/InP interface, together with the stacking fault diagonally crossing the InSb NW. We computed the strain and displacement fields in that inset (fig. 4.II.4, red square), and we created an InP supercell to capture the effects of these displacement fields (fig. 4.II.4, red margin). Interestingly enough, a strong jump discontinuity emerges in the x displacement map at the interface of the geometrical phase change. Remarkably, we applied the strain transfer methodology to discover a huge gap where the InSb/InP interface should sit. This substantial displacement jump is attributed to the heterojunction mismatch and the separation between the planes in reciprocal space, as detected by the automated GPA mask. Intriguingly, this effect can be regarded as the strength or energy contained in the heterojunction. In this experimental case, this energy is plastically relaxed by the formation of misfit dislocations (fig. 4.I.25).

This same pattern could be extended to situations where elastic relaxation mechanisms play a dominant role in maintaining the stability of the heterostructure. Interestingly, we deliberately eliminated the jump discontinuity by vertically scanning the x displacement map and shifting the displacement levels of the nanowire to match those of the substrate. Surprisingly, this departure from the natural behaviour, forcing continuity, results in patterns resembling misfit dislocations. Thus, we observe oscillatory displacement curves with a period corresponding to the misfit dislocation spacing. On the one hand, to achieve this correction, we applied a constant shift to region-wise displacements, resulting in a wavy x displacement map with a period precisely matching the dislocation spacing. Alternatively, we can adapt the

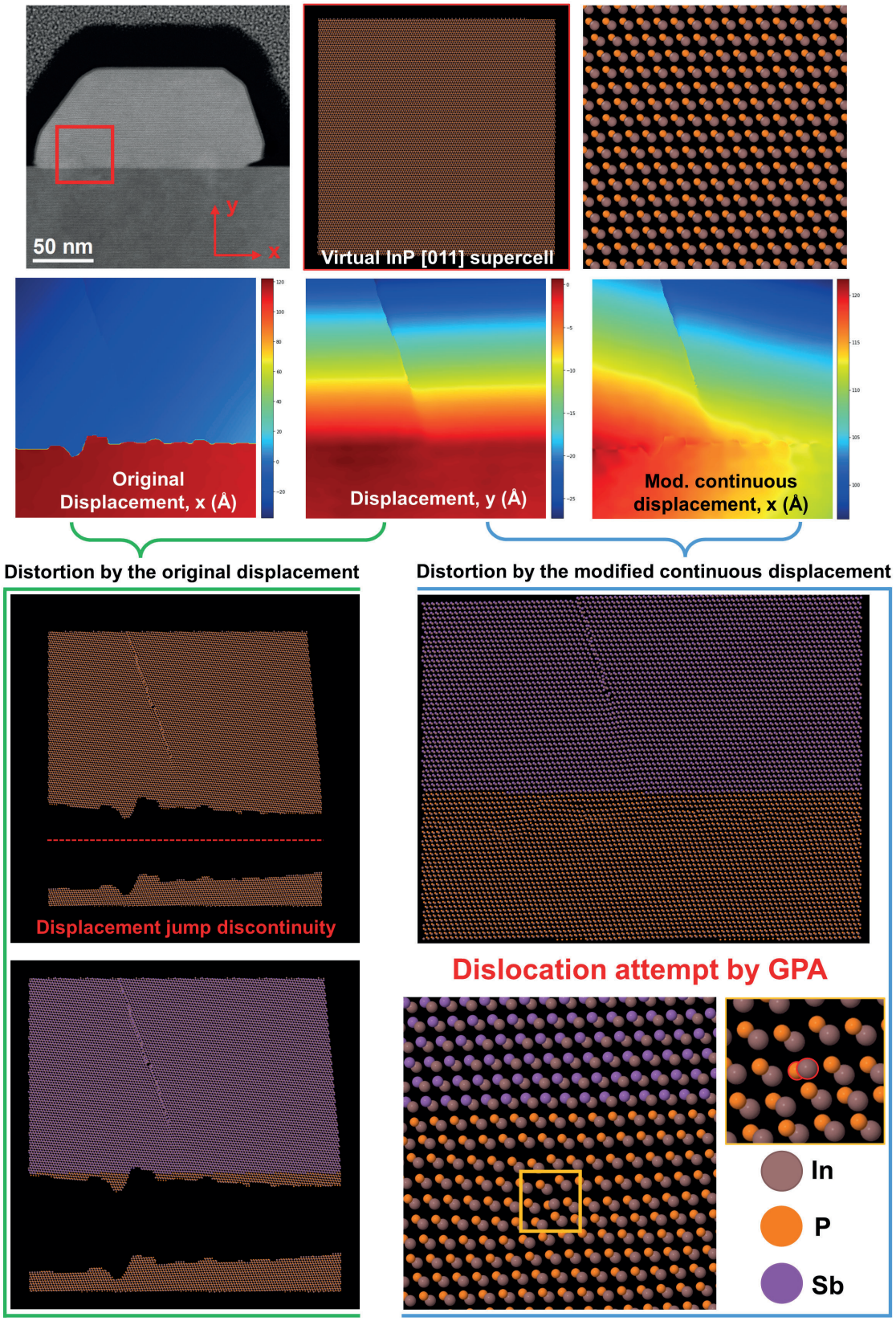


FIGURE 4.II.4: Implications of applying a strain field into a 3D Atomic Model (3DAM) for building realistic experimental models. Starting from the ongoing InSb nanowire on InP, we compute the virtual supercell to distort, and the displacement fields that do so. [37] The displacement field component in the x direction meets a jump discontinuity, which generates a huge unphysical gap in the supposed interface of the model (green margins). Removing the discontinuity from the x displacement map removes this gap and attempts to create the dislocations arising from the materials mismatch (blue margins).

vertical shifting condition at the pixel level, leading to a correction curve oscillating with the same period as the dislocations. The adaptive curve effectively smooths the displacement map, as demonstrated in Figure 4.II.4, where the modified continuous displacement map is showcased. It is important to note that the y displacement map remains continuous by default and undergoes no alteration. However, when we apply the modified distortion to the same initial InP virtual supercell, we generate a continuous InSb/InP interface. Notably, this interface exhibits noticeable deformations at the positions where dislocations are expected. The process inherently strives to create the necessary dislocations where they experimentally occur, as highlighted in the yellow inset in the same Figure 4.II.4. The orientation of the dumbbells changes due to the spiky displacement associated with dislocations. Nevertheless, since we manually mitigated the mismatch, the applied displacement field cannot adequately distort the supercell to naturally accommodate the dislocation array.

The creation of perfect models, especially of interfaces, in highly mismatched heterostructures, still needs further investigation. Nevertheless, this example strengthens how a custom and conscious use of this methodology can enhance the interpretation of how strain affects materials and its consequences on the device's performance. Therefore, the strain transfer process might also serve as an independent module with its standalone objective perfectly surviving outside the realm of its mother workflow.

4.II.3.4 Compositional information: Spectroscopic mapping and quantification

A complementary step addressing compositional adjustments can be applied in certain scenarios. Concretely, when the atomic model is built up from a single detected phase, either combining segmented regions with coincident phases or just considering a single region. Interestingly, if spectroscopic data is available and quantitative information on the present elements is reachable, we can use it to significantly enhance the accuracy of the atomic models. The workflow asks whether there are quantitative compositional maps available, and uses them to refine the model. It is worth noting that the way of obtaining the quantitative maps is out of the scope of this manuscript. In the studied examples, the extraction was done manually, although there are existing solutions with a certain extent of implemented automation. [118, 119, 120] Additionally, note that EELS is the spectroscopic source in all the provided examples. However, EDX would work the same way as long as quantitative elemental maps can be accessed.

The implementation of this step is very similar to what is explained in section 4.I.4. Once again, we make use of multi-scale template matching to locate the area where the spectrum image was taken from. For that, the (typically) STEM- Dark Field (DF) signal acquired in parallel to the spectroscopic signal, which acts as its spatial survey, is used as the template. A template which is matched in the main analysed

image, which acts as its query. Logically, the spectrum image must include a portion of the area depicted in the micrograph to adequately represent it. In fact, the template matching process is not unidirectional. Instead, it evaluates the sizes of the mapped areas, embedding the smaller one (acting as the template) within the larger one (acting as the query). Therefore, the size of the spectroscopically refined atomic model is defined by this template-query embedding. In other words, the optimal situation occurs when the spectrum image fully captures the area displayed in the micrograph.

After the template matching, the actual compositional information is applied to the model. Thus, the survey informs us on where the maps must be allocated. Additionally, the element-wise maps display the local probability that an atom from the model belongs to these element types. As a result, the quantitative compositional information is spatially interpolated to the atomic coordinate. Then, this atom receives a uniform probability distribution informing on the element type, for all the possible elements. In other words, the occupancy of that atom in that given coordinate for that given element type corresponds to the value obtained from its quantitative spectroscopic measurement. Interestingly, we can express this result as a non-collapsed probability distribution with as many atomic coordinate entries as element types can locally appear. That is, the elements spatially coexist with a certain probability. On the other hand, we can collapse these entries into a single one with complete occupancy (i.e., occupancy = 1) based on these probabilities. Thus, only an atom exists in every atomic position and the decision on the element type depends on its probability.

$$Occupancy_{x,y,z} = 1 = \sum_{i=1}^{N} P^{i}_{x,y,z} \tag{4.II.46}$$

where $P^{i}_{x,y,z}$ is the probability to find the element i from all the possible N elements in the spatial coordinate x, y, z.

Regarding the atomic substitutions, two scenarios need to be distinguished. Firstly, if the elements found in the spectroscopic analysis and the ones identified in the phase identification step are not coincident. Thus, no correlation between the atomic species and their preferred position within the unit cell can be established. As a result, the substitution is only ruled by the probability established by the quantitative maps. An example appears in figure 4.II.5, where the phase identification finds Ge as the best phase in the QW and in the two surrounding $Ge_{0.7}Si_{0.3}$ buffer layers (which makes sense given the proximity to the cell parameter of Ge) (device GSQ_4 from Chapter 3). Nevertheless, the EELS analysis is capable of discovering the actual binary distribution between Ge and Si (fig. 3.6). Consequently, as Si is not present in the identified phases, the substitution is just based on the local probability explained by the Si and Ge maps. In other words, randomness compensates for a knowledge

gap originated by the intrinsic limitations of the phase identification based on reciprocal space analysis. Interestingly, this particular case is not negatively affected by this fact. That is, the Si and Ge with diamond cubic structure share their Wyckoff positions, for which a position-aware substitution and a random one output the same result. Importantly, although models are created as translationally invariant systems, this random substitution and the collapsing of the probabilities breaks this translational invariance. As a result, the thicker the model the closer it will be to reality, as the calculated local probability of the occupancies is spread through a larger distance (i.e., if the thickness was infinite, the occupancies would exactly match the probabilities). This effect is visible in figure 4.II.12 from section 4.II.5.1, where the 50 Å atomic model realistically resembles the experimental case, whereas the 12 Å cannot compensate the HAADF contrast of the Si substitutions due to the few atoms per column scenario.

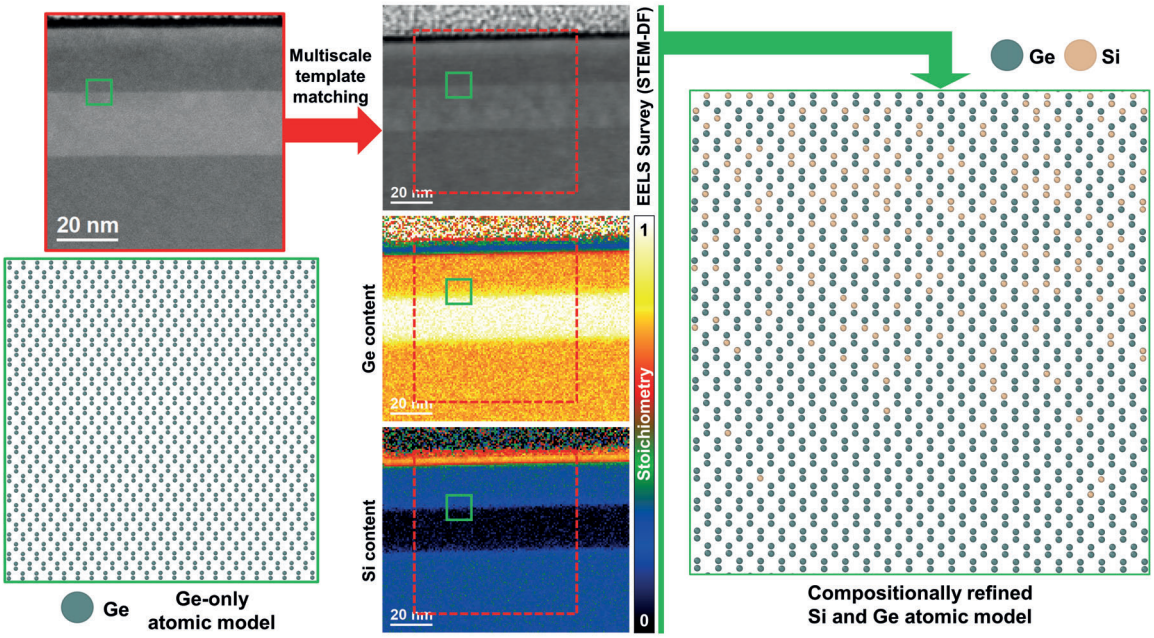


FIGURE 4.II.5: Workflow of the compositional refinement of the atomic models based on spectroscopic data. The figure showcases a Ge quantum well embedded in two $Ge_{0.7}Si_{0.3}$ buffer layers (device GSQ_4 from Chapter 3). [38] The phase identification process finds Ge in the three main regions, and builds an atomic model only based on Ge atoms. The available complementary Electron Energy Loss Spectroscopy (EELS) maps are used to allocate, through a multi-scale template matching step, the local stoichiometry of both Si and Ge. Eventually, this information is mapped back into a compositionally accurate atomic model.

On the other hand, the substitution can be refined if the elements identified in the spectroscopic analysis and the ones identified in the phase identification step are indeed coincident. Specifically, each atomic species in the original model built, based on phase identification, should correspond to a spectroscopically identified element. The pathway unfolds as follows:

1. Selecting an atom in the base model: We begin by choosing an atom from the base model, coming from the phases identified. Its specific element type is crucial for subsequent steps.

2. Unveiling Wyckoff Multiplicity: We evaluate the Wyckoff multiplicity of the selected atom's unit cell base positions. The Wyckoff multiplicity reveals how many atoms contribute to building the conventional unit cell from the original primitive unit cell. Keep in mind that the conventional unit cell results from applying all characteristic symmetry operations of the space group to the primitive unit cell.

3. Comparison and Compensation: Next, we compare the multiplicity of the selected atom with the multiplicity of the elements to which it should be substituted. Additionally, we factor the quantitative spectroscopic information from their maps. This comparison guides the substitution process, ensuring the best agreement between the cell symmetry and the spectroscopy results.

In summary, aligning phase identification with spectroscopic data enhances the precision of atomic substitutions, leading to more reliable models.

Thanks to these considerations, the elemental substitution cares about the preferential position of individual elements in their crystal phases. For instance, ionic materials will only substitute the cations for cationic species and anions for anionic species. Thus, the substitution is filtered by whether it is physically realistic that the element would sit in that position or not. Hence, the probability of substituting is only weighted by the spectroscopically quantified elements that can fit into the relative position within the unit cell (equation 4.II.47).

$$P^{i}_{x,y,z} = \frac{P^{i}_{spectr;x,y,z}}{\sum_{j=1}^{M} P^{j}_{x,y,z}} \tag{4.II.47}$$

where $P^{i}_{x,y,z}$ is the probability to substitute the atom in x, y, z by the element i from all the possible M elements spatially equivalent between crystal phases, and $P^{i}_{spectr;x,y,z}$ is the quantification of element i from the spectroscopy experiment, interpolated at x, y, z.

This position-aware substitution is especially relevant and useful when spectroscopic information is not available. This concept is developed further in section 4.II.3.5, where we discuss the elemental substitutions when the only source of compositional awareness is the identified phases.

4.II.3.5 Compositional information: Segmentation and symmetry equivalences

When no spectroscopic data is accessible, the only compositional information lies in the materials recognised by the phase identification. In section 4.II.3.3, we specified how the different epitaxy scenarios rule the way the atomic models are engineered. In situations with a tight epitaxy (i.e., paired spots) but still fairly resolved mismatch to allow a distinctive phase identification, an additional model is built similarly to the perfect epitaxy and single identified phase case. That is, the supercell occupies all its segmented regions with the same crystal phase. This crystal phase is the virtual unit cell belonging to the region considered as the reference for the single GPA calculation. Afterwards, we use our only source of compositional data, the additional identified phases, to substitute the base model with the new elements from these phases. This approach is particularly accurate when two crystal phases with the same space group participate in the heteroepitaxy. Remarkably, this is the habitual case in the semiconductor industry, where most of the epitaxies are formed by pairs of II-VI, IV-IV, III-V, or IV-VI compounds.

In section 4.II.3.4, the idea of positionally-aware substitution is first presented. When spectroscopy is involved, we check the spatial equivalences of the elements to select the ones that weigh the probability of locally substituting the atoms. However, this procedure shines when only identified phases are available. Every atom in the supercell is spatially evaluated to determine its expected phase (i.e., in which segment it stands). Then, its element type is inspected in the reference unit cell file to figure out the relative positions of that element and their respective occupancies. As a result, we are aware of which symmetry operations could potentially allocate that atom in that position. In fact, as symmetry operations are directly encoded from the primitive unit cell description, we characterise the Wyckoff symbol of these relative positions, and more importantly, their multiplicity. Therefore, the total number of atoms N^i of that particular element type i contained in the reference conventional unit cell is defined by equation 4.II.48,

$$N^i = \sum_{j=1}^{M} Wyck_j^i Occ_j^i \tag{4.II.48}$$

where M is the number of relative positions of the element i in the unit cell, and $Wyck_j^i$ and Occ_j^i are, respectively, the Wyckoff multiplicities and occupancies of the element i in the relative position j.

The next step is mapping the relative positions j found for that particular element type i back to the crystal phase that is supposed to fill that coordinate. Then, the candidate elements that could act as substitutes for the analysed atom are all the elements which share a relative position j with the reference phase. Consequently, the probability that the atom, being an element i, is substituted by element k out of

the K elements in the target unit cell with a relative position l coincident with the relative position j of the reference unit cell, is defined by equation 4.II.49,

$$P^k = \frac{\sum_{l=1}^{L_k} Wyck_l^k Occ_l^k}{\sum_{l=1}^{L_k} \sum_{k=1}^{K} Wyck_l^k Occ_l^k} \tag{4.II.49}$$

where L_k is the number of coincident relative positions for the same element type k in the target unit cell.

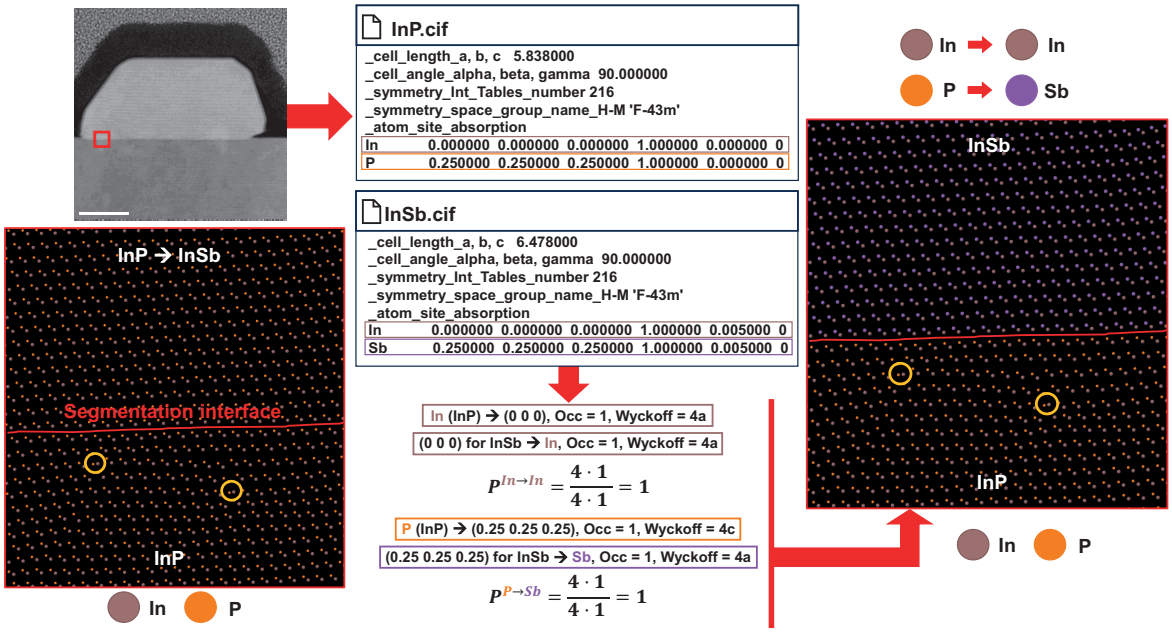


FIGURE 4.II.6: Workflow of the substitution of element types in the atomic models based on compositional information from identified phases. The workflow is showcased by an InSb on InP nanowire, which is originally fully built with an InP supercell. [37] The In and P atoms are evaluated within their unit cell of origin to assess their relative positions and their equivalence with the target unit cell, the InSb. The probabilities of substitution are computed following the equation 4.II.49. As a result, we end up with the adequate atomic arrangement of InSb and InP in a sharp interface. To prove that the before and after correspond to the exact same region, the light yellow circles pinpoint the dislocations that are wrongly generated by the strain transfer methodology (note that dislocations are not optimally considered in the single phase supercell approach (section 4.II.3.3)). The scale bar in the micrograph stands for 50 nm.

Much like in the spectroscopy case study, we can either collapse these probabilities into a unitary occupancy or leave them as the actual non-collapsed occupancies. Again, the decision on which one to turn to must depend on the final purpose of the generated model and its requirements. In any case, outlining the probabilities with the mentioned criteria enables, for example, the substitution of cations and anions, in any binary, ternary or multinary semiconductor alloy, keeping the global stoichiometry as originally encoded in the unit cell. This can be seen in figure 4.II.6, with a InSb on InP nanowire. Interestingly, the initial distorted model is entirely built of a single InP supercell. Then, the segmentation limits the In and P atoms that require

a substitution to the crystal phase identified in that specific region. The substitution proceeds as explained previously, by computing the probabilities of substitution based on their relative positions within the unit cells, their occupancies and Wyckoff multiplicities (equation 4.II.49). Importantly, this case is a simple binary case, but a multinary one would follow the same path by substituting the cations only by the cations and the anions by the anions, and ensuring a final global stoichiometry as initially defined in the target unit cell.

4.II.4 Proofs of the full workflow

In line with Part I, where we established the general nature of phase identification and the complete structural characterisation (Section 4.I.8), we provide instances that validate the entire workflow. Starting from segmentation and extending to the creation and refinement of 3D models, we omit simulations in the cases where they have been performed, focusing instead on the fully automated process as a whole. Further discussion specifically on these simulations can be found in the next section 4.II.5.

- (Figure 4.II.7) The recurring outstanding GSQ_4 device (Chapter 3) constitutes the most successful and fruitful proof of the whole developed workflow. In addition, interesting simulations are further extracted from it. The next sections explain this in detail.

  Figure 4.II.7 wraps together the results we have already seen for this $Ge_{0.7}Si_{0.3}/Ge/Ge_{0.7}Si_{0.3}$ quantum well. [38] The displayed mid-magnification micrograph simultaneously provides an accurate segmentation and region-wise phase identification. The three main crystalline regions pinpoint the same identified Ge Fd-3m phase and spatial orientation. Therefore, heteroepitaxy can be claimed. In fact, the knowledge retrieved so far indicates a "homoepitaxy", as the crystal phases in every region are supposed to be the same Ge diamond cubic structure. The complementary EELS analysis will later locally define the binary stoichiometry of Si and Ge (fig. 3.6). Importantly, the "homoepitaxy" awareness enables the automated strain analysis utilizing GPA. As detailed in figure 4.II.3, the strain maps highlight any distortion or imperfection represented by the micrograph. These may arise either from the real physical deformation of the sample, or from the imaging process. The latter is the cause of the waviness displayed in the present GPA maps, due to the varying on-axis condition along the sampled region of interest. This is visible in the magnified inset (square with red margins) from figure 4.II.3. However, its consequences reach beyond the strain mapping. Thus, the 3DAM creation based on GPA eventually showcases these distortions as unphysical artefacts which restrict the usability of this 3D model to only its defect-free segments (outside the dark-blue box and the bottom right edge of the model).

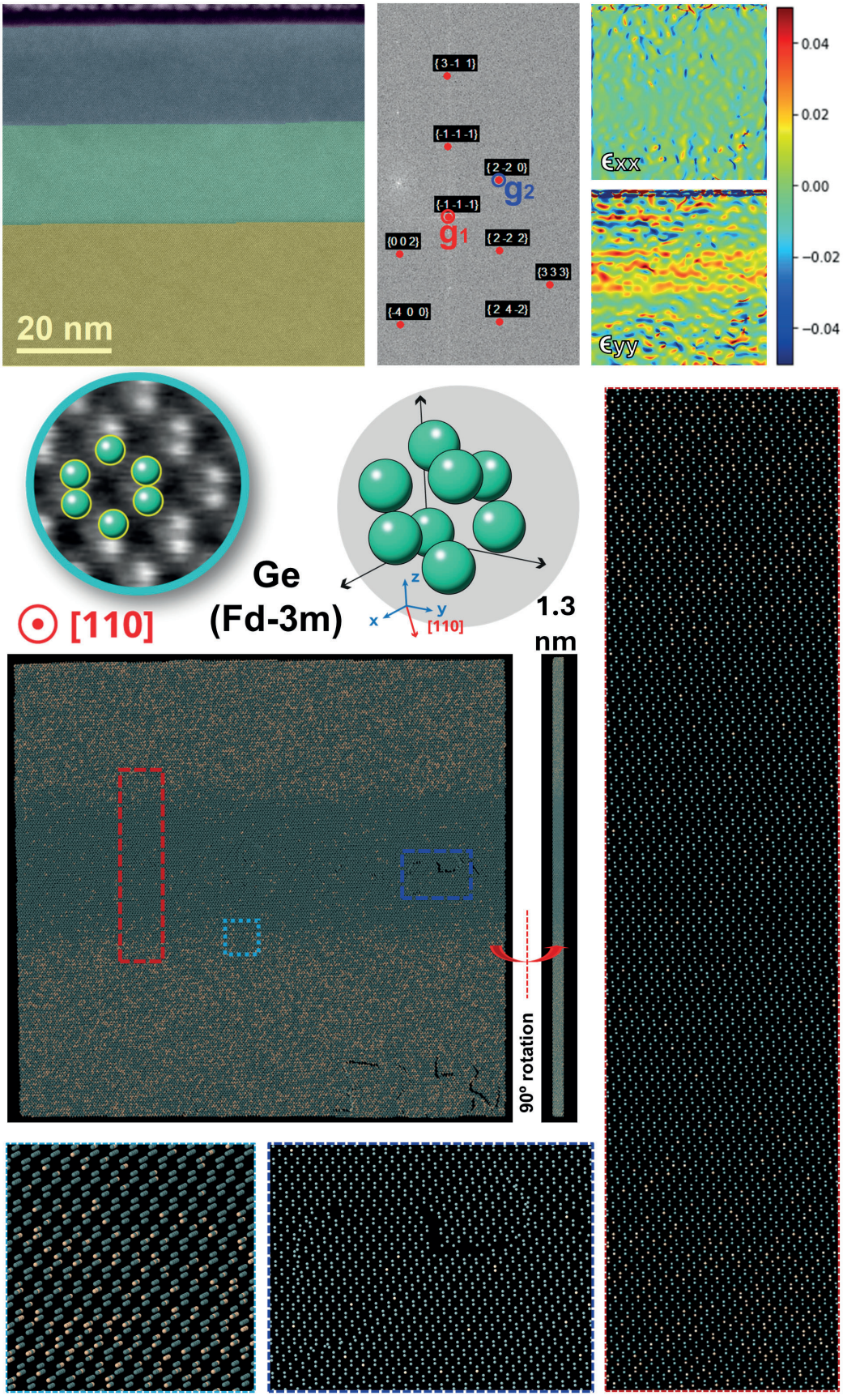


FIGURE 4.II.7: Proof of the full workflow, from the segmentation to the atomic model creation, of a Ge quantum well embedded in two $Ge_{0.7}Si_{0.3}$ buffer layers (device GSQ_4 from Chapter 3). [38] The selected micrograph could be considered a large field of view one, although still containing atomic resolution.

The intermediate FOV displayed in this micrograph makes it a suitable candidate to exploit the also retrieved FEM for large-scale simulations.

- (Figure 4.II.8) Same device as in the previous example (fig. 4.II.7), GSQ_4, but through a higher-magnification STEM-HAADF micrograph. [38] While the phase identification is benefited by a pristine representation of the atoms (i.e., dumbbells of the identified Ge [011]), the segmentation is not as sharp as in the previous lower-magnification micrograph. In any case, as we identify the same "homoepitaxy", we create a single supercell for the three segmented regions out of the reference virtual unit cell extracted from the GPA calculation. The GPA reveals a homogeneous and smoother strain field. After transferring the strain field into the base supercell, and using the EELS quantitative compositional information (fig. 3.6), we end up with a seamless and artefact-free 3DAM model. Constructing a 1.3 nm and a 5 nm thick device demonstrates that accessing either one is simply a matter of computational power and resources. Remarkably, a single unit cell is sufficient to represent translational invariance along z. By enforcing the non-collapsed atom-wise stoichiometry to match the EELS experimental stoichiometry, we emulate an infinitely thick model with collapsed occupancies.

- (Figure 4.II.9) Same scenario as in figure 4.II.8, but in a Si QW on $Ge_{0.3}Si_{0.7}$ representing the SSQ_3 device. [54] The segmentation and phase identification work smoothly. However, the phase identification selects Ge over Si, despite the interplanar distances being closer to Si than to Ge. Consequently, the virtual supercell is built upon Ge, although, crucially, with the selected interplanar distances corresponding to the relaxed $Ge_{0.3}Si_{0.7}$ taken as reference for a smooth GPA calculation. Interestingly, the false positive Ge identification is reverted to Si when applying the experimental EELS stoichiometry to a final seamless model which resembles a $Ge_{0.3}Si_{0.7}/Si/Ge_{0.3}Si_{0.7}$.

- (Figure 4.II.10) The whole workflow applied to the recurring HRTEM micrograph of three CuTe nanoparticles. [39] The SAM-based segmentation unlocks particle-wise phase identification and FFT indexing. The result is three almost identical FFTs coming from each particle. The difference is a subtle relative rotation of the found [001] zone axis of the CuTe Pm-3n phase. Remarkably, the workflow autonomously finds the relative rotation between the particles and opens a mask to compute the relative strain concerning the bottom right particle, considered as the reference. The chosen g vectors are centred in the central spot of the threefold peak group per equivalent crystal plane. The reference particle defines the rotation reference (purple dashed line), which is catalogued in the central image of the panel. This central image informs us about the experimental rotation between particles (in absolute value, 4.5º and 11.5º).

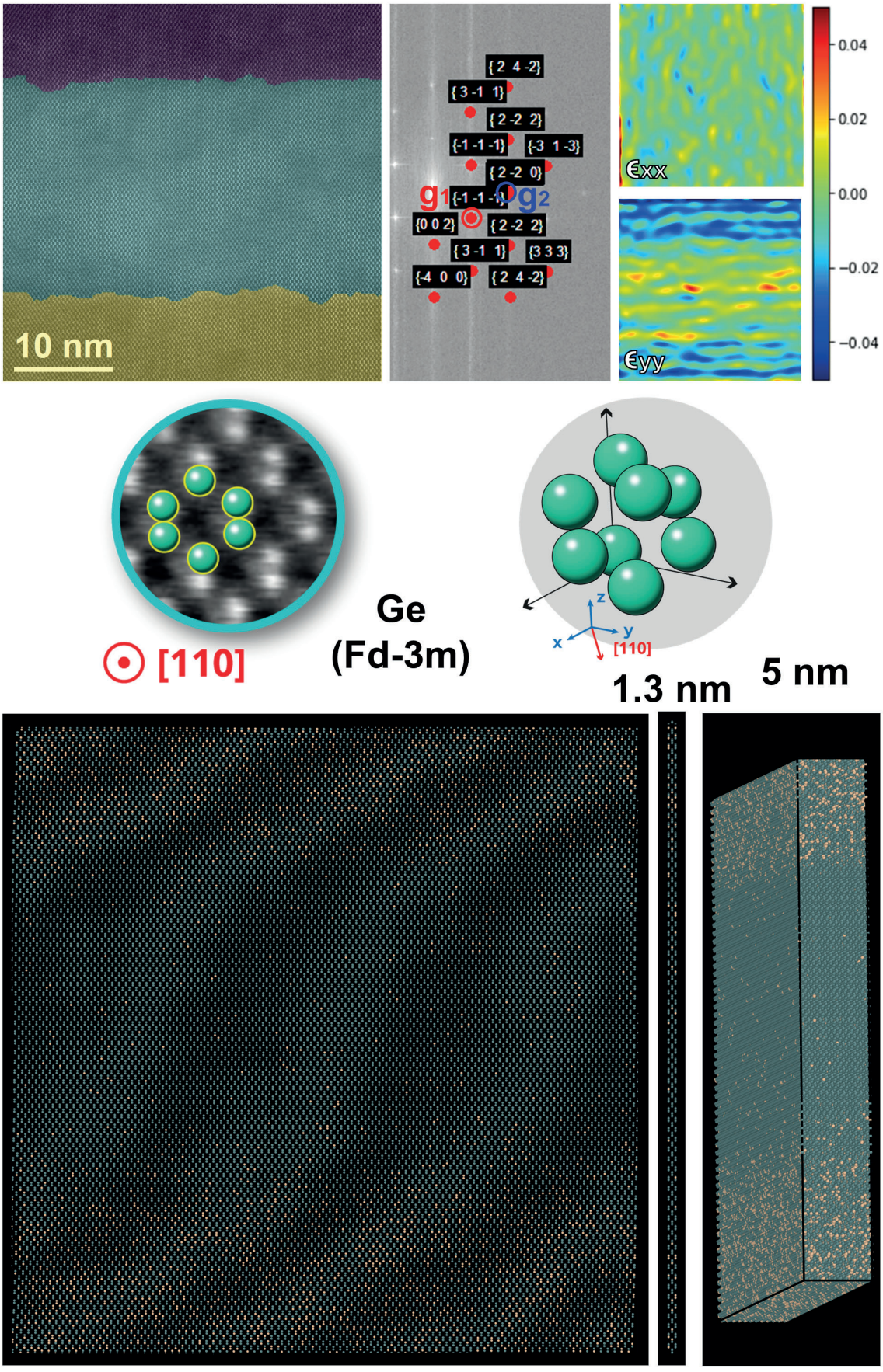


FIGURE 4.II.8: Proof of the full workflow, from the segmentation to the atomic model creation, of a Ge quantum well embedded in two $Ge_{0.7}Si_{0.3}$ buffer layers (device GSQ_4 from Chapter 3). [38] The selected micrograph could be considered a reduced field of view one, with crystal-clear atomic resolution.

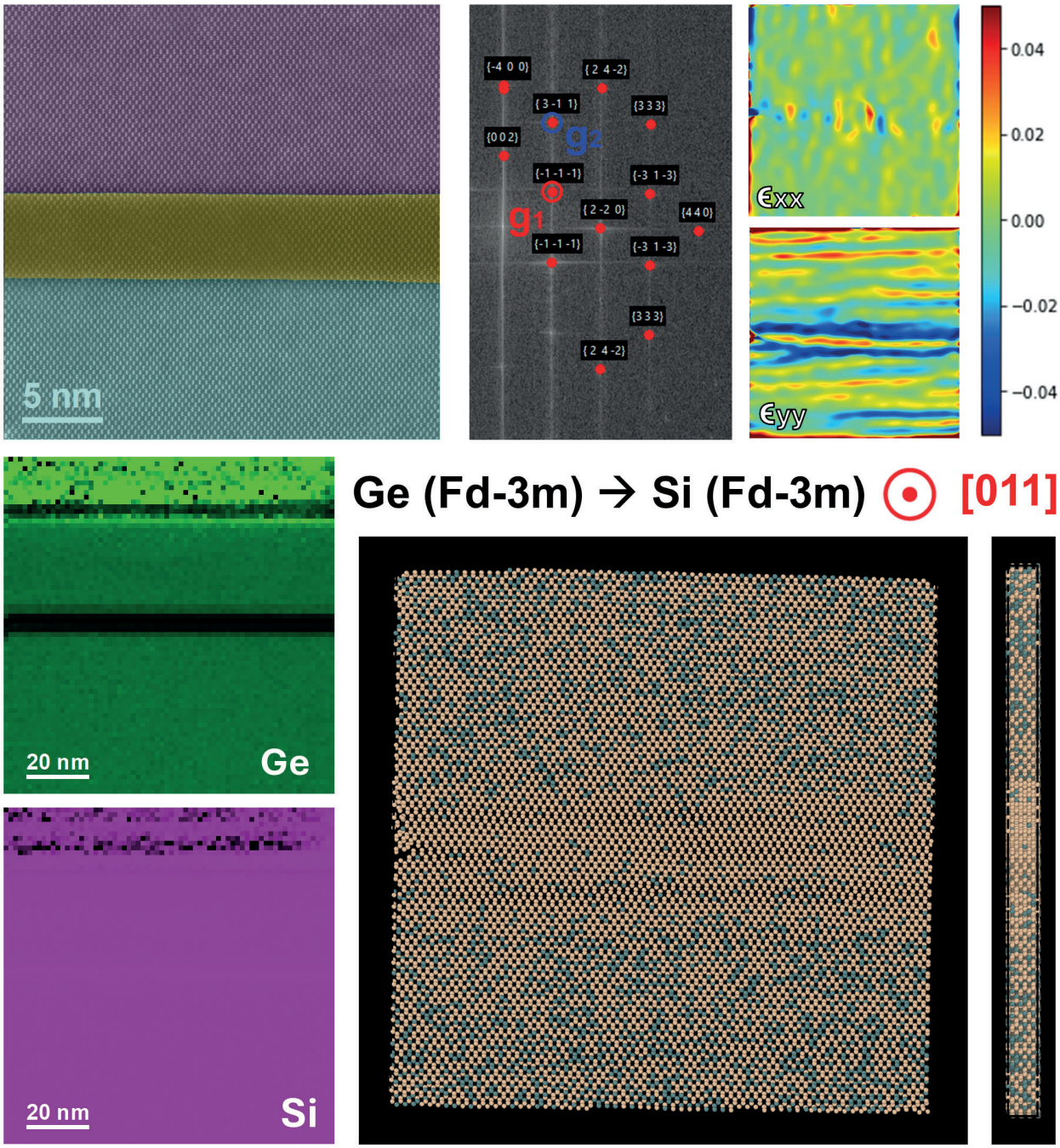


FIGURE 4.II.9: Proof of the full workflow, from the segmentation to the atomic model creation, of a Si quantum well embedded in two $Ge_{0.3}Si_{0.7}$ buffer layers (device SSQ_3 from Chapter 3). [54] The selected micrograph could be considered a reduced field of view one, with crystal-clear atomic resolution.

Interestingly, the algorithm finds that the bottom right (reference) and bottom left particles are comparably oriented, as if they were forming an ordered pattern resembling a misoriented heteroepitaxy (in this case, a "homoepitaxy", again). As a result, its joint model shares the same original virtual unit cell with the same initial spatial orientation. The separation of two particles rises

from the segmentation. Conversely, the relative rotation angle of the top particle lies beyond the epitaxy threshold. As a result, its original building block is independent and builds its orientation based on its FFT. However, we find a rotation not reaching the 11.5° found experimentally, but 9°. Indeed, this remarkable fact cannot hide the fact that major distortions are still present in the global model.

As mentioned, the starting point for the two lower particles involves a perfectly aligned in-(image)-plane rotation originating from a shared building block. The observed 3° rotation in the 3DAMs occurs after strain transfer, which is remarkable given the complexity of accurately correlating independent and *a priori* uncorrelated nanoparticles using strain maps. However, this unique observation does not conceal the nonphysical distortion present on the edges of the bottom-left particle model. The distortion likely arises due to beam delocalisation effects at the interface, impacting the strain maps. On a different level, the top particle model is adversely affected by an inaccurately represented strain field. The issue arises because the GPA mask only partially overlaps with the FFT peaks corresponding to this particle. Therefore, the GPA strain considers only a part of the particle, resulting in a collection of distortions that ultimately deform the 3D model.

The discrepancy between the experimental angles and the 3DAM ones can be attributed to the segmentation-peak finding intercommunication choice. Figure 4.II.10 relies in the "crop" setting. As previously mentioned, if the crop significantly reduces the total number of pixels by isolating a small region, it can negatively impact its FFT calibration and any measurements derived from it. Thus, even if the peaks are well-identified, the resulting undersampled reciprocal space might shift the obtained values the same way a calibration drift would do. It can happen both in distances and angles, affecting the identification of phases and the consequent model created.

Last but not least, this is the only full workflow proof we provide in a system without translational invariance. In the present example, it would be reasonable to hypothesize that the particles exhibit a cubic-like morphology. In this case, expanding the model along the z dimension to form a cube could be potentially useful as an approximation of a 3D model which could be morphologically refined in 3D via the voxel segmentation of a tomogram. In any case, the obtained result builds on how promising the workflow can be to accurately model rotations and deformations in complex systems.

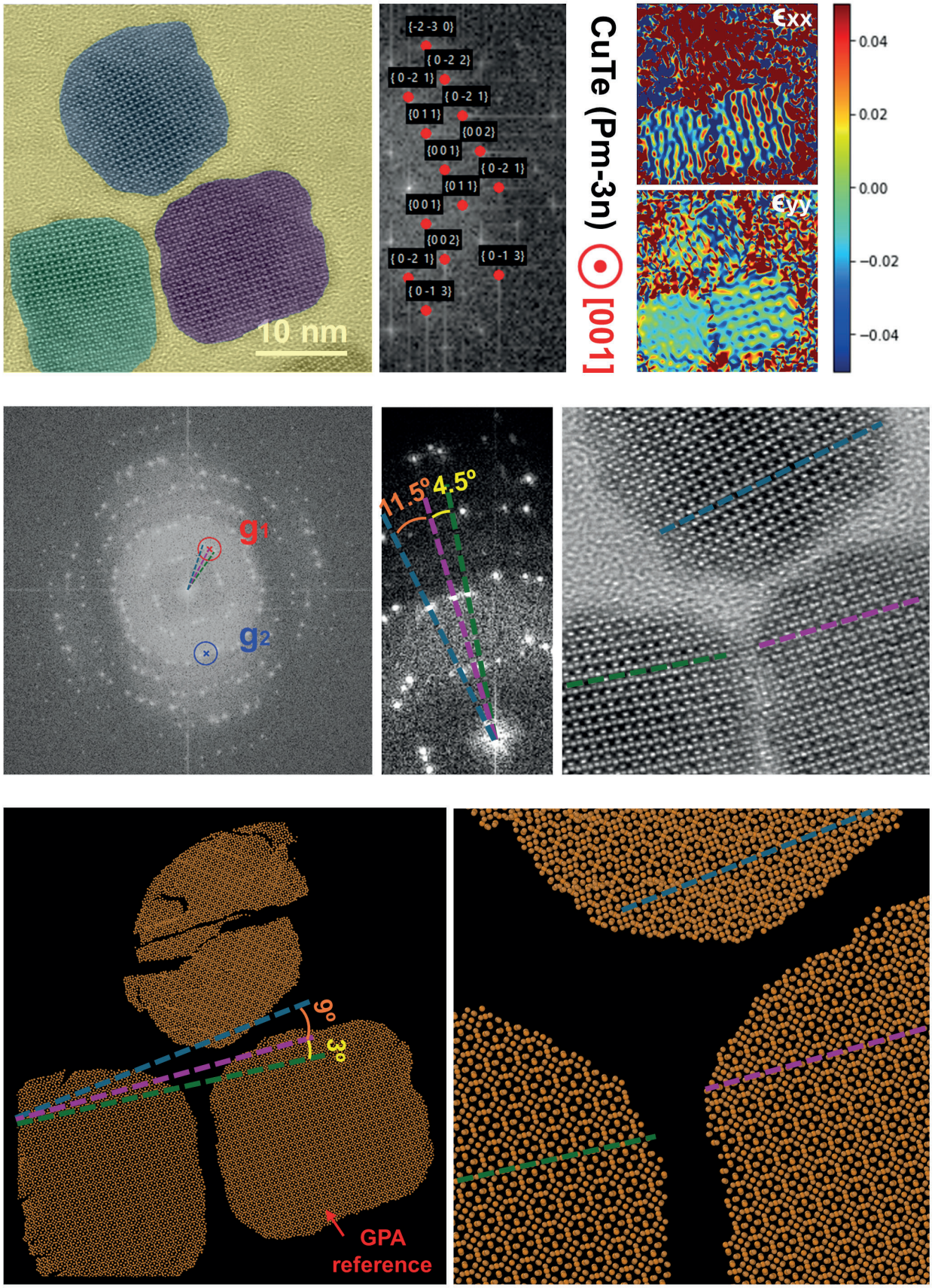


FIGURE 4.II.10: Proof of the full workflow, from the segmentation to the atomic model creation, of a parallel beam Transmission Electron Microscopy (TEM) image of three CuTe nanoparticles. [39]

- (Figure 4.II.11) Finally, we also present the recurring InSb NW on InP. Although its 3DAM has already been extensively discussed in Figure 4.II.4, we now provide an overview of the overall process leading to the presented 3D models. The segmentation process effectively collaborates with the phase identification to distinguish the two crystal phases: InSb and InP in a zinc-blende structure. These phases are oriented along [011], with a diagonal arrangement of the dumbbells. A reliable FEM model is extracted based on this segmentation, and it proves to be the most suitable simulation strategy for a system of this size.

  The initial step of the workflow involves identifying the heteroepitaxy condition, which prompts the GPA mask to open and encompass the intensity peaks from both phases. We obtain the atomic model of the entire nanowire, which consists of approximately 2.2 million atoms for a thickness of 2 nm. To promote discussion, we focus on a specific region of the device marked by the white inset. In this region, the rotation map highlights the stacking fault and the array of misfit dislocations. Therefore, we limit the strain transfer to that specific region of the device and its corresponding fragment of the 3DAM. Our goal is to replicate the stacking fault and the dislocations. We explore two approaches, as discussed in subsection 4.II.3.3. First, we employ the regular single strain calculation approach, resulting in the model shown in the left column of Figure 4.II.11. Interestingly, this model is also detailed in Figure 4.II.4, after modifying the x displacement map to ensure its continuity. However, the dislocations do not emerge as expected; instead, we observe traces of a preliminary displacement that could generate them. Nonetheless, a major issue arises: the absence of lattice expansion from the virtual InP cell, which would create the experimental mismatch. This absence is likely due to the forced continuity condition applied to the displacement. Additionally, we encounter a discrepancy between the segmentation interface and the geometrical phase, which depends on the choice of g vectors and the masked reciprocal space. Furthermore, the strain transfer results in lattice overlapping of the stacking fault, causing a flip in the orientation of the dumbbells. The obtained pattern does not perfectly match the ideal one with perfectly rotated and ordered dumbbells. However, we do observe a preliminary ordering of the atoms that attempts to emulate the expected arrangement. Interestingly, manually deleting some overlapping atoms reveals a pattern resembling the desired order. This incipient version of the stacking fault could be perfectly imitated if an atomic identification routine working in real space could reposition challenging atoms to their actual locations.

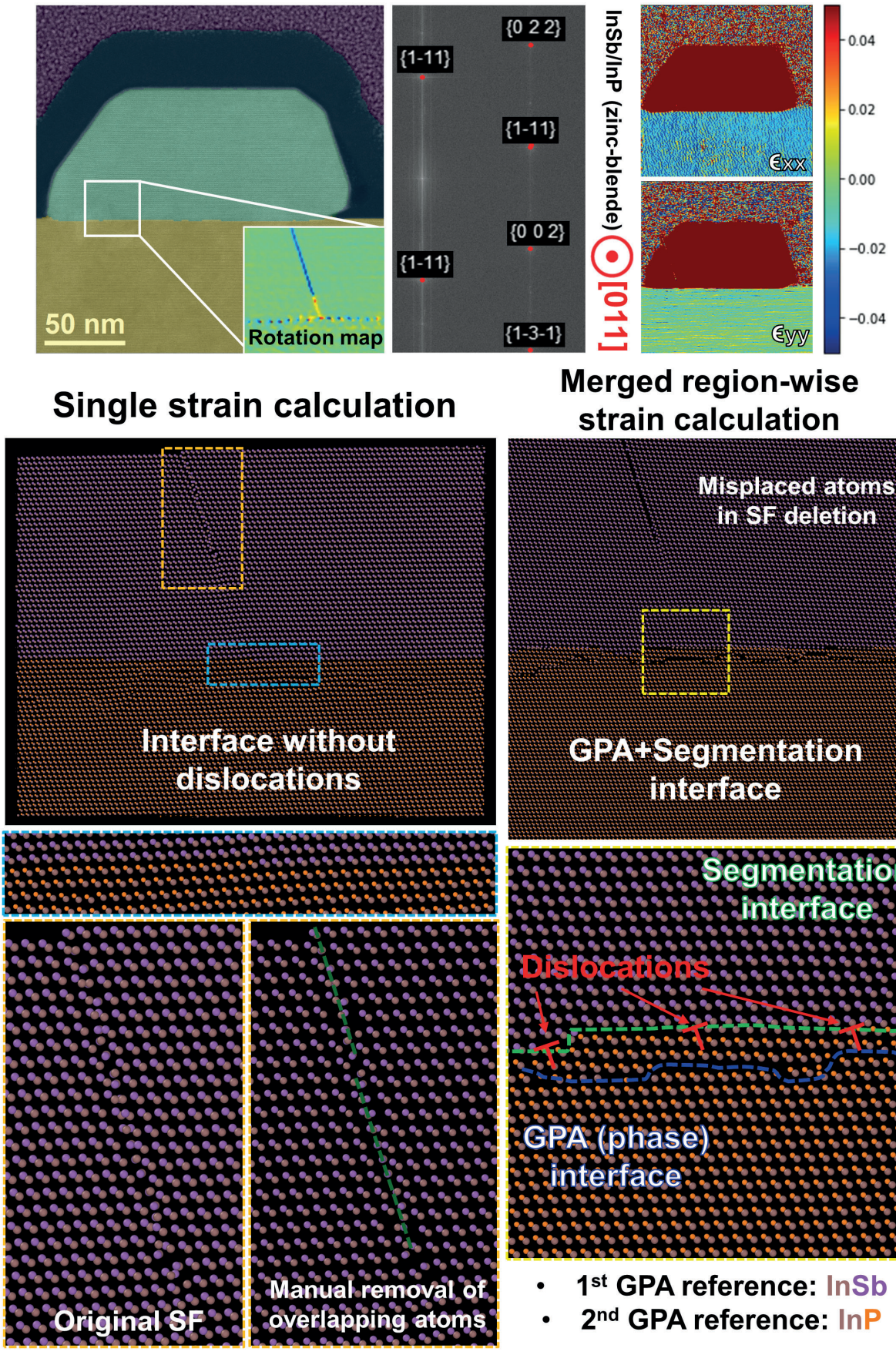


FIGURE 4.II.11: Proof of the full workflow, from the segmentation to the atomic model creation, of a transversal view of an InSb nanowire grown on InP. [37]

Alternatively, the strain field is transferred into the atomic model through region-wise strain calculation (fig. 4.II.11, right column). We merge the atomic models created, first with the reference in the InSb, and second with the reference in the InP. The interplanar distances are now well captured as both models raise from separate virtual unit cells built upon the experimental information extracted in each material domain. The merging is then entirely ruled by the interface found by the segmentation algorithm. The challenge rises in the slight rotation misalignment between the interface of the segmentation and the horizontal planes of the semiconductor. They cause a stepped profile which blurs the interface and even locally cuts the dumbbells by half. Conversely, the correct reproduction of the InSb and InP lattices inherently reveals the dislocations and their spacing, despite the stepped interface and the still mismatched distortion of the geometrical phase interface.

An essential aspect that requires attention is the polarity of the dumbbells and how they are represented in the 3DAM. Unfortunately, this information cannot be solely extracted from reciprocal space. While it may not be critical for apolar Si and Ge diamond cubic structures, it becomes relevant for II-VI and III-V semiconductors like InSb and InP, where polarity certainly matters. [46, 121] Currently, in our workflow, we manually introduce this information by inducing a 180º in-(image)-plane rotation of the model based on the identified polarity. As a potential future solution, we could explore an algorithm that draws an intensity profile in real space along the dumbbell direction, automating the process (see section 4.I.5 for further discussion).

4.II.5 Experimental simulation of key functional properties

The interest of the derived 3D models, including both the FEM and 3DAM, extend beyond providing a comprehensive 3D view of the device and its structural building blocks. These models serve a crucial purpose: enabling what we refer as experimental simulations. These simulations rely on a certain level of theory, but are entirely based on experimental data. Remarkably, the fidelity and accuracy achieved result from our end-to-end workflow, which unlocks possibilities that would be practically unattainable through manual methods. Despite relying on theoretical principles, our approach allows us to explore the functional properties of experimental devices, even during operation. These devices exhibit fine and characteristic features that set them apart from the theoretically perfect, ideal devices initially conceived and targeted. By comparing simulations of the theoretically envisioned device with those of the experimental one generated through our workflow, we can identify discrepancies in the results. These discrepancies provide valuable insights into the subtle factors affecting the optimal performance of the device. As a result, the materials research cycle becomes more aware of the experimental nuances, their origins, and how to address them effectively to achieve the ultimate device performance.

4.II.5.1 STEM simulations

The atomic models can also be used to simulate electron micrographs, mainly for model validation and hypothesis testing. Figure 4.II.12 showcases the linear STEM simulation of two atomic models of the same device, but with different thicknesses (12 and 50 angstroms). These simulations were conducted using the STEM_CELL software, employing linear STEM mode at 300 kV, with a C aperture of 15 μm, and a defocus of 0 nm. [41, 42] The input file containing the atomic model has a universal format, although it must be adapted for the specific software being used. For instance, in STEM_CELL, the file should specify the supercell dimensions and present the element types in atomic number format.

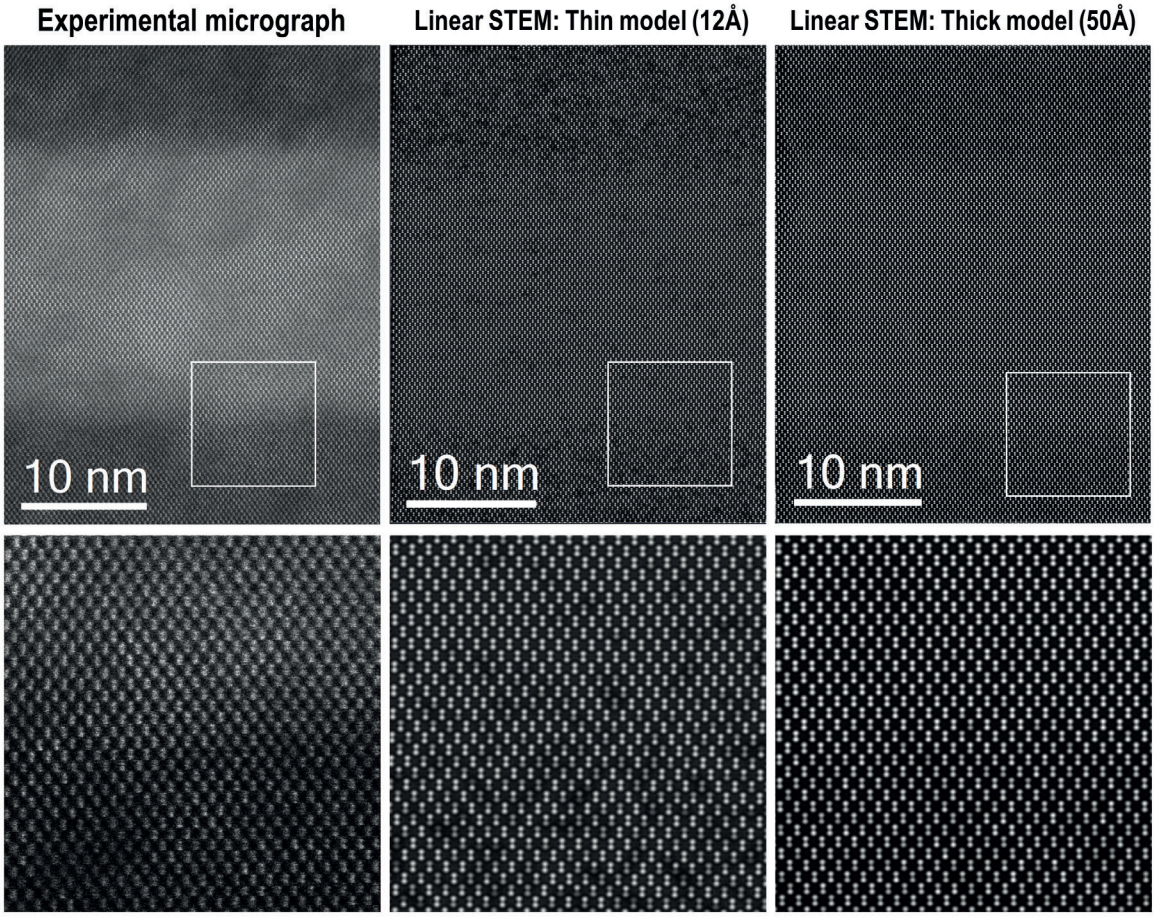


FIGURE 4.II.12: Scanning Transmission Electron Micrograph (STEM) of a Ge quantum well embedded in two $Ge_{0.7}Si_{0.3}$ buffer layers, and two linear STEM simulations of the system (device GSQ_4 from Chapter 3). [38] The models from which the simulations are processed are automatically obtained from the workflow and compositionally refined with Electron Energy Loss Spectroscopy (EELS) data. The thin one extends for 12 angstroms while the thick one is 50 angstroms thick. The simulation highlights how, when the translational invariance scenario is valid, the thicker the sample the finer the convergence towards the experimental observation. The scale bars stand for 10 nm.

The same strategy would apply to a more complex and higher fidelity simulation, for instance involving the multislice formalism. Therefore, all the benefits of multislice

are directly accessible through the resulting output model, too. In the presented case, though, we opted for a simpler approximation. Thus, the results would be comparable in the present system, and if any, the differences would not be worth the excessive computational time required for a multislice simulation of the present system unless quantitative (S)TEM is aimed.

4.II.5.2 FEM simulations

In the present section, we validate the workflow's end through FEM. We demonstrate how the segmentation step is essential to capture the morphological essence of devices and samples, and how this information can be used as the core of FEM simulations. Particularly, we used the models extracted automatically to assess the effect of lamellae thinning in the measured strain. The path is equivalent as in Chapter 3, as displayed in figure 3.7 for device GSQ_4. [38] We compare the 3D model of a thinned lamella with an infinitely thick one to distil whether the process of thinning affects the experimental strain we retrieve, and importantly how as a function of lamellae thickness. The concept is illustrated by using the InSb NW on InP sample, already widely discussed in the dissertation. [37] Intriguingly, this sample provides a rich platform to deploy multiple FEM considerations. They outshine the possibilities of the whole workflow we propose and the ease of capturing nuances that could only be identified by considering experimental data.

The main steps of the process are detailed in figure 4.II.13. Step 1 refers to the analytical outcome retrieved from Part I of the workflow, with special emphasis on the segmentation process (as discussed in Section 4.I.1). In this step, the (S)TEM micrograph is segmented into regions labelled by integer. Specifically, the HAADF STEM intensity is converted into a set of vertices, contour vectors, and boundary elements. The calibration of the micrograph allows us to convert these boundary elements into spatially meaningful vertices, which are then encoded into a .gds file (as shown in Figure 4.II.13, step 2). As highlighted in Section 4.II.2, the .gds format is universally understood by FEM software. It provides an intuitive way to load pre-built shapes from devices, particularly those related to semiconductor chip manufacturing involving lithography steps. [106, 107, 122, 123] For our calculations, we rely on the COMSOL Multiphysics software. [107] Properly calibrating the micrograph ensures that the FEM software captures the experimental dimensions of the sample, aligning with the morphological features delineated during segmentation. It is essential to address pixelation (aliasing) carefully to prevent future artefacts arising from unreal pixel shapes (as discussed in Section 4.II.2).

The third step creates the volume out of the 2D closed contour profile. We consider translational invariance as any transversal cut must end up with the same shape (fig. 4.II.13, step 3). Conveniently, this assumption perfectly fits most of TEM lamellae and is especially interesting for the scientific problem we are trying to address.

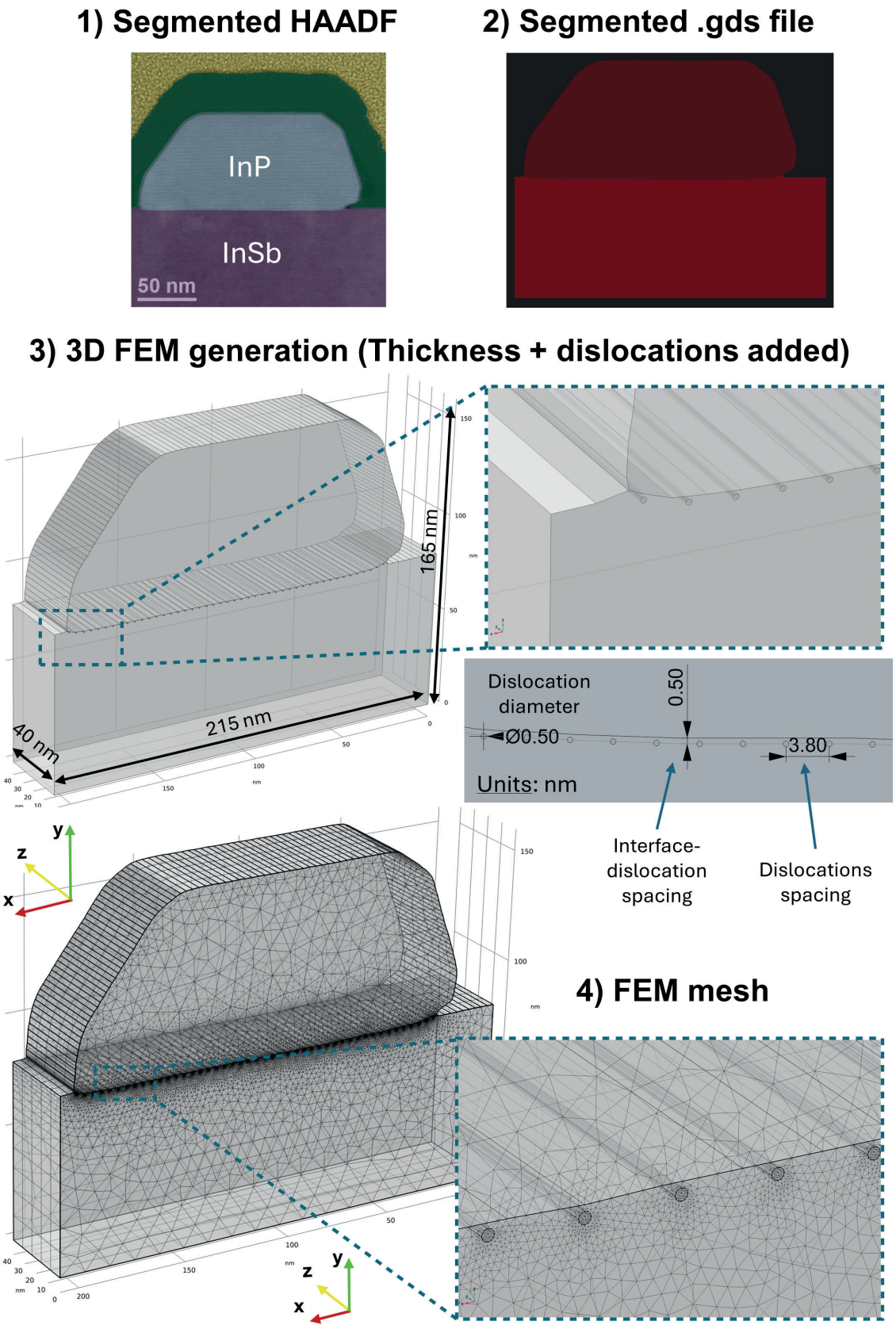


FIGURE 4.II.13: Generation of the 3D Finite Element Models (FEMs) from the automated data analysis of the workflow. 1) The segmented High-Angle Annular Dark-Field (HAADF) image of the InSb nanowire on InP 2) is transformed into a set of coordinates encoded in a .gds file. 3) Under the assumption of translational invariance, these coordinates can create a volume. 4) The volume is meshed to allocate the FEM calculations.

Eventually, we mesh the FEM to encode the physics we want to retrieve in each geometrical element (fig. 4.II.13, step 4). Steps 1 and 2 are done automatically by the workflow, but steps 3 and 4 need to be manually done within the environment of the FEM software. Particularly, the mesh and its particularities, as well as the simulation conditions need to be carefully judged to fit our necessities.

The mentioned process has many particularities in our chosen InSb NW on InP substrate system. Its high mismatch of 10.3% makes it a structure in which strain is minimised by creating an array of dislocations in the InSb/InP interface. This effect is visible in Figure 4.I.25. Dislocations are a purely atomic defect caused by the fact that atoms are discrete units. Consequently, FEM models will by default not consider these defects as they treat the medium as continuous. In fact, as a first consideration, we impose a thermal strain of 10 %. We set the thermal expansion coefficient to 10 %, as if the change from 0K to 1K effectively caused an expansion of the NW by 10 %, therefore matching the real lattice mismatch. In addition, we can circumvent the discrete limitation of dislocations by imposing conditions that individually resemble each dislocation. Thus, by only imposing the mismatch by defining the materials in both regions as found by the phase identification algorithm, we would simulate how the relaxation would happen if dislocations could not be formed. In other words, a thoroughly elastic relaxation.

However, the flexible base supported by the workflow provides a ground from which to intuitively modify our simulation system. Indeed, we model these dislocations by adding additional elements within the already-drawn interface. We create an array of circumferences spatially coinciding with their expected positions, by manually adding them in the COMSOL environment. We will discuss automation possibilities a few paragraphs later, but for the time being this step is fully manual. The spacing between dislocations is 3.8 nm, while their diameter and spacing to the interface - they fully sit on the InP substrate - is 0.5 nm. The spacing of 3.8 nm is obtained from inverting the mismatch (i.e., the repeated units of the total dislocation cycle) and multiplying it by the experimental interplanar distance of the epitaxial 111 planes (e.g., 0.40 and 0.36 nm for InSb and InP respectively). The 0.5 nm of diameter are drawn to allocate the mismatch of 0.4 nm between the InSb and InP dislocated planes. In other words, when a dislocation arises the planes coincide and the added plane of 0.4 nm vanishes. In fact, this will be crucial when defining the boundary conditions that represent the physics we want to study. The dislocations are added within the InP to avoid interfering with the interface between regions.

Eventually, we mesh our geometry according to the particularities of the system. Our key region is the interface with the array of dislocations. Therefore, we define a higher mesh density in the dislocations and their surroundings, and we reduce it in the centre "bulk" of both the substrate and the NW, where the expected response is more homogeneous.

The arbitrary lamella thickness of 40 nm serves as a representative value, capturing the typical thickness of a lamella for TEM analysis. However, due to the lack of an experimental lamella thickness specific to the present InSb NW on InP, direct one-to-one comparisons with exact experimental values are challenging. By defining this general thickness, we gain insights into how a typical lamella would behave within a system like the one under consideration. This approach opens up exciting discussions and allows us to explore the implications of varying lamella thicknesses in the context of our study.

Independently, we require specific boundary conditions to collate a thin lamella with an infinitely thick "lamella". Like in Chapter 3 for device GSQ_4, we define the surfaces in which we force their displacement components to be null. A surface with zero displacement is fixed and any expansion or compression is forbidden. Consequently, we emulate the scenario of an infinitely extended surface which by definition does not expand or shrink due to its fully relaxed state. Figure 4.II.14 wraps these conditions up by showcasing the surfaces constrained to represent both physical situations. For the thin system, we define three boundary conditions (fig. 4.II.14, top row):

1. The bottom surface of the lamella, to mimic an infinitely large substrate along y. This is a realistic consideration as the substrate is expected to be fully relaxed due to its effective "infinite" extent.

2. We add a consideration or constraint to our simulations, making the lateral surfaces (yz surface as presented in figure 4.II.14) of the lamella infinite, too. This means that we constrain the component perpendicular to this surface (x in the system) to have null displacement. Essentially, we create a simulated system where the lamella is thin and has an infinitely extended substrate. This choice reflects a realistic situation because regular lamellae are typically much larger transversely than the FOV captured in the present micrograph. Furthermore, we can conclude that the lamella will not relax transversely as long as its width is not comparable to the thickness. To verify this, we conducted simulations without limiting the lateral displacement of the substrate. Not surprisingly, even when simulating a lamella as wide as the micrograph, we found no significant changes in strain relaxation. Therefore, we confirm that if the FOV significantly exceeds the lamella thickness, no transversal relaxation effects are expected.

3. Displacement of the dislocations array. Each dislocation within the array exhibits a fixed displacement of 0.4 nm, as explained above. Thus, every circumference (representing the region around each dislocation) experiences a consistent displacement in the x direction, which aligns with the direction of plane epitaxy. This displacement corresponds to the plane that is either added or missed when a dislocation appears.

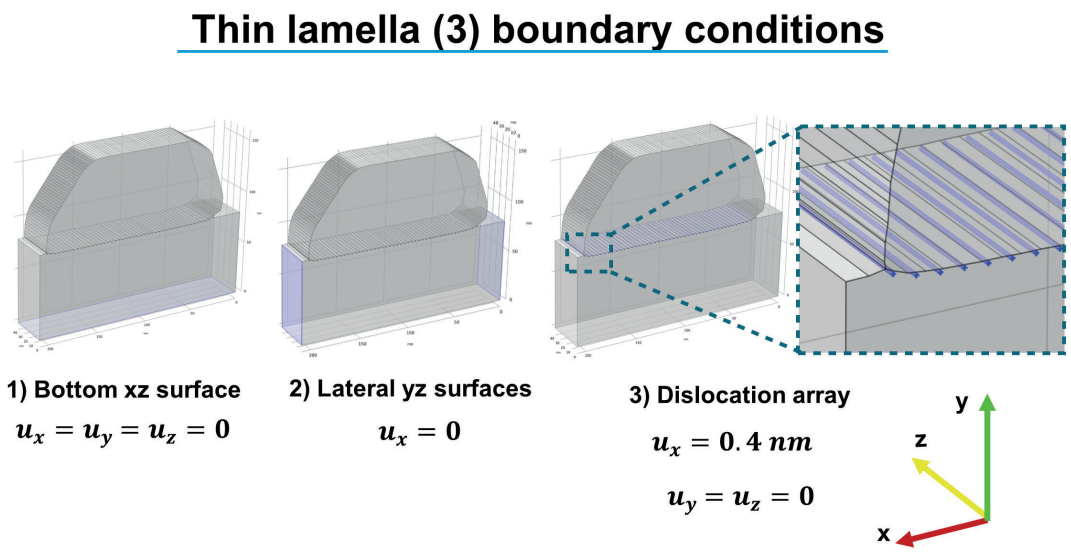


FIGURE 4.II.14: Boundary conditions comparing the thin lamella with the infinitely extended "lamella" in Finite Element Model (FEM) simulations. In the top row, the thin lamella scenario requires three boundary conditions to enable transversal relaxation. On the other hand, in the bottom row, the infinitely thick "lamella" adds a fourth boundary condition to prohibit any transversal displacement.

In this case, we allow the surfaces perpendicular to the electron beam to expand or compress, simulating the behaviour of a thinned electron-transparent sample. However, if we wish to restrict this degree of freedom and create a system that is infinitely large in the translationally invariant direction, we would impose null displacement on these facets. The boundary conditions are four (fig. 4.II.14, bottom row):

1. Infinitely extended bottom surface of the lamella. Same as for the thin lamella scenario.

2. Infinitely extended lateral surfaces of the lamella (substrate). Same as for the thin lamella scenario.

3. Displacement of the dislocations array. Same as for the thin lamella scenario.
4. Transversal facets, perpendicular to the electron beam, are kept with null displacement in the longitudinal component (z as presented in figure 4.II.14). Effectively, the lamella cannot be displaced perpendicularly to this facet, for which we accurately represent an infinitely extended NW with its substrate.

We merge these conditions and apply the boundary conditions in the meshed system. We compute the strain relaxation in both situations to figure out the effect that thinning can have on its components. Each simulation takes around five minutes when parallelised using six cores. [107] These results are captured in Figure 4.II.15, where we compare the experimental components extracted from GPA with the theoretical relaxations of both the thinned and unthinned scenarios.

The first and most obvious observation is the overall larger component in the unthinned system. Specifically, when considering the infinitely thick "lamella", we obtained average values of 11-12 % dilatation in the nanowire. However, in the thinned version, the dilatation only reaches around 9 % dilatation in both components. This implies that the effect of the lamella thinning reduces the components by approximately 22 % $(1 - \frac{9\%}{11.5\%})$. In other words, thinning a 40 nm lamella in the present system releases 22 % of the expected transversal strain through longitudinal relaxation. Consequently, our experimental measurements would reflect 22 % less than the unmodified value we ideally seek. Interestingly, in Chapter 3 (specifically in Figure 3.7), we observe that the decay in transversal strain rounds up to 23 % $(1 - \frac{1.35\%}{1.75\%})$. In fact, the simulated lamella in the GSQ_4 Ge QW device has a thickness of 50 nm. Surprisingly, despite the difference in lamellae thickness, the reduction in measurement is equivalent in both scenarios. Intuitively, the thicker lamella is closer to the physical limit of the infinitely thick "lamella". Therefore, we might expect that the thicker lamella of the GSQ_4 Ge QW would release less strain and better resemble the infinite case. However, there's a critical distinction: GSQ_4 releases strain fully elastically, whereas the major contribution in the present system is plastic. This consideration is key when we delve into the full picture behind experiment and theory comparisons.

Furthermore, the x component is slightly smaller than the y component due to the effect of dislocations only acting in the x direction. The dislocations create a progressively increasing expansion from the interface to the top of the NW because the vertical component of the epitaxial planes adapts to the dislocation fixed displacement in x. At the very top of the NW, the values are comparable. It is also interesting to observe from the 3D views of the components, how the shape of the NW, especially at the kinks on both edges of the interface, is unevenly modulating the strain. Remarkably, this strengthens the importance of actually having the experimental morphology of the sample instead of just a schematic simplification of the ideal system.

We compare the experimental GPA relative strain maps with the calculated ones, and in this case, we observe some discrepancies. We use the same colour scale for all the elements to simplify the comparison. Firstly, although it is not a huge difference, we see a slight difference between the x and y components, but opposed to the theory, as experimentally there is a 0.5 % larger dilatation in x than in y. Note that in x, it is 0.078 - (-0.019) = 0.097, while in y it is 0.09 - (-0.002) = 0.092. One would expect the same average value in both experimental components as the relaxation is purely plastic. This small added percentage in x can correspond to the remaining elastic component of the strain relaxation, which happens in x due to the vertical epitaxial planes. This component is seen as the gradual increase from the interface to the top of the NW that reduces the apparent average value. Alternatively, it could also be the lamella thinning effect slightly relaxing the y component not ruled by the epitaxy.

On the other hand, the colour ranges of the experimental maps differ from the simulations in all cases. The experimental values are significantly smaller than the simulated ones, and only when comparing the y component with the thinned lamella we find similarities. The understanding of this discrepancy obligatorily encompasses a volumetric understanding of the system. The relaxation of the structure originally happens in all three directions of space. Thus, the relaxation will be led along the x axis, where the epitaxy is leading the plastic relaxation, and will be complemented in the other two axes according to the Poisson modulus. Therefore, in the case of the infinitely thick "lamella", we are forcing that the transversal facets are immutable in displacement terms. In other words, the longitudinal z axis cannot be relaxed. As a result, we are assuming that our simulated conditions in that axis perfectly coincide with the fully relaxed scenario, which is the real infinite translational invariant system. Therefore, any volumetric relaxation component not considered in that axis will eventually be summed in the transversal components. A direct consequence is the possibility of finding relaxations and relative strains surpassing the theoretical maximum imposed by the mismatch itself. In fact, both x and y strain components present an average in the NW larger than 10 %. The reason for the observed behaviour lies in the absence of a global relaxation mechanism along the longitudinal z axis. In fact, Figure 4.I.28 showcases the expected appearance of a longitudinal cut and its strain analysis in the present system. In this context, dislocations also emerge in the z direction, forming dislocation lines perpendicular to the modelled ones (fig. 4.II.13 and 4.II.14). It is important to recognize that epitaxy not only happens laterally but also longitudinally. However, our simulation remains incomplete in terms of capturing this longitudinal relaxation. As a consequence, the calculated relaxation along the z-axis is purely elastic and, therefore, residual. The remaining unrelaxed z strain is compensated by the x and y components. Notably, the y-component weighs a larger percentage of this relaxation, given that the x-component is tightly constrained by the epitaxial conditions.

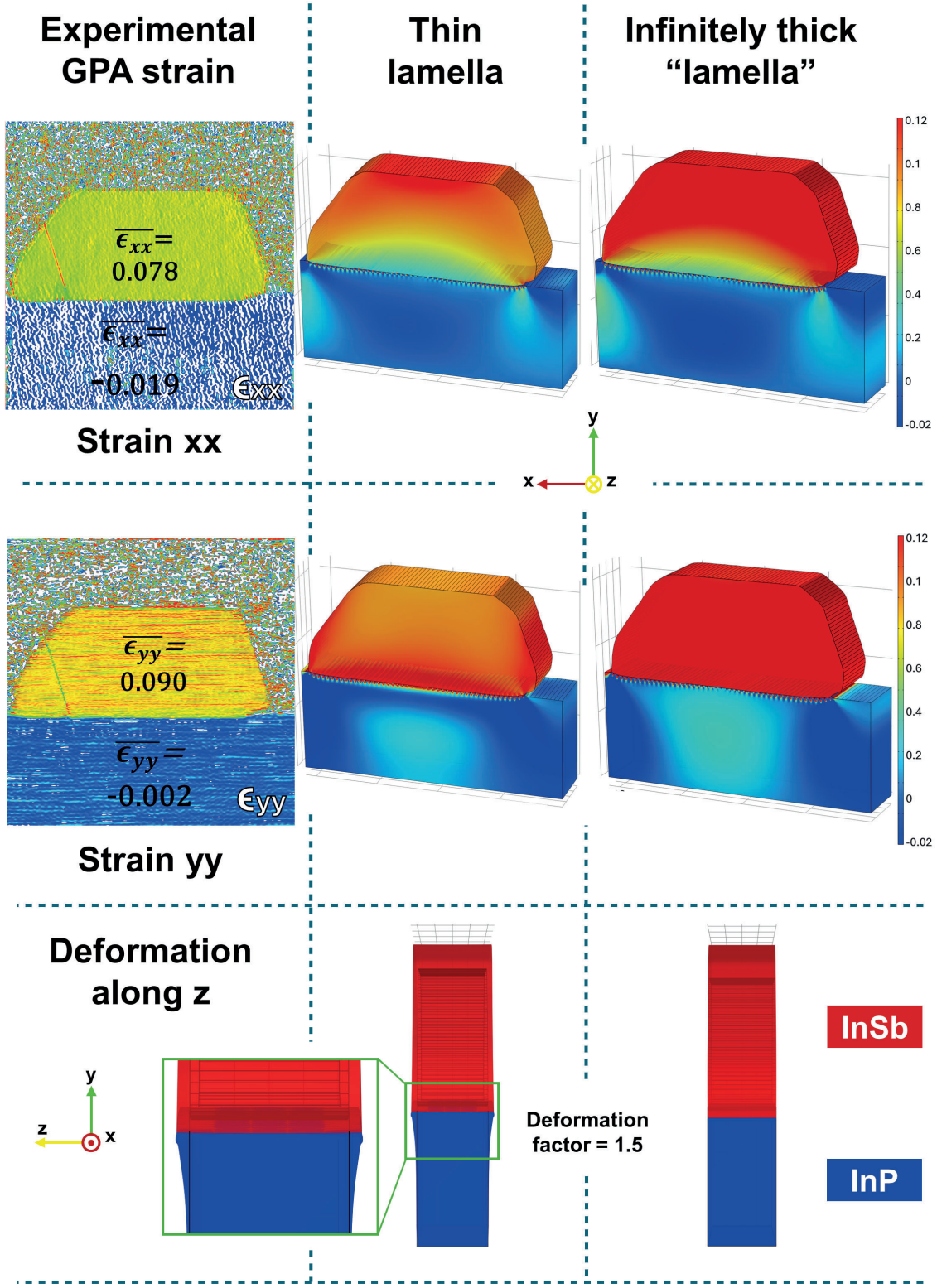


FIGURE 4.II.15: Comparison of the Finite Element Model (FEM) strain relaxation, considering a thinned lamella of 40 nm, an infinitely thick "lamella", and the experimental relative strain maps computed with Geometrical Phase Analysis (GPA). The colour scale applies to the three elements.

Figure 4.II.15 also showcases the expected deformation of the lamella in the z direction because of the thinning. We see that the effect is relatively small, especially considering that we amplify the deformation by a factor of 1.5 for visibility. However, for accurate quantitative experimental strain measurements, this effect must be considered despite its absolute module. Therefore, more precise and complete simulations ought to be conducted to ensure a proper experiment-theory correlating unlocking experimental data correction. For instance, a more meticulous approach would not ignore the protective polycrystalline Pt shell adding an interface not being considered yet. Although neither simulated scenarios consider the Pt protective layer on top of the NW, it is possibly a more noticeable effect in the thinned lamella case. Thus, in the unthinned one, the system undergoes a strong transversal constraint which minimises the effect of a possible free surface (the top of the NW). In any case, the scale of the NW is larger than the thickness of the lamella, for which no major elasticity should be found in the free facets parallel to the electron beam.

In summary, we would need further simulations to give a complete answer to the problem and fully understand the source of the discrepancy between the theory raised so far and the experiments. Having the experimentally measured thickness, for instance, by EELS, would be an important starting point to systematically address this concern. In addition, scanning simulated thicknesses, as well as the complementary constraints exposed could offer a complementary understanding as well. Intriguingly, some of these considerations could be automatically implemented with ease. For example, the boundaries that directly coincide with the image borders could be immediately set to extend infinitely (fixed total null displacement). In the InSb/InP case, this would constrain the full rectangle representing the substrate but not the NW. Besides, the Pt protective layer could be added as well in case the square of the image is not complete and some area is unlabelled or lacking an identified material. Furthermore, the dislocations could also be encoded in the FEM simulation software by identifying them and centring a circle in their position, which by default would entail a local tighter and denser predefined mesh. Eventually, a more realistic atomic nature could be provided by considering the projection of the Poisson and Young moduli along the real zone axis and crystallographic orientation. Indeed, the knowledge of the materials is already computed (see section 4.II.2) and the missing step is the anisotropic strain implementation within the FEM software.

4.II.5.3 Simulations of atomic forces and relaxations

One of the current limitations in the methodology is the lack of a secondary verification process for the convergence criteria to determine the validation of our 3DAMs. In this section, we showcase that simulations from the 3DAMs can be used to identify artefacts in these models. Therefore, beyond the extraction of physical properties, we can combine simulations to distinguish whether the displacement of a

single atom is physically realistic or driven by a nonphysical artefact distorting the GPA outcomes (as discussed in sections 4.I.6 and 4.II.3.3). We prove this capability with the particular example of the GSQ_4 device (see Chapter 3). [38] Let us dive into the context making it possible.

The ideal scenario is accessing *ab initio* simulations, as they provide the highest available level of theory. First-principles, *ab initio* calculations describe the system, considering all parameters with high accuracy. However, their two main limitations are the computation time and the level of complexity needed to attain some key properties. One of the most common first-principle solutions for many-body calculations, if not the most, is Density Functional Theory (DFT), which uses functionals to model the electrons. From the practical point of view, several codes allow running DFT calculations of large systems within a reasonable amount of time, among which we highlight SIESTA. [124] Interestingly, its efficiency stems from the use of a basis set of strictly localised atomic orbitals. Additionally, an important feature of the code is that its accuracy and cost can be tuned over a wide range, from quick exploratory calculations to highly accurate simulations matching the quality of other approaches, such as plane-wave methods. Note that it scales up well until tens of thousands of atoms. In fact, DFT demands unitary occupancies. In other words, non-collapsed models are not relevant except as the base of eventually thick collapsed models which represent systems worth simulating. As a consequence, the repetition opens the balance between model fidelity and calculation viability.

Therefore, our 3DAMs lie between the tens of thousands of atoms to the few millions. Consequently, we can consider DFT a potentially beneficial tool to extract physical properties of interest with the accuracy and precision first-principle calculations entail. For example, we can limit the number of atoms to simulate by manually cropping the model to a specific region of interest. Interestingly, these cuts can be done along specific crystallographic planes to facilitate fulfilling the periodicity and translational symmetry constraints required for DFT. In the case of our large disordered systems, translational symmetry can only be attained after careful surface passivation with hydrogen atoms. The purpose is to avoid the creation of new unphysical bonds when incorporating the periodic boundary conditions.

Importantly, the DFT description of the energy band gap in semiconductors tends to underestimate the value when using most of the functionals. [125] The best DFT approaches for band gap calculation would mean cubical scaling of the computing time with the number of atoms. In addition, DFT-based band gap calculations for Si and Ge not only underestimate the positioning of the band structure but often shift the band gap type. [126] In summary, a DFT-based band gap calculation for the GSQ_4 system would be too resource-consuming computationally and would very likely result in a wrong estimation of the band structure. Thus, DFT is not the best option to compute band gaps and Densities Of States (DOS), for which other models can be more interesting, as we will see in the next section. Note that our

interest in computing this key property of the semiconductors will be experimentally reestablished in Chapter 5.

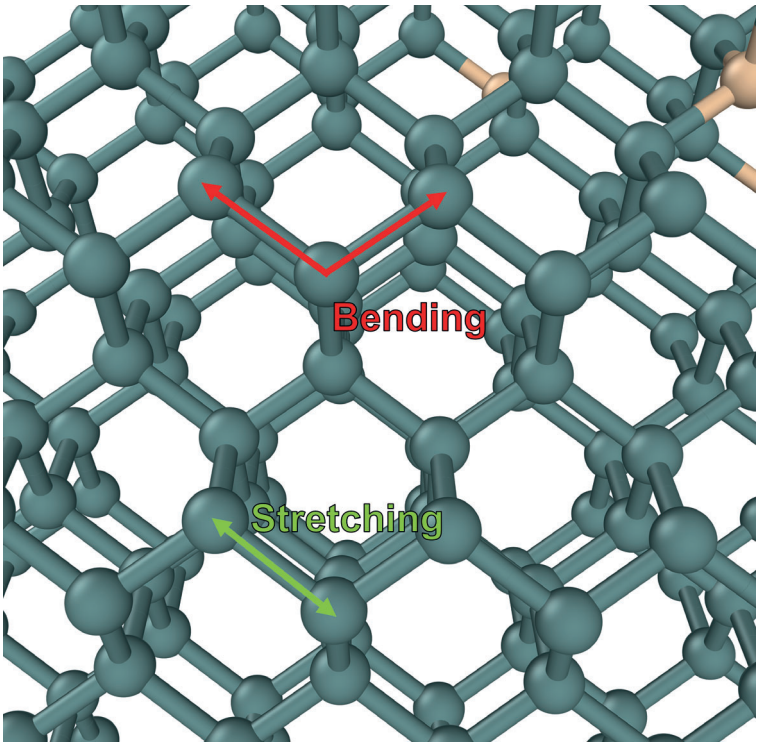


FIGURE 4.II.16: Scheme of the bending and stretching in the Keating model applied to a SiGe atomic model.

The explained limitation and the use of an alternative model to compute the band gap led us to model the atoms differently. Although DFT can be adjusted to compute global strain and structure relaxation (i.e., force and energy minimisation), the following complementary modelling underscores the versatility the 3DAMs offer with different levels of theory. Unlike DFT, electrons will now be neglected and nuclei will be classically encoded instead. We employed the long-reported Keating model to compute the atomic forces for all atoms in the model. [127, 128] The Keating model is a valence force field model particularly developed for crystals with a cubic diamond structure (i.e., group IV semiconductors). It defines the interaction between first neighbours using a harmonic approximation: we can understand the atomic arrangement as a combination of springs which can be stretched and bent given bond distance, d, and stretching a and bending b spring constants of group IV alloys (fig. 4.II.16). It has been shown that with appropriately optimized parameters, this model can produce satisfactory results. Remarkably, as the Keating model describes the system by its stretching and bending, it can also be used to study vibrational modes (i.e., direct link to phonon spectroscopy), as reported in the literature. [129] We implemented a custom version to compute the forces, perform structure relaxations, and reveal the phonon Local DOS (LDOS), with the parameters summarized in tables 4.II.1 and 4.II.2. [127, 128] Importantly, the relaxation of structures

unlocks an additional layer of confidence to our calculations, as detailed in a few paragraphs below.

Stretching	**Si-Si**	**Si-Ge**	**Ge-Ge**
a (N/m)	47.207	42.855	38.212
d (Å)	2.3516	2.4008	2.4500

TABLE 4.II.1: Keating model constants for stretching in SiGe alloys.

Bending	**Si-Si-Si**	**Si-Si-Ge**	**Ge-Si-Ge**	**Si-Ge-Si**	**Si-Ge-Ge**	**Ge-Ge-Ge**
b **(N/m)**	14.320	13.674	13.083	13.083	12.725	12.387

TABLE 4.II.2: Keating model constants for bending in SiGe alloys.

We retrieved the atomic model from a STEM micrograph of GSQ_4 (fig. 4.II.7) and used its atomic coordinates to compute the forces with the provided Keating model. We limited the thickness to just two atomic planes to ensure the unit cell is fully captured in the translational invariant dimension. Remarkably, periodicity constraints do not apply in Keating modelling. We start from the non-collapsed version of the model, with fractional occupancies for every atom. From there, the structure input for the calculation of the forces can be as large as desired, as we could collapse the structure for any number of in-z repetitions of the first (110) planes. The lateral size of the plane was limited by the field of view of the experiment but consisted of around 15k atoms per plane. In the computation of the forces, adding layers just slightly impacts the in-(image)-plane components of these forces, mostly due to the collapsing of the fictional occupancies. However, we will see that the thickness choice is important for both DFT and the calculations reported in the next section.

Figure 4.II.17 summarises the calculation of the in-(image)-plane force vectors of GSQ_4. In this case, we emphasise the particular image from which it is computed, as its quality directly impacts the force diagram found. As seen in section 4.II.4, in figure 4.II.7, the presence of distortions in the GPA map from locally blurry regions in the micrograph created artefacts in the atomic models. We deliberately cropped the model to get rid of these apparent flaws to pursue the present simulations. As a result, we end up with an apparently flawless model fragment which will serve as input file for our Keating simulation code.

The forces distribution calculation showcases a map of small forces implying no or minimal expected atom-wise distortion. In fact, most of these slightly increasing forces (electric blue-dark blue transitions) correspond to the randomly allocated Si atoms (fig. 4.II.17). Thus, Si atoms in a supercell whose average cell parameter is much closer to Ge than to Si will experience a tendency to be closer to their neighbours. The effect will then be stronger the more Si atoms are collapsing together. Importantly, this effect should be averaged if infinite random collapses of the fractional occupancies are done. This effect is almost unnoticeable in the Ge QW itself,

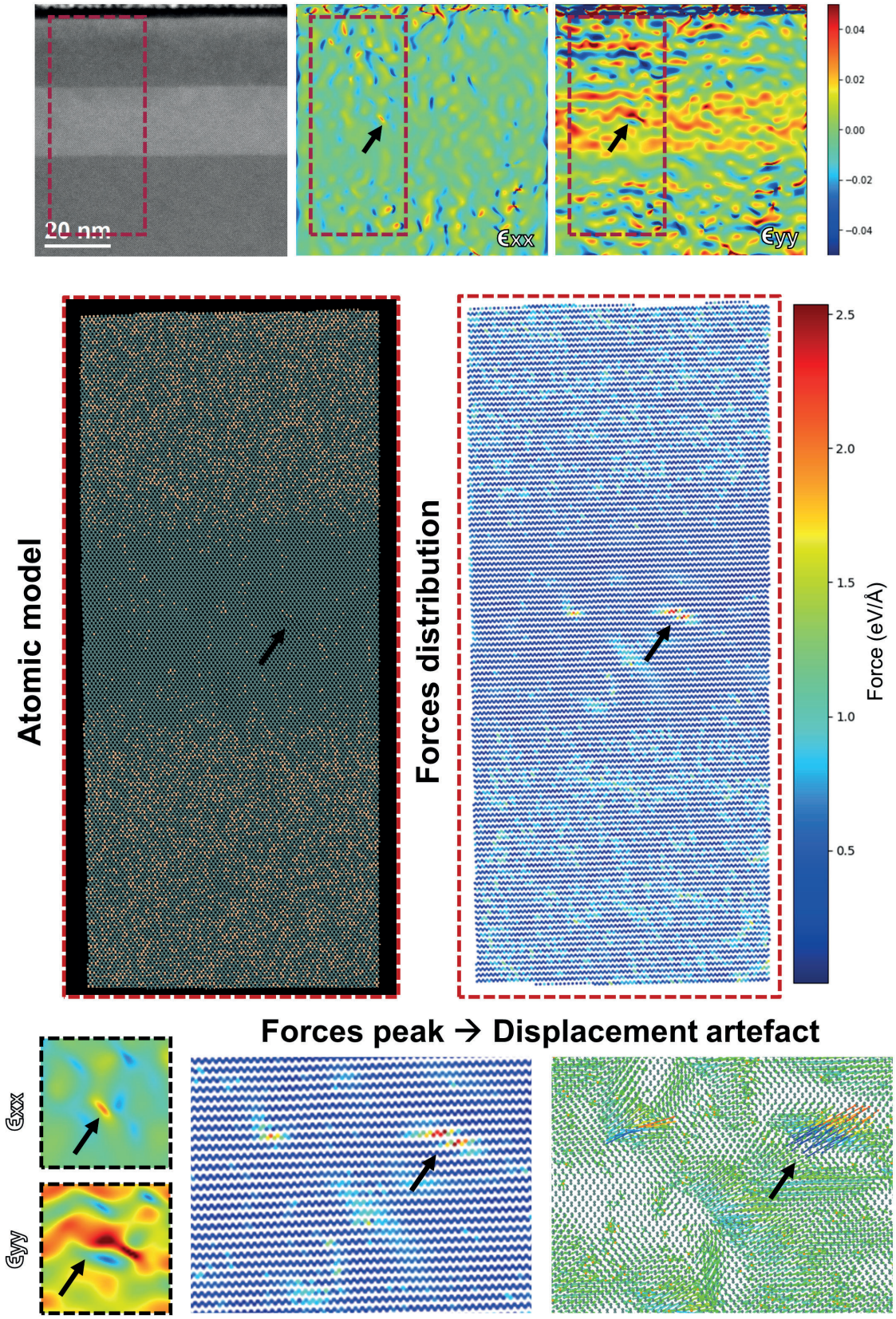


FIGURE 4.II.17: Simulation of the forces distribution of a 3D Atomic Model (3DAM) of a Ge quantum well embedded in two $Ge_{0.7}Si_{0.3}$ buffer layers (GSQ_4 device from Chapter 3). The black arrow points to the presence of an image distortion and its effect is highlighted in the strain maps. The position coincides with a nonphysical jump in the forces calculated through Keating's model. The local effect can only be explained by the sensitivity provided by the algorithm displacing the atoms of the 3DAMs.

as the presence of Si atoms is minimal. However, the effect that would remain is the force (intensity) peak pinpointed by the black arrow. Not surprisingly, it coincides with an intensity peak in the strain maps. Indeed, the large forces detected from the atomic force calculation are not related to the chemical composition of the sample. Instead, they can be directly attributed to structural distortions originating from the construction of the 3DAM. Remarkably, the force calculation unveils an artefact that remained hidden to the naked eye, fitting seamlessly within what was initially a flawless 3DAM fragment.

Further exploration of the 3DAM in that specific region revealed an excessive bending of the planes which nonphysically curved them. The effect, originally caused by the local loss of quality in the image, cannot be physically meaningful due to the already reported flawless nature of the GSQ_4 device. Moreover, a sudden plane wriggling confined in three atomic planes does not have any thermodynamic justification. Remarkably, we discover and reinforce the condition of this artefact through a complementary approach: the calculation of the DOS through tight-binding explained in the following section.

Interestingly, the calculation of the atom-wise forces and their vectors also allows us to relax the structure. Thus, we can iteratively compute the forces and displace the atoms according to the displacement vectors outputted until the resulting atomic forces are minimised. The convergence criteria is reached when the forces become smaller than a threshold value of 0.01 eV/Å. Relaxation is a theoretical calculation widely adopted in literature to work with more physically realistic models. However, in the present case, a relaxation calculation is detrimental to the model fidelity and goes against the purpose of the whole workflow. As discussed, the forces mainly emerge from the collapsed Si atoms and bring them closer to neighbours. That is, certain atoms of the experimental 3DAM will be out of their optimal position just considering the forces. However, the non-optimal position initially set by the strain transfer procedure (see section 4.II.3.3) will hardly ever coincide with the optimal coordinate based on a dynamic basis. The strain fields will generate positions out of equilibrium, as they indeed represent out-of-equilibrium structures. Therefore, working with a strained structure should output force heterogeneity as a sign of model success. Furthermore, forces in the isolated Si atoms are expected, as the GPA displacement maps allocate Si in a unit cell that is not its default one. In other words, the calculated strain annihilates the forces that would otherwise relax the 3DAM.

As indicated before, these forces should average when computing different collapse settings and different 3DAM thicknesses. The artefacts would then be highlighted as force outliers. The beautiful part is that identifying these anomalies could pinpoint where our local relaxations must be directed. For instance, the forces map can be represented as a distribution of forces based on their module. Conveniently, we could limit the relaxation to only these atoms lying in the upper tail of this distribution, which might correspond to strain transfer artefacts. In addition, the used

convergence criteria would be the force average or a physically meaningful variation around this value. As a result, we end up with a model that keeps the essence of the experimental strain throughout, and discards displacement outliers. Of course, this approach would work in the smooth device from figure 4.II.17, but would flatten the models if a real and physically meaningful strain peak is present in the calculation.

This underscores the sensitivity of our method, capable of detecting minute effects imperceptible to the naked eye. Conversely, it can identify any physically-originating effect visible in the GPA maps, showcasing unparalleled sensitivity. In fact, in the next section devoted to electronic tight-binding simulations, we compare a model with artefacts from the image acquisition, such as the depicted one in Figure 4.II.17, and a trustworthy model from a flawless micrograph. Importantly, we will see the importance of computing the forces through the proposed Keating model as a key although not unique step to validate our 3DAMs.

4.II.5.4 Tight-binding simulations

4.II.5.4.1 Remarks and considerations of the simulations

We describe the electronic structure of the system using an $sp^3d^5s^*$ tight-binding Hamiltonian for SiGe alloys in the presence of strain and including spin-orbit coupling. [130]. Therefore, we use 10 orbitals and 2 spins, totalling 20 spin-orbitals. The number of total orbitals used will be important as it is the reference with which the computing time scales, rather than the number of atoms. Particularly, the parameters of this tight-binding Hamiltonian were adjusted on *ab initio* data with quasiparticle corrections, or, when available, experimental data, thus correcting for the gap issues of pure DFT. Within this tight-binding framework, the electronic structure of the system is described at the atomic level. Importantly, tight-binding simulations match the workflow's philosophy well as the system's electronic Hamiltonian is calculated from the knowledge of atomic positions and the superposition of the indicated orbitals. The construction of this tight-binding Hamiltonian from atomic positions data is done through the TB_Sim code. [131]

Let us recall that the tight-binding Hamiltonian, $\hat{H}$, is naturally expressed and efficiently calculated in a basis of atomic orbitals, $\{|\mathbf{n}, \lambda\rangle\}_{\mathbf{n},\lambda}$, where the $\mathbf{n}$ run over the atomic positions, and the λ are multi-indices labelling the spin-orbitals. However, the sizes of the samples considered in the present study make a full diagonalisation of the tight-binding Hamiltonian impractical. Indeed, the computational effort in time for exact diagonalisation of dense matrices scales poorly with system size, typically with $O(N^3)$ where N is the number of spin-orbitals in the total system. This diagonalisation would allow us to obtain the system's eigenstates, the LDOS.

To sidestep this difficulty, we make use of linear scaling techniques as implemented in the LSQUANT code. [132, 133] The idea of such methods is that one can eschew the computation of the system's eigenstates to compute directly a physical property

of interest, typically at a much lower computational cost. In this study, we target the computation of averaged LDOS over spatial regions of interest in the sample. These regions will divide the GSQ_4 QW into segments of different compositions along the growth direction. This can be seen in figure 4.II.18. Using the aforementioned linear scaling methods, this computation typically scales, as its name indicates, linearly as $O(N)$.

Formally, the DOS can be expressed as:

$$\rho(E) \propto \mathrm{Tr}[\delta(\hat{H} - E\mathbb{1})] \tag{4.II.50}$$

$$\propto \sum_{\mathbf{n},\lambda} \langle \mathbf{n}, \lambda | \delta(\hat{H} - E\mathbb{1}) | \mathbf{n}, \lambda \rangle \tag{4.II.51}$$

The individual matrix elements, $\rho_{\mathbf{n},\lambda}(E) = \langle \mathbf{n}, \lambda | \delta(\hat{H} - E\mathbb{1}) | \mathbf{n}, \lambda \rangle$ can be understood as LDOS: each individual element is the contribution of a single orbital to the total DOS at a given energy E. By this method, the contribution of a specific spatial area within a sample to the total Density of States (DOS) can be accurately calculated by aggregating the contributions of all orbitals from all atoms within that area. If $\mathcal{R}$ is a set of atoms in the sample, its contribution to the DOS is then given by:

$$\rho_{\mathcal{R}}(E) = \sum_{\mathbf{n} \in \mathcal{R}, \lambda} \underbrace{\langle \mathbf{n}, \lambda | \delta(\hat{H} - E\mathbb{1}) | \mathbf{n}, \lambda \rangle}_{\rho_{\mathbf{n},\lambda}(E)} \tag{4.II.52}$$

If $\mathcal{R}$ is chosen such that it includes all atoms in a given spatial region, then $\rho_{\mathcal{R}}(E)$ can be thought of as the average LDOS in that region of space. This concept is crucial as it is our direct link with the microscopy from which our models arise. In other words, we can compute this quantity at any given spatial resolution. Therefore, we can grab a single atom with its spin-orbitals or all the atoms from the model. Fruitful discussion would arise to give adequate physical meaning to each, as going from a crystal to a single atom progressively dilutes the concept of DOS to raise its origin of discrete electronic states in single atoms.

Furthermore, the theoretical calculation of the DOS and the band gap gives us a tangible explanation of how the experimental structural properties affect the electronic structure. Moreover, they serve as a basis to compare with experimental band gap measurements. Indeed, being able to spatially map the band gap both theoretically and experimentally is a correlation never achieved before. Indeed, with the present tight-binding simulations we build the theoretical side. Intriguingly, in Chapter 5 we build the experimental side of the comparison. [4]

Importantly, tight-binding models are an excellent choice for computing LDOS in a spatially resolved way. There are several compelling reasons for this:

- Cost-Effective Computation: Tight-binding models can be computed relatively inexpensively compared to other observables that are accessible using linear scaling, such as the conductivity. Moreover, in a first approximation, the size of the box in which to calculate the LDOS does not modify a computational cost dependent on the total number of spin-orbitals of the system. In the next paragraphs, we will see that this claim comes at a cost when relying on the random phase states approximation.
- System Information: The LDOS obtained from tight-binding models contains valuable information about the system. It can be used to infer direct correlations with the structural properties, which aligns with the central goal of the PhD thesis.
- Valence and Conduction Edges Estimation: Tight-binding models allow us to estimate the effective valence and conduction edges within a specific spatial region. This is particularly relevant in QW systems.
- Effective Local Bandgap: By analysing the difference between the valence and conduction edges, tight-binding models can also provide an estimate of the effective local band gap. This estimate can potentially be compared with experimental results.

The tight-binding calculations can be characterised by their energy resolution and the noise. In other words, resolution and noise trade off the computational resources, often in a complementary way. Although high energy resolution and low noise simulations are possible, they reach impractical computational complexities. Therefore, the improvement of one comes to the detriment of the other in benefit of the computing time. [133, 134]

For the sake of further discussions, it is important to specifically define how the energy resolution and the noise are encoded in our simulations. For that, we make use of the computationally efficient method to map the DOS by evaluating matrix elements of the form $\langle\Psi|\delta(\hat{H}-E\mathbb{1})|\Psi\rangle$ known as the Kernel Polynomial Method. In our context, it proceeds essentially using a Tchebyschev expansion of $\delta(\hat{H}-E\mathbb{1})$:

$$\langle\Psi|\delta(\hat{H}-E\mathbb{1})|\Psi\rangle \approx \sum_{m=0}^{M} c_m(E)\,\langle\Psi|T_m(\hat{H})|\Psi\rangle \tag{4.II.53}$$

where the $c_m(E)$ are system-independent expansion coefficients which are known analytically, the $\mu_m = \langle\Psi|T_m(\hat{H})|\Psi\rangle$ are so-called Tchebyschev moments which can be calculated very efficiently using recursion relation techniques typical for these Tchebyschev polynomials, and M is a truncation order which sets the precision of the calculation: the energy resolution. Effectively, precision is characterized by a broadening parameter, denoted as η. This parameter is inversely proportional to the number of computed moments M, which directly affects the computational effort.

Moreover, the precision is also proportional to the spectral range of the system. The spectral range corresponds to the energy interval within which the tight-binding Hamiltonian is defined. In other words, the energy domain of our tight-binding Hamiltonian, along with its base orbitals, determines the range of energies we can map. Moreover, the local energy sampling must align with our specific requirements. For instance, if we aim to map the LDOS around the band gap, the tight-binding model must be well-defined and sampled appropriately in the vicinity of that energy of interest.

Fortunately, our tight-binding Hamiltonian is well-defined around the band gap energies, making it suitable for our purposes. [130] Specifically, within our SiGe tight-binding description, the global spectral range spans approximately 60 eV. To compute the LDOS, we employ two different broadenings, $\eta \sim 15$ meV and $\eta \sim 4$ meV. These choices correspond to using $M = 6000$ and $M = 25000$ moments, respectively. The finer broadening (4 meV) serves a specific purpose: it enhances the resolution of certain wide features observed with the 15 meV broadening. Notably, the calculation of moments constitutes the primary computational cost in LDOS calculations via tight binding. As a result, we achieve a time scaling of $O(MN)$ to efficiently compute the DOS.

To evaluate noise, we introduce the concept of random phase states. This concept allows us to perform calculations that would otherwise be computationally impractical. The idea is that the LDOS is computed using the Equation 4.II.50. However, for large systems like ours, this calculation scales with the number of spin-orbitals, resulting in a complexity of $O(N^2)$ ($O(MN^2)$ if moments are used). Specifically, each term in the trace of the matrix with N elements contributes with a computational cost of $O(N)$ ($O(NM)$ when moments are computed). To circumvent this limitation, we set a basis of states:

$$|R\rangle \propto \sum_{\mathbf{n},\lambda} e^{i\phi^{\mathcal{R}}_{n,\lambda}} |n,\lambda\rangle \tag{4.II.54}$$

where the $\phi^{\mathcal{R}}_{\mathbf{n},\lambda}$ are random numbers drawn uniformly in $[0, 2\pi]$ if $\mathbf{n} \in \mathcal{R}$ and 0 otherwise. As we use these random states to approximate the trace and therefore the LDOS, our process becomes stochastic and as a result, we introduce random noise in the calculation. The stochastic evaluation of the trace converges slowly as $O(1/\sqrt{N_{\mathcal{R}}})$, where $N_{\mathcal{R}}$ is the number of spin-orbital comprised in $\mathcal{R}$ (i.e., the mapped spatial domain). Note this is not a time scaling but an error scaling associated with the approximation of the trace through random phase states. This is the noise we expect from the calculations. Practically, we average the results over a number n_r of realisations of the random vector in equation 4.II.54. Indeed, this requires a compromise in the size of the regions $\mathcal{R}$ we map as small regions with few atoms will require multiple random phase states for convergence. n_r between 1-10

random phase states are recommended despite convergence being expected sooner for larger systems like ours. Independently, the more stochastic calculations we do the closer to the direct trace evaluation we get:

$$\rho_{\mathcal{R}}(E) \approx \sum_{r=1}^{n_r} \langle R_r|\delta(\hat{H} - E\mathbb{1})|R_r\rangle \tag{4.II.55}$$

where a new set of random phases $\phi^{\mathcal{R}}_{\mathbf{n},\lambda}$ are used for each realization r, and typically $n_r \ll N$. Importantly, combining this technique with the polynomial expansion of equation 4.II.53, the averaged LDOS in the region $\mathcal{R}$, $\rho_{\mathcal{R}}(E)$, can be evaluated in $O(n_r MN)$ time. The take-home message is that our impractical starting dense diagonalisation complexity of $O(N^3)$ becomes simplified to $O(n_r MN)$, where we have control over n_r and M to trade-off precision, noise and computing time. Therefore, given this final computational complexity, we need to tune the resolution and noise according to these three variables keeping in mind that some features will only be highlighted if a minimum of precision is achieved.

In any case, the primary factor to consider is the number of spin-orbitals, which is generally related to the total number of atoms in our supercell. Preparing the supercell for tight-binding models involves specific considerations that we should revisit. We begin with the 3DAM directly outputted from the workflow. However, manual intervention is necessary to tailor it for tight-binding simulations. Firstly, we only need a single period in the translational invariant direction. This simplifies the computational setup. Secondly, we use the full fractional occupancies of the atoms. Consequently, we replicate the supercell along the translational invariant direction. This replication process introduces a collapse of occupancies, as explained in Sections 4.II.3.4 and 4.II.3.5. The collapse of occupancies occurs when the system reaches the desired longitudinal size. Notably, the atomic column-wise collapse of occupancies aligns with experimental values only if the occupancies were collapsed in an infinitely thick system. As a consequence, this step introduces random noise to our calculation. Each occupancy collapse setting is unique. For instance, Si atoms may occupy different positions in each iteration. Ideally, we would average over an infinite sum of collapses to mimic the exact local stoichiometry observed experimentally. With sufficiently large spatial regions being probed, averaging effects should reduce the dependence on specific collapsing settings. However, for finer spatial investigations, we may need to carefully consider collapsing methods, especially if the intricate chemical composition of the system influences the results.

The use of a single unit cell along the electron beam direction involves a not fully obvious consideration. In practice, we must replicate the cell in this direction (and virtually in any direction where we have a reduced default number of atoms). Specifically, in our case, this consideration occurs longitudinally. Although we assume and impose periodic boundary conditions in this direction, tight-binding calculations are

effectively centred around the Gamma point of the Brillouin zone. As a result, the larger our supercell, the closer it approximates the exact Hamiltonian, which inherently extends over the entire reciprocal space. Thus, by adding more atoms to the supercell, we effectively introduce more bands for mapping the Brillouin zone, even when formally computing only at the Gamma point. To achieve this, we replicate from two atomic planes to a total of 60 planes and then apply the boundary conditions. This specific number is justified by programmatic limitations encountered when simulating the widest systems. Additionally, we ensure that the sizes in each direction exceed the confinement of the device. In the case of the GSQ_4 QW, this confinement effectively occurs within a few nanometers.

Furthermore, to avoid edge effects, we passivate the system using pseudo-hydrogen atoms. This is handled in the following way:

1. Remove all Si and Ge atoms with less than four neighbours.
2. For all bonds suppressed due to the removal of these atoms, we place a pseudo-H atom along this bond, three-quarters of the way to the now missing atom.

We note that, in our tight-binding setting, these pseudo-H atoms do not interact with one another, and serve only for this passivation purpose. That is, we want to ensure that all Si and Ge atoms have four first nearest neighbours. If we define the translational invariant direction as z, these pseudo-H atoms are essentially added to the xz and yz facets of the sample. They coincide with the red margins of the atomic model in figure 4.II.18.

4.II.5.4.2 Results of the experimental/tight-binding/Keating correlation

We simulate the local Density of States (DOS) by dividing the sample along its transversal symmetry axis, maintaining a constant distance in the x-direction to ensure uniform composition. To address the limitations of smaller cuts and reduce stochastic noise, we further divide the sample into 10 regions (coloured boxes in figure 4.II.19) for DOS calculations. We present in Figure 4.II.19 the preliminary results of the LDOS calculations discussed above over its whole spectral range. We used here only one occupancy collapsing. In total, the system has around a million atoms and the number of spin-orbitals is $N \sim 16 \times 10^6$. The results are averaged results over $n_r = 2$ random phases. The parallelised computation of one random phase density with $M \sim 8000$ moments takes on the order of 2 hours. The full spectrum range LDOS in Figure 4.II.19 is computed with a characteristic broadening of ~ 15 meV. At this scale, the typical feature of the Si and Ge DOS can be recognized. Spin qubit physics, however, involve mostly states close to the band gap: essentially valence states since the system holds a 2D hole gas. Therefore, as indicated, we will narrow the spectral range and reduce the broadening to resolve these features close to the band gap.

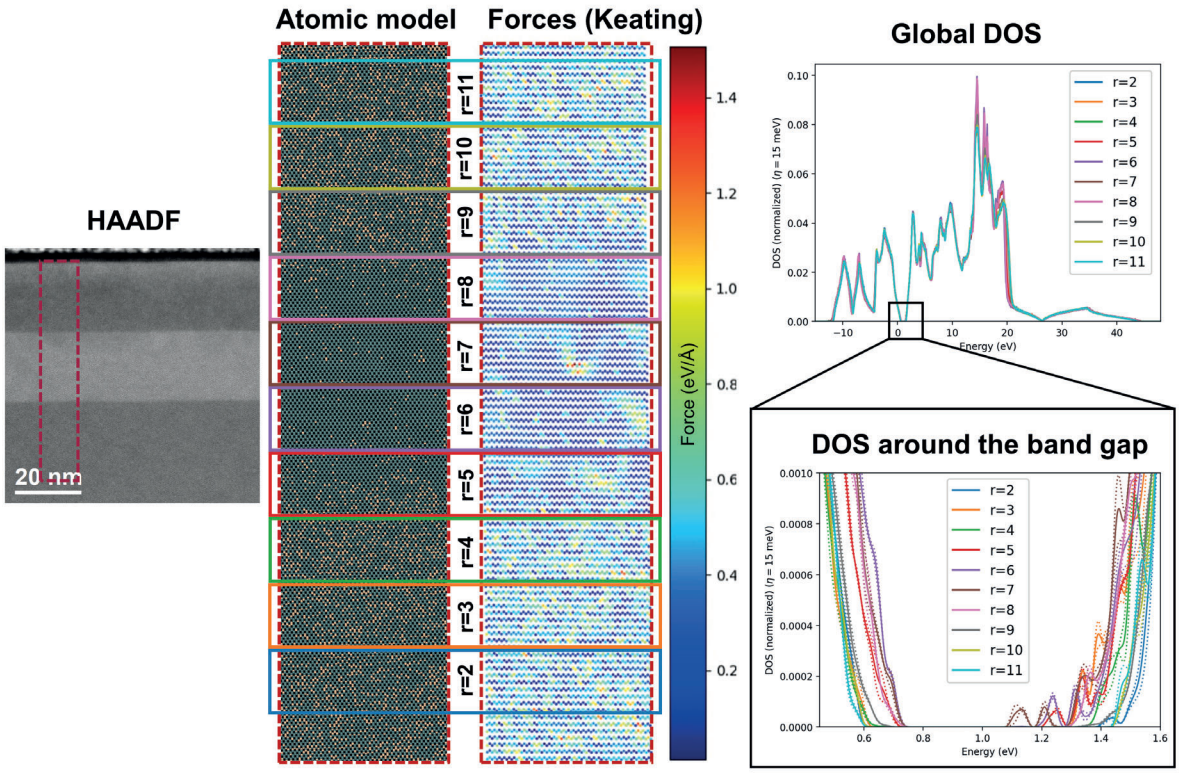


FIGURE 4.II.18: On the left, the High-Angle Annular Dark Field (HAADF) micrograph of the Ge Quantum Well (QW) GSQ_4 device. [38] The dotted red line depicts the area from where the atomic model is extracted to perform tight-binding simulations. We compute the Local Density of States (LDOS) in each coloured box cutting the atomic model and the atomic forces calculated with Keating model. On the top right, we find the LDOS along the full spectral range, while in the bottom we zoom-in to find the LDOS around the band gap. The spectral broadening is of $\sim$ 15 meV, and still allows us to resolve the peaks corresponding to mid-gap states.

This zoom-in is represented in Figure 4.II.19, as well. The solid lines represent the average of the two random phases displayed as dotted lines. The effective valence band edges for the different regions appear relatively clean, and the effective Valence Band Maximum (VBM) can thus be mapped for each of the scanned regions. The situation appears more complex next to the conduction band side, where we observe several mid-gap peaks, especially in regions 5, 6 and 7, which are inside the nearly pure Ge QW region. Interestingly, we can see the variability from one random phase to the other, which suggests that to obtain a fully converged calculation, we would need to average over more random phases, as explained above (eq. 4.II.55).

Therefore, our averaged solid line does not include as many random phases as we would like due to computational limitations. However, the presence of these peaks is not a random noise-related artefact from these limitations and has an underlying physical source worth understanding. Let us evaluate it further.

These mid-gap peaks are isolated, as they appear at single energies instead of being part of a clear continuum corresponding to the valence or conduction bands. To

verify this, we run a calculation with better resolution. But as introduced, narrowing the spectral range from ($\sim$ 60 eV) to our region of interest of ($\sim$ 1 eV) while improving the energy broadening requires a very large number of moments (M). More precisely, we compute about $M \sim 25000$ moments, resulting in a broadening of $\eta \approx 4$ meV. However, to lower the cost of the calculation, we compute only one random phase, as we prioritise energy resolution than noise to infer the nature of their spectral isolation. This calculation is presented in Figure 4.II.19, for regions $r = 6, 7, 8, 9$. It appears at once very likely that the peaks from regions $r = 6$ and $r = 7$ are energetically isolated. This is important information that will allow us to infer their origin.

On the first hand, we discarded the effect of collapsing as their physical origin. To verify this, we ran the same calculation on another collapsing of the same structural model, using the same random phase states for each region. For the regions under investigation, which are inside the well and therefore mostly Ge with very little variation, it appears that collapsing has rather little impact on the mid-gap peaks. Moreover, we forced the QW to be 100 % and discarded the compositional effect of the isolated Si atoms in the QW. The result is persistent mid-gap peaks hinting at a structural origin rather than a compositional one.

The definitive test, though, comes from extracting the exact wavefunction relative to these peaks with sparse diagonalisation. As tight-binding Hamiltonians are expressed on the basis of localised atomic orbitals, they are sparse matrices (i.e., most elements are zero) which can be diagonalised efficiently. This allows us o calculate a small number of eigenvectors and eigenvalues in the neighbourhood of a given energy. In this way, such methods are complementary to the linear scaling methods discussed above: the latter provide information on the full energy spectrum of the system, but do not provide access to the eigenstates. Instead, sparse diagonalisation techniques can only target a few selected energies, but provide corresponding eigenstates, and therefore have spatial resolution to, for instance, map these mid-gap states. [135] In other words, if the averaged linear scaling LDOS of a region exhibits significant signals that we suspect originate from a small number of states, we can directly investigate the states at these energies and achieve atomic resolution using sparse diagonalisation methods.

Using the low-broadening LDOS results from Figure 4.II.19, we consider the strong peak in region 7 at an energy of $E_w \approx 1.13$ eV. We use the value of E_w estimated through the linear scaling LDOS calculation to seek one single eigenvalue due to the high computational cost. Thus, We must note that this is still an expensive calculation. Comparatively, while the evaluation of one random phase LDOS in one region takes on the order of hours, the sparse diagonalisation discussed here required a calculation time on the order of days (about five days for the system in Figure 4.II.19).

To reduce the complexity of the up to 20 considered orbitals per atom, we can reduce the full wavefunction, $|\Psi^w\rangle = \sum_{\mathbf{n},\lambda} \Psi^w_{\mathbf{n},\lambda} |\mathbf{n},\lambda\rangle$, to an atomic reduced density proportional to $\sum_{\mathbf{n},\lambda} \left|\Psi^w_{\mathbf{n},\lambda}\right|^2$. This probability is spatially maximised in the origin of these mid-gap states when centring the calculation in E_w.

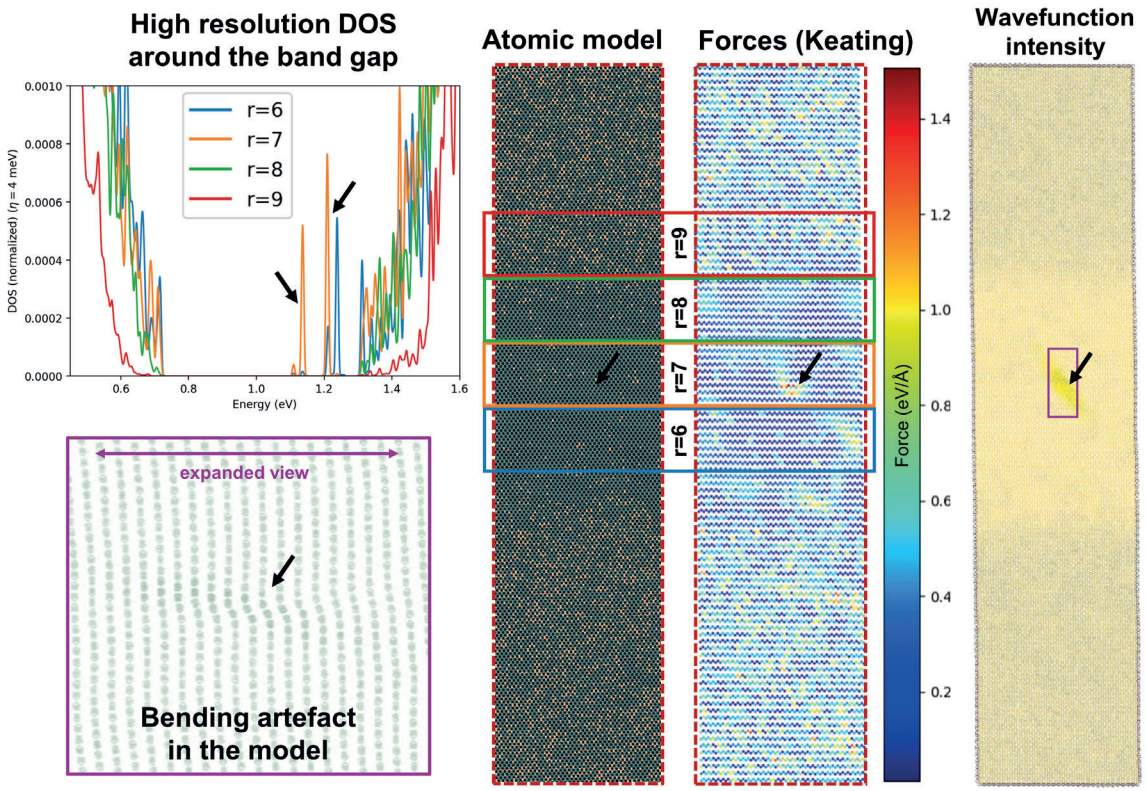


FIGURE 4.II.19: The top-left panel showcases tight-binding simulations via linear scaling to compute the Local Density of States (LDOS) with a broadening of $\sim$ 4 meV. They are computed by grabbing the portion of atoms within the coloured boxes in the 3D Atomic Model and its consequent Keating atomic forces in the central panel. The broad mid-gap states seen in Figure 4.II.18 for a broadening of $\sim$ 15 meV are now sharp and spectrally isolated. Sparse diagonalisation enables to locally map their source wavefunction and map its abundance, as in the right panel. It correlates the mid-gap states with the artefact structural bending observed in the zoomed-in and laterally expanded view within the quantum well (bottom-left panel), within the purple rectangle.

We observe that the state is localized in the xy plane and extended in z. This is consistent with the way the workflow builds the structure, assuming translational invariance, where the only changes in z come from local Si/Ge substitutions. The area where the state is localized in (x, y) corresponds to a region within the QW where the atomic planes in (y, z) show an abnormal bending. It coincides with the maximum of the wavefunction intensity, as showcased in the right panel of Figure 4.II.19. Inside it, in the purple rectangle, we expand the view in the x direction to illustrate this bending which otherwise would be impossible to notice by the naked eye. Intriguingly, force calculations using the Keating model show very high forces

on the atoms in the same area. This spatially coincides with the artefacts found during the atomic forces calculations displayed in Figure 4.II.17. The region in the same 3DAM of the same micrograph of the same QW where the forces are abnormally high makes us conclude that the mid-gap peak under consideration is caused by this abnormal bending of the planes. Such bending appears to be an artefact of the strain calculation and its transfer to the 3DAM, as indicated in the previous section, and as can be highlighted in Figures 4.I.26, 4.II.7 and 4.II.17. We extrapolate this knowledge and generalise by saying that the additional mid-gap states outside E_w also come from the same origin.

Similarly, we conducted an initial set of simulations with a 3DAM as wide as the one presented in Figure 4.II.17. Interestingly, in addition to the reported mid-gap states, we observed similar peaks but localised in the valence band. After cutting the model to ensure it was not an effect of the edge cutting, the added valence band states disappeared. Remarkably, this cut removed the structural artefact pinpointed with the black arrow in Figure 4.II.17. Thus, given the found correlation between the first artefact and the mid-gap states, we can easily say these valence band states raised because of the larger artefacts.

One would argue that the peaks and effects are nonphysical as they are not originally present in the experimental device. Although this statement is not false, it would mask the applicability of the overall workflow for knowledge and physics discovery. The current validation demonstrates the full potential of the workflow in terms of accuracy, precision, and sensitivity, showcasing what can be achieved if a real or experimental effect were to be faced. In other words, if a real bending appeared in the QW we would be able to see real modulations in the mid-gap region and in the valence band. Eventually, this would potentially affect the pursued hole gas and the quantum and transport properties of the device.

At this point, we stress the importance of performing strain calculations on the model structures as a first check, since these are computationally rather cheap compared to the electronic calculations. They might point a finger at local structural features that differ from the average behaviour, allowing us to verify if they are artefacts or real structural defects. Hence, they should be limited to informing rather than relaxing the structure, as we would otherwise lose the experimental significance gained with the workflow. At this point, we want to highlight how the combination of the three knowledge sources converged into a solution to validate the models. Until now, we were missing a tool to validate the confidence in the 3DAMs and their consequent simulations. By leveraging the strain analysis and 3DAM building, the Keating model for the atomic forces, and tight-binding models for the electronic structure, we can validate if a structural feature is artificial or not.

To finish with the simulations section and therefore with Part II of Chapter 4, we want to present a fully physically meaningful phenomenon which strengthens what

was found for the previous 3DAM cut. For the same GSQ_4 device, we used a higher-quality image, with minor acquisition noise. Figure 4.II.20 displays the high-quality HAADF micrograph. Its purple inset captures how a trustworthy high-quality micrograph should depict a crystal lattice: a uniform representation of all the FFT peaks with no clear preferential lattice fringes, together with the absence of scanning artefacts or drift. This idea is strengthened by the smooth GPA maps extracted from this image (fig. 4.II.20, Strain analysis).

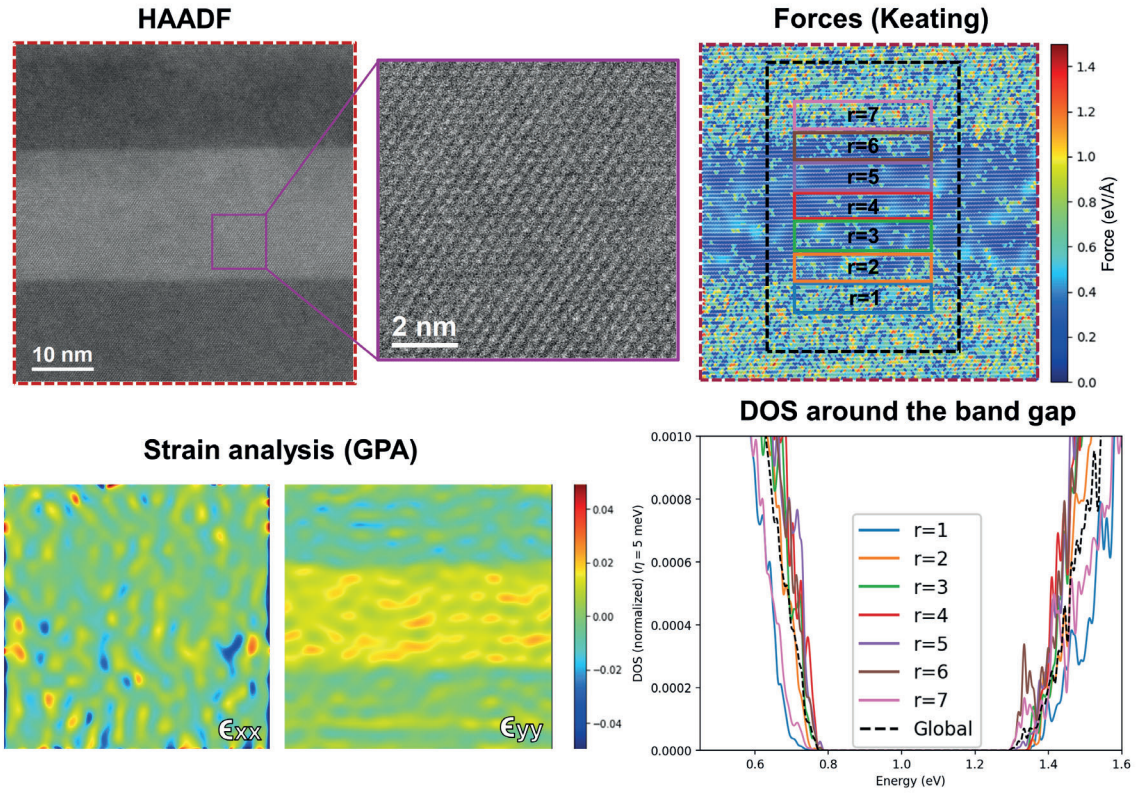


FIGURE 4.II.20: Tight-binding simulations in a high-quality High-Angle Annular Dark Field (HAADF) micrograph of the Ge Quantum Well (QW) GSQ_4 device. The Geometrical Phase Analysis (GPA) strain maps underscore the smoothness of the micrograph in terms of acquisition noise and artefacts. With a broadening of $\eta = 5$ meV, the Local Density of States (LDOS) is computed in the coloured regions depicted in a Keating atomic forces map displaying no major forces heterogeneity. The bottom-right panel showcases the LDOS and the absence of mid-gap states due to the absence of abrupt structural bending, contrary to Figure 4.II.19.

We used Keating model to compute the atomic forces in the model of the whole image and to seek signatures of possible artefacts. The result is a homogeneous force map where the only peaks rise from the expected presence of Si atoms in a Ge-like crystal lattice (fig. 4.II.20, Forces (Keating)). As a result, we do not require a strict cut removing artefact candidates. Therefore, as a first approximation, the edges of the 3DAM used are rough and have not been carefully recut. We just used a single random phase state and a broadening of $\eta = 5$ meV. We then compute the LDOS in every coloured box and over the global system to identify possible edge

effects observed in the global calculation vs the local ones. Remarkably, no visible difference appears between the curves, indicating that the smoothing of the edges may not be critical at the present level of precision. Therefore, the passivation with pseudo-Hydrogen atoms may be sufficient to avoid unwanted edge effects. Keep in mind that since we have used only one random state, further evaluation would be necessary to fully validate this remark.

Moreover, we observe a slight reduction in the band gap of the pure Ge region, outputting around 0.6 eV versus the expected 0.66 eV of Ge at room temperature, 298 K (0.74 eV at 0 K). [136] The reduction is plausible given the compressive strain Ge is experiencing and the tightening of the local cell parameter from its bulk relaxed reference. However, this reduction goes against the expected band gap increase due to the quantum confinement effects of the QW, and the consequent discretisation of the states. Therefore, we miss further investigations and systematic proofs to fully confirm whether this large shift is indeed physical or if unperceived strain as the one discussed in these sections is leading to unrealistic effects. Moreover, tackling other complementary QW thicknesses would also contribute to unveiling the dependence and reach of the quantum confinement effect. As a result, to fully answer concerns like this one, being able to locally map the band gap with high spatial and energy resolution would be a breakthrough for the community. This would provide a real reference to correlate with the simulations and verify their feasibility. Importantly, we precisely address this knowledge gap in Chapter 5.

However, the most thought-provoking result from these simulations is that the mid-gap states at energies around 1.13 eV and 1.2 eV are no longer seen in this structure, which re-enforces the idea that they were linked to structural artefacts (fig. 4.II.20, DOS plot). However, we are not yet capable of clarifying some aspects of the simulations still unknown. For instance, we still notice peaks in the DOS near the conduction band edge. It would be of interest to understand if these are physical of not, especially if they contribute to the calculation of the local band gap from the difference between the conduction and the valence edges in the LDOS. Furthermore, if we aim to see signatures relative to the devices in their working conditions, or similarly, heterogeneity in the bands, higher precisions will be needed, as well as possibly other simulation paradigms or target observables such as the conductivity.

Part III

Conclusions

4.III.1 Conclusions

We have introduced an innovative workflow that autonomously translates electron micrographs into models (FEMs and 3DAMs) suitable for theoretical analysis. This novel approach provides a rapid, accurate, and comprehensive structural description of imaged heterostructures and devices, addressing a critical gap in (S)TEM data analysis. Furthermore, it utilises this structural information to generate realistic models, either finite elements or atomic, that empower theoretical simulations aimed at extracting functional properties in the final material or device configuration. Our work represents a significant achievement with demonstrated reproducibility and validity.

Crucially, we believe that this workflow establishes a new paradigm in automated data analysis for experimental techniques. The underlying concept can be extended to various methods, with scanning probe and atomic force microscopies being prominent examples, especially for 2D materials. [28, 137, 138, 139, 140, 141, 142] We envision its potential widespread adoption within the microscopy community, both in academia and industry. The automation and user-friendly nature of this tool make it capital in overcoming the main long-standing limitation of TEM: its small throughput and lack of reproducibility in analysing large regions of interest. Thus, we regularly validate the knowledge extracted from a single or a limited number of lamellae as representative knowledge of the device. These sometimes precipitate conclusions result from the usually impractical replication of the experiments to enhance the statistical significance of the measurements. As a response, our workflow removes the data analysis bottleneck ensuring that an unlimited flow of acquired data could be analysed within minutes or hours, depending on device complexity. In that sense, combining the present workflow with ongoing developments in automation for data acquisition is pivotal in boosting the high analytical throughput our workflow can offer, thus accelerating scientific discoveries.

Through parts I (data analysis) and II we have demonstrated the workflow's exceptional ability to generalise and adapt to diverse samples or devices. This adaptability renders it suitable for retrieving a wide range of functional properties pertinent to the analysed material or device. The main proofs have targeted the simulation of the density of states and band structure for quantum devices, and their relaxation and interatomic forces. Therefore, our initial goal (section 1.2) of developing the workflow and validating it in a broad range of possible device configurations, but with special emphasis on quantum devices (Chapter 3), has been met. Furthermore, the validated versatility of the workflow could extend its reach to carrier densities and potential barriers for the same transport devices, but also beyond this application with, for instance, oxidation potentials and reaction mechanisms for catalytic particles.

We offer unprecedented value even beyond the pure automation of the (S)TEM data analysis which already skyrockets the currently possible throughput of the technique. The ultimate 3D models we engineer, both FEM and 3DAM, are unlikely to be manually engineered as they imply a vertex-wise or atom-wise coordinate modification, respectively. Therefore, its accomplishment could only happen through an automated workflow like the one we propose. Consequently, the direct link it establishes between structure and properties opens up possibilities for generating even more complex models incorporating additional inputs, physical or chemical constraints, or even inversely linking functional properties to retrieve the best structure. The modular nature of the workflow facilitates the integration of new functionalities, improvement of existing ones, and collaboration within the community, leveraging its open-source nature. [31, 143]

Despite the potential for enhancement of individual modules, the strength of the entire workflow lies in its robustness. Each modular unit, when perfected, transforms into a powerful analytical tool in its own right, contributing to an even more resilient ensemble. In fact, we 1) improved the segmentation of STEM micrographs, and to the best of our knowledge we 2) created the best FFT peak finding model to date, 3) enabled a general crystal phase identification usable for a broad range of situations (space groups, material systems, device and sample configurations...), 4) linked high and low magnification micrographs to access high-resolution information at unprecedented FOVs, 5) automated GPA providing awareness of the way the device spatially interconnects its building blocks, 6) built FEM models and 7) 3DAM with accuracy and precision only achievable by the present workflow, and 8) experimentally simulated key functional properties from these outputted models. Importantly, the modular nature unlocks numerous opportunities for integrating complementary models into the workflow. These may include data cleaning routines such as denoising, drift correction, or scanning artefacts correction, as well as final model refinement processes like atom counting for non-translational invariant models or model tuning. For instance, the parallel extraction of the real space atomic coordinates could help untangle a still not completely solved challenge of the workflow, which is the 3DAM validation. Additionally, there is ample room to expand into spectroscopy-related functionalities. [23, 144, 145, 146] While our current study primarily focuses on micrograph processing, future work can explore the automation of full model extraction directly from images and spectrum images as a unified input source. This could encompass retrieving compositional and physical information from the low-loss regime, [118, 119, 120, 147, 148] possibly incorporating advanced AI models that optimise simulation conditions for property retrieval. [50, 51, 149, 150, 151] While these considerations extend beyond the present scope, they represent fascinating avenues for future research.

4.III.2 Contributions

Although still crucial, Marc Botifoll Moral's role in the research presented in Chapter 3 supported the leading roles of the scientific collaborators, as described in the corresponding Contributions section of the chapter. However, in the present chapter, Marc Botifoll Moral fully lead the research, from its inception to its reported stage in Chapter 4. Specifically, Marc Botifoll Moral, supervised by Prof. Dr. Jordi Arbiol Cobos, contributed with:

- Conceived the idea of the workflow as a whole, of its independent parts, and their interconnection. The initial ambitious idea was as exposed in the central research goal (section 1.2): from the image to directly retrieve the 3D models. I strongly pushed for a vision of universality, generalisation and reproducibility as a central feature, which was until now the missing and pending task of AI in electron microscopy. Moreover, it was also pivotal to prosecute innovation beyond automation, which we achieved with the unprecedented accuracy of the compiled 3D models, among other contributions.
- Coded or supervised the coding of the entire workflow, except for some features detailed next. The whole workflow is built upon around 70000 lines of code. The size of the program is mainly due to 1) making it as much general as possible, and 2) ensuring it is fully free of human intervention. It can perfectly run with no more human intervention than just its mere execution after specifying the image to analyse.
- Acquired or participated in the data acquisition of the devices used to prove and develop the method. As most of the used devices appear in Chapter 3, the contributions lie in the decision on how to use them to develop and validate the methodology. That is, to transfer the needs identified in their characterisation detailed in Chapter 3 to a workflow capable of answering them.

The attained complexity of the final workflow would have been impractical without both the continuous and timely support of many scientific partners. Through its leading role, Marc Botifoll Moral actively supervised and coordinated their contributions, which are the following:

- Together with Prof. Dr. Jordi Arbiol Cobos, I co-supervised Ivan Pinto-Huguet in his final maths degree project, master thesis, and currently in his PhD thesis. During these projects, we developed the final version of the custom segmentation routine based on Canny edge detection and its post-processing. In addition, we entirely trained the CNN-based peak-finding model, and its post-processing based on the 1DScan model. Ivan led the creation of an intuitive graphical user interface that integrates every independent module, but also the whole workflow. The visual representation of the indexation of FFT with their corresponding family of planes (seen in sections 4.I.8 and 4.II.4) serves as

an illustration of this concept. Furthermore, Ivan's initiative extended to the evaluation of the SAM segmentation model, as well as the manual labelling of experimental FFTs to provide metrics on the peak-finding models.

- Dr. Enzo Rotunno and Dr. Vincenzo Grillo from CNR-IMEM in Italy extensively helped throughout the complete workflow development. They shared code to perform kinematical diffraction simulations, which is the base of data generation used to train the peak finding CNN, and the backend of the phase identification routine for model cell testing. They also shared their GPA code already implemented in STEM_CELL. [41, 42] Importantly, their contributions were key to creating the 3DAM strain transfer routine. Moreover, they built the foundations of the interconnection between their C code with the workflow backbone written in Python. Finally, after agreeing on the most interesting examples to simulate, Dr. Vincenzo Grillo and Payam Habibzadeh Kavkani conducted the FEM simulations using the COMSOL simulating software. [107]

- Dr. Thomas Jean-Fraçois Galvani and Prof. Stephan Roche, also from Catalan Institute of Nanoscience and Nanotechnology, fully deployed the tight-binding model to simulate the Si/Ge systems presented all over Chapters 3 and 4. They counted on the fundamental help of Dr. Yann-Michel Niquet and his precise tight-binding Hamiltonian for Si/Ge alloys. [130]

- Dr. Catalina Coll and Prof. Pablo Ordejón, also from Catalan Institute of Nanoscience and Nanotechnology, fully deployed the Keating model of the Si/Ge alloys to relax the structure and see the force distribution of its atoms. [128]

- Dr. Maria Chiara Spadaro, from Catalan Institute of Nanoscience and Nanotechnologyand the University of Catania, Italy, helped with the linear STEM simulations and their validation. [41]

- Dr. Martin Børstad Eriksen and Prof. Gonzalo Merino, from Punt d'Informació Científica (PIC), assisted in the training of the CNN devoted to peak finding. Their contribution was both through scientific discussion and by granting access to the AI-devoted computing facilities of PIC, where the generation of the training data and the training itself could be comfortably done.

- Everybody involved in the growth and conception of the devices studied in this chapter. As most of them come from Chapter 3, we refer to section 3.4 to assess these particular contributions and the researchers leading their conception.

References

[1] Roser Fernández-Climent et al. "Highly Durable Nanoporous Cu2–xS Films for Efficient Hydrogen Evolution Electrocatalysis under Mild pH Conditions". In: *ACS Catalysis* 13.15 (2023), pp. 10457–10467. DOI: 10.1021/acscatal.3c01673. eprint: https://doi.org/10.1021/acscatal.3c01673. URL: https://doi.org/10.1021/acscatal.3c01673.

[2] Sabbir A. Khan et al. "Epitaxially Driven Phase Selectivity of Sn in Hybrid Quantum Nanowires". In: *ACS Nano* 17.12 (2023). PMID: 37317984, pp. 11794–11804. DOI: 10.1021/acsnano.3c02733. eprint: https://doi.org/10.1021/acsnano.3c02733. URL: https://doi.org/10.1021/acsnano.3c02733.

[3] Marco Valentini et al. "Majorana-like Coulomb spectroscopy in the absence of zero-bias peaks". In: *Nature* 612 (2022), pp. 442–447. DOI: 10.1038/s41586-022-05382-w. eprint: https://doi.org/10.1038/s41586-022-05382-w. URL: https://doi.org/10.1038/s41586-022-05382-w.

[4] Sara Martí-sánchez et al. "Sub-nanometer mapping of strain-induced band structure variations in planar nanowire core-shell heterostructures". In: *Nature Communications* 13 (2022), p. 4089. DOI: /10.1038/s41467-022-31778-3. eprint: https://doi.org/10.1038/s41467-022-31778-3. URL: https://doi.org/10.1038/s41467-022-31778-3.

[5] Mathieu Kociak and Odile Stéphan. "Mapping plasmons at the nanometer scale in an electron microscope". In: *Chem. Soc. Rev.* 43 (11 2014), pp. 3865–3883. DOI: 10.1039/C3CS60478K. URL: http://dx.doi.org/10.1039/C3CS60478K.

[6] Cheng Chang et al. "Surface Functionalization of Surfactant-Free Particles: A Strategy to Tailor the Properties of Nanocomposites for Enhanced Thermoelectric Performance". In: *Angewandte Chemie International Edition* 61.35 (2022), e202207002. DOI: https://doi.org/10.1002/anie.202207002. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.202207002. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.202207002.

[7] Maria Ibáñez et al. "Tuning Transport Properties in Thermoelectric Nanocomposites through Inorganic Ligands and Heterostructured Building Blocks". In: *ACS Nano* 13.6 (2019). PMID: 31185159, pp. 6572–6580. DOI: 10.1021/acsnano.9b00346. eprint: https://doi.org/10.1021/acsnano.9b00346. URL: https://doi.org/10.1021/acsnano.9b00346.

[8] Peter D. Nellist Stephen J. Pennycook. *Scanning Transmission Electron Microscopy*. Springer New York, 2011. ISBN: 978-1-4419-7199-9. DOI: /10.1007/978-1-4419-7200-2. eprint: https://doi.org/10.1007/978-1-4419-7200-2. URL: https://doi.org/10.1007/978-1-4419-7200-2.

[9] David B. Williams and C. Barry Carter. *Transmission Electron Microscopy*. 2nd ed. Springer New York, 2009. ISBN: 978-0-387-76502-0.

[10] David A. Muller. "Structure and bonding at the atomic scale by scanning transmission electron microscopy". In: *Nature Materials* 8 (2009), pp. 263–270. DOI: /10.1038/nmat2380. eprint: https://doi.org/10.1038/nmat2380. URL: https://doi.org/10.1038/nmat2380.

[11] Marc Botifoll, Ivan Pinto-Huguet, and Jordi Arbiol. "Machine learning in electron microscopy for advanced nanocharacterization: current developments, available tools and future outlook". In: *Nanoscale Horiz.* 7 (12 2022), pp. 1427–1477. DOI: 10.1039/D2NH00377E. URL: http://dx.doi.org/10.1039/D2NH00377E.

[12] Rama K. Vasudevan et al. "Materials science in the artificial intelligence age: High-throughput library generation, machine learning, and a pathway from correlations to the underpinning physics". In: *MRS Communications* 9 (3 2019), pp. 821–838. DOI: 10.1557/mrc.2019.95. eprint: https://doi.org/10.1557/mrc.2019.95. URL: https://doi.org/10.1557/mrc.2019.95.

[13] Sergei V. Kalinin et al. "Machine learning in scanning transmission electron microscopy". In: *Nature Review Methods Primers* 2 (2022), p. 11. DOI: /10.1038/s43586-022-00095-w. eprint: https://doi.org/10.1038/s43586-022-00095-w. URL: https://doi.org/10.1038/s43586-022-00095-w.

[14] A. Gloter et al. "Improving energy resolution of EELS spectra: an alternative to the monochromator solution". In: *Ultramicroscopy* 96.3 (2003). Proceedings of the International Workshop on

Strategies and Advances in Atomic Level Spectroscopy and Analysis, pp. 385–400. ISSN: 0304-3991. DOI: https://doi.org/10.1016/S0304-3991(03)00103-7. URL: https://www.sciencedirect.com/science/article/pii/S0304399103001037.

[15] Stephen Jesse and Sergei V Kalinin. "Principal component and spatial correlation analysis of spectroscopic-imaging data in scanning probe microscopy". In: *Nanotechnology* 20.8 (Feb. 2009), p. 085714. DOI: 10.1088/0957-4484/20/8/085714. URL: https://dx.doi.org/10.1088/0957-4484/20/8/085714.

[16] Pavel Potapov, Paolo Longo, and Eiji Okunishi. "Enhancement of noisy EDX HRSTEM spectrum-images by combination of filtering and PCA". In: *Micron* 96 (2017), pp. 29–37. ISSN: 0968-4328. DOI: https://doi.org/10.1016/j.micron.2017.02.001. URL: https://www.sciencedirect.com/science/article/pii/S0968432816303821.

[17] Lewys Jones et al. "Improving the SNR of Atomic Resolution STEM EELS
amp; EDX Mapping while Reducing Beam-damage by using Non-rigid Spectrum-image Averaging". In: *Microscopy and Microanalysis* 21.S3 (Sept. 2015), pp. 1215–1216. ISSN: 1431-9276. DOI: 10.1017/S1431927615006868. eprint: https://academic.oup.com/mam/article-pdf/21/S3/1215/48334791/mam1215.pdf. URL: https://doi.org/10.1017/S1431927615006868.

[18] J.M.P. Nascimento and J.M.B. Dias. "Vertex component analysis: a fast algorithm to unmix hyperspectral data". In: *IEEE Transactions on Geoscience and Remote Sensing* 43.4 (2005), pp. 898–910. DOI: 10.1109/TGRS.2005.844293.

[19] Jacob Madsen et al. "A Deep Learning Approach to Identify Local Structures in Atomic-Resolution Transmission Electron Microscopy Images". In: *Advanced Theory and Simulations* 1.8 (2018), p. 1800037. DOI: https://doi.org/10.1002/adts.201800037. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adts.201800037. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/adts.201800037.

[20] Kohei Aso et al. "Detection of picometer-order atomic displacements in drift-compensated HAADF-STEM images of gold nanorods". In: *Microscopy* 65.5 (Sept. 2016), pp. 391–399. ISSN: 2050-5698. DOI: 10.1093/jmicro/dfw018. eprint: https://academic.oup.com/jmicro/article-pdf/65/5/391/8038086/dfw018.pdf. URL: https://doi.org/10.1093/jmicro/dfw018.

[21] Maxim Ziatdinov et al. "Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations". In: *ACS Nano* 11.12 (2017). PMID: 29215876, pp. 12742–12752. DOI: 10.1021/acsnano.7b07504. eprint: https://doi.org/10.1021/acsnano.7b07504. URL: https://doi.org/10.1021/acsnano.7b07504.

[22] Robert S. Pennington et al. "Neural-network-based depth-resolved multiscale structural optimization using density functional theory and electron diffraction data". In: *Phys. Rev. B* 97 (2 Jan. 2018), p. 024112. DOI: 10.1103/PhysRevB.97.024112. URL: https://link.aps.org/doi/10.1103/PhysRevB.97.024112.

[23] Feng Wang et al. "Noise2Atom: unsupervised denoising for scanning transmission electron microscopy images". In: *Applied Microscopy* 50 (2020), p. 23. DOI: /10.1186/s42649-020-00041-8. eprint: https://doi.org/10.1186/s42649-020-00041-8. URL: https://doi.org/10.1186/s42649-020-00041-8.

[24] Yuncheng Jin et al. "Machine learning guided rapid focusing with sensor-less aberration corrections". In: *Opt. Express* 26.23 (Nov. 2018), pp. 30162–30171. DOI: 10.1364/OE.26.030162. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-26-23-30162.

[25] E. Rotunno et al. "Alignment of electron optical beam shaping elements using a convolutional neural network". In: *Ultramicroscopy* 228 (2021), p. 113338. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2021.113338. URL: https://www.sciencedirect.com/science/article/pii/S0304399121001224.

[26] Enzo Rotunno et al. "Convolutional neural network as a tool for automatic alignment of electron optical beam shaping devices". In: *Microscopy and Microanalysis* 27.S1 (Aug. 2021), pp. 822–824. ISSN: 1431-9276. DOI: 10.1017/S143192762100324X. eprint: https://academic.oup.com/

mam/article-pdf/27/S1/822/47965307/mam0822.pdf. URL: https://doi.org/10.1017/S143192762100324X.

[27] Benjamin P. Cumming and Min Gu. "Direct determination of aberration functions in microscopy by an artificial neural network". In: *Opt. Express* 28.10 (May 2020), pp. 14511–14521. DOI: 10.1364/OE.390856. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-28-10-14511.

[28] Sergei V. Kalinin et al. "Automated and Autonomous Experiments in Electron and Scanning Probe Microscopy". In: *ACS Nano* 15.8 (2021). PMID: 34269558, pp. 12604–12627. DOI: 10.1021/acsnano.1c02104. eprint: https://doi.org/10.1021/acsnano.1c02104. URL: https://doi.org/10.1021/acsnano.1c02104.

[29] M.J. Hÿtch, E. Snoeck, and R. Kilaas. "Quantitative measurement of displacement and strain fields from HREM micrographs". In: *Ultramicroscopy* 74.3 (1998), pp. 131–146. ISSN: 0304-3991. DOI: https://doi.org/10.1016/S0304-3991(98)00035-7. URL: https://www.sciencedirect.com/science/article/pii/S0304399198000357.

[30] Martin J. Hÿtch, Jean-Luc Putaux, and Jean-Michel Pénisson. "Measurement of the displacement field of dislocations to 0.03 Å by electron microscopy". In: *Nature* 423 (2003), pp. 270–273. DOI: /10.1038/nature01638. eprint: https://doi.org/10.1038/nature01638. URL: https://doi.org/10.1038/nature01638.

[31] Suhas Somnath et al. *USID and Pycroscopy – Open frameworks for storing and analyzing spectroscopic and imaging data*. 2019. arXiv: 1903.09515 [physics.data-an].

[32] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4.

[33] Robbie Sadre et al. "Deep Learning Segmentation of Complex Features in Atomic-Resolution Phase-Contrast Transmission Electron Microscopy Images". In: *Microscopy and Microanalysis* 27.4 (2021), pp. 804–814. DOI: 10.1017/S1431927621000167.

[34] Rama K. Vasudevan et al. "Phases and Interfaces from Real Space Atomically Resolved Data: Physics-Based Deep Data Image Analysis". In: *Nano Letters* 16.9 (2016). PMID: 27517608, pp. 5574–5581. DOI: 10.1021/acs.nanolett.6b02130. eprint: https://doi.org/10.1021/acs.nanolett.6b02130. URL: https://doi.org/10.1021/acs.nanolett.6b02130.

[35] Artem Maksov et al. "Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS2". In: *npj Computational Materials* 5 (2019), p. 12. DOI: /10.1038/s41524-019-0152-9. eprint: https://doi.org/10.1038/s41524-019-0152-9. URL: https://doi.org/10.1038/s41524-019-0152-9.

[36] R. Kannan et al. "Deep data analysis via physically constrained linear unmixing: universal framework, domain examples, and a community-wide platform". In: *npj Computational Materials* 4 (2018), p. 6. DOI: /10.1186/s40679-018-0055-8. eprint: https://doi.org/10.1186/s40679-018-0055-8. URL: https://doi.org/10.1186/s40679-018-0055-8.

[37] Pavel Aseev et al. "Ballistic InSb Nanowires and Networks via Metal-Sown Selective Area Growth". In: *Nano Letters* 19.12 (2019). PMID: 31730748, pp. 9102–9111. DOI: 10.1021/acs.nanolett.9b04265. eprint: https://doi.org/10.1021/acs.nanolett.9b04265. URL: https://doi.org/10.1021/acs.nanolett.9b04265.

[38] Daniel Jirovec et al. "A singlet-triplet hole spin qubit in planar Ge". In: *Nature Materials* 20 (2021), pp. 1106–1112. DOI: /10.1038/s41563-021-01022-2. eprint: https://doi.org/10.1038/s41563-021-01022-2. URL: https://doi.org/10.1038/s41563-021-01022-2.

[39] Wenhua Li et al. "CuTe Nanocrystals: Shape and Size Control, Plasmonic Properties, and Use as SERS Probes and Photothermal Agents". In: *Journal of the American Chemical Society* 135.19 (2013). PMID: 23647089, pp. 7098–7101. DOI: 10.1021/ja401428e. eprint: https://doi.org/10.1021/ja401428e. URL: https://doi.org/10.1021/ja401428e.

[40] Debangshu Mukherjee and Raymond Unocic. "STEMTooL: An Open Source Python Toolkit for Analyzing Electron Microscopy Datasets". In: *Microscopy and Microanalysis* 26.S2 (2020), pp. 2960–2962. DOI: 10.1017/S143192762002334X.

[41] Vincenzo Grillo and Enzo Rotunno. "STEMCELL: A software tool for electron microscopy: Part I—simulations". In: *Ultramicroscopy* 125 (2013), pp. 97–111. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2012.10.016. URL: https://www.sciencedirect.com/science/article/pii/S0304399112002732.

[42] Vincenzo Grillo and Francesca Rossi. "STEMCELL: A software tool for electron microscopy. Part 2 analysis of crystalline materials". In: *Ultramicroscopy* 125 (2013), pp. 112–129. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2012.10.009. URL: https://www.sciencedirect.com/science/article/pii/S0304399112002665.

[43] Y. Chandra et al. "Advances in finite element modelling of graphene and associated nanostructures". In: *Materials Science and Engineering: R: Reports* 140 (2020), p. 100544. ISSN: 0927-796X. DOI: https://doi.org/10.1016/j.mser.2020.100544. URL: https://www.sciencedirect.com/science/article/pii/S0927796X20300024.

[44] Jan Pomplun et al. "Adaptive finite element method for simulation of optical nano structures". In: *physica status solidi (b)* 244.10 (2007), pp. 3419–3434. DOI: https://doi.org/10.1002/pssb.200743192. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pssb.200743192. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/pssb.200743192.

[45] P. Havu et al. "Finite-element implementation for electron transport in nanostructures". In: *The Journal of Chemical Physics* 124.5 (Feb. 2006), p. 054707. ISSN: 0021-9606. DOI: 10.1063/1.2162900. eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.2162900/15379852/054707_1_online.pdf. URL: https://doi.org/10.1063/1.2162900.

[46] María de la Mata et al. "The Role of Polarity in Nonplanar Semiconductor Nanostructures". In: *Nano Letters* 19.6 (2019). PMID: 31039314, pp. 3396–3408. DOI: 10.1021/acs.nanolett.9b00459. eprint: https://doi.org/10.1021/acs.nanolett.9b00459. URL: https://doi.org/10.1021/acs.nanolett.9b00459.

[47] Raquel Nafria et al. "Growth of Au–Pd2Sn Nanorods via Galvanic Replacement and Their Catalytic Performance on Hydrogenation and Sonogashira Coupling Reactions". In: *Langmuir* 34.36 (2018). PMID: 30096238, pp. 10634–10643. DOI: 10.1021/acs.langmuir.8b02023. eprint: https://doi.org/10.1021/acs.langmuir.8b02023. URL: https://doi.org/10.1021/acs.langmuir.8b02023.

[48] Jacob Madsen and Toma Susi. "abTEM: ab Initio Transmission Electron Microscopy Image Simulation". In: *Microscopy and Microanalysis* 26.S2 (2020), pp. 448–450. DOI: 10.1017/S1431927620014701.

[49] Toma Susi et al. "Efficient first principles simulation of electron scattering factors for transmission electron microscopy". In: *Ultramicroscopy* 197 (2019), pp. 16–22. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2018.11.002. URL: https://www.sciencedirect.com/science/article/pii/S0304399118302481.

[50] J. S. Smith, O. Isayev, and A. E. Roitberg. "ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost". In: *Chem. Sci.* 8 (4 2017), pp. 3192–3203. DOI: 10.1039/C6SC05720A. URL: http://dx.doi.org/10.1039/C6SC05720A.

[51] Nongnuch Artrith and Alexander Urban. "An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO2". In: *Computational Materials Science* 114 (2016), pp. 135–150. ISSN: 0927-0256. DOI: https://doi.org/10.1016/j.commatsci.2015.11.047. URL: https://www.sciencedirect.com/science/article/pii/S0927025615007806.

[52] Pavel Aseev et al. "Selectivity Map for Molecular Beam Epitaxy of Advanced III–V Quantum Nanowire Networks". In: *Nano Letters* 19.1 (2019). PMID: 30521341, pp. 218–227. DOI: 10.1021/acs.nanolett.8b03733. eprint: https://doi.org/10.1021/acs.nanolett.8b03733. URL: https://doi.org/10.1021/acs.nanolett.8b03733.

[53] Brian Paquelet Wuetz et al. “Reducing charge noise in quantum dots by using thin silicon quantum wells”. In: *Nature Communications* 14.1 (Mar. 2023), p. 1385. ISSN: 2041-1723. DOI: 10.1038/s41467-023-36951-w. URL: https://doi.org/10.1038/s41467-023-36951-w.

[54] Brian Paquelet Wuetz et al. “Reducing charge noise in quantum dots by using thin silicon quantum wells”. In: *Nature Communications* 14 (2023), p. 1385. DOI: /10.1038/s41467-023-36951-w. eprint: https://doi.org/10.1038/s41467-023-36951-w. URL: https://doi.org/10.1038/s41467-023-36951-w.

[55] Lingfeng Wang and Chunhong Pan. “Robust level set image segmentation via a local correntropy-based K-means clustering”. In: *Pattern Recognition* 47.5 (2014), pp. 1917–1925. ISSN: 0031-3203. DOI: https://doi.org/10.1016/j.patcog.2013.11.014. URL: https://www.sciencedirect.com/science/article/pii/S0031320313005001.

[56] Toufiq Parag et al. “A context-aware delayed agglomeration framework for electron microscopy segmentation”. en. In: *PLoS One* 10.5 (May 2015), e0125825.

[57] Randy M Sterbentz, Kristine L Haley, and Joshua O Island. “Universal image segmentation for optical identification of 2D materials”. en. In: *Sci. Rep.* 11.1 (Mar. 2021), p. 5808.

[58] Concepción Paz et al. “Development of an active contour based algorithm to perform the segmentation of soot agglomerates in uneven illumination TEM imaging”. en. In: *Powder Technol.* 400.117260 (Mar. 2022), p. 117260.

[59] Muller Andreas C and Sarah Guido. *Introduction to machine learning with python*. O’Reilly Media, Sept. 2016.

[60] Francisco de la Pena et al. “Electron Microscopy (Big and Small) Data Analysis With the Open Source Software Package HyperSpy”. In: *Microscopy and Microanalysis* 23.S1 (Aug. 2017), pp. 214–215. ISSN: 1431-9276. DOI: 10.1017/S1431927617001751. eprint: https://academic.oup.com/mam/article-pdf/23/S1/214/48301678/mam0214.pdf. URL: https://doi.org/10.1017/S1431927617001751.

[61] Francisco de la Peña et al. *hyperspy/hyperspy: v2.0*. Version v2.0. Dec. 2023. DOI: 10.5281/zenodo.10412190. URL: https://doi.org/10.5281/zenodo.10412190.

[62] Fabian Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12.85 (2011), pp. 2825–2830. URL: http://jmlr.org/papers/v12/pedregosa11a.html.

[63] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. " O’Reilly Media, Inc.", 2008.

[64] *2.3. Clustering*. en. https://scikit-learn.org/stable/modules/clustering.html. Accessed: 2023-11-14.

[65] *Sklearn.Cluster.KMeans*. en. https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html. Accessed: 2023-11-14.

[66] *Sklearn.Cluster.AgglomerativeClustering*. en. https://scikit-learn.org/stable/modules/generated/sklearn.cluster.AgglomerativeClustering.html. Accessed: 2023-11-14.

[67] *Sklearn.Cluster.DBSCAN*. en. https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html. Accessed: 2023-11-14.

[68] *OpenCV: Canny edge detection*. https://docs.opencv.org/4.x/da/d22/tutorial_py_canny.html. Accessed: 2023-11-14.

[69] Kushagra Aggarwal et al. “Enhancement of proximity-induced superconductivity in a planar Ge hole gas”. In: *Phys. Rev. Res.* 3 (2 Apr. 2021), p. L022005. DOI: 10.1103/PhysRevResearch.3.L022005. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.3.L022005.

[70] Alberto Tosato et al. “Hard superconducting gap in germanium”. In: *Communications Materials* 4 (2023), p. 23. DOI: /10.1038/s43246-023-00351-w. eprint: https://doi.org/10.1038/s43246-023-00351-w. URL: https://doi.org/10.1038/s43246-023-00351-w.

[71] Alexander Kirillov et al. *Segment Anything*. 2023. arXiv: 2304.02643 [cs.CV].

[72] John Canny. “A Computational Approach to Edge Detection”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8.6 (1986), pp. 679–698. DOI: 10.1109/TPAMI.1986.4767851.

[73] N. Kanopoulos, N. Vasanthavada, and R.L. Baker. "Design of an image edge detection filter using the Sobel operator". In: *IEEE Journal of Solid-State Circuits* 23.2 (1988), pp. 358–367. DOI: 10.1109/4.996.

[74] Girish N. Chaple, R. D. Daruwala, and Manoj S. Gofane. "Comparisions of Robert, Prewitt, Sobel operator based edge detection methods for real time uses on FPGA". In: *2015 International Conference on Technologies for Sustainable Development (ICTSD)*. 2015, pp. 1–4. DOI: 10.1109/ICTSD.2015.7095920.

[75] Lei Yang et al. "An improved Prewitt algorithm for edge detection based on noised image". In: *2011 4th International Congress on Image and Signal Processing*. Vol. 3. 2011, pp. 1197–1200. DOI: 10.1109/CISP.2011.6100495.

[76] D. Marr, E. Hildreth, and Sydney Brenner. "Theory of edge detection". In: *Proceedings of the Royal Society of London. Series B. Biological Sciences* 207.1167 (1980), pp. 187–217. DOI: 10.1098/rspb.1980.0020. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspb.1980.0020. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.1980.0020.

[77] Jialin Peng, Jiajin Yi, and Zhimin Yuan. "Unsupervised Mitochondria Segmentation in EM Images via Domain Adaptive Multi-Task Learning". In: *IEEE Journal of Selected Topics in Signal Processing* 14.6 (2020), pp. 1199–1209. DOI: 10.1109/JSTSP.2020.3005317.

[78] Lin Chen, Ruba Jebril, and Kamal Al Nasr. "Segmentation-Based Feature Extraction for Cryo-Electron Microscopy at Medium Resolution". In: *Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*. BCB '20. Virtual Event, USA: Association for Computing Machinery, 2020. ISBN: 9781450379649. DOI: 10.1145/3388440.3414711. URL: https://doi.org/10.1145/3388440.3414711.

[79] Alberto Garcia-Garcia et al. "A survey on deep learning techniques for image and video semantic segmentation". In: *Applied Soft Computing* 70 (2018), pp. 41–65. ISSN: 1568-4946. DOI: https://doi.org/10.1016/j.asoc.2018.05.018. URL: https://www.sciencedirect.com/science/article/pii/S1568494618302813.

[80] Fuzhen Zhuang et al. "A Comprehensive Survey on Transfer Learning". In: *Proceedings of the IEEE* 109.1 (2021), pp. 43–76. DOI: 10.1109/JPROC.2020.3004555.

[81] Debangshu Mukherjee et al. "mpfit: a robust method for fitting atomic resolution images with multiple Gaussian peaks". In: *Advanced Structural Chemical Imaging* 6.1 (2020), pp. 1199–1209. DOI: /10.1186/s40679-020-0068-y. URL: https://doi.org/10.1186/s40679-020-0068-y.

[82] Pauli Virtanen et al. "SciPy 1.0: fundamental algorithms for scientific computing in Python". In: *Nature Methods* 17.3 (Mar. 2020), pp. 261–272. ISSN: 1548-7105. DOI: 10.1038/s41592-019-0686-2. URL: https://doi.org/10.1038/s41592-019-0686-2.

[83] *The OpenCV Reference Manual*. 2.4.13.7. OpenCV. Apr. 2014.

[84] D.G Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". In: *International Journal of Computer Vision* 60 (2004), pp. 91–110. DOI: /10.1023/B:VISI.0000029664.99615.94. URL: https://doi.org/10.1023/B:VISI.0000029664.99615.94.

[85] Herbert Bay et al. "Speeded-Up Robust Features (SURF)". In: *Computer Vision and Image Understanding* 110.3 (2008). Similarity Matching in Computer Vision and Multimedia, pp. 346–359. ISSN: 1077-3142. DOI: https://doi.org/10.1016/j.cviu.2007.09.014. URL: https://www.sciencedirect.com/science/article/pii/S1077314207001555.

[86] Ethan Rublee et al. "ORB: An efficient alternative to SIFT or SURF". In: *2011 International Conference on Computer Vision*. 2011, pp. 2564–2571. DOI: 10.1109/ICCV.2011.6126544.

[87] Shaharyar Ahmed Khan Tareen and Zahra Saleem. "A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK". In: *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. 2018, pp. 1–10. DOI: 10.1109/ICOMET.2018.8346440.

[88] Amil Merchant et al. "Scaling deep learning for materials discovery". In: *Nature* 624.7990 (Dec. 2023), pp. 80–85. ISSN: 1476-4687. DOI: 10.1038/s41586-023-06735-9. URL: https://doi.org/10.1038/s41586-023-06735-9.

[89] Ute Kolb, Yaşar Krysiak, and Sergi Plana-Ruiz. "Automated electron diffraction tomography – development and applications". In: *Acta Crystallographica Section B* 75.4 (Aug. 2019), pp. 463–474. DOI: 10.1107/S2052520619006711. URL: https://doi.org/10.1107/S2052520619006711.

[90] Ron Lifshitz. "Quasicrystals: A Matter of Definition". In: *Foundations of Physics* 33.12 (Dec. 2003), pp. 1703–1711. ISSN: 1572-9516. DOI: 10.1023/A:1026247120031. URL: https://doi.org/10.1023/A:1026247120031.

[91] Ottavio Raul D. R. Carmignano et al. "Serpentinites: Mineral Structure, Properties and Technological Applications". In: *Journal of the Brazilian Chemical Society* 31.1 (Jan. 2020), pp. 2–14. ISSN: 0103-5053. DOI: 10.21577/0103-5053.20190215. URL: https://doi.org/10.21577/0103-5053.20190215.

[92] Catherine M. Oikonomou and Grant J. Jensen. "The development of cryo-EM and how it has advanced microbiology". In: *Nature Microbiology* 2.12 (Dec. 2017), pp. 1577–1579. ISSN: 2058-5276. DOI: 10.1038/s41564-017-0073-7. URL: https://doi.org/10.1038/s41564-017-0073-7.

[93] E. Alnabati and D. Kihara. "Advances in structure modeling methods for cryo-electron microscopy maps". In: *Molecules* 25 (2020), pp. 1–13.

[94] R. Henderson and P. N. T. Unwin. "Three-dimensional model of purple membrane obtained by electron microscopy". In: *Nature* 257.5521 (Sept. 1975), pp. 28–32. ISSN: 1476-4687. DOI: 10.1038/257028a0. URL: https://doi.org/10.1038/257028a0.

[95] Shawn Q. Zheng et al. "MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy". In: *Nature Methods* 14.4 (Apr. 2017), pp. 331–332. ISSN: 1548-7105. DOI: 10.1038/nmeth.4193. URL: https://doi.org/10.1038/nmeth.4193.

[96] Yadira Arroyo Rojas Dasilva et al. "Structural defects in cubic semiconductors characterized by aberration-corrected scanning transmission electron microscopy". In: *Ultramicroscopy* 176 (2017). 70th Birthday of Robert Sinclair and 65th Birthday of Nestor J. Zaluzec PICO 2017 – Fourth Conference on Frontiers of Aberration Corrected Electron Microscopy, pp. 11–22. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2016.09.015. URL: https://www.sciencedirect.com/science/article/pii/S0304399116302121.

[97] A. M. Sanchez et al. "Stable Defects in Semiconductor Nanowires". In: *Nano Letters* 18.5 (2018). PMID: 29624404, pp. 3081–3087. DOI: 10.1021/acs.nanolett.8b00620. eprint: https://doi.org/10.1021/acs.nanolett.8b00620. URL: https://doi.org/10.1021/acs.nanolett.8b00620.

[98] E. Rotunno et al. "Three dimensional analysis of the composition in solid alloys by variable probe in scanning transmission electron microscopy". In: *Ultramicroscopy* 146 (2014), pp. 62–70. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2014.07.003. URL: https://www.sciencedirect.com/science/article/pii/S0304399114001260.

[99] V. Grillo, E. Carlino, and F. Glas. "Influence of the static atomic displacement on atomic resolution Z-contrast imaging". In: *Phys. Rev. B* 77 (5 Feb. 2008), p. 054103. DOI: 10.1103/PhysRevB.77.054103. URL: https://link.aps.org/doi/10.1103/PhysRevB.77.054103.

[100] Tom Willhammar et al. "Structure and vacancy distribution in copper telluride nanoparticles influence plasmonic activity in the near-infrared". In: *Nature Communications* 8.1 (Mar. 2017), p. 14925. ISSN: 2041-1723. DOI: 10.1038/ncomms14925. URL: https://doi.org/10.1038/ncomms14925.

[101] Chao Yue Zhang et al. "Identifying the Role of the Cationic Geometric Configuration in Spinel Catalysts for Polysulfide Conversion in Sodium–Sulfur Batteries". In: *Journal of the American Chemical Society* 145.34 (2023). PMID: 37603793, pp. 18992–19004. DOI: 10.1021/jacs.3c06288. eprint: https://doi.org/10.1021/jacs.3c06288. URL: https://doi.org/10.1021/jacs.3c06288.

[102] Santhanu Panikar Ramanandan et al. "Control of Ge island coalescence for the formation of nanowires on silicon". In: *Nanoscale Horiz.* (2024), pp. -. DOI: 10.1039/D3NH00573A. URL: http://dx.doi.org/10.1039/D3NH00573A.

[103] Meshuggah. *Dancers To A Discordant System*. https://open.spotify.com/intl-es/track/2qbQLga60JnISUYwAG9aDx. Accessed: 08/04/2024. 2008.

[104] Ricardo Martins, Nuno Lourenço, and Nuno Horta. "LAYGEN II—Automatic Layout Generation of Analog Integrated Circuits". In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 32.11 (2013), pp. 1641–1654. DOI: 10.1109/TCAD.2013.2269050.

[105] Wim Bogaerts and Lukas Chrostowski. "Silicon Photonics Circuit Design: Methods, Tools and Challenges". In: *Laser & Photonics Reviews* 12.4 (2018), p. 1700237. DOI: https://doi.org/10.1002/lpor.201700237. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/lpor.201700237. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.201700237.

[106] Stefan Birner et al. "nextnano: General Purpose 3-D Simulations". In: *IEEE Transactions on Electron Devices* 54.9 (2007), pp. 2137–2142. DOI: 10.1109/TED.2007.902871.

[107] COMSOL and AB. *COMSOL Multiphysics® v. 6.2*. Stockholm, Sweden. URL: www.comsol.com.

[108] C. Koch. *QSTEM: Quantitative TEM/STEM Simulations*. Strukturforschung / Elektronenmikroskopie.

[109] J. Barthel. "Dr. Probe: A software for high-resolution STEM image simulation". In: *Ultramicroscopy* 193 (2018), pp. 1–11. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2018.06.003. URL: https://www.sciencedirect.com/science/article/pii/S0304399118301402.

[110] S Bernal et al. "The interpretation of HREM images of supported metal catalysts using image simulation: profile view images". In: *Ultramicroscopy* 72.3 (1998), pp. 135–164. ISSN: 0304-3991. DOI: https://doi.org/10.1016/S0304-3991(98)00009-6. URL: https://www.sciencedirect.com/science/article/pii/S0304399198000096.

[111] S. Bernal et al. "High-resolution electron microscopy investigation of metal–support interactions in Rh/TiO2". In: *J. Chem. Soc., Faraday Trans.* 92 (15 1996), pp. 2799–2809. DOI: 10.1039/FT9969202799. URL: http://dx.doi.org/10.1039/FT9969202799.

[112] Ask Hjorth Larsen et al. "The atomic simulation environment—a Python library for working with atoms". In: *Journal of Physics: Condensed Matter* 29.27 (June 2017), p. 273002. DOI: 10.1088/1361-648X/aa680e. URL: https://dx.doi.org/10.1088/1361-648X/aa680e.

[113] Aaron Meurer et al. "SymPy: symbolic computing in Python". In: *PeerJ Computer Science* 3 (Jan. 2017), e103. ISSN: 2376-5992. DOI: 10.7717/peerj-cs.103. URL: https://doi.org/10.7717/peerj-cs.103.

[114] J. M. Ede and R. Beanland. "Improving electron micrograph signal-to-noise with an atrous convolutional encoder-decoder". In: *Ultramicroscopy* 202 (2019), pp. 18–25.

[115] A. Stevens et al. "The potential for bayesian compressive sensing to significantly reduce electron dose in high-resolution STEM images". In: *Microscopy* 63 (2014), pp. 41–51.

[116] L. Jones et al. "Smart Align—a new tool for robust non-rigid registration of scanning microscope data". In: *Adv. Struct. Chem. Imaging* 1 (2015), pp. 1–16.

[117] L. Jones and P. D. Nellist. "Identifying and correcting scan noise and drift in the scanning transmission electron microscope". In: *Microsc. Microanal.* 19 (2013), pp. 1050–1060.

[118] Jo Verbeeck et al. "Advancing EELS into an Unsupervised Quantification Method". In: *Microscopy and Microanalysis* 29.Supplement 1 (July 2023), pp. 397–398. ISSN: 1431-9276. DOI: 10.1093/micmic/ozad067.186. URL: https://doi.org/10.1093/micmic/ozad067.186.

[119] J. Verbeeck and S. Van Aert. "Model based quantification of EELS spectra". In: *Ultramicroscopy* 101.2 (2004), pp. 207–224. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2004.06.004. URL: https://www.sciencedirect.com/science/article/pii/S0304399104001433.

[120] J. Verbeeck, S. Van Aert, and G. Bertoni. "Model-based quantification of EELS spectra: Including the fine structure". In: *Ultramicroscopy* 106.11 (2006). Proceedings of the International Workshop on Enhanced Data Generated by Electrons, pp. 976–980. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2006.05.006. URL: https://www.sciencedirect.com/science/article/pii/S0304399106000945.

[121] M. de la Mata et al. "Polarity assignment in ZnTe, GaAs, ZnO and GaN-AlN nanowires from direct dumbbell analysis". In: *Nano Letters* 12 (2012), pp. 2579–2586.

[122] H.G. Weller et al. “A Tensorial Approach to Computational Continuum Mechanics Using Object Orientated Techniques”. In: *Computers in Physics* 12 (Nov. 1998), pp. 620–631. DOI: 10.1063/1.168744.

[123] A. Trellakis et al. “The 3D nanometer device project nextnano: Concepts, methods, results”. In: *J. Comput. Electron.* 5.4 (2006), pp. 285–289. DOI: 10.1007/s10825-006-0005-x.

[124] Alberto García et al. “Siesta: Recent developments and applications”. In: *The Journal of Chemical Physics* 152.20 (May 2020), p. 204108. ISSN: 0021-9606. DOI: 10.1063/5.0005077. eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0005077/16745439/204108_1_online.pdf. URL: https://doi.org/10.1063/5.0005077.

[125] Manish Jain, James R. Chelikowsky, and Steven G. Louie. “Reliability of Hybrid Functionals in Predicting Band Gaps”. In: *Phys. Rev. Lett.* 107 (21 Nov. 2011), p. 216806. DOI: 10.1103/PhysRevLett.107.216806. URL: https://link.aps.org/doi/10.1103/PhysRevLett.107.216806.

[126] Xue-ke Wu et al. “The first principle study of the electronic structure of SixGe(1x) alloy films”. In: *Physics Letters A* 382.47 (2018), pp. 3418–3422. ISSN: 0375-9601. DOI: https://doi.org/10.1016/j.physleta.2018.09.035. URL: https://www.sciencedirect.com/science/article/pii/S0375960118309952.

[127] P. N. Keating. “Effect of Invariance Requirements on the Elastic Strain Energy of Crystals with Application to the Diamond Structure”. In: *Phys. Rev.* 145 (2 May 1966), pp. 637–645. DOI: 10.1103/PhysRev.145.637. URL: https://link.aps.org/doi/10.1103/PhysRev.145.637.

[128] H. Rücker and M. Methfessel. “Anharmonic Keating model for group-IV semiconductors with application to the lattice dynamics in alloys of Si, Ge, and C”. In: *Phys. Rev. B* 52 (15 Oct. 1995), pp. 11059–11072. DOI: 10.1103/PhysRevB.52.11059. URL: https://link.aps.org/doi/10.1103/PhysRevB.52.11059.

[129] F. Pezzoli et al. “Strain-induced shift of phonon modes in Si1-xGex alloys”. In: *Materials Science in Semiconductor Processing* 9.4 (2006). Proceedings of Symposium T E-MRS 2006 Spring Meeting on Germanium based semiconductors from materials to devices, pp. 541–545. ISSN: 1369-8001. DOI: https://doi.org/10.1016/j.mssp.2006.08.046. URL: https://www.sciencedirect.com/science/article/pii/S1369800106001909.

[130] Y. M. Niquet et al. “Onsite matrix elements of the tight-binding Hamiltonian of a strained crystal: Application to silicon, germanium, and their alloys”. In: *Phys. Rev. B* 79 (24 June 2009), p. 245201. DOI: 10.1103/PhysRevB.79.245201. URL: https://link.aps.org/doi/10.1103/PhysRevB.79.245201.

[131] *TB_Sim webpage*. https://www.mem-lab.fr/en/Pages/L_SIM/Softwares/TB_Sim.aspx. Accessed: 2024-05-03.

[132] Jose H Garcia and Tatiana G Rappoport. “Kubo–Bastin approach for the spin Hall conductivity of decorated graphene”. In: *2D Materials* 3.2 (2016), p. 024007.

[133] Zheyong Fan et al. “Linear scaling quantum transport methodologies”. In: *Physics Reports* 903 (2021), pp. 1–69.

[134] Alexander Weiße et al. “The kernel polynomial method”. In: *Reviews of modern physics* 78.1 (2006), p. 275.

[135] Dean Moldovan. “Electronic properties of strained graphene and supercritical charge centers”. PhD thesis. University of Antwerp, 2016.

[136] S.C. Jain and D.J. Roulston. “A simple expression for band gap narrowing (BGN) in heavily doped Si, Ge, GaAs and GexSi1x strained layers”. In: *Solid-State Electronics* 34.5 (1991), pp. 453–465. ISSN: 0038-1101. DOI: https://doi.org/10.1016/0038-1101(91)90149-S. URL: https://www.sciencedirect.com/science/article/pii/003811019190149S.

[137] Rama K. Vasudevan et al. “Autonomous Experiments in Scanning Probe Microscopy and Spectroscopy: Choosing Where to Explore Polarization Dynamics in Ferroelectrics”. In: *ACS Nano* 15.7 (2021). PMID: 34228427, pp. 11253–11262. DOI: 10.1021/acsnano.0c10239. eprint: https://doi.org/10.1021/acsnano.0c10239. URL: https://doi.org/10.1021/acsnano.0c10239.

[138] Kyle P. Kelley et al. "Fast Scanning Probe Microscopy via Machine Learning: Non-Rectangular Scans with Compressed Sensing and Gaussian Process Optimization". In: *Small* 16.37 (2020), p. 2002878. DOI: https://doi.org/10.1002/smll.202002878. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smll.202002878. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/smll.202002878.

[139] Alexey G. Okunev et al. "Nanoparticle Recognition on Scanning Probe Microscopy Images Using Computer Vision and Deep Learning". In: *Nanomaterials* 10.7 (2020). ISSN: 2079-4991. DOI: 10.3390/nano10071285. URL: https://www.mdpi.com/2079-4991/10/7/1285.

[140] Maxim Ziatdinov et al. "Quantifying the Dynamics of Protein Self-Organization Using Deep Learning Analysis of Atomic Force Microscopy Data". In: *Nano Letters* 21.1 (2021). PMID: 33306401, pp. 158–165. DOI: 10.1021/acs.nanolett.0c03447. eprint: https://doi.org/10.1021/acs.nanolett.0c03447. URL: https://doi.org/10.1021/acs.nanolett.0c03447.

[141] Yue Liu et al. "General Resolution Enhancement Method in Atomic Force Microscopy Using Deep Learning". In: *Advanced Theory and Simulations* 2.2 (2019), p. 1800137. DOI: https://doi.org/10.1002/adts.201800137. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adts.201800137. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/adts.201800137.

[142] Benjamin Alldritt et al. "Automated structure discovery in atomic force microscopy". In: *Science Advances* 6.9 (2020), eaay6913. DOI: 10.1126/sciadv.aay6913. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.aay6913. URL: https://www.science.org/doi/abs/10.1126/sciadv.aay6913.

[143] Francisco de la Pena et al. "Electron Microscopy (Big and Small) Data Analysis With the Open Source Software Package HyperSpy". In: *Microscopy and Microanalysis* 23.S1 (2017), pp. 214–215. DOI: 10.1017/S1431927617001751.

[144] Lewys Jones et al. "Smart Align - a new tool for robust non-rigid registration of scanning microscope data". In: *Advanced Structural Chemical Imaging* 1 (2015), p. 8. DOI: 10.1186/s40679-015-0008-4. eprint: https://doi.org/10.1186/s40679-015-0008-4. URL: https://doi.org/10.1186/s40679-015-0008-4.

[145] A. De Backer et al. "Atom counting in HAADF STEM using a statistical model-based approach: Methodology, possibilities, and inherent limitations". In: *Ultramicroscopy* 134 (2013). 65th Birthdays of W Owen Saxton, David J Smith and Dirk Van Dyck / PICO 2013 From Multislice to Big Bang, pp. 23–33. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2013.05.003. URL: https://www.sciencedirect.com/science/article/pii/S0304399113001228.

[146] Lewys Jones and Peter D. Nellist. "Identifying and Correcting Scan Noise and Drift in the Scanning Transmission Electron Microscope". In: *Microscopy and Microanalysis* 19.4 (2013), pp. 1050–1060. DOI: 10.1017/S1431927613001402.

[147] David Rossouw et al. "A New Method for Determining the Composition of Core–Shell Nanoparticles via Dual-EDX+EELS Spectrum Imaging". In: *Particle & Particle Systems Characterization* 33.10 (2016), pp. 749–755. DOI: https://doi.org/10.1002/ppsc.201600096. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ppsc.201600096. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ppsc.201600096.

[148] Sang-Hyeok Yang et al. "Deep Learning-Assisted Quantification of Atomic Dopants and Defects in 2D Materials". In: *Advanced Science* 8.16 (2021), p. 2101099. DOI: https://doi.org/10.1002/advs.202101099. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/advs.202101099. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/advs.202101099.

[149] Anton V. Sinitskiy and Vijay S. Pande. *Deep Neural Network Computes Electron Densities and Energies of a Large Set of Organic Molecules Faster than Density Functional Theory (DFT)*. 2018. arXiv: 1809.02723 [physics.chem-ph].

[150] Jörg Behler. "First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems". In: *Angewandte Chemie International Edition* 56.42 (2017), pp. 12828–12840. DOI: https://doi.org/10.1002/anie.201703114. eprint:

https : / / onlinelibrary . wiley . com / doi / pdf / 10 . 1002 / anie . 201703114. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201703114.

[151] Nongnuch Artrith, Tobias Morawietz, and Jörg Behler. "High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide". In: *Phys. Rev. B* 83 (15 Apr. 2011), p. 153101. DOI: 10.1103/PhysRevB.83.153101. URL: https://link.aps.org/doi/10.1103/PhysRevB.83.153101.



CHAPTER 5: SUB-NANOMETER MAPPING OF STRAIN-INDUCED BAND STRUCTURE VARIATIONS IN SEMICONDUCTOR HETEROSTRUCTURES

Chapter 5

Sub-Nanometer Mapping of Strain-Induced Band Structure Variations in Semiconductor Heterostructures

The complete structural and compositional characterisation of semiconductors at the atomic level is crucial for their functional understanding, as showcased in Chapters 3 and 4. However, reaching atomic or at least sub-nanometer spatial resolution when accessing physical properties like their most characteristic one, the band gap, is still an unresolved challenge. In this chapter, we aim to answer this open question by pushing state-of-the-art methodologies for band gap mapping at the high spatial and energy resolutions Electron Energy Loss Spectroscopy can offer. The goal is to understand the spatial distribution of this attribute along devices to unveil how interfaces, defects, strain, or other local effects can affect the optoelectronic properties of devices. Through our proposed methodology, we circumvent the challenges that have previously hindered the advancement of this knowledge. Additionally, following the principles outlined in Chapter 4, we introduce a semi-automated solution that bridges the gap between experimentally measured band gaps and those simulated from models extracted in Chapter 4, thereby closing the loop initiated in Chapter 3.

5.1 Motivation and context

The complete structural and compositional characterisation of semiconductors at the atomic level is crucial for comprehending their functional properties in operational devices. Chapter 4 created a functional automated workflow for the comprehensive structural characterisation of devices. Importantly, its central focus was on

semiconductor devices devoted to quantum applications, as extensively reviewed in Chapter 3. Importantly, the proposed workflow ultimately outputs trustworthy 3D Atomic Models (3DAMs) of the experimental devices from which simulations can be extracted. Moreover, Chapter 3 justified that band gap is a fundamental property to know when addressing the functional performance of quantum devices, among manifold applications, of semiconductor platforms. This is why we demonstrate the simulation of band gap and Density Of States (DOS) out of the models extracted from Chapter 4 from devices in Chapter 3. Nevertheless, there is still a major scientific void in the correlation of these knowledge sources: the experimental measurement of the band gap. The intention is not just measuring it with energetically precise methods like photoluminescence, but doing it also with a high spatial resolution.

A high spatial resolution is key to understanding the spatial distribution of the band gap throughout devices. Thus, to unveil how local structural features, such as interfaces, defects, and strain, among many others, can affect their optoelectronic properties. [1, 2, 3] However, reaching sub-nanometer spatial resolution when accessing physical properties like the band gap is still an unresolved matter. As briefly justified in the Methodology, we make use of low-loss Electron Energy Loss Spectroscopy (EELS) to overcome this challenge. EELS allows us to exploit the spatial resolution of aberration-corrected STEM while still accessing the necessary energy resolution to gauge the band gap of most semiconductors. In this way, by combining the spatial and energy resolution records in STEM, 14 pm and 5 meV, we should be able to potentially translate it into fascinating correlations between imaging and spectroscopy. [4, 5, 6] The energy realm associated with band gaps, typically in the eV range, indeed offers exciting possibilities. Therefore, band gap measurements do not strongly constraint the experimental setup as ultra-high resolutions are not mandatory. For instance, determining the nominal band gap of Si, 1.1 eV does not require an overkill resolution of 5 meV. Instead, such precision could be more relevant for studying phonons. [7, 8] However, the real challenges in band gap mapping extend beyond spatial or energy resolution attainable by instruments. These challenges stem from the fundamental physics governing measurements involving fast electrons.

These challenges are known as Cherenkov radiation and delocalisation effects. [9, 10] On the one hand, Cherenkov radiation arises when the velocity of charged particles, v, exceeds the velocity of light in that medium (equation 5.1).

$$nv > c \tag{5.1}$$

where n is the refractive index of the material, v is the velocity of the fast electrons, and c is the velocity of light in vacuum. As it will be important for posterior relations between quantities, the refractive index n is the real part of the complex refractive

index, $N = n + ik$, which correlates with the complex dielectric function $\epsilon = \epsilon_1 + i\epsilon_2$ as follows:

$$\epsilon_1 = n^2 - k^2 \quad ; \quad \epsilon_2 = 2nk \tag{5.2}$$

Cherenkov radiation is the cause of the glowing blue light in nuclear reactors, due to the nuclear emission of charged particles, or equivalently for sound, in supersonic jets and their sonic explosion. In terms of electromagnetic waves, Cherenkov rises because the polarisation of the dielectric medium caused by the incoming electric field from the fast charged particle is retarded. Thus, the photon-emitting relaxation of the polarised particles in the medium cannot progressively recover in an ordered fashion from its polarised state, locally accumulating and causing a radiation emission which now constructively interferes. This is more likely to happen below and around interband transitions, where the materials are transparent and with a high refractive index. As materials become absorptive at larger spectral ranges, the smaller refractive index impedes satisfying this condition (equation 5.1). Therefore, we do not appreciate Cherenkov losses in core-loss experiments.

The problem arises when the Cherenkov radiation interacts with the electrons inducing them energy losses. These losses will be captured in EELS, hazardously masking the signals arising from meaningful physical interactions with the specimen and preventing the direct extraction of trustworthy conclusions from the data. In other words, if a given device presents Cherenkov radiation in the energy ranges where physical excitations like plasmons or the band gap itself are expected, we might be measuring the Cherenkov losses instead of these excitations. [11, 12, 13, 14] In addition, Cherenkov radiation wavefronts might scatter with the dielectric interfaces of the device forming the so-called guided modes. As their name indicates, these light modes are guided along the interfaces between materials and can act as localised peaks in the spectra. Importantly, this scattering is thoroughly dependent on the device morphology and its configuration. In summary, Cherenkov radiation and its interfacial scattering in the form of waveguide modes can hamper the experimental gauge of the band gap and any other intrinsic property of the device. Interestingly, specific and tailored experimental setups like off-axis EELS have minimised the presence of Cherenkov to measure band gaps by leveraging its limited angular distribution. [15] Moreover, leveraging the angular range by increasing the collection angle to reduce Cherenkov can only be applied if the loss function does not depend on the momentum transfer. [11] However, to date, no definitive answer could circumvent the presence of Cherenkov radiation in regular experimental conditions and in a regular and reproducible basis.

On the other hand, delocalisation means that the energy-loss signal recovered in a spatial position is affected by the beam interaction with neighbouring regions. Delocalisation does not affect core losses, facilitating the common atomic resolution

compositional maps nowadays. Nonetheless, it is an important issue in low-loss, and therefore for us to measure the band gap. In other words, locally measuring the band gap in a heterostructure might be everywhere influenced by both materials simultaneously.

In the present chapter, we propose a novel methodology which allows for assessing the influence of Cherenkov when performing an EELS experiment, and how to correct it when present. Moreover, we indirectly tackle the effect of delocalisation to give statistical validity to a resulting local band gap mapping of sub-nanometer spatial resolution. [10, 16]

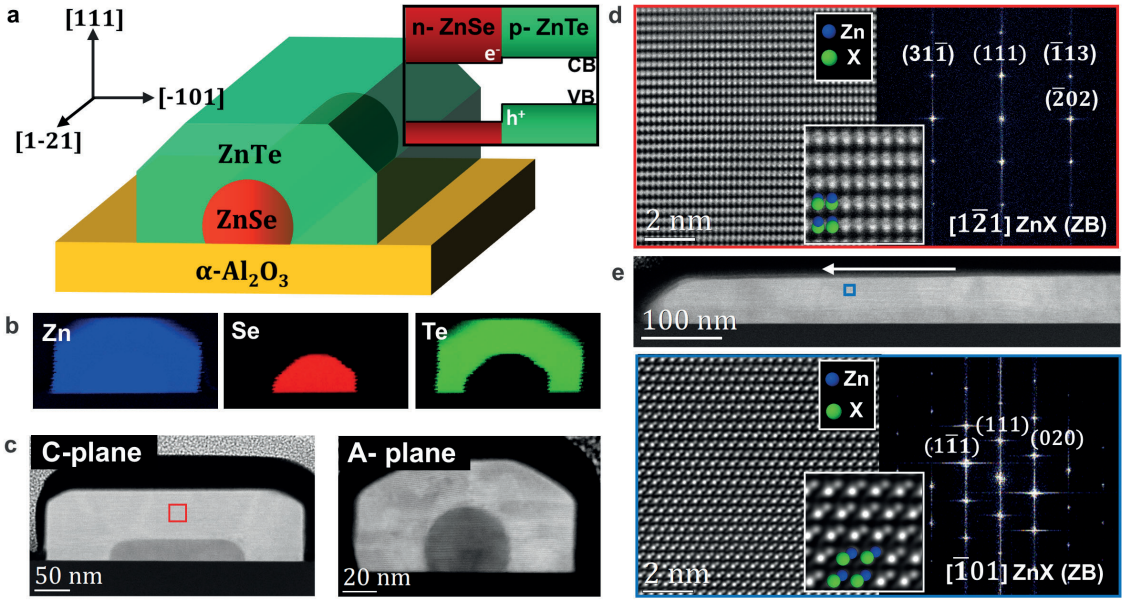


FIGURE 5.1: Structure and crystallographic orientations of the sapphire-supported core-shell nanowire (NW). (a) Schematics of NW configuration, ZnX crystal directions (X = Se/Te) and core-shell band alignment. (b) Electron Energy Loss Spectroscopy (EELS)-based elemental maps for Zn, Te and Se. (c) Low magnification High-Angle Annular Dark-Field - Scanning Transmission Electron Microscopy (HAADF-STEM) micrograph of NW transversal cross-sections on A-plane and C-plane sapphire. (d) Atomic resolution HAADF-STEM detail and its corresponding Fast Fourier Transform (FFT) of the transversal cross-section shown in (c). The inset reveals the atomic arrangement in more detail for precise atomic identification through the Z-contrast. (e) Low magnification HAADF-STEM micrograph of a longitudinal cut on a NW grown on C-plane sapphire, and higher magnification detail (inset) with its corresponding FFT of the area highlighted with a blue square.

Importantly, we developed the reported methodology with a device not seen in the thesis yet. Si and Ge have narrow bulk band gaps of 1.1 eV and 0.66 eV at 300K, respectively. Both represent indirect transitions. The small nominal value and its default indirect nature entail a challenging configuration to start building upon. It would require pushing the experimental conditions to an extent where the first proof of concept of our methodological idea would be difficult to validate. As a result, we develop the methodology for a system which allows loosening the experimental conditions to attain 1) enough energy resolution, 2) high pixel-wise Signal-to-Noise

Ratio (SNR) and 3) high spatial resolution to characterise key local effects within a complex device.

The selected system is an optoelectronic device, specifically a photodetector, with the band gap being its most significant and characteristic property. It is a ZnSe@ZnTe core-shell nanowire grown on α-Al_2O_3, sapphire, through the horizontal growth approach called guided growth (fig. 5.1). [16, 17, 18, 19] The nominal bulk direct band gaps of ZnSe and Znte, 2.67 [20] and 2.26 [21, 22] eV at 300K, stand as a trade-off between complexity and attainability. Both lie in the visible range, and they form a radial type-II p-n heterojunction (both valence and conduction bands are above the neighbour ones) when interfaced, which is highly advantageous for photovoltaic applications. [23, 24, 25] Let us briefly review the structural implications of the system, which make it interesting to deploy the methodology, and will provide the basis to later push it further to our ambitious central systems in Chapter 3.

5.1.1 Guided-grown ZnSe@ZnTe core-shell nanowires on α-Al_2O_3

The presented ZnSe@ZnTe core-shell NWs are based on a horizontal growth approach called guided growth. It is the natural response to the original Vapour-Liquid-Solid (VLS) growth mechanism of NWs, which generated vertical nanowires only contacting the substrate by its base. [26, 27, 28, 29, 30, 31] This translated into high-quality NWs where the strain was easily relaxed in that initial interface. Nevertheless, the primary challenge was their integration into wafers, circuits, NW networks and generally horizontal devices, which often proved to be a prohibitive, non-reproducible and non-scalable process. [32, 33] Consequently, alternative approaches that grew the devices directly horizontally and on their final substrate gained popularity. Examples of these are selected area growth or the guided growth which gave birth to our system of choice. [34, 35, 36, 37]

This new paradigm based on horizontally grown NWs demands a closer look at the strong interactions between the substrate and different materials along the whole length of the grown nanostructures. The high strain induced by the lattice mismatch at the interfaces of these horizontal NWs is not as easily released as in vertical ones. [38] The resulting strain profiles can substantially influence the interface properties of these materials and thus also their final emerging electronic, photonic or quantum performance. Therefore, to obtain desired material properties, it is important to monitor in detail the structure of these interfaces and correlate them to the local changes of physical properties.

In particular, crystal lattice deformations caused by strain have been exploited for tuning the electronic band alignment in such material heterostructures. [1, 2] These effects are accentuated in radial core-shell nanostructures compared with axial heterostructures resulting from the increase in the interaction area between the interfaced materials. [2, 38] Therefore, we would expect, along these large interfaces,

complex strain relief mechanisms, in which both plastic and elastic deformations occur simultaneously and take place at the nanoscale. [38, 39] All previous arguments illustrate the potential interest of the discussed system to correlate its structural features of interest (i.e., strain) with the potentially measured band gap. Remarkably, it holds the potential to be established as a reference or proof of concept in the band gap mapping of a system with physical significance, in this case, within optoelectronics.

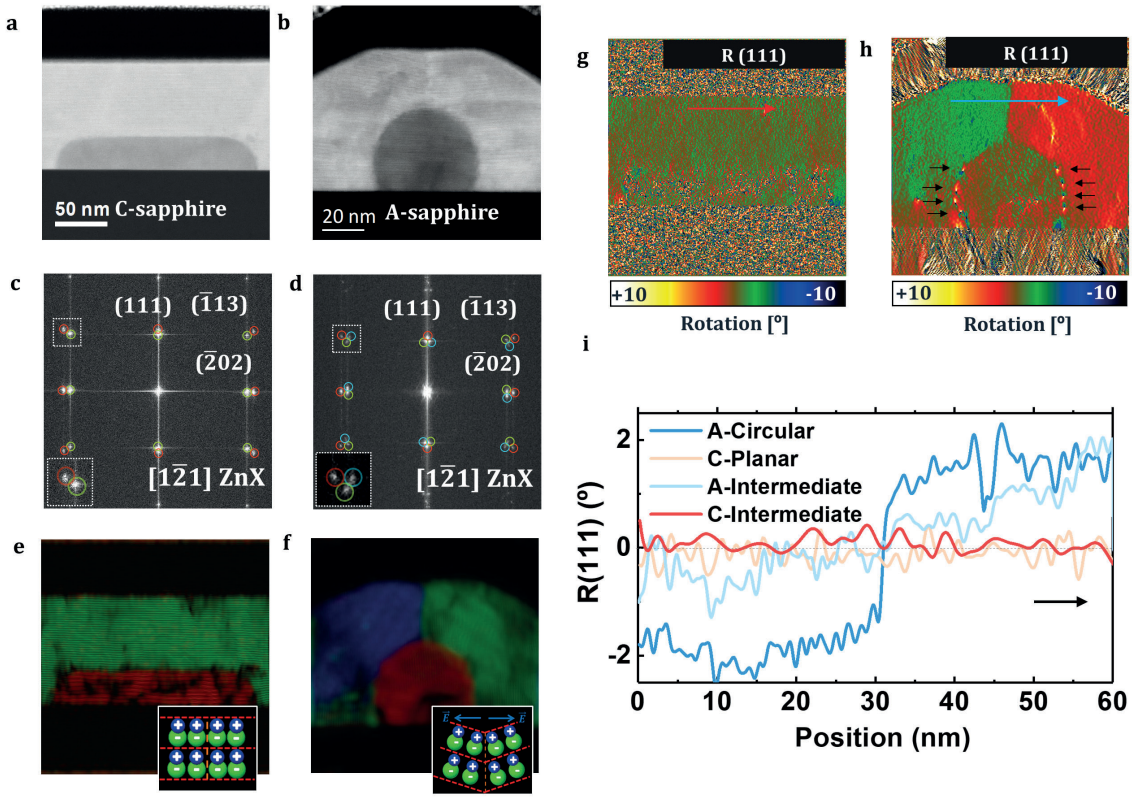


FIGURE 5.2: Core-shell ZnSe@ZnTe NanoWires (NWs) on C-plane sapphire (a,c,e,g)) and A-plane sapphire (b,d,f,h). (a,b) Low-magnification atomically resolved High-Angle Annular Dark-Field - Scanning Transmission Electron Microscopy (HAADF-STEM) image showing a representative NW cross-section. (c,d) The corresponding plane indexed Fast Fourier Transform (FFT) highlighting ZnSe core planes (red) and ZnTe shell planes (green/blue to differentiate orientations) in these structures. (e,f) Bragg-filtered RGB structural maps with colour codes as indicated in c and d. The insets schematically show dumbbell orientations in the central part of the ZnTe shell. (g,h) Geometrical Phase Analysis (GPA) (111) plane rotation maps relative to the NW core planes, which are taken as a reference of the horizontal core-shell epitaxy. Black arrows in h indicate dislocation positions. (i) GPA (111) plane rotation profiles across the centre of the NW shell, following the red and blue arrows direction in g and h, respectively for relevant core morphologies in both C-plane and A-plane substrates.

The complete structural characterisation of all the casuistic around this system is beyond the scope of the present chapter and of the thesis as a whole. However, let us pinpoint some crucial structural details to understand the system we analysed, which can be more extensively detailed in the literature. [16] These characteristics will be key to the discussion on the features that can affect the results of the band gap profiling through spectroscopy.

The sapphire α-Al_2O_3 substrates can arise from cutting it in two different orientations (A-plane (11-20) and C-plane (0001)) (fig. 5.2). Interestingly, the growth on different planes leads to changes in the morphology of the NWs and consequently in the way the strain is released when they come into contact. Figure 5.2 underscores the limits of the NW morphologies one can find when growing the ZnSe@ZnTe (both F-43m, space group number 216) core-shell nanowires on each plane. There are even differences in the specific non-equivalent growth directions within the same substrate plane, raising a broad range of morphology possibilities. The key takeaway is that the C-plane tends to produce nanowires with a flatter core, while the A-plane generates nanowires with a cylindrical core shape.

Interestingly, the growth occurs in two main stages. Firstly, the ZnSe core is grown by depositing a gold seed in the substrate or facet of choice. The interaction between the core and the substrate will define the morphology of that core. Importantly, this will be definitive for the overall device, as the growth and morphology of the shell will be mainly ruled by the initial growth of the core and its adaption with the shell's crystallographic structure. In addition, other effects such as the presence of atomic steps in the substrate may alter the whole combined core-shell morphology. In summary, the central strain source is therefore localised at the interface between the ZnSe core and the ZnTe shell.

Depending on their shape, a combination of elastic and plastic relaxation mechanisms will compete to relax the structure. For instance, in the rounded cylindrical ZnSe@ZnTe core-shell nanowires, a combination of rotations and misfit dislocations relax the shell. This can be understood as the way the ZnTe shell planes have to reduce the original high mismatch of 7.6 % with the perfectly horizontal planes of the core. Figure 5.2 showcases the differences in plane rotation between the shell planes when growing a flat and a cylindrical nanowire. Thus, we learn that the observed rotation is sideways $\pm 2^o$ to minimise the number of dislocations generated in the core-shell interface (fig. 5.2, d,f,h,i). Interestingly, the resulting effective mismatch from the rotated shell planes increases with the angle as the effective distance also increases by a factor of $cos\theta^{-1}$. As a result, flat wires prioritise the absence of rotation. However, radial cores promote the shell rotation because the curvature of the core creates local atomic steps in its horizontal planes which reduce the mismatch by raising the effective spacing of the core planes. The effectively larger interplanar spacing of the horizontal planes of the core now matches well with an effective distance of the shell planes which also increases due to rotation but is compensated by shortening the atomic steps. This idea is represented in figure 5.3, where atomic models of the rounded nanowire grown on A-sapphire (fig. 5.2, b) are engineered to geometrically compare a hypothetical 0^o core planes rotation and the experimental 2^o one. As stated, the GPA maps on the simulated STEM micrographs of the models allocate more dislocations and plane tension in the 0^o interface rather than in the

experimental 2º case. The attentive reader is referred to the bibliography for a detail that would blur the scope of the chapter. [16]

These strain mechanisms are rich enough to potentially induce modulations in the band gap, and consequently, on its photoluminescence spectra. Indeed, the photoluminescence spectrum of the core-shell system is broadened concerning the individual spectral peaks obtained from the isolated components of the wire. [17] This signature resembles a locally changing band gap. That is why checking the local distribution of the band gap across our optoelectronic device is key to directly correlating any visible band structure change with the reported strain mechanisms.

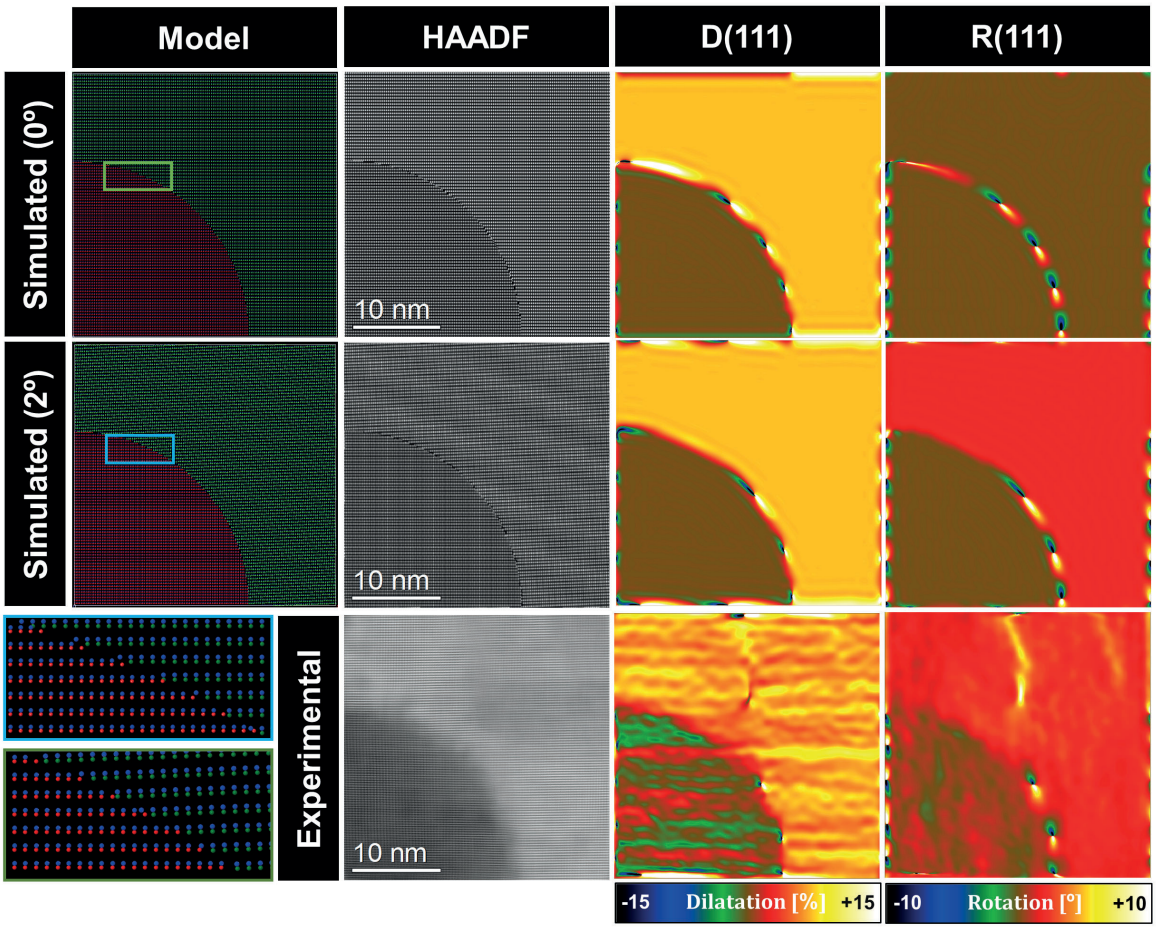


FIGURE 5.3: (Model column) Atomic modelling of a NW top-right section reproducing the experimental morphology (ZnTe presenting 2º rotation along [1-21] axis), along with a hypothetical configuration with ZnTe planes completely parallel to ZnSe ones (0º). (HAADF Column) High-Angle Annular Dark-Field - Scanning Transmission Electron Microscopy (HAADF-STEM) micrograph simulations of both cases and GPA dilatation (D(111)) and rotational (R(111)) maps of the (two upper rows) simulated models in comparison with (bottom row) the equivalent experimental analysis in the same NW section.

The particular system in which we developed the methodology and conducted the band gap maps is a case in-between the flat and cylindrical NWs in figure 5.2. Despite growing on the C-plane of sapphire, the core morphology is not fully flat. However, it does not reach the high curvature seen for the NW grown on A-plane. Contrarily, its curvature is a mixture of both, which interestingly presents curvature at the edges and a flatter interface on top. In this morphology, the optimal relative orientation between core and shell planes is 0°, which is especially beneficial for the top flat interface portion. In other words, the smaller curvature makes a rotation detrimental to strain release. Then, we expect dislocations to be concentrated at the interface as the main (plastic) strain relaxation mechanism. Dislocations like the ones seen in figures 5.2 and 5.3 must affect the band structure and the optoelectronic properties of this intermediate NW. Therefore, let us review the band gap mapping methodology we propose, as it indeed allowed us to identify noticeable features in the band gap profile of the core-shell interface.

5.2 Band gap measurements through Valence EELS

Low-loss electron energy-loss spectra contain information about the allowed optical and electronic transitions of the material, including collective excitations such as phonons and plasmons, as well as single-electron transitions. [7, 40, 41] Importantly, the low-energy region of the loss spectra, Valence EELS (VEELS), (i.e., 0-10 eV), reflects the Joint DOS (JDOS), whose dependence on the energy losses is $(E - E_g)^{1/2}$ in a direct band gap material, and $(E - E_g)^{3/2}$ in an indirect band gap material, where E is the energy lost by an electron inelastically scattered by the specimen, and E_g is the local band gap energy at the electron-sample interaction site. [42] Consequently, VEELS is a useful technique for band gap evaluation with the advantage of offering high spatial resolution. Nevertheless, as widely discussed before, the complexities of the required data analysis, in addition to the presence of parasitic effects such as Cherenkov radiation losses and delocalisation effects limit direct extraction of the property from the unprocessed spectra. As pointed out, these limitations can be dramatic in semiconductors and, generally, in high refractive index materials, below and around the band gap energy. [21, 22] As a result, we propose a pioneering way to identify whether these effects are interfering with our data, and if so, to overcome their limiting effect. To do so, it is pivotal to delve deep into how the theory explains the rising of the Cherenkov radiation when using fast electrons to extract properties from materials.

5.2.1 Low-loss EELS: theoretical background

Manifold theory levels can be used to explain the interaction of a fast electron with a homogeneous material. One can encode the information about the material in terms of its frequency-dependent dielectric function in classical electrodynamics. On the other hand, one can consider both the quantum nature of the electron beam and

of the excitations to gauge. The latter offers a more accurate formulation of the interaction, although it has been proved that it can be equivalent to semiclassical formalisms. However, it comes at the expense of a more complex calculation which in most cases would be an overkill and would just entail subtle differences with the classical approach. Consequently, we opt for the classical electrodynamics formalism, which gives us flexibility, accuracy and computational viability at once. [9]

According to this formalism, the electric field at position **r** inside a homogeneous material, produced by a swift electron travelling straight with constant velocity, **v**, can be extracted from Maxwell's equations and can be expressed as in equation 5.3. [9] Bold characters represent vectors.

$$\mathbf{E}(\mathbf{r},\omega) = \frac{ie}{\pi}\int d^3\mathbf{q}\frac{\mathbf{q}/\epsilon - k\mathbf{v}/c}{q^2 - k^2\epsilon}e^{i\mathbf{q}\cdot\mathbf{r}}\delta(\omega - \mathbf{q}\cdot\mathbf{v}) \tag{5.3}$$

where $k = \omega/c$ is the light wave number in free space, $\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega)$ is the complex dielectric function of the homogeneous material, $\mathbf{q}$ is the momentum transfer of the electron and e is the electron charge. Significantly, figure 5.4 represents the complex dielectric function of both ZnSe and ZnTe, where the real part, $\epsilon_1(\omega)$, informs about the polarisability and the way the velocity of light changes in the medium, while the imaginary part, $\epsilon_2(\omega)$, is correlated with the absorption of radiation by the medium. In general, the larger $\epsilon_1(\omega)$, the more retardation effects we expect as equation 5.1 is more easily fulfilled. On the other hand, as $\epsilon_2(\omega)$ refers to absorbance, the band gap can be gauged by inferring the spectral point where the curve becomes non-zero and starts growing. In other words, the band gap and the density of states are encoded in $\epsilon_2(\omega)$.

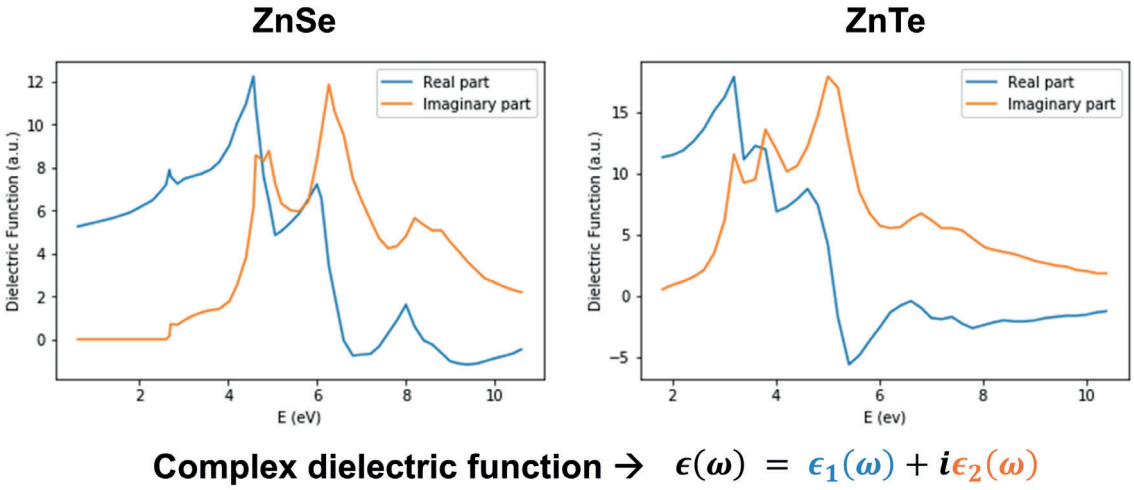


FIGURE 5.4: ZnSe and ZnTe complex dielectric functions. [43, 44]

This last expression is of utmost importance to start the analysis of how the electric field induced by the particle generates the losses back on the moving particle. In

fact, retardation effects can be seen already in this expression from its k terms. [9] The momentum transfer is separated into transversal and longitudinal components for convenience, as its term in the Dirac delta only contributes when aligning with the velocity vector along the beam axis. Eventually, the transversal momentum, expressed in polar terms for convenience, will be radially cut as when experimentally defining the spectrometer aperture.

We can now correlate the energy loss suffered by the same electron, which is the signal we measure experimentally in EELS. This loss can be linked to the force exerted by the induced electric field that acts back on the electron, along a straight line, $\mathbf{r} = \mathbf{r}_e(t)$. This is translated into equation 5.4. [9]

$$\Delta E = e \int dt \mathbf{v} \cdot \mathbf{E}^{ind}[\mathbf{r}_e(t), t] = \int_0^\infty \hbar\omega d\omega \Gamma_{EELS}(\omega) \tag{5.4}$$

where the negative charge is already considered in e, for which we get positive losses. The integrated distribution probability tells us the energy lost by the electron. Experimentally, this is equivalent to counting the number of electrons in the spectrometer channel corresponding to that energy loss. However, our interest is on $\Gamma_{EELS}(\omega)$, the loss probability, given in units of transferred frequency or energy $(1/\text{eV})$. This energy-dependent function will inform us of the expected EELS curve that can be directly compared with the experimental data (eq. 5.5). [9]

$$\Gamma_{EELS}(\omega) = \frac{e}{\pi\hbar\omega} \int dt Re\{e^{-i\omega t} \mathbf{v} \cdot \mathbf{E}^{ind}[\mathbf{r}_e(t), \omega]\} \tag{5.5}$$

With these considerations on the table, the problem boils down to understanding the electric field induced by the electron within the device. However, our expression for the induced electric field inside the material assumes homogeneity (eq. 5.3). Unfortunately, our theory is not yet sufficiently general to account for real heterogeneous materials with interfaces and complex morphologies. In other words, applying this model to devices similar to the ones analysed in Chapters 3, and 4, as well as our target ZnSe@ZnTe NW, would yield incomplete results. Let us keep this idea in mind, as it will serve as a cornerstone for our present analysis.

In any case, the homogeneous material approximation can be considered valid if we look at the material locally. In the presence of real interfaces, the mathematical limit would lie at an infinite distance from the interface. As a result, we can express the loss probability as a combination of a term considering the material as homogeneous, or bulk, and another one referring to the interfaces, which will modulate the first one. To compute the term regarding retarded bulk losses, which we will refer to as $\Gamma^R_{Bulk}(\omega)$, we combine equations 5.3 and 5.5. [9, 45]

$$\Gamma^R_{Bulk}(\omega) = \frac{e^2 L}{\pi\hbar v^2} Im\{(\frac{v^2}{c^2} - \frac{1}{\epsilon}) ln(\frac{q_c^2 - k^2\epsilon}{(\omega/v)^2 - k^2\epsilon})\} \tag{5.6}$$

where L is the thickness of the crossed material, arising when integrating along the crossed section. For us, this is the thickness of the lamella or grid. We assume that the dielectric function does not depend on the momentum transfer. We also assume that we cut the momentum transfer at a given point, which experimentally acts as if we cut the beam with an aperture, defining a collection semi-angle in the spectrometer, φ_{out} (assuming the convergence angle is significantly smaller than the collection one). Indeed, q_c accounts for this effect (eq. 5.7).

$$\hbar q_c \approx \sqrt{(m_e v \varphi_{out})^2 + (\hbar\omega/v)^2} \tag{5.7}$$

The expression derived from equation 5.6 encompasses losses arising directly from the dielectric response of the material, as well as the Cherenkov losses from the retarded response. This expression plays a crucial role in our methodology, allowing us to estimate the total EELS signal from a material without considering interactions with interfaces. However, we can extend our analysis beyond this expression by leveraging our understanding of Cherenkov losses. Equation 5.6 inherently accounts for retardation effects. Yet, we can explore a physical limit by setting $c \to \infty$ within the same equation. In this limit, particle velocities cannot exceed the speed of light in the medium. Consequently, neither Cherenkov losses nor guided modes would be expected. This exploration leads us to equation 5.8, which specifically quantifies energy losses arising solely from the material itself, devoid of retardation effects, namely, the non-retarded bulk losses $\Gamma^{NR}_{Bulk}(\omega)$.

$$\Gamma^{NR}_{Bulk}(\omega) = \frac{2e^2 L}{\pi\hbar v^2} Im\{(-\frac{1}{\epsilon})ln(q_c v/\omega)\} \tag{5.8}$$

Wrapping up, we now have two bulk losses, $\Gamma^{R}_{Bulk}(\omega)$ and $\Gamma^{NR}_{Bulk}(\omega)$. The first one includes the full bulk contribution and the second removes the presence of parasitic radiation. Nevertheless, the complete picture of the losses misses the term considering the surface effects of the induced electric field. Remarkably, huge efforts in the community provided analytical expressions of the surface term for different morphologies (sphere, cone, parallelepiped, among others). Nevertheless, these cases are limited to high symmetry configurations. To generalise for any morphology, heterostructure, or device complexity, we need to rely, as done in Chapters 3 and 4, in FEM. Specifically, we will use Boundary Element Methods (BEM) to encode our missing surface/interface component. [45] Our current knowledge of the bulk term justifies the sole encoding of the interfaces. As we miss the contribution of the surface, we assume the changes and the scattering of the Cherenknov radiation as guided modes only occur at the interfaces. Then, by solving the electric field in every boundary element, we can extract a term dependent on the material and the probe position. Importantly, this encoding is only valid if we assume translational invariance, as in the modelling from Chapter 4, part II. Therefore, we can encode

interfaces, either transversely or longitudinally to the beam as long as the symmetry allows translational invariance in this representation. Except for high symmetry conditions, we will be forced to encode transversal interfaces. In these cases, we compute the losses per unit of thickness length and eventually multiply over the thickness, L. Therefore, every term in the proposed formalism depends linearly on the thickness, making the total terms, the retarded and non-retarded, keep this dependency.

Thus, the full picture of the loss probability is embedded in equations 5.9 and 5.10, for the retarded and non-retarded case, respectively.

$$\Gamma^{R}_{Total}(\omega) = \Gamma^{R}_{Bulk}(\omega) + \Gamma_{Surf}(\omega) = \Gamma_{Total}(\omega) \tag{5.9}$$

$$\Gamma^{NR}_{Total}(\omega) = \Gamma^{NR}_{Bulk}(\omega) \tag{5.10}$$

The origin of the non-retarded scenario was neglecting every retardation effect, including the intrinsically retarded surface term. For the retarded case, the surface term, $\Gamma_{Surf}(\omega)$, influences the total signal, $\Gamma_{Total}(\omega)$, over the bulk term. In fact, an intuitive interpretation of its effect is a frequency-dependent addition or subtraction to the main bulk term (eq. 5.9). That is, the surface term can be negative as long as the total losses remain positive. [45] In practice, a negative term suggests that the bulk loss, calculated assuming the absence of interfaces, was not as significant as anticipated. This is because its contribution is obscured or overshadowed by the presence of an interface portrayed with BEM. This effect is known as *begrenzung* effect, or limit in German. [45, 46, 47] Furthermore, negative losses in the simulated $\Gamma_{Total}(\omega)$ are not catastrophic if they remain relatively small compared to the overall simulated intensity. Indeed, it might happen due to local numeric divergences especially when cutting the transversal momentum transfers (fig. 5.6).

Importantly enough, we disclosed a way to assess whether our EELS spectra are affected by Cherenkov and guided modes. By comparing the total retarded and non-retarded terms, we can gauge if parasitic radiation is present in the acquisition and if so, to which extent and from which contribution, surface or bulk. Therefore, at this point, we have enough ingredients to cook the recipe for overcoming the challenges of retardation. The next sections underscore the pivotal use of this information at our convenience.

5.2.2 VEELS band gap mapping methodology

The energy resolution of approximately 150 meV provided by the Nion UltraSTEM 'HERMES' (refer to Chapter 2), is characterized by a Zero-Loss Peak (ZLP) Full Width at Half Maximum (FWHM) of the same value and a dispersion of 50 meV/channel. This level of resolution results in a noisy, yet discernible pixel-wise

representation of the Local Density of States (LDOS) of the materials involved, which can be attributed to the performance of the energy filter and the intensity reduction caused by the monochromator. This energy resolution is far from the best achievable resolution. [6]. However, it offers a good trade-off between a resolution capable of resolving band gaps larger than 2 eV, and a high SNR in every pixel of the spectral image. Indeed, as we want to see the local changes of the band gap, we acquire a spectral image in which every probe position in space is a spectrum from which a band gap value can be retrieved. As a result, the pixel-wise extraction of the band gap would output a band gap map with a spatial resolution ruled by the pixel size.

The used operating accelerating voltage of the microscope was 60 kV. It is the most convenient choice as the energy resolution improves when decreasing the accelerating voltage, but still with a spatial resolution which allows atomic resolution. Thus, the energy spread of the beam is narrowed but the wavelength of the beam increases. In addition, we hinder the emergence of Cherenkov radiation, as the velocity of the particles in the medium reaches a smaller fraction of the velocity of light in vacuum (i.e., equation 5.1 becomes harder to fulfil). The counterpart is the necessity to perform the measurements in thin samples, ideally thinner than 1 t/λ. The log-ratio method in GMS was used to compute the relative thickness map. [48] The obtained map shows that, within the studied regions, the thickness computed is uniform (fig. 5.5). Therefore, the variations in the colour scale between materials, observed in figure 5.5 are attributed to differences in the inelastic mean free path, which is material dependent, and not thickness. Moreover, the absolute thickness computation, performed for the main materials ZnSe and ZnTe, revealed comparable average thicknesses, 44$\pm$ 3 nm and 46$\pm$ 1 nm, respectively. Then, we can say that the average thickness is around 45 nm, which exceeds the optimal value of 1 t/λ. Interestingly, we will claim we have achieved our goal despite not being in the exact sweet spot of every experimental condition.

The data analysis begins by extracting the spectrum image from the Gatan Microscopy Suite software to a format readable by Python. [49] A custom script is prepared to apply a generalized power law background model to the full spectrum image as a reliable tool for removing the ZLP. Moreover, each pixel is smoothed down by a Savitzky-Golay filter with a second-order polynomial, and the resulting data is extracted as tabbed text. Importantly, the data analysis now shifts to the Python programming environment to access the tools indicated in the Methodology chapter and the flexibility they grant.

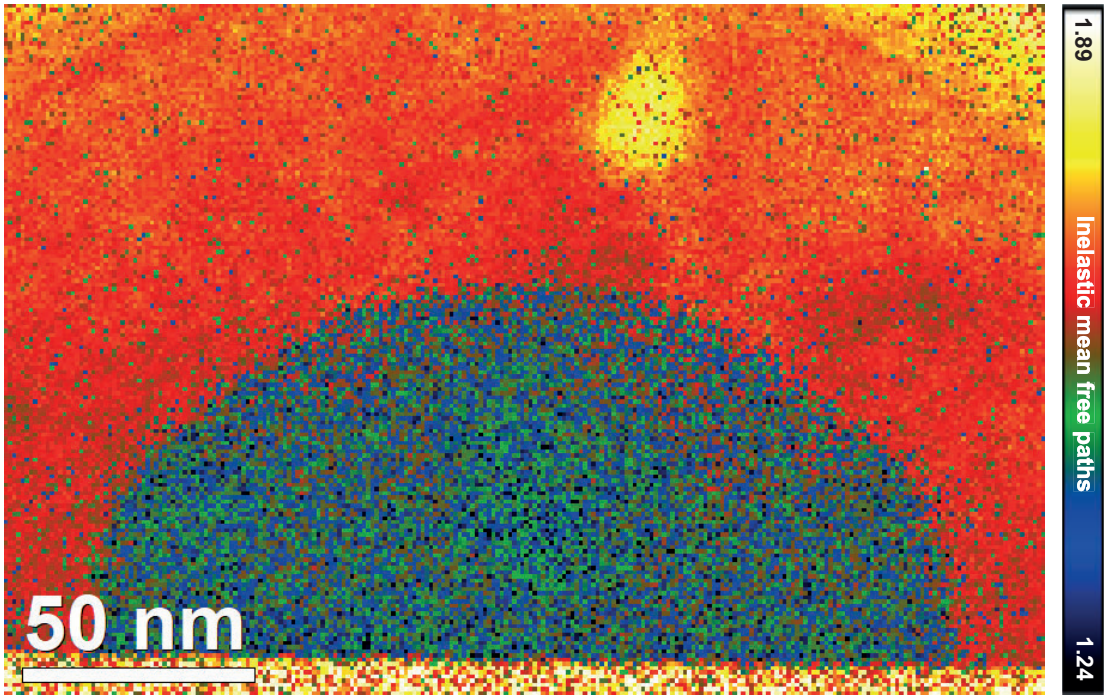


FIGURE 5.5: Relative thickness map of the ZnSe@ZnTe core-shell nanowire computed with Gatan Microscopy Suite's solution for t/λ estimation via log-ratio method, in units of inelastic mean free paths. [48]

We employed several approaches to distinguish the different regions of interest across the spatially resolved EELS data. Indeed, we aim for segmentation as in Chapter 4 (section 4.I.1), but through a completely different approach which exploits the added complexity and hidden knowledge present in the spectral image. In this case, our segmentation method allows us to independently identify each region of the device, adaptating to the proposed fitting routine detailed later. To achieve this, we used the main plasmon peaks of both ZnTe (centred at 15.3 eV) and ZnSe (centred at 16.9 eV), as well as the Al_2O_3 peak (centred at 25.4 eV). These plasmon peaks helped us to map the local distribution of these materials. By applying appropriate thresholds, we ensured that these regions aligned with the domains observed in the spectral image, mitigating the effects of plasmon delocalisation. Additionally, we employed unsupervised machine learning through PCA of the spectrum image to obtain its components as spatial maps. The main output of this analysis is a standalone map highlighting the interface pixels between ZnTe and ZnSe. Remarkably, the spatial distribution of these components agrees with the Middle-Angle ADF (MAADF) STEM image presented in figure 5.8, which highlights the regions containing strain accumulation, particularly near interfaces and volumetric defects. [50, 51] Furthermore, our approach fills any potential pixel-wise gaps in the ZnTe-ZnSe segmented interface after thresholding the plasmons.

The presented results rely on the synergistic effort between PCA and plasmon-based segmentation sources. Importantly, we end up accurately distinguishing every material in a pixel-wise manner. Interestingly, this necessity envisions a possible future refinement based on more advanced machine learning techniques, or complementary experimental setups involving compositional analysis. In fact, we discuss this aspect in the paragraphs below, as it serves as the cornerstone for partially automating the methodology. The ability to spatially identify pixels corresponding to each of the materials of interest is crucial. It enables tailored data processing based on the STEM-VEELS simulations of the device. Our discussion in section 5.2.1 becomes a pillar in the global methodology and an indispensable tool for achieving our goals. Now, let us revisit the simulations we performed to design the methodology, focusing specifically on the present system.

5.2.2.1 VEELS simulations

As detailed in section 5.2.1, the performed simulations are based on classical electrodynamics, with the materials described by their frequency-dependent, local dielectric functions. The energy-loss probability (Γ_{Surf}) is the sum of a bulk contribution, which is independent of geometry and proportional to the path travelled in each material, and a surface term determined by the interfaces and their distribution. Following this formalism, we performed both retarded and non-retarded calculations and compared them to assess the role of Cherenkov radiation emitted by the fast electrons in their interaction with the bulk materials.

Importantly, our TEM lamella can be modelled as a translationally invariant system, with encoded interfaces perpendicular to the electron beam. In other words, the way we encode these interfaces must faithfully mimic the experimental morphology observed when imaging the device. To achieve this, we consider both bulk contribution (Γ^{R}_{Bulk}), which corresponds to an electron in an infinitely extended crystal, and interface corrections (Γ_{Surf}). These corrections are determined by solving Maxwell's equations using the BEM described in section 5.2.1. [9, 45] The spatial resolution of the simulations is 0.3 and 0.5 nm between parametrisation points, aligning well with the spatial resolution of an experimental spectral image that does not reach atomic resolution and can only convolve the effects caused by dislocations. The BEM resolution varies across the simulated system, emphasizing key interfaces such as ZnSe-ZnTe. Notably, we do not account for quantum confinement effects due to the nanowire's relatively large radius (around 200 nm). While our simulations assume perfect crystals, complementary simulations have demonstrated the method's robustness against hypothetical strain and band gap changes smaller than ±1 eV (see section 5.2.2.4). The simulation results, depicted in figure 5.6, on the top, showcase paradigmatic examples of all four simulated components and their spatial origins within the simulated spectral image resembling the nanowire device. The graphs highlight the crucial role of retarded effects in the three materials. Specifically, the

ZnSe and ZnTe materials contribute from the bulk term, while Al_2O_3 derives its effects from the surface term. Consequently, directly retrieving the band gap is challenging, as it would involve measuring its convolution with the retardation effects.

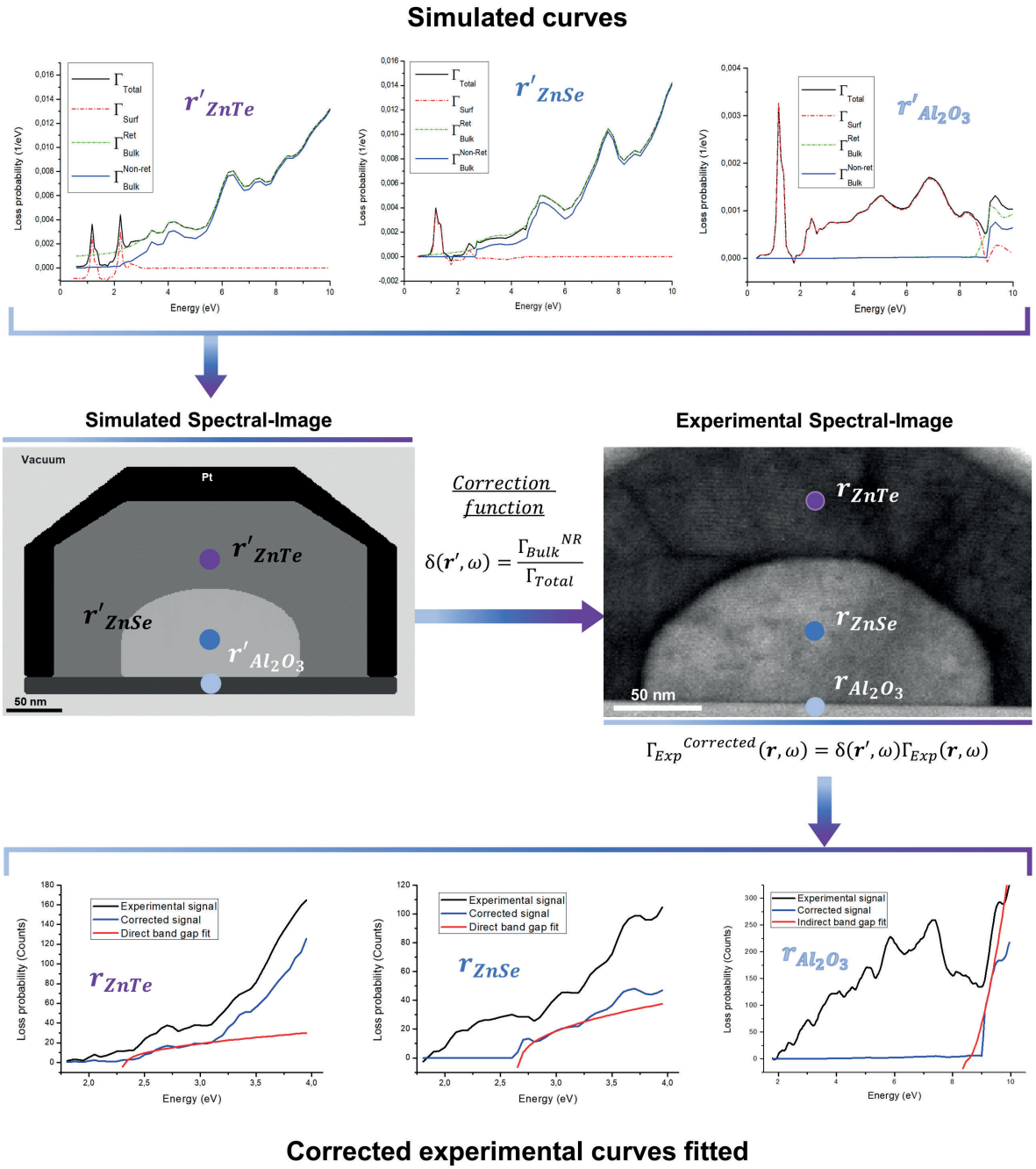


FIGURE 5.6: Workflow of the developed method for correcting experimental low-loss Electron Energy-Loss Spectroscopy (EELS) spectra affected by Cherenkov radiation and surface modes. Classical electrodynamics simulations are used to get a spectral image of both the non-retarded and retarded loss probabilities, which can be split into the bulk and surface contributions. Next, a correction function that compares the simulated non-retarded and retarded losses is used to map the effect of parasitic radiation and to clean the experimental spectral image. Finally, the corrected experimental signal is suitable to meaningfully fit either a direct or indirect band type depending on the material and its spatial position, and to extract a band gap value out of it. [16]

At this point, the use of the simulations becomes the critical step to get rid of these additional parasitic contributions to the energy loss spectrum, The following considerations have been taken into account:

1. Every pixel in the experimental spectrum-image has been labelled with the material from which it was acquired, thanks to the thresholded mapping of their corresponding plasmons.
2. In the simulated spectrum-image, every pixel is also labelled with its corresponding material. Additionally, each pixel has its own calculation for each of the following components: total signal (Γ_{Total}), surface term (Γ_{Surf}), retarded bulk term (Γ^{R}_{Bulk}) and non-retarded bulk term (Γ^{NR}_{Bulk}).

Furthermore, for each pixel in the simulated spectrum-image, a parameter called the "Correction ratio", denoted as $\delta(\mathbf{r'},\omega)$, is computed:

$$Correction\ ratio = \delta(\mathbf{r'},\omega) = \frac{\Gamma^{NR}_{Bulk}}{\Gamma_{Total}} \ ;\ where\ \ \Gamma_{Total} = \ \Gamma^{R}_{Bulk} + \Gamma_{Surf} \tag{5.11}$$

This ratio is always a value between 0 and 1, and it represents the amount of signal or loss probability that comes intrinsically from the material at a given energy, ω, and position of the simulated spectrum-image, $\mathbf{r'}$, (in other words, signal free of all kinds of additional contributions), compared to the total signal or loss probability that also includes Cherenkov radiation and waveguide modes arising from the interfaces.

Section 5.2.1 discussed the thickness dependence of the loss probabilities: the correction ratio divides terms that depend linearly on the thickness of the considered device. Then, the results in this level of theory do not depend on thickness. Nonetheless, as in reality it is a particularly important quantity disregarded by our model, we need to consider that a discrepancy in the experiment-theory collation might be caused by overlooking it. In other words, when a lamella is cut, the two perpendicular surfaces, one for entrance and one for exit, relative to the electron beam, facilitate the propagation of guided modes. Not surprisingly, the thinner the lamellae are, the more pronounced the effect is. This obvious example is, among others, an omitted contribution to solve in a future methodology upgrade.

5.2.2.2 Theory-based experimental spectrum image cleaning

Next, each pixel from the experimental spectrum-image, defined by its material and position $\mathbf{r}$, is correlated to its representative pixel of the same material in the simulated spectrum-image, defined by $\mathbf{r'}$. Equation 5.12 is then applied to get the experimental EELS signal with the accessible LDOS information, free of all the previously mentioned unwanted additional contributions, as seen in figure 5.6 and reinforced in figure 5.7.

$$\Gamma_{Exp}{}^{Corrected}(\mathbf{r},\omega) = \delta(\mathbf{r'},\omega)\Gamma_{Exp}(\mathbf{r},\omega) \tag{5.12}$$

With this approach, we can subtract the parasitic contributions from the experimental EELS data and perform meaningful fittings to extract physically representative information. The same idea applies to delocalisation: since we have local knowledge of the expected shape of losses without contributions from neighbouring regions, we can apply a correction factor. In the case of BEM simulations, this correction factor evaluates the localised signal in the numerator relative to a partially delocalised signal (only the surface term) in the denominator. This approach allows us to refine our understanding and enhance the accuracy of our results.

Nevertheless, a methodological challenge emerges when correlating **r'** with **r**. The difference in shape between the experimental spectrum image and the perfectly symmetrical simulation collides and hampers the direct one-to-one spatial correlation. Figure 5.7 showcases the solution we adopted. A relative position-finding algorithm is used to solve these morphological variations. This routine allows us to correlate the simulated and experimental positions so that the curves in the shell and core are correlated between both spectrum images. Conveniently, the routine separates the three main regions (i.e., 1) ZnSe core, 2) ZnTe shell, and 3) Al_2O_3 substrate) as given by the experimental segmentation explained above. Let us review how the case-wise separation facilitates the correlation and enhances the result:

1. For the core, we first find the relative horizontal position by normalizing the experimental position over the full length of the core. Given this relative position, the range of possible vertical values is found. For every horizontal position, there is a different range of possible vertical values, which reaches its maximum exactly in the central pixels of the core (where the core is the highest) and its minimum on the left and right edges of the core. This relative vertical position as a function of the horizontal position is translated to the simulated spectrum image in the same order: first, we find the equivalent horizontal pixels, and between these, pick the one that fits the relative vertical position to be the final result, **r'**.

2. The process in the shell is equivalent, but with minor changes to adapt the symmetry. First, the shell is divided into three regions: pixels on top of the core, on its right and on its left. Then, the same process of the core applies for each division but finding the relative vertical position first in this case, and then the horizontal ones. This division into several regions is made to ensure that the curvature is considered, as the vertical position finds the height at which to consider the radius, and the horizontal component computes the radius at that height. Despite finding a good agreement between both shapes, it can be seen in the final maps how the noisiest area corresponds to the shell pixels located

at, thinking in the core as a clock's top semicircle, 11 and 1 o'clock, where the deviation of the shapes is highest.

3. For the sapphire substrate, the process is simpler due to the simpler geometry. Comparing the rectangular sections and setting the dimensions to lie between 0 and 1 in both horizontal and vertical directions already allows us to find the proper equivalent position.

As said, the corrections apply independently for each independent region, all labelled by a single and univocal material defined by the segmentation. Thus, we ensure that the relative pixel counting starts with the real first pixel separating the interfaces. Nevertheless, the implementation of these correlations constitutes a particularly time-consuming and tedious step. Hence, it makes the process impractical in many scenarios where the geometry is irregular and difficult to model. This point of criticism will serve as a strong driving force to improve the methodology, particularly in terms of automation and accessibility to transform it into a tool for daily use.

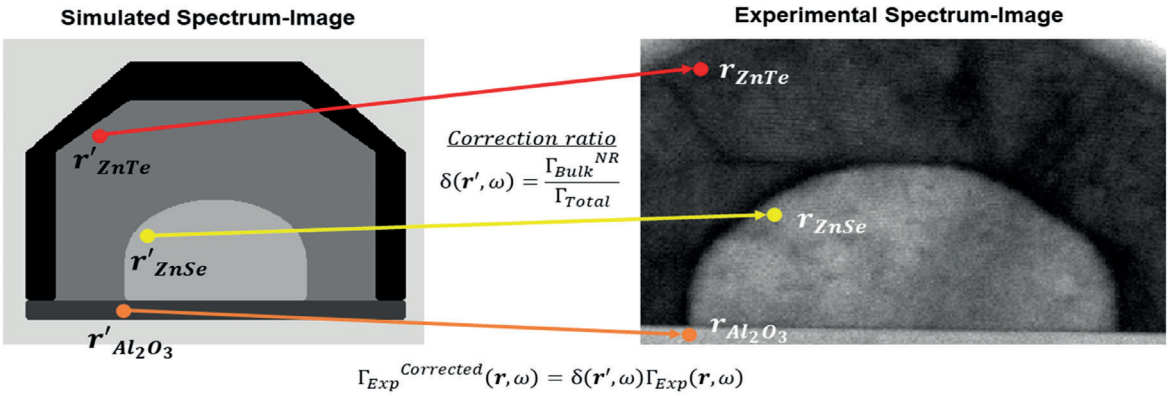


FIGURE 5.7: Correction and relative position finding algorithm. For each material, every simulated pixel is linked to its representative experimental pixel ($\mathbf{r'} \rightarrow \mathbf{r}$), and the correction ratio for that simulated pixel, $\mathbf{r'}$, is applied to the experimental data, $\mathbf{r}$, to obtain a new experimental spectrum-image free of Cherenkov and other additional losses.

5.2.2.3 Adaptive data fitting and resulting band gap mapping

Finally, the experimental data, free of interfering signal, is ready to be meaningfully analysed. Adaptive and material-dependent curve fittings are pixel-wise performed to distinguish the mathematical pattern, direct or indirect, that better represents the nature of the band structure to map. Remarkably, the JDOS curves become distinguishable thanks to the correction method, which despite removing a given percentage of the signal, allows to keep the nature of the acquired curves and to be sensitive to local band gap changes. Examples comparing the signal before and after the correction process can be seen in figure 5.6, but particularly in figure 5.9. These instances demonstrate key correction milestones, such as the effective removal of the surface loss probability in sapphire, of the waveguide mode immediately before

the expected ZnSe band gap energy, or the overall reduction in the ZnTe signal due to the high contribution of bulk Cherenkov losses. Indeed, the correction procedure could be refined by the deconvolution of the simulated curves with the experimental for a better adaption of the simulated features in the final corrected signal. This routine would complement the current percentage-based routine and could be of great interest if finer spectrum images, such as atomically resolved EELS, were analysed.

With the spectrum-image corrected in this fashion, the fitting of the functions that model the JDOS can be done reliably, as mentioned previously, by

- Fitting a $(E - E_g)^{1/2}$ function in the case of a direct band gap material.
- Fitting a $(E - E_g)^{3/2}$ function in the case of an indirect band gap material.

To do that, a fitting algorithm evaluates, for every pixel and its material, all the possible energy ranges within a range of values centred at each material's reported bulk band gap. This adaptive range is a trade-off between tractable computational cost and the lessening of outliers from unphysical fittings. A goodness-of-fit analysis is performed in order to identify the numerically optimal spectral position at which the JDOS represent the characteristic patterns of the band gap. We gauge the goodness of fit through its representative coefficient of determination, r^2, defined as:

$$r^2 = 1 - \frac{\sum_i^n (y_i - \hat{y}_i)^2}{\sum_i^n (y_i - \bar{y})^2} \tag{5.13}$$

where y_i denotes each dependent variable's value (i.e. electron counts) of the considered segment of data, $\hat{y}_i$ is the fitted value of the model at the same independent variable's value, and $\bar{y}$ is the mean value of the dependent variables of the considered segment of data. The fitting capacity of the algorithm has been proven by the evaluation of individual pixels over the full spectrum image. It extensively showcased a good ability to find the adequate curve, as can be seen in figure 5.9

The band gap error is estimated by the following considerations. Given a curve of the form $Counts = A(E - E_g)^n$, where n=1/2 or 3/2 depending on the band gap type, we fit a curve like $Counts^{1/n} = AE - AE_g$, to work with a linear dependence on the energy. This allows us to fit a linear polynomial $y = mx + c$, where $y = Counts$, $m = A$, $x = E$, and $c = AE_g$. The condition to extract the band gap energy from this fitting is that $y = 0$, allowing us to compute $E_g = -c/m$. The regression coefficients are computed with a regular fitting algorithm, and their respective standard errors are computed with the following equations:

$$s_m = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{(n-2)\sum_{i=1}^n (x_i - \bar{x})^2}} \tag{5.14}$$

$$s_c = \sqrt{\frac{\sum_{i=1}^{n} x_i{}^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}{n(n-2)\sum_{i=1}^{n} (x_i - \bar{x})^2}} \tag{5.15}$$

By propagating the errors, the statistical error of the retrieved band gap would be:

$$s_f = \sqrt{\sum_{i}^{l} \left(\frac{\partial f}{\partial x_i} s_i\right)^2} \rightarrow s_{E_g,\ stat} = \sqrt{\left(\frac{1}{m} s_c\right)^2 + \left(\frac{c}{m^2} s_m\right)^2} \tag{5.16}$$

Then, the final error associated with the measure of the band gap is the combination (i.e., square root of the sum of squares) of the experimental error, defined by the FWHM of 150 meV, and the statistical error, s_f, computed as before. We have observed that our statistical error is typically larger or similar in magnitude to the experimental error, for which the magnitude of the global error will resemble the FWHM experimental error, $s_f \simeq s_{E_g,stat} \simeq s_{E_g}$. It is important to note that these uncertainties refer to pixel-wise measurements. Thus, the standard deviation of the computed band gap value per pixel will refer to a deviation closer to the FWHM. However, the average values of each region are better descriptors of the tendency of the band gap in the specific areas. Thus, mapping several thousands of pixels per region during the spectrum image acquisition allows us to raise the statistical significance of the average values computed, as indicated in Appendix A.2. Generally, the standard error of the average band gap is a factor of $1/\sqrt{(n)}$, where n is the number of pixels in the specific region contributing to the band gap average.

The nature of performed fittings, based on the literature, is the following: direct band gap for ZnSe and ZnTe, and indirect for Al_2O_3 (fig. 5.8, e, f). [43, 44, 52] The strained ZnTe interface pixels (as identified by the MAADF-STEM – PCA correlated component thresholding analysis described above) have also been fitted with the indirect gap type function (fig. 5.8, a,b). Thus, we assume ZnTe could undergo a band-type transition under compressive strain, as reported in the literature with DFT calculations. [53] Significantly, statistical tests (i.e., $t \rightarrow z$ since n, the number of pixels, is always huge) have been performed by comparing the average band gap energy values per region. Therefore, we statistically validate the significance of the observed small changes between the band gap values of 'bulk' and interfacial pixels. Moreover, statistical tests have been carried out with the coefficients r^2 that discriminate the band gap type in the fitting to support the band gap type evaluation further.

The obtained average band gap energy values and the standard error of their mean for each region are as follows: 'bulk' ZnTe pixels 2.3162±0.0012 eV ; interface ZnTe pixels = 2.28±0.01 eV ; interface ZnSe pixels = 2.755±0.011 eV ; 'bulk' ZnSe = 2.7733±0.0012 eV ; Al_2O_3 = 8.271±0.016 eV (fig. 5.8, tab. 5.1). Interestingly, there is a slight decrease in the band gap values at the interfaces in both materials (order of

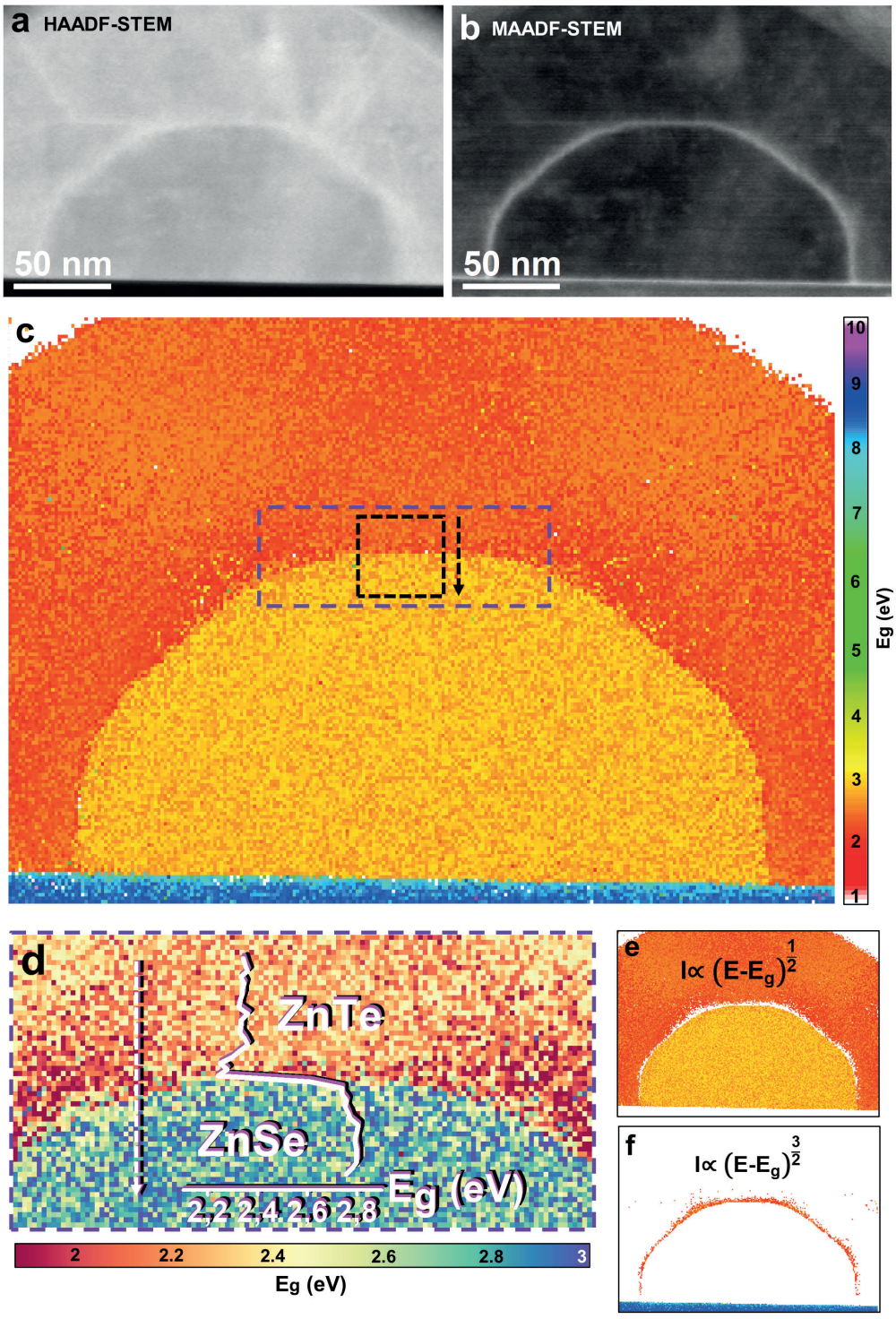


FIGURE 5.8: a) High-Angle Annular Dark-Field Scanning Transmission Electron Microscopy (HAADF-STEM) and b) Middle-AADF (MAADF-STEM) micrographs of a ZnSe-ZnTe core-shell nanowire grown on C-plane sapphire to which the band gap mapping methodology has been applied. c) Valence Electron Energy Loss Spectroscopy (VEELS)-band gap map of the system. The dotted purple square indicates the zoomed-in area displayed in d), in which the 2-3 eV scale allows for a better qualitative representation of the theoretically predicted decrease of the band gap in the strained ZnSe-ZnTe interface. This tendency is quantitatively observed in the black curve superimposed in d), which has been horizontally averaged through the pixels included in the black dotted box and along the black arrow direction in c). VEELS-band gap map decomposition into the regions in which we observe either e) a direct band structure (all ZnSe pixels and ZnTe 'bulk' pixels), or f) an indirect band gap (ZnTe shell pixels at the core interface, and sapphire pixels). The pixels corresponding to the metallic polycrystalline Pt protective layer have been manually set to 0 eV (metallic behaviour).

tens of meV) (fig. 5.8, d). It is still statistically significant given the large number of pixels ($\approx 10^3$ at the interface region and $> 210^4$ at the bulk region) used for their calculation. Remarkably, this agrees with the reported effects calculated from theory in the band structure of ZnTe (ZnSe) under compressive (tensile) strain. [53]

Furthermore, the indirect fitting performs significantly better than the direct one at the ZnTe interface pixels. As a control case, the same comparison of fitting coefficients is done with the bulk ZnTe pixels. Fitting an indirect gap in these pixels appears to result in better goodness of fit as well, compared to a direct gap fit. However, in this case, the increase in the r^2 parameter can be solely attributed to an unphysical fit of the indirect curve to the remains of the tails of the waveguide mode clearly visible around 2.3 eV. Additionally, it outputs a much smaller and more dispersed (2.1±0.3eV) result than the direct one. All the r^2 values, and the corresponding measurement errors, can be consulted in the Appendix A.2. The worse performance of the direct curve fit in purely statistical terms (i.e., r^2) can be explained by the strong effect of the aforementioned waveguide mode. Furthermore, the pre-correction basal bulk Cherenkov radiation also notably affects around these energies, which produce an uneven curve that hinders the fitting.

Region	**Band gap type**	**Mean±SEM**
Pt (protective)	Metallic	Set to 0 eV
Al_2O_3	Indirect	8.271±0.016 eV
ZnSe (bulk)	Direct	2.7733±0.0012 eV
ZnTe (bulk)	Direct	2.3162±0.0012 eV
ZnSe (interface)	Direct	2.755±0.011 eV
ZnTe (interface)	Indirect	2.28±0.01 eV

TABLE 5.1: Band gap average values with their Standard Error of the Mean (SME), for every distinctive region of the ZnSe@ZnTe core-shell nanowire. The nature of the best band gap type fitting the Electron Energy Loss Spectroscopy (EELS) curves is also specified. The underlining highlights the pixels corresponding to the strained ZnTe interface, where the band gap type transition is measured. [16]

It is crucial to note that waveguide modes manifesting as peaks present a significant challenge in identifying them in the experimental system compared to the simulated one. Differences such as the presence of various factors such as subtle geometry discrepancies, assumed translational invariance, local thickness variations, or delocalization, make the waveguide modes particularly challenging to identify. Hence the residuals of the tails of the most intense and shifted peaks can remain in the final corrected curves, although most of their contribution is erased by the corrections (fig. 5.9). For instance, the convolution of the tails of the waveguide mode centred around 2.2 eV, visible in the ZnTe domain as in figure 5.9, c and d, leads us to expect a higher fitting error, especially at the edges of the nanowire shell where the deviation between the simulated and experimental geometry is greater. Nonetheless, the

resulting direct fittings are more uniform and result in a standard deviation of almost half of that obtained from the indirect fitting. In addition, there is no apparent physical mechanism in our core-shell nanowire that would explain an almost 200 meV drop of the ZnTe reference bulk band gap. Similarly, the direct fitting applied to the interface ZnTe pixels (shown to be significantly worse than the indirect) leads to an average band gap value of 2.6$\pm$0.3 eV, which we believe is a too large shift given the expected strain.

Therefore, these arguments support the band type transition of ZnTe under compressive strain that has been reported theoretically. [53] It is crucial to note that the experimental pixel size is 0.9 nm. Moreover, the resolution of the BEM simulation (0.3-0.5 nm) aligns with these experimental conditions to not restrict it further. Consequently, we cannot resolve the atomic nature of the strain contained at the interface. In other words, the location of the dislocations and the distribution of the plane bending is lost in the spectral image. With figures 5.2 and 5.3 we demonstrated the presence of dislocations in similar systems, which extrapolate to the present symmetry. Consequently, we can assert that the average decrease in the band gap, along with the type transition, results from the cumulative impact of the global strain mechanism. Thus, it is not feasible to isolate individual contributions and attribute the effect solely to either the elastic or plastic components, especially, within the context of the demonstrated experimental conditions. To address the localised band gap behaviour near dislocations, an atomic-level approach with enhanced spatial resolution, complemented by validated channelling simulations, would be necessary.

Figure 5.9 presents some of the most common scenarios encountered regarding the global performance of the method. The presence of local negative signal (counts) can be spotted in individual pixels of the experimental spectrum image, but only anecdotally and following no apparent spatial pattern. Since the ZLP extraction follows the same power law model throughout the entire spectrum image, any artefacts causing a deviation from this general model in the corresponding spectral region can result in a locally misfitted ZLP. Nevertheless, by studying the full spectrum image we can confirm that this is not a common nor relevant issue. Only 1.59% of the total pixels, randomly distributed throughout the spectrum-image, had 30% of channels with negative counts between 2 and 3 eV. Moreover, the average of these negative counts only approaches, in absolute terms, 30% of the average of the remaining positive counts. Even when changing from 30% to 20% of channels in both considerations, which is much less harmful for the fitting, the value only rises to 4.69%. Moreover, this issue would require special caution in the percentage corresponding to ZnTe pixels, as the negative counts are typically centred around 2.17eV.

Figure 5.9, a-c, show individual spectra from the ZnTe shell. The first two represent a good enough starting point regarding the quality of the experimental signal, which, after correction, leads to a quite accurate fitted function in the spectral range where

the band gap lies (between 2 and 3 eV). Nevertheless, figure 5.9, c shows how the waveguide mode at around 2.3 eV can make the correction troublesome and induce a shifted fitting. The d-f panels in figure 5.9 are examples of ZnTe interface pixels. The main feature is the appearance of the indirect pattern close to the band gap energies, which is particularly clear in figure 5.9, d. Figure 5.9, f displays the difficulty of choosing a pixel-wise perfect ZLP removing tool, and, despite working well overall, in some isolated pixels the background removal might be too aggressive, while in some other cases, the removal may lead excessive residuals. In this case, this is translated into getting a higher-than-expected local band gap energy value. Figure 5.9, g-i presents core ZnSe pixels. Here, the spectrum presents the remains of the JDOS of the used reference dielectric function, visible as the abrupt increase of states around the ZnSe band gap energy. However, the fitting process still captures the essence of the experimental signal and can be adapted to the direct curve confirming this trend. Keep in mind that pre-band gap residuals might also come from Urbach energy instead, especially in the interfaces where the expected disorder is higher. [54, 55]

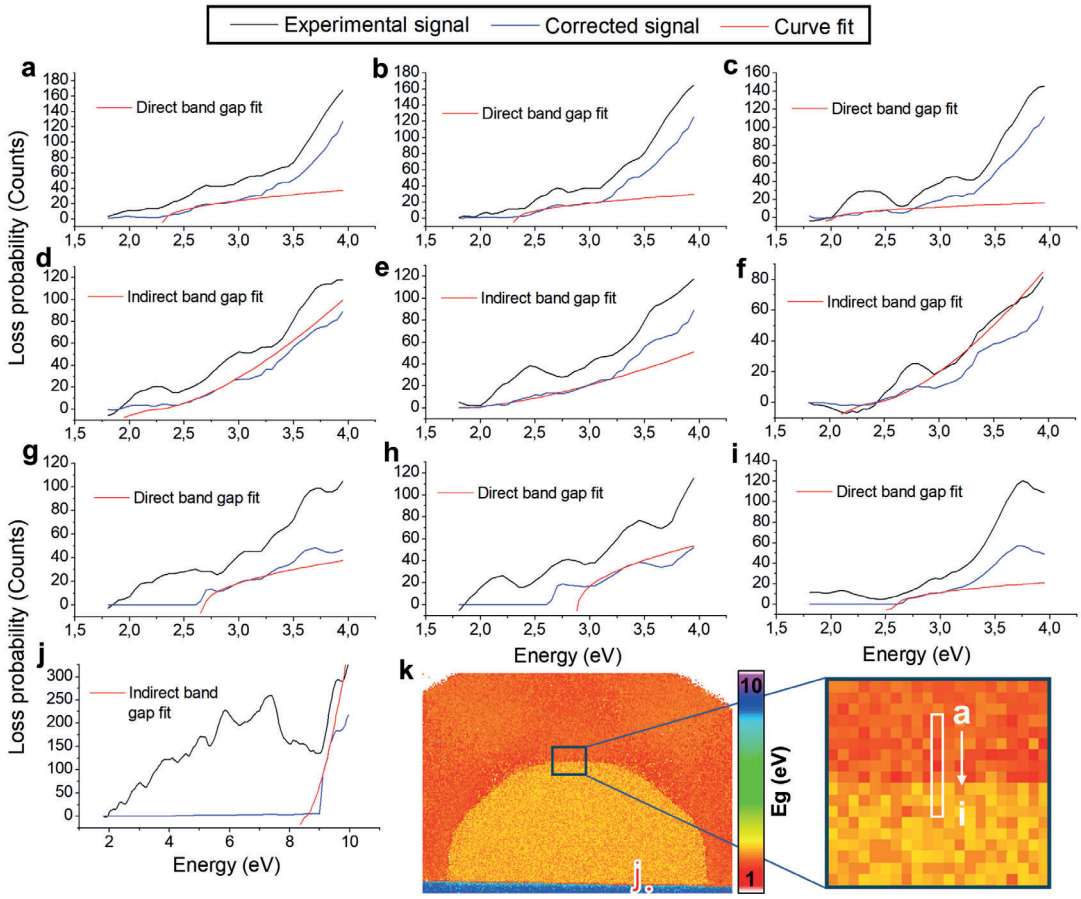


FIGURE 5.9: Pixel-wise analysis of the correction method through a single-pixel interfacial profile, indicated in the map and its zoom-in, k). a) to c) show ZnTe pixels and their direct curve fit; d) to f) show ZnTe interfacial pixels and their indirect curve fit; g) to i) show ZnSe pixels and their direct curve fit. j) shows an isolated sapphire pixel from its corresponding region, and its indirect curve fit. [16]

Finally, panel j belongs to the sapphire region. Here we see how the surface modes observed in panel d are completely removed. In addition, the curve indicating the real states becomes clear together with its indirect band structure fitting. Concerning curve fitting and the method's tolerance to changes in the band gap, the pixel-wise evaluation of the spectral ranges that were fitted revealed a compromise between overfitting and noise handling at 1 eV. This means that we expect that any band gap shift of ± 1 eV for tabulated values should be properly captured by the model. Indeed, this range would be enough to consider most of the phenomena that can induce band shifts. For more substantial increases, and particularly for larger decreases in the band gap, it would be challenging to guarantee the validity of the method. The limitation lies in keeping the tabulated value as the reference one. In order to generalise it and adapt it to larger variations, the simulations should also consider band modulation by strain (e.g., *ab initio*). [16]

5.2.2.4 Strain considerations

The regions where the biggest changes in the band gap are expected are the surroundings of the interfaces, where the strain is located. The strain release mechanisms in this sample are fundamentally based on the creation of misfit dislocations and bending of the shell structure, which makes the resulting spatial distribution of the elastic deformation very localised and only spanning a few nanometres (between 3 nm and 7 nm depending on the radial position along the interface). The nature of the BEM simulations used does not take into account the possibility that strain affects the system, mainly by modifying the dielectric function of the materials at the interfaces.

Although strain is not considered in the BEM simulations, an alternative experiment has been done to show that, given the expected strain, the used level of theory is enough to provide a trustworthy agreement with the experimental data. VEELS-STEM simulations have been performed to check whether an arbitrary variation in the band gap of ZnSe (ZnTe) due to strain, would affect the dielectric response (i.e., energy loss probability) of the ZnTe (ZnSe) in the first nanometres of the core-shell interface. For that purpose, the dielectric function of one of the materials is modified by making the material transparent to radiation at energies below approximately 5 eV (cancelling the imaginary part of the dielectric function), which simulates an enormous band gap deviation (increase) from the original conditions of the bulk material. While doing this, the other materials' properties are kept unchanged. These two variations capture all the possible scenarios of a hypothetical increase of either ZnTe or ZnSe band gap from its bulk reference value up to reach a value of 5 eV.

This is repeated by modifying the material left unchanged and keeping the original dielectric function of the first material. With this change and keeping the geometries and materials distribution exactly as in the main simulation, two new simulated data sets are obtained. The unmodified original simulation (the one used for the

corrections) can now be compared pairwise with each of the two new simulated data sets to map the deviations that the deliberate/artificial change in the band gap induced. The new simulated data is compared with the one obtained by the unmodified dielectric functions by performing a χ^2 analysis of the spectra in those pixels that surround the interface:

$$\chi^2\left(\mathbf{r'},\omega\right) = \frac{\left(\Gamma_{T,R}\left(\mathbf{r'},\omega\right) - \Gamma_{T,M}\left(\mathbf{r'},\omega\right)\right)^2}{\Gamma_{T,R}\left(\mathbf{r'},\omega\right)} \tag{5.17}$$

where $\Gamma_{T,R}(\mathbf{r'},\omega)$ is the simulated total EELS signal with the unmodified/original dielectric functions at a certain position $\mathbf{r'}$ and energy ω; and $\Gamma_{T,M}(\mathbf{r'},\omega)$ is the simulated total EELS signal considering the modified dielectric function for one of the materials.

The results can be observed in figure 5.10. We have evaluated the effect that a change of the band gap energy in the shell would have at the core region near the interface and vice versa, by analysing vertical profiles crossing the core-shell interface. Figure 5.10,a evaluates the consequences of an increase in the band gap energy of the ZnSe core to the dielectric response of the nearby ZnTe shell. The null χ^2 curve at the targeted energies between 2 and 3 eV illustrates that no additional contributions to the low loss response are present when this hypothetical variation occurs. This means that the standard dielectric response would be the same as with a hypothetical increase in the band gap energy of ZnSe.

On the other hand, figure 5.10,b evaluates the influence of an increase in the band gap energy of the ZnTe shell to the dielectric response of the ZnSe core region closer to the interface. In this case, the χ^2 curve is not zero through the spectral region of interest, meaning both the standard and modified responses would differ. Therefore, we will consider in our models that a hypothetic increase in ZnTe's band gap would induce the appearance of an extra peak around 2.4 eV in the core. This would happen because this peak corresponds to a waveguide mode in the shell, which is strongly absorbed by the ZnTe in the real band gap scenario due to a nominal value of 2.26 eV (with the experimentally obtained average being 2.3162 ± 0.0012 eV). Nonetheless, the forced band gap increase allows the radiation that was absorbed at the energies below the new band gap to propagate through the rest of the heterostructure, therefore inducing an extra contribution to the core. In fact, the radiation coming from the shell to the core converges in the centre of the core from all sides of the shell. As a result, it raises the signal the closer we get to the centre of the core (i.e., intensity increase from curves 2 to 4 in fig. 5.10,b). Despite this result, we do not expect or observe any increase in the ZnTe band gap, but rather a decrease due to compressive strain at the interface. This means that the 2.4 eV extra peak would be absorbed by the ZnTe itself without causing any undesired perturbation. Importantly, this supports the validity of the obtained simulations and corrections.

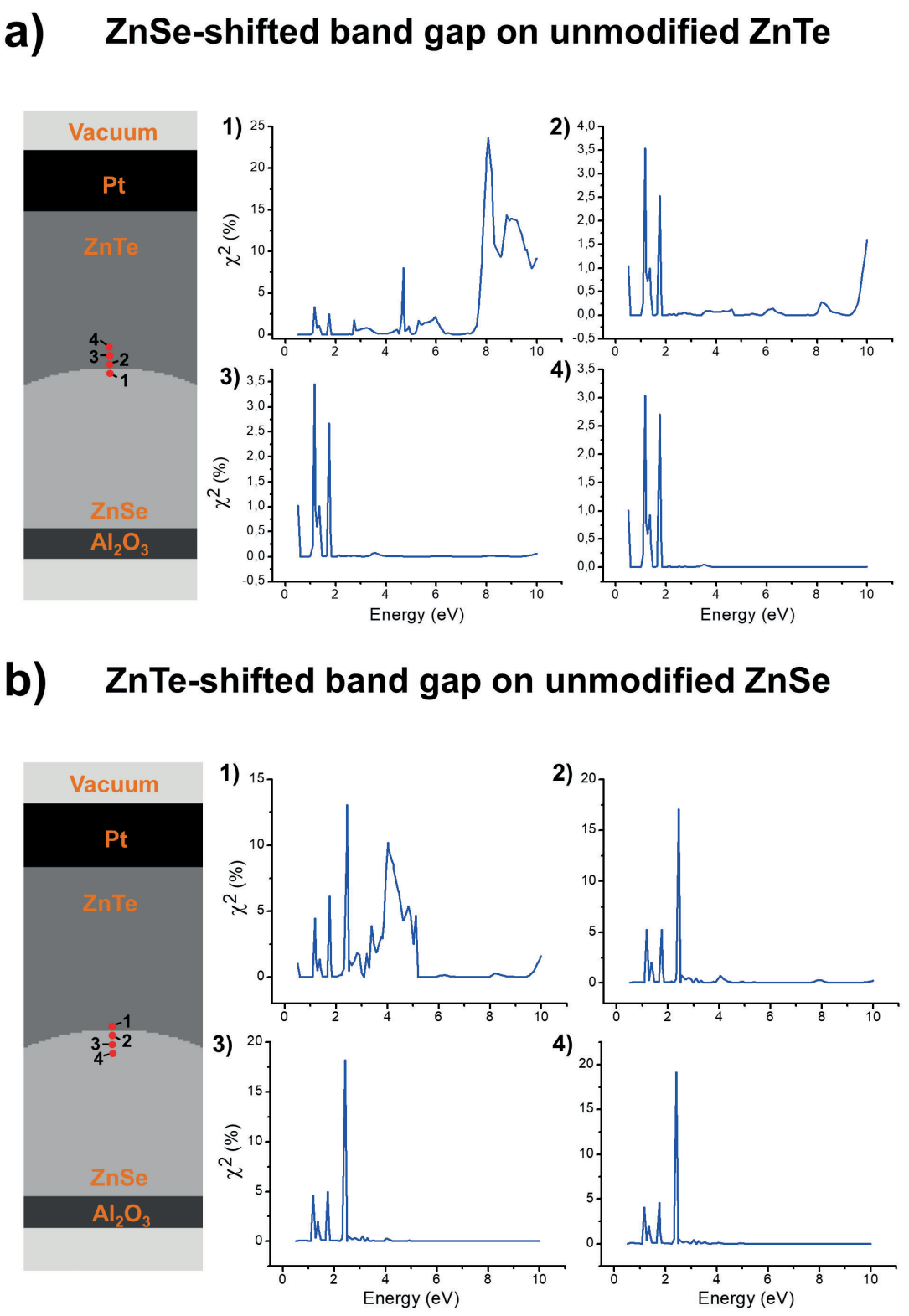


FIGURE 5.10: a) χ^2 tests between unaltered simulations and ZnSe-shifted band gap simulations to check the influence of ZnSe's band gap modification on ZnTe's dielectric response. Four pixels are considered for this evaluation: the closest-to-the-interface ZnSe pixel (1), and the three ZnTe closest ones (2, 3 and 4). It shows that ZnTe would experience no dielectric deviation in the spectral region of interest (>2 eV) due to a hypothetical increase of the band gap of the interfacial ZnSe. b) χ^2 tests between unaltered simulations and ZnTe-shifted band gap simulations to check the influence of ZnTe's band gap modification on ZnSe's dielectric response. Four pixels are considered for this evaluation: the closest-to-the-interface ZnTe pixel (1), and the three ZnSe closest ones (2, 3 and 4). An important peak appears around 2.4 eV, meaning that a variation (increase) of ZnTe's band gap would modify the loss probability of ZnSe in this specific energy region.

5.3 Automation of the methodology

Throughout the text, we have discussed certain aspects of the developed methodology that are specifically customised for the targeted core-shell nanowire. While some essential routines, such as the positional correlation between simulations and experiments, could apply to core-shell nanowires with similar shapes, any significant morphological variation would require fine-tuning or, in most cases, a complete adaptation of the simulation-experiment correlation routine. The first successful proof of concept for this method required scientific and methodological development, and the tailoring did not imply a major disadvantage. However, the broader scientific community would greatly benefit from further improving the methodology and automating it. Hence, in this section, we delve into the details of automating the previously explained band gap mapping methodology.

The proposed automation process follows a similar logic to the workflow outlined in Chapter 4. Both approaches involve an initial segmentation step, rely on simulations based on an analytical physical model, and utilise finite elements as a fundamental theory source. Notably, the knowledge gained from Chapter 4 plays a pivotal role in shaping the proposed automation. As a result, the fusion and cross-fertilisation of these two methods naturally give rise to the pipeline discussed in the following sections.

Before implementing any automation, we conduct manual assessments of both the ZLP removal and denoising steps. Specifically, we adhere to a power-law-based approach for ZLP removal and employ PCA for denoising. [49, 56] The ZLP removal is particularly critical for subsequent processing, warranting meticulous manual evaluation. Similarly, PCA denoising requires careful attention. Although automatic optimisation of the number of components to limit the reconstruction remains challenging, [57, 58] automating these two steps would significantly expedite the process and enhance the reproducibility. Notably, the general ZLP extraction recipe remains non-robust and lacks generality, despite community efforts. [58, 59]

The initial step in our automated pipeline focuses on segmenting the spectral image. Despite the similar starting point, this represents a significant departure from Chapter 4, where spectral images served merely as complementary information rather than the primary source. Leveraging the additional spectral dimension, we fully exploit it for sample segmentation. Unlike relying on HAADF STEM images or specific imaging intensity, we treat the spectral data as if we were segmenting a dataset of 2D images corresponding to each spectrometer channel. This added complexity enables us to employ more sophisticated algorithms, including unmixing routines described in the Introduction. Specifically, we tested NMF, ICA and VCA as unsupervised methods to extract spectral endmembers representative of specific regions. We then compare these endmembers across the entire spectral image to identify regions of maximum coincidence (fig. 5.11, step 1). [60, 61, 62, 63, 64] These identified

regions will serve as the independent segments of the map, effectively discretising and separating areas that exhibit distinct spectral characteristics.

Keep in mind that NMF, ICA and VCA are all unsupervised machine learning methods. Consequently, we will need a certain degree of manual intervention for hyperparameter tuning to achieve meaningful results. The first manual tuning significantly impacts the subsequent pipeline: the number of segments. These defined segments will serve as the basis for subsequent manual inputs, which correlate with the corresponding materials. Since no compositional knowledge is injected into the workflow, this information must be manually provided. Our prior understanding of the sample becomes crucial in selecting the appropriate number of segments that represent distinct regions within the device. Interestingly, starting with a prior characterisation using the automated workflow proposed in Chapter 4 would enhance the process. Additionally, even a manual core-loss-based analysis could offer further insights into the sample's composition. This step is pivotal because it profoundly influences the choice of the unmixing method.

In our automated pipeline, the choice between Independent Component Analysis (ICA) and Vertex Component Analysis (VCA) hinges on the expected statistical independence between segments. ICA is preferred when we anticipate that certain excitations or edges from one region will not appear in others or remain unaffected by neighbouring regions. For example, edges dependent on the chemical environment might exhibit this behaviour. The independence criteria in ICA facilitate the extraction of components directly linked to single independent materials. On the other hand, VCA works well when we have prior knowledge about the number of spectrally distinct regions. Unlike ICA, where endmembers represent linear combinations, VCA's endmembers correspond to real spectra in the spectral image. Therefore, VCA can effectively separate regions even if they are statistically dependent. In the provided example, we utilized VCA because the areas we aim to distinguish share common elements (i.e., Si). This would make the independence criteria struggle to output more regions than purely statistical methods. Remarkably, VCA can even distinguish similar spectral patterns with subtle shifts, such as the 1 eV shift in the plasmon peaks of Si and Ge (fig. 5.11, step 2). [58, 65]

After identifying the endmembers, the unmixing process allows us to map the abundance of each extracted component (fig. 5.11, step 2). Specifically, every pixel in the spectral image is unmixed into a linear combination of the endmembers. The segmentation then groups pixels based on the relative weight of the corresponding endmember. The outcome is a coefficient map ranging from 0 to 1. For instance, when using VCA, where endmembers are derived from actual pixels, we find at least as many pixels as endmembers, each with a weight of 1 corresponding to themselves. These abundance maps are subsequently thresholded on an endmember-wise basis to discretise their respective abundance distributions. While manual individual thresholding is an option, a robust general choice is to set the threshold between

25 and 30%. In other words, the pixels are kept if the abundance intensity is larger than 0.25 or 0.3. This effectively removes pixels where the endmember representation falls below that percentage. Next, we stack the segmented regions sequentially, prioritizing the most significant ones towards the end. This approach ensures that the key areas are well represented, especially if the threshold overlaps with edges. Despite its established robustness, this flexible threshold parameter facilitates adaptability. In cases where missing edges appear, we apply the same completion routine used during the intensity-based segmentation in Chapter 4 (section 4.I.1). A manual definition of this ordering is also necessary.

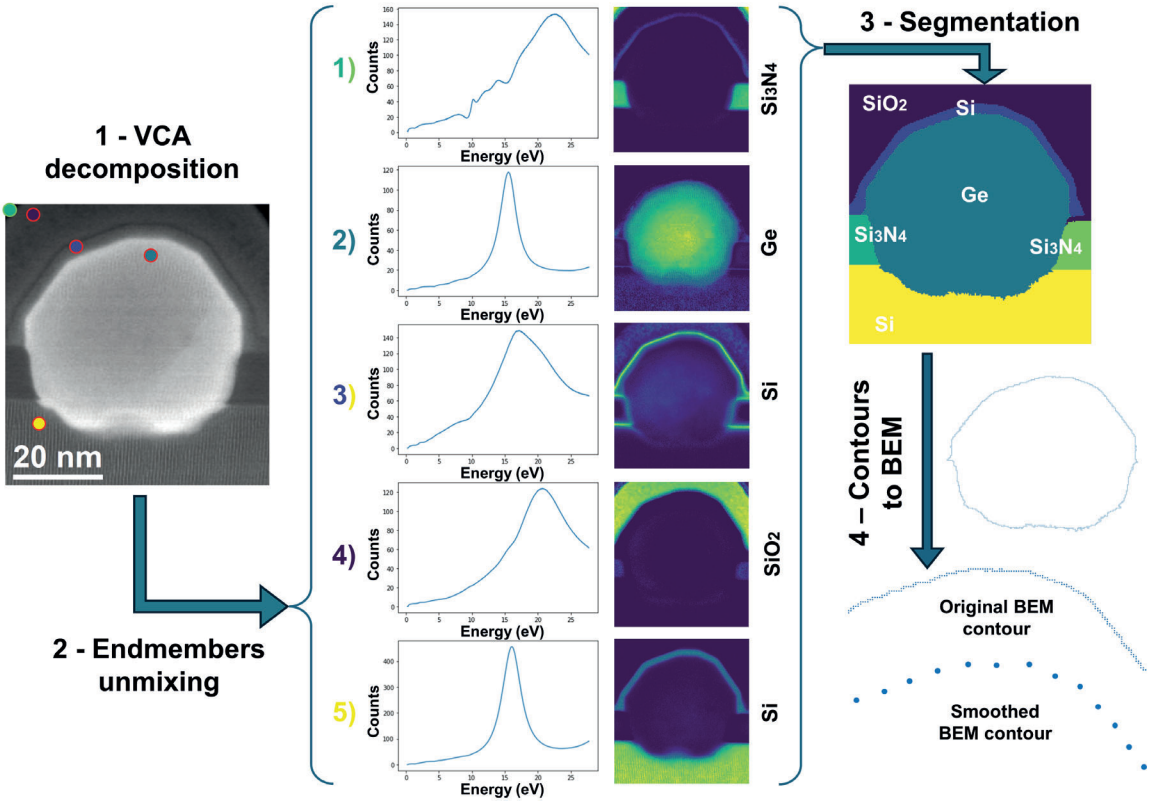


FIGURE 5.11: Semi-automated workflow for band gap mapping through Valence Electron Energy Loss Spectroscopy (VEELS). 1) Unsupervised machine learning methods like Vertex Component Analysis (VCA) are used to estimate the endmembers representing the experimental data, which are 2) later unmixed to distil their spectral abundance. 3) The thresholded maps are combined to segment the image and 4) to create a smoothed Boundary Element Method (BEM) representation from its contours.

Figure 5.11, step 2, displays a particular unmixing detail. The showcased third endmember is a spectral vertex mapping the SiX-Si interface. No additional distinctive material should be found beyond the four main ones, although a different stoichiometry of SiO_x, with higher oxygen content is expected immediately between these interfaces. Indeed, this endmember arises from the combination of the spectral response of both contacting regions. In fact, figure 5.11 indicates that their plasmons are spectral neighbours, which combined with their delocalised nature originates

a convolution of the signals from endmembers 4 and 5, understood by VCA as an independent vertex elucidating this nuance about the increased oxygen content. [66]

To finalise the segmentation process, a denoising routine ensures that the contour followed by each segment remains unique. The algorithm examines the pixel-wise anticlockwise path, and if any bifurcations are detected, they are unified by removing the pixels along the edges (fig. 5.11, step 3). Although seemingly minor, this detail is crucial to ensure the success of the subsequent automation steps.

Once our segmented image is prepared, the next step involves establishing a formal association between each region and a specific material or crystal phase. We achieve this by assigning an integer to each segment and linking it to the corresponding material along with its dielectric function. Using equations 5.6 and 5.8 we can then map the retarded and non-retarded bulk losses (fig. 5.11, step 3). However, the real breakthrough occurs when we seamlessly integrate the BEM simulation, necessary for extracting the surface term, (Γ_{Surf}), directly into the spectral-based segmentation process. This innovative approach draws inspiration from the FEM engineering principles discussed in Chapter 4, section 4.II.2.

In Chapter 4, we validated that FEM representations of segmented images can be effectively used for large-scale simulations of key properties in (often massive) devices. Retrieving a segmentation from the spectral image also facilitates the creation of boundary elements necessary for calculating the surface term. In our previous approach in section 4.II.2, we required manual steps to mesh the interior of the encoded region and specify simulation conditions. However, the present BEM simulations solely rely on boundary elements, making any manual refinement unnecessary. Following the approach detailed in figure 4.II.1 from Chapter 4 and defining standard simulation parameters (such as accelerating voltage), we can automatically generate a file directly usable as raw input for BEM software, which computes the surface term, (Γ_{Surf}). Importantly, we leverage the contour denoising not only to achieve smoother boundaries but also to control the resolution and accuracy of the simulation (fig. 5.11, step 4). By adjusting the smoothing process, we can eliminate vertices more aggressively, resulting in larger BEM spacing but potentially poorer resolution. Since resolution directly impacts computing time, our choice indirectly allocates computational resources for the calculation. The BEM simulation represents the most time-consuming step in the process, sometimes taking weeks even with consciously minimised resolution. Notably, the BEM simulation for the ZnSe@ZnTe core-shell NW took nearly a month on our workstation (for detailed information, refer to the Methodology chapter). Therefore, gaining better control over this crucial parameter constitutes a significant improvement in our approach.

In the next step, we material and region-wise add the retarded bulk term based on the segmentation. The outcome is a full spectrum image for each of the four loss

contributions, including the total loss, which we can then compare with the experimental data. Remarkably, this step becomes pivotal because the comparison - specifically, computing the correction factor (eq. 5.11) to subsequently correct and clean the experimental data (eq. 5.12) - can be achieved through a one-to-one correlation between the simulated and experimental pixels. In other words, the painstaking and highly customized algorithm we previously coded for pixel correlation in the ZnSe@ZnTe NW (as described in section 5.2.2.2), is now simplified to an immediate and universally applicable direct matrix multiplication.

Eventually, by applying the pixel-wise data correction method to the experimental spectrum image, we obtain a signal free of retardation effects. This allows us to proceed with the final fitting step, as discussed in section 5.2.2.3. The process aligns seamlessly, as we define the fitting range, minimal energy range, and the nature of the band gap transition to map. Mathematically, it follows the same adaptive fitting routine used for the ZnSe@ZnTe NW. Although the computation may run for more than a day to output the band gap maps, standard errors, and goodness-of-fit coefficients, parallelisation is feasible. However, the primary computational bottleneck remains the BEM simulation, a challenge we are addressing by managing the smoothing process.

The semi-automated approach is still limited and currently far from the automation and complexity levels reached in Chapter 4. Nevertheless, its potential is clear to serve as the foundation for the fully automated solution. Indeed, it could already provide preliminary results in a challenging system already tackled in the previous chapters, as discussed in the next section holding the conclusions and outlook of the chapter.

5.4 Conclusions

In this chapter, we have introduced a novel methodology for mapping band gaps, even in scenarios where retardation effects significantly obscure the JDOS curves for direct or indirect transitions. Our workflow serves as an ideal complement to get rid of contributions from Cherenkov and waveguide modes. By considering not only the bulk losses in the theory but also incorporating surface losses through BEM calculations, our proposed routine stands out as the most comprehensive approach available in the literature. [11] We achieved remarkable sensitivity, enabling the detection of band gap-type transitions induced by strain. Moreover, we developed a semi-automated approach that bypasses the arduous manual tailoring process. This streamlined workflow requires minimal human intervention and can be computed within hours, limited primarily by BEM simulations. As a result, we have successfully achieved the third and final research goal initially set for this thesis (section 1.2).

The accomplished milestones do hide the fact that the method is in its first stages. It can therefore be further improved, both by adding existing knowledge from literature, and by generating new concepts and processing routines, for sure crossing the path of more complex supervised routines. For instance, the reported iterative approach involving Kramers-Kronig analysis for removing retardation effects can be used as a convergence criterion complementing the additional theory constraints of our method. [11] In addition, it would allow us to sum the dielectric function to the outputs we can already extract. Furthermore, the currently partial consideration of delocalisation effects could be comprehensively accounted for by, for instance, Monte Carlo simulations collapsing the spatial delocalisation into its possible effects. However, the result would be a probability distribution convolving the experimental signal and resulting in an added uncertainty of the measurement.

An important aspect to consider in our workflow refers to its validity when experimental values significantly deviate from the bulk reference. While the observed band gap shifts for the ZnSe@ZnTe NW fall within the meV range, scenarios involving variations of fractions of an eV or beyond remain unexplored. If a strong strain or confinement effect displaces the band gap into this range, the validity of the correction based on the bulk dielectric function becomes uncertain. The theoretical pattern representing the band gap might be too spectrally distant to accurately induce a correction. To address this, we need to incorporate a further level of theory. For instance, including quantum effects or strain from an atomic model would allow us to evaluate cases where additional physical effects deform the bands. Additionally, the workflow presented in Chapter 4 provides an exciting opportunity to access theoretical and local band gap calculations directly from these atomic models. This additional layer of validation enhances the robustness of our results. Notably, the interconnected outputs from each chapter offer a promising outlook for revolutionizing material science research. By implementing these solutions on a daily basis, we can drive significant advancements in the field.

While we have qualitatively assessed the impact of one of these effects - strain - we have yet to explore it quantitatively. Specifically, we lack answers regarding the separate contributions of elastic and plastic strain and their correlation with band gap changes. For instance, this was evident at the ZnSe-ZnTe interface, where dislocations and plane bending coexisted and played a role. Nevertheless, the proposed workflow for band gap mapping opens up several scientific questions. How do different levels of elastic strain precisely tune this property? To address this, we can turn to quantum wells and devices discussed in Chapter 3. These systems offer an excellent opportunity to evaluate the method and take it to the next level by quantitatively measuring the correlation between strain and band gap. In addition, the functional properties of most devices in Chapter 3 are strongly affected by disorder. This disorder might come from strain itself, predominantly if plastic components

exist, or by local stoichiometry modulations. In these cases, the optoelectronic response will be impacted and unveiling its diffuse physical origin would entail a major achievement. These effects can come from Urbach energy levels or excitons potentially interfering and shifting the onset of the apparent EELS LDOS growth to smaller energies. [54, 55] The atomic models extracted from Chapter 4 can also be fundamental to address these question marks as they require an atomic approach. This ongoing work benefits from the enabling capabilities of the semi-automated approach, which not only increases throughput in results but also enhances robustness.

Exploring more complex experimental setups or devices would be invaluable for uncovering the influence of various factors. For instance, the presence of metals in the dielectric response of the device as a whole remains an intriguing area. Specifically, we can investigate the impact of surface plasmon polaritons at metal-dielectric interfaces on losses and corrections. To achieve this, the levels of theory used must adapt and match the specific interactions or excitations we aim to map. In principle, any property or excitation masked by another signal, which can be theoretically modelled, can be removed using an equivalent methodology. For example, a momentum-resolved equivalent would allow us to gauge the dispersion relation. However, achieving this would require a deeper level of theory, akin to the tight-binding model exploited in Chapter 4. Interestingly, this model could be deployed once again, leveraging the automatically predicted 3DAMs from its workflow. On the other end of the spectrum, theory could be intentionally left behind to address situations entirely through experimentation. Ambitiously, sample preparation could play a pivotal role by designing protective layers that function as photonic crystals, effectively preventing parasitic signals (such as Cherenkov radiation) from interfering. [67, 68]

Looking ahead, the next crucial step involves achieving a thoroughly autonomous workflow. Although our semi-automated approach, which is currently under development, already offers a more intuitive and user-friendly experience, we envision taking it even further. Preliminarily, we have deployed and validated the methodology for a characteristic sample described in Chapter 3. Specifically, this sample is a Ge@Si core-shell nanowire grown on Si (Chapter 3, section 3.2.3). Interestingly, the same device appears in Chapter 4, specifically in figures 4.I.47 and 4.I.48, where it validates the multiscale template matching-based phase identification for large devices.

The proposed semi-automated approach now benefits from the robust methodology developed for the ZnSe@ZnTe case (fig. 5.12). [66] With this knowledge, we aim to address expected nominal values around the eV scale and below (1.1 eV and 0.66 eV for Si and Ge, respectively). Importantly, our methodology is not inherently limited to any specific energy range. As long as the defined energy resolution allows for ZLP removal, freeing the spectral range before the band gap, we can apply it

effectively. For this purpose, we have chosen a resolution of 30 meV (equivalent to 5 meV per channel). Next, we manually fine-tuned the hyperparameters for this specific sample configuration:

- We set the number of endmembers to five.
- We assigned their composition (including Si, Ge, SiO_x and Si_3N_4, which appear in the device).
- We configured their dielectric functions.
- We defined the possible fitting ranges around their nominal bulk band gap values and their band gap types (indirect, indirect, direct, and direct, respectively).

This tailored approach ensures that we can accurately analyze the spectral data and extract meaningful information for our Ge@Si core-shell NW.

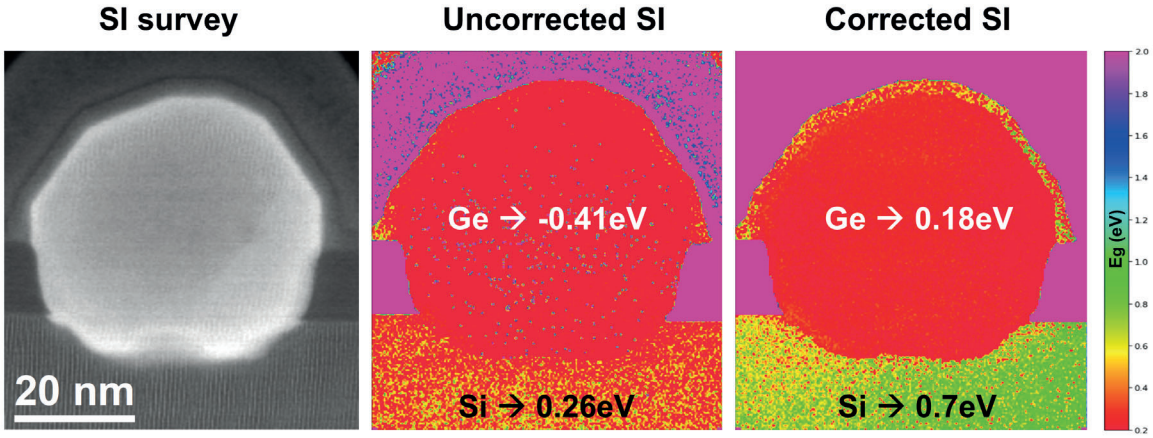


FIGURE 5.12: Band gap maps of a Ge@Si core-shell nanowire obtained from the semi-automated workflow. [66] Its automated nature facilitates comparing the fitting done in the uncorrected version with the corrected one. The differences are significant, indicating the strong effect of retardation effects. Both maps present unphysical average band gap values for both Si and Ge. However, the improvement is major when correcting despite no case-specific tuning encoded in the fitting routine.

Interestingly, we did not customize the fitting ranges or the goodness of fit to generate the resulting band gap maps. We compare the fitting results from the uncorrected case with those from the corrected one, along with the mean band gap values for each region. The uncorrected data produces entirely nonphysical results, yielding average negative band gaps. However, after applying the correction, the trend changes (fig. 5.12). The corrected average band gaps per region highlight that the fitting is not accurate: the shift from bulk reference values is too large to represent a physically meaningful change in a relaxed system like this Ge NW. Importantly, no fine-tuning was applied to this case. An iterative manual evaluation of the fitted curves and adjustments must enhance the physical significance of the process. Furthermore, this system poses greater challenges than the ZnSe@ZnTe case, as no

VEELS band gap map has been ever achieved previously. Despite these challenges, these initial steps hold promise for fully automatically and systematically retrieving band gap information and density of states without being hindered by noise, retardation effects, or narrow band gaps.

Ultimately, we propose integrating the knowledge gained from all three chapters and seamlessly concatenating the workflow from Chapter 4 with the automated band gap mapping routine presented here. Specifically, the template matching techniques used for phase identification (sections 4.I.5, 4.I.8 and 4.I.4) and for modelling local stoichiometry gradients into atomic models (section 4.II.3.4) could be extrapolated to correlate the identified phases with the spectral segmentation. As a result, manual intervention to link each region with its material would become unnecessary. Remarkably, this is one of the many necessary initiatives to fully automate the nowadays limited semi-automated routine, which has already demonstrated the capability to output results that are still in development and constitute the main future research line derived from this chapter.

In summary, the core of the spectral data cleaning and band gap mapping method, along with its initial steps towards automation, combined with the ideas for strengthening its foundation and integrating it with a larger-scale workflow (as seen in Chapter 4) lays a promising foundation. This approach aims to consistently and reproducibly map band gaps at the nanoscale in a standardized and democratized manner.

5.5 Contributions

Marc Botifoll Moral led the proposed research, from its inception to its reported stage in the present chapter. Specifically, Marc Botifoll Moral, supervised by Prof. Dr. Jordi Arbiol Cobos, contributed with:

- Conceived the idea of the methodology to answer the requirements of the ambitious third goal presented in section 1.2. That is, I conceived and coded the methodology combining both GMS's scripting language and Python, conducted and optimised the BEM simulations from which the spectrum image is corrected, and analysed the results outputted from the methodology.
- Discussed and performed the structural data analysis of a broad range of ZnSe@ZnTe core-shell nanowires on sapphire, as reported in the first part of the current chapter.
- Fully developed the semi-automated solution out of the main methodology. The semi-automated approach holds the philosophy of Chapter 4 to democratise and make the methodology more universal and accessible to everybody. The result is not as general as the workflow in Chapter 4, but it points the way towards accomplishing it.

- Under the supervision of Prof. Quentin Ramasse at SuperSTEM, acquired the spectral data used to develop the semi-automated approach, including the raw data generating the band gap maps of the Si@Ge NW in section 5.3. The microscope used is the Nion UltraSTEM 100MC 'HERMES' (see Chapter 2), located at SuperSTEM. [69]

The attained global complexity would have been impractical without the support of many scientific partners. Through its leading role, Marc Botifoll Moral actively coordinated their contributions, which are the following:

- Dr. Sara Martí-Sánchez, from Catalan Institute of Nanoscience and Nanotechnology, led the acquisition of (S)TEM data to structurally characterise the broad range of ZnSe@ZnTe core-shell nanowires. In addition, Dr. Sara Martí-Sánchez led the data analysis of the structural data.
- Dr. Eitan Oksenberg and Prof. Ernesto Joselevich, from the Weizmann Institute of Science, conceived and performed the growth of the broad range of ZnSe@ZnTe core-shell nanowires.
- Dr. Valerio Di Giulio and Prof. F. Javier García de Abajo, from Institut de Ciències Fotòniques (IFCO), guided and supported the theoretical analysis beyond the implementation of the theory-based data cleaning routines. Remarkably, Prof. F. Javier García de Abajo originally conceived the theoretical foundation exploited in the methodology, as well as the coding of the BEM software to compute the surface losses.
- Prof. Quentin Ramasse, from SuperSTEM, acquired the spectral data of the main ZnSe@ZnTe NW. Prof. Quentin Ramasse also supervised and participated in the acquisition of the data employed in the development of the semi-automated routine. The data was acquired operating the Nion UltraSTEM 100MC 'HERMES' (see Chapter 2), located at SuperSTEM. [69]

References

[1] G. Signorello et al. "Tuning the Light Emission from GaAs Nanowires over 290 meV with Uniaxial Strain". In: *Nano Letters* 13 (2013), pp. 917–924.

[2] K.H. Hong et al. "Strain-Driven Electronic Band Structure Modulation of Si Nanowires". In: *Nano Letters* 8 (2008), pp. 1335–1340.

[3] Martin Heiss et al. "Direct correlation of crystal structure and optical properties in wurtzite/zinc-blende GaAs nanowire heterostructures". In: *Phys. Rev. B* 83 (4 Jan. 2011), p. 045303. DOI: 10.1103/PhysRevB.83.045303. URL: https://link.aps.org/doi/10.1103/PhysRevB.83.045303.

[4] Wenfeng Yang et al. "Local-orbital ptychography for ultrahigh-resolution imaging". In: *Nature Nanotechnology* (Jan. 2024). ISSN: 1748-3395. DOI: 10.1038/s41565-023-01595-w. URL: https://doi.org/10.1038/s41565-023-01595-w.

[5] O.L. Krivanek et al. "Progress in ultrahigh energy resolution EELS". In: *Ultramicroscopy* 203 (2019). 75th Birthday of Christian Colliex, 85th Birthday of Archie Howie, and 75th Birthday of Hannes Lichte / PICO 2019 - Fifth Conference on Frontiers of Aberration Corrected Electron Microscopy, pp. 60–67. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2018.12.006. URL: https://www.sciencedirect.com/science/article/pii/S0304399118303668.

[6] Ondrej L Krivanek et al. "Towards sub-10 meV energy resolution STEM-EELS". In: *Journal of Physics: Conference Series* 522.1 (June 2014), p. 012023. DOI: 10.1088/1742-6596/522/1/012023. URL: https://dx.doi.org/10.1088/1742-6596/522/1/012023.

[7] F. S. Hage et al. "Phonon Spectroscopy at Atomic Resolution". In: *Phys. Rev. Lett.* 122 (1 Jan. 2019), p. 016103. DOI: 10.1103/PhysRevLett.122.016103. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.016103.

[8] Fredrik S. Hage et al. "Nanoscale momentum-resolved vibrational spectroscopy". In: *Science Advances* 4.6 (2018), eaar7495. DOI: 10.1126/sciadv.aar7495. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.aar7495. URL: https://www.science.org/doi/abs/10.1126/sciadv.aar7495.

[9] F. J. García de Abajo. "Optical excitations in electron microscopy". In: *Rev. Mod. Phys.* 82 (1 Feb. 2010), pp. 209–275. DOI: 10.1103/RevModPhys.82.209. URL: https://link.aps.org/doi/10.1103/RevModPhys.82.209.

[10] Wei Zhan et al. "Band gap maps beyond the delocalization limit: correlation between optical band gaps and plasmon energies at the nanoscale". In: *Scientific Reports* 8.1 (Jan. 2018), p. 848. ISSN: 2045-2322. DOI: 10.1038/s41598-017-18949-9. URL: https://doi.org/10.1038/s41598-017-18949-9.

[11] M. Stöger-Pollach. "Optical properties and bandgaps from low loss EELS: Pitfalls and solutions". In: *Micron* 39.8 (2008), pp. 1092–1110. ISSN: 0968-4328. DOI: https://doi.org/10.1016/j.micron.2008.01.023. URL: https://www.sciencedirect.com/science/article/pii/S0968432808000334.

[12] M. Stöger-Pollach et al. "Čerenkov losses: A limit for bandgap determination and Kramers–Kronig analysis". In: *Micron* 37.5 (2006). Proceedings of the International Workshop on Enhanced Data Generated with Electrons (EDGE), pp. 396–402. ISSN: 0968-4328. DOI: https://doi.org/10.1016/j.micron.2006.01.001. URL: https://www.sciencedirect.com/science/article/pii/S0968432806000047.

[13] M. Stöger-Pollach, A. Laister, and P. Schattschneider. "Treating retardation effects in valence EELS spectra for Kramers–Kronig analysis". In: *Ultramicroscopy* 108.5 (2008), pp. 439–444. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2007.07.003. URL: https://www.sciencedirect.com/science/article/pii/S0304399107001751.

[14] M. Stöger-Pollach and P. Schattschneider. "The influence of relativistic energy losses on bandgap determination using valence EELS". In: *Ultramicroscopy* 107.12 (2007), pp. 1178–1185. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2007.01.015. URL: https://www.sciencedirect.com/science/article/pii/S0304399107000198.

[15] Svetlana Korneychuk et al. "Exploring possibilities of band gap measurement with off-axis EELS in TEM". In: *Ultramicroscopy* 189 (2018), pp. 76–84. ISSN: 0304-3991. DOI: https://doi.org/10.1016/j.ultramic.2018.03.021. URL: https://www.sciencedirect.com/science/article/pii/S0304399117304229.

[16] Sara Martí-sánchez et al. "Sub-nanometer mapping of strain-induced band structure variations in planar nanowire core-shell heterostructures". In: *Nature Communications* 13 (2022), p. 4089. DOI: /10.1038/s41467-022-31778-3. eprint: https://doi.org/10.1038/s41467-022-31778-3. URL: https://doi.org/10.1038/s41467-022-31778-3.

[17] Eitan Oksenberg et al. "Surface-Guided Core–Shell ZnSe@ZnTe Nanowires as Radial p–n Heterojunctions with Photovoltaic Behavior". In: *ACS Nano* 11.6 (2017). PMID: 28505415, pp. 6155–6166. DOI: 10.1021/acsnano.7b02199. eprint: https://doi.org/10.1021/acsnano.7b02199. URL: https://doi.org/10.1021/acsnano.7b02199.

[18] E. Oksenberg et al. "Guided Growth of Horizontal ZnSe Nanowires and their Integration into High-Performance Blue-UV Photodetectors". In: *Adv. Mater.* 27 (2015), pp. 3999–4005.

[19] G. Reut et al. "Guided Growth of Horizontal P-Type ZnTe Nanowires". In: *J. Phys. Chem. C* 120 (2016), 1708717100.

[20] Mahalingam T. et al. "Electrochemical deposition of ZnTe thin films". In: *Semicond. Sci. Technol.* 17 (2002), pp. 465–470.

[21] Keller D. et al. "Local Band Gap Measurements by VEELS of Thin Film Solar Cells". In: *Microsc. Microanal.* 20 (2014), pp. 1246–1253.

[22] Zamani R.R. et al. "Electronic band structure of narrow-band gap p-n nanojunctions in heterostructured nanowires measured by electron energy loss spectroscopy". In: *arXiv:2103.04632* (2021).

[23] H. Jeon et al. "Blue-Green Injection Laser Diodes in (Zn,Cd)Se/ZnSe Quantum Wells". In: *Appl. Phys. Lett* 59 (1991), 36193621.

[24] Q. Wu, M. Litz, and X. C. Zhang. "Broadband Detection Capability of ZnTe Electro-Optic Field Detectors". In: *Appl. Phys. Lett.* 68 (1996), 29242926.

[25] J. Bang et al. "ZnTe/ZnSe (Core/Shell) Type-II Quantum Dots: Their Optical and Photovoltaic Properties". In: *Chem. Mater.* 22 (2010), 233240.

[26] R. S. Wagner and W. C. Ellis. "Vapor-Liquid-Solid Mechanism of Single Crystal Growth". In: *Appl. Phys. Lett.* 4 (1964), p. 89.

[27] H. Y. Hui et al. "Low-Temperature Growth of Axial Si/Ge Nanowire Heterostructures Enabled by Trisilane". In: *Chemistry of Materials* 29 (2017), pp. 3397–3402.

[28] E. Russo-Averchi et al. "High yield of GaAs nanowire arrays on Si mediated by the pinning and contact angle of Ga". In: *Nano Letters* 15 (2015), pp. 2869–2874.

[29] M. Zamani et al. "3D Ordering at the Liquid–Solid Polar Interface of Nanowires". In: *Advanced Materials* 32 (2020), p. 2001030.

[30] María de la Mata et al. "The Role of Polarity in Nonplanar Semiconductor Nanostructures". In: *Nano Letters* 19.6 (2019). PMID: 31039314, pp. 3396–3408. DOI: 10.1021/acs.nanolett.9b00459. eprint: https://doi.org/10.1021/acs.nanolett.9b00459. URL: https://doi.org/10.1021/acs.nanolett.9b00459.

[31] M. de la Mata et al. "Polarity assignment in ZnTe, GaAs, ZnO and GaN-AlN nanowires from direct dumbbell analysis". In: *Nano Letters* 12 (2012), pp. 2579–2586.

[32] K. Storm et al. "Spatially Resolved Hall Effect Measurement in a Single Semiconductor Nanowire". In: *Nature Nanotechnology* 7 (2012), pp. 718–722.

[33] S. Gazibegovic et al. "Epitaxy of advanced nanowire quantum devices". In: *Nature* 548 (2017), pp. 434–438.

[34] D. Tsivion et al. "Guided Growth of Millimeter-Long Horizontal Nanowires with Controlled Orientations". In: *Science* 333 (2011), pp. 1003–1007.

[35] Y. Zi et al. "Understanding Self-Aligned Planar Growth of InAs Nanowires". In: *Nano Letters* 13 (2013), pp. 2786–2791.

[36] Martin Friedl et al. "Template-Assisted Scalable Nanowire Networks". In: *Nano Letters* 18.4 (2018). PMID: 29579392, pp. 2666–2671. DOI: 10.1021/acs.nanolett.8b00554. eprint: https://doi.org/10.1021/acs.nanolett.8b00554. URL: https://doi.org/10.1021/acs.nanolett.8b00554.

[37] S. Vaitiekėnas et al. "Selective-Area-Grown Semiconductor-Superconductor Hybrids: A Basis for Topological Networks". In: *Physical Review Letters* 121 (2018), p. 147701.

[38] M. de la Mata et al. "Atomic Scale Strain Relaxation in Axial Semiconductor III-V Nanowire Heterostructures". In: *Nano Letters* 14 (2014), pp. 6614–6620.

[39] S. Conesa-Boj et al. "Plastic and Elastic Strain Fields in GaAs/Si CoreShell Nanowires". In: *Nano Letters* 14 (2014), pp. 1859–1864.

[40] R.F. Egerton. *Electron Energy-Loss Spectroscopy in the Electron Microscope*. 3rd ed. Springer, 2011. ISBN: 978-1-4419-9583-4.

[41] A. Genç et al. "Tuning the Plasmonic Response up: Hollow Cuboid Metal Nanostructures". In: *ACS Photonics* 3 (2016), pp. 770–779.

[42] Rafferty B. and Brown L. M. "Direct and indirect transitions in the region of the band gap using electron-energy-loss spectroscopy". In: *Phys. Rev. B* 58 (1998), p. 10326.

[43] Van Stryland E. W. et al. "Energy band-gap dependence of two-photon absorption". In: *Opt. Lett.* 10 (1985), p. 490.

[44] H. Morkoç et al. "Large-band-gap SiC, III-V nitride, and II-VI ZnSe-based semiconductor device technologies". In: *Journal of Applied Physics* 76 (1994), p. 1363.

[45] F. J. García de Abajo and A. Howie. "Retarded field calculation of electron energy loss in inhomogeneous dielectrics". In: *Phys. Rev. B* 65 (2002), pp. 1–17.

[46] A. Howie and R. H. Milne. "Electron energy loss spectra and reflection images from surfaces". In: *Journal of Microscopy* 136.2 (1984), pp. 279–285. DOI: https://doi-org.are.uab.cat/10.1111/j.1365-2818.1984.tb00534.x. eprint: https://onlinelibrary-wiley-com.are.uab.cat/doi/pdf/10.1111/j.1365-2818.1984.tb00534.x. URL: https://onlinelibrary-wiley-com.are.uab.cat/doi/abs/10.1111/j.1365-2818.1984.tb00534.x.

[47] B.G. Mendis. "Surface Core Hole Electron Energy-Loss Fine Structure in MgO: Experiment and Theory". In: *Microscopy and Microanalysis* 27.6 (2021), pp. 1316–1327. DOI: 10.1017/S1431927621012691.

[48] T. Malis, S. C. Cheng, and R. F. Egerton. "EELS log-ratio technique for specimen-thickness measurement in the TEM". In: *Journal of Electron Microscopy Technique* 8.2 (1988), pp. 193–200. DOI: https://doi.org/10.1002/jemt.1060080206. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jemt.1060080206. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jemt.1060080206.

[49] *Gatan Microscopy Suite (GMS)*. 2021.

[50] A. Howie D.D. Perovic C.J. Rossouw. "Imaging elastic strains in high-angle annular dark field scanning transmission electron microscopy". In: *Ultramicroscopy* 52 (1993), pp. 353–359.

[51] P.J. Phillips et al. "Atomic-resolution defect contrast in low angle annular dark-field STEM". In: *Ultramicroscopy* 116 (2012), pp. 47–55.

[52] R. French. "Electronic Band Structure of Al2O3 with Comparison to Alon and AlN". In: *J. Am. Ceram. SOC.* 13.3 (1990), pp. 471–489.

[53] Peköz R. and Raty J. Y. "Band structure modulation of ZnSe/ZnTe nanowires under strain". In: *Physical Review B* 84 (2011), p. 165444.

[54] Esma Ugur et al. "Life on the Urbach Edge". In: *The Journal of Physical Chemistry Letters* 13.33 (2022). PMID: 35960888, pp. 7702–7711. DOI: 10.1021/acs.jpclett.2c01812. eprint: https://doi.org/10.1021/acs.jpclett.2c01812. URL: https://doi.org/10.1021/acs.jpclett.2c01812.

[55] S. M. Wasim et al. "Effect of structural disorder on the Urbach energy in Cu ternaries". In: *Phys. Rev. B* 64 (19 Oct. 2001), p. 195101. DOI: 10.1103/PhysRevB.64.195101. URL: https://link.aps.org/doi/10.1103/PhysRevB.64.195101.

[56] S. Lichtert and J. Verbeeck. "Statistical consequences of applying a PCA noise filter on EELS spectrum images". In: *Ultramicroscopy* 125 (2013), pp. 35–42.

[57] P. Potapov and A. Lubk. "Optimal principal component analysis of STEM XEDS spectrum images". In: *Adv. Struct. Chem. Imaging* (2019). DOI: 10.1186/s40679-019-0066-0.

[58] Marc Botifoll, Ivan Pinto-Huguet, and Jordi Arbiol. "Machine learning in electron microscopy for advanced nanocharacterization: current developments, available tools and future outlook". In: *Nanoscale Horiz.* 7 (12 2022), pp. 1427–1477. DOI: 10.1039/D2NH00377E. URL: http://dx.doi.org/10.1039/D2NH00377E.

[59] L. I. Roest et al. "Charting the low-loss region in electron energy loss spectroscopy with machine learning". In: *Ultramicroscopy* 222 (2021).

[60] L. Daniel D. and S. H. Sebastian. "Learning the parts of objects by non-negative matrix factorization". In: *Nature* 401 (1999), pp. 788–791.

[61] J. Hachtel et al. "Beyond NMF: Advanced Signal Processing and Machine Learning Methodologies for Hyperspectral Analysis in EELS". In: *27* (2021), pp. 322–324.

[62] N. Bonnet and D. Nuzillard. "Independent component analysis: A new possibility for analysing series of electron energy loss spectra". In: *Ultramicroscopy* 102 (2005), pp. 327–337.

[63] F. de la Peña et al. "Mapping titanium and tin oxide phases using EELS: An application of independent component analysis". In: *Ultramicroscopy* 111 (2011), pp. 169–176.

[64] J.M.P. Nascimento and J.M.B. Dias. "Vertex component analysis: a fast algorithm to unmix hyperspectral data". In: *IEEE Transactions on Geoscience and Remote Sensing* 43.4 (2005), pp. 898–910. DOI: 10.1109/TGRS.2005.844293.

[65] A. Tharwat. "Independent component analysis: An introduction". In: *Appl. Comput. Informatics* 17 (2018), pp. 222–249.

[66] Santhanu Panikar Ramanandan et al. "Control of Ge island coalescence for the formation of nanowires on silicon". In: *Nanoscale Horiz.* (2024), pp. -. DOI: 10.1039/D3NH00573A. URL: http://dx.doi.org/10.1039/D3NH00573A.

[67] Robert L. Olmon et al. "Optical dielectric function of gold". In: *Phys. Rev. B* 86 (23 Dec. 2012), p. 235147. DOI: 10.1103/PhysRevB.86.235147. URL: https://link.aps.org/doi/10.1103/PhysRevB.86.235147.

[68] Hao Hu et al. "Surface Dyakonov–Cherenkov radiation". In: *eLight* 2.1 (Jan. 2022), p. 2. ISSN: 2662-8643. DOI: 10.1186/s43593-021-00009-5. URL: https://doi.org/10.1186/s43593-021-00009-5.

[69] SuperSTEM. *Uncovering atomic secrets - Instrumentation — superstem.org*. https://www.superstem.org/facility/instrumentation. [Accessed 21-02-2024].



CHAPTER 6: GENERAL CONCLUSIONS

Chapter 6

General conclusions

6.1 General conclusions

The research goals we initially proposed were three: 1) the structural characterisation of quantum devices, 2) the automation of the data analysis and modelling of STEM data, and 3) the band gap mapping at the nanoscale. Importantly, the research proposed in the three chapters revealed an, at least, partial achievement of these goals. Scientific goals are indeed rarely thoroughly attained due to the never-ending spread any topic can reach. However, to the best of our knowledge, the status reported for each chapter significantly pushes the state of the art. Indeed, a representative sign of pushing the state-of-the-art is the massive generation of new research scopes and perspectives.

The research work in Chapter 3 complemented the progress towards advanced quantum devices. As a result, the research lines it opened are fundamentally about the quantum nature of the measurements. However, the discussion about structural features like strain or defects, and their correlation with the electronic properties, and even with the effect of lamellae thinning is innovative and still not convincingly answered. Importantly, this concern made Chapters 4 and 5 shine. Importantly, they suppose the main methodological breakthroughs of the PhD thesis, and open an endless range of possibilities. Thus, the tools created in these chapters enable the automated and optimized generation of detailed 3D models depicting millions of atoms within the analyzed devices with unprecedented fidelity and accuracy. Beyond the not minor comprehensive structural understanding, we additionally prove their value as simulations representing the experimental device. Indeed, we even provide the experimental basis to compare these simulations with experimental evaluation of properties like the band gap, even when parasitic signals affect

the measurements. The research cycle is completed by experimentally gauging key properties for the devices that originally conceived the theoretical modelling needs.

Each chapter includes its own conclusion section. Therefore, reiterating the main conclusions here would be redundant. However, we want to emphasise one of the central driving forces behind our work: universality and democratisation, while still advancing the state of the art. In other words, anyone can replicate and exploit the proposed scientific methods without requiring cutting-edge or the latest microscopes. Additionally, data processing can be performed on any standard PC, and although atomic modelling is resource-intensive, having 32 GB of RAM should suffice in most cases. The only experimental limitation would arise in mapping narrow band gap materials, where a beam monochromator becomes essential.

In addition to the chapter-specific conclusions, it is worth delving into the broader implications and potential of this research. For example, the quantum devices that served as the inspiration for the automated workflow for generating atomic models could potentially serve as the building blocks of a quantum processor capable of simulating the atomic models generated from that basis. This would replace classical computer calculations with quantum calculations, as atomic simulations are one of the well-known applications which would greatly benefit from the power of quantum algorithms. Let us discuss and articulate more realistic insights, but also the perspectives of a dreamer.

6.2 Future work and outlook

The perspectives offered by the research proposed are endless. Specifically, Chapter 4 delivers an almost white canvas to develop new knowledge based on the fundamentals proposed. The basis is robust enough to, as a whole, provide a reproducible workflow. However, its individual modules can be further improved in what could mean a full thesis chapter on its own.

For instance, the structural analysis demonstrated in Chapter 3 could be repeated and reproduced with the routines presented in the other chapters. To some extent, we already did this, as most examples in 3 are somehow processed in Chapter 4 as proofs to validate specific modules, groups of them, or the workflow as a whole. The same happened for Chapter 3, but on a reduced scale. Importantly, the devices in Chapter 3 are the ones that motivated the creation of the workflow capable of what it is. Therefore, further structural-electronic correlative investigations on devices alike would underscore new challenges which would eventually be new additions to the workflow. Thus, the ultimate aim is to link a measured band gap with a simulated one from the experimental model.

Regarding Chapter 4, manifold improvements and future work can be derived. As said, some work would imply improving the already present pipeline and its elements, and some would come from developing it further in unimaginable ways. For instance, the segmentation and the peak finding should eventually be entirely based on a supervised model (sections 4.I.1 and 4.I.2). In this supervised model, an input image would be perfectly segmented as the output of a CNN. Intermediately, the segmented image would go through a decision tree to unveil the optimal segmentation-peak finding intercommunication (section 4.I.3). Moreover, an added step could be focused on the detailed and careful characterisation of the interfaces between these segmented regions, for instance, to quantify their sharpness and disorder through models equivalent to Muraki segregation. [1] Finally, the peak finding would also involve a fine-tuned unique supervised model to directly output the single-pixel representation of the peaks. This would be similar to the reported SAM or peak finding CNN, but with the added benefits of being flawless and specifically tailored for (S)TEM data. Crucially, not requiring post-processing. [2, 3] Importantly, post-processing hardly ever becomes a general procedure. Consequently, post-processing is detrimental to the universal character we want to ultimately achieve after addressing the presented limitations (section 4.I.5). Nevertheless, this entails the combination of a deeper understanding of the data, a more sound command of AI methods, and smart brute force as the last resource.

Indeed, we want to reinforce the idea that scientific ML or AI cannot rely on the mechanical Turk approach. It is a philosophy stressed throughout the PhD thesis, and it defends not allocating talent to label training data for supervised models. Talent is required to exploit its talent, for which repetitive and valueless tasks like these cannot be justified for the sake of application performance. The alternative is the strategy followed in this dissertation. Enquiring physical models to create synthetic data based on a deep understanding of the data is the winning path in the long term. The generated knowledge can then go beyond the application. In some cases, the performance of the attained supervised model would not be as outstanding as would be by manually labelling data. However, it will probably underscore an incomplete understanding of the underlying nature behind the data. As a result, it can be the motivation towards new knowledge and a better model representing it.

Flexibility in science is primordial. Therefore, completely rejecting brute force labelling is not always justified, as certain scenarios may warrant it. For example, once a supervised model is prepared, manual labelling can enhance the process by creating a test and validation set to compute the metrics that evaluate the model's performance. Alternatively, a supervised model built using multiple knowledge sources could be fine-tuned manually if achieving a specific performance threshold is necessary. In such cases, a small manually labelled training set can complement the validation and test sets, allowing us to achieve the desired metric. However, it is

important to recognise that brute force labelling is not the core but a supplementary force, and it should be used accordingly with responsibility.

Remarkably, the modules listed in Chapter 4 can serve as independent tools and as interconnected parts of a bigger whole. In the latter case, the interconnection of the modules is as pivotal as the modules themselves. In that sense, the advancements made in the template matching for the inter-correlation of datasets have been hindered by the extreme intricacy of the correlation between steps, as detailed in section 4.I.4. Although crystal phase identification is successfully reported in this context, the advancements do not go much further than that. For instance, further progress not yet reported allowed us to segment huge FOVs (>1 μm) based on the smaller ones. However, the results are not as strong as required to continue the workflow and extrapolate the results obtained for the single image case.

The immense complexity within the realm of classical programming provides an ideal stage for AI to shine. Generative models like GANs would be excellent candidates for processing data groups and assembling them into new, previously unconsidered content. [4, 5] The potential benefits are substantial as well, potentially enabling to construct models for even larger devices. First, these models would set records for the sizes that can be simulated. Second, and more importantly, these larger atomic models, representing the full working device, could offer simulations directly under operational conditions. In the case of single-image analysis, this capability is limited to small devices embedded in a single image with atomic resolution. However, the challenge lies in organizing all the outputs from the dataset cohesively, which can be so significant that the process of analyzing vast 4D-STEM datasets is no longer a disadvantage. In any case, further insights and work would be necessary to reveal its ambitious viability while evaluating the possibilities of creating an equivalent workflow for 4D-STEM.

GPA and its automated environment would also easily benefit from the direct application of an AI approach (section 4.I.6). The current solution is fully based on classical permutation evaluation between AI-identified FFT peaks. However, a neural network could be trained with the same underlying principle as our peak finding. Indeed, we would not even need a generative AI model here, but just regular supervised training. Interestingly, the training data could be streamlined like in our peak finding model, but constraining the labels as needed. For instance, instead of labelling the individual peaks, only peak pairs being close together could constitute the labels considering the same physical criteria limiting the current permutation method.

In any case, the current limiting step in terms of computing time is the modelling, particularly the atomic modelling. In the case of FEM modelling, the evident improvement would lie in achieving the equivalent pipeline as in the EELS automation (sections 4.II.2 and 5.3). Thus, automatically extracting an input file for the

FEM/BEM software, requiring no further modification. The answer to this is not obvious, though. The enormous spread of possibilities the FEM can offer out of our experimental models is endless and strictly dependent on the needs of the scientist. The true value lies in having the ability to flexibly adapt the models to suit the needs of the investigation loop at any given moment. Therefore, probably automating this step further lacks real interest unless a more powerful AI takes part. Indeed, a generative model could compare similar device types and the applications they were designed for, and automatically design simulations to test equivalent properties. This advanced AI model would be the most advanced one in the proposed pipeline so far, and would need to be based on transformer architectures and rely on a solid training database (i.e., like ChatGPT or Copilot). The challenges are manifold, although, for instance, the database could be the already available research literature. Indeed, ChatGPT or Copilot excel at pinpointing research articles on the demanded topic. Therefore, the same principle could be the starting point to demand specific simulations based on the successful research so far.

Intriguingly, this same AI can also be applied to atomic simulations, with the necessary modifications made to accommodate the unique requirements of discretisation (section 4.II.3). The current bottleneck in the workflow is the creation of atomic models, which is the module that demands the most resources and time. Models of few million atoms with intricate structural features like defects or heterostructures might take weeks to compute. Moreover, it appears that the process of building atomic models may not be well-suited for AI assistance, but rather, the focus should be on optimizing the requirements for better exploitation of symmetry constraints and memory allocation. For instance, migrating the atomic model generation step to C++, a compiled language intrinsically faster, would boost an algorithm which is currently not optimised at all.

These ideas constitute dreaming awake advancements to the already presented backbone of the workflow in Chapter 4. But dreams give rise to incredible advancements, and we dream of automated comprehensive (S)TEM analysis and acquisition. Specifically, the knowledge generated in the first part (i.e., part I) can be easily transferred into routines for automating the acquisition. Indeed, preliminary steps have been already made for that, in collaboration with the same contributing authors who helped to develop the data analysis workflow. Regarding the data analysis, the possible add-ons are endless as well. From adding support to higher complexity imaging modes like 4D-STEM, to the most obvious missing element: the completion of the spectroscopic compositional analysis. It could set the fundamentals to connect the compositional spectroscopy with the low-loss signal to retrieve band gaps as in Chapter 5. In addition, an excitingly promising research line based on this work has just started with the same collaborating team to push for the automation of the simulations as well. That is, not only delving into the perfect simulation to get our desired properties or information, but also to automate

and process the underlying physics of the simulation tools with AI. For instance, training a NN on DFT or tight-binding simulations of the band gap so the NN learns to do the calculations much faster and with the extrapolating capabilities defined by adequate and representative training.

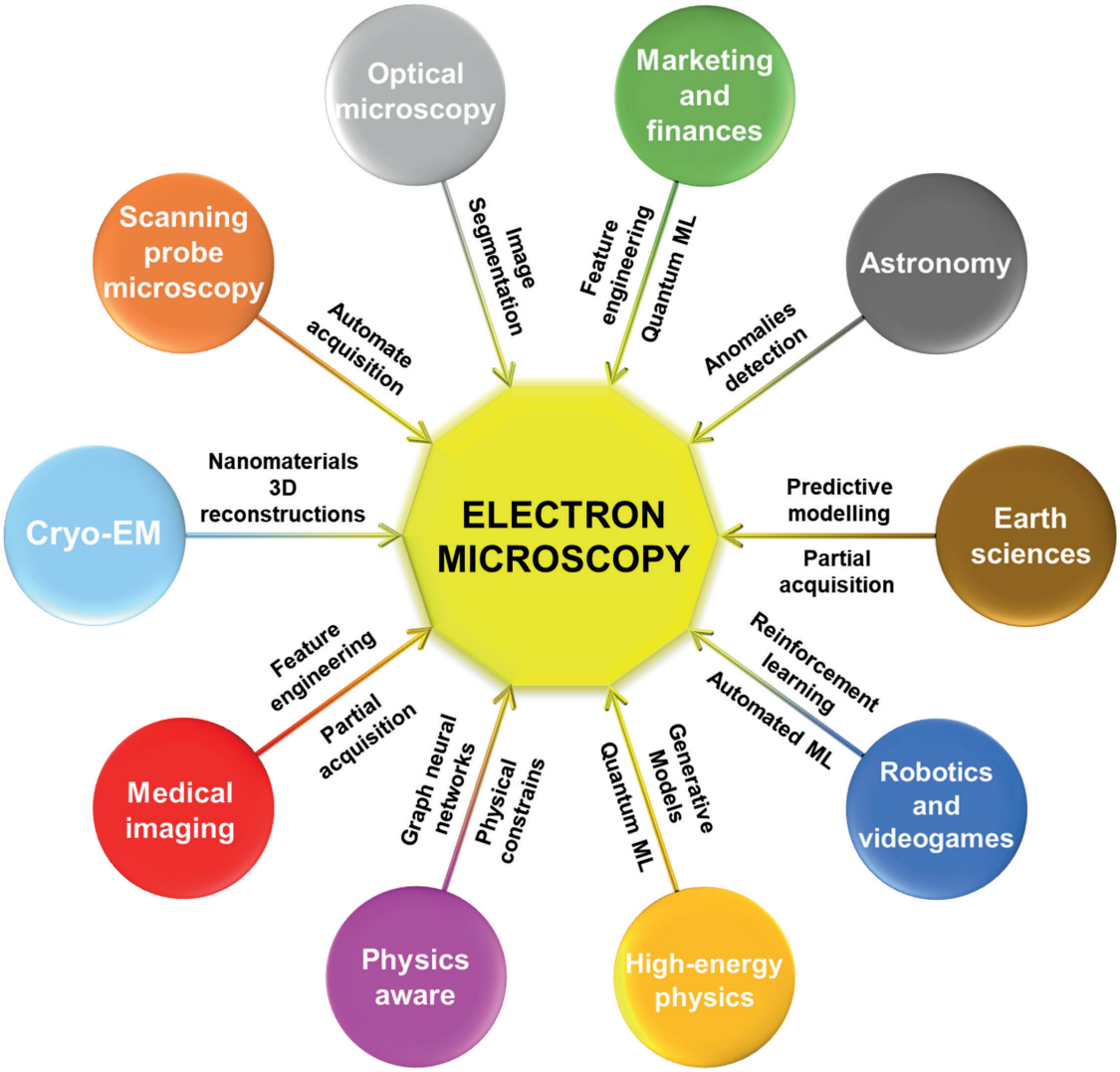


FIGURE 6.1: Cross-fertilization perspectives of data analysis and automation between electron microscopy and other fields. [6]

Ultimately, these AI-enabled band gap calculations could be paired with the experimental measurements. As widely discussed in Chapter 5, the band gap mapping methodology and its automation can reach further lengths by, for instance, incorporating elements of the iterative correction based on Kramers-Kronig analysis. [7] However, the major breakthrough would suppose integrating it into the main workflow. As said, connecting it with the compositional spectroscopy could automate the only real manual intervention in the workflow so far: the quantitative compositional analysis. For example, when a stoichiometry modulation is expected (e.g. elemental diffusion or alloying), simultaneous core-loss EELS and VEELS (in a dual-EELS

system) or simultaneous EDX and VEELS could be useful to correlate the exact composition/stoichiometry (obtained by core-loss EELS or EDX chemical maps) of the material in every pixel, with the electronic structure we are interested in (collecting the inelastically scattered electrons, i.e., VEELS). In parallel, knowing the exact composition in every segment or region would delimit the possible crystal structures to be tested during the phase identification process and allow for better accuracy when selecting the final crystal structure. In summary, a precise material phase identification based on core-loss compositional quantification could be done, while allowing also for more accurate VEELS simulations to eliminate parasitic contributions in the experimental EELS signal.

Generally speaking, this PhD thesis consolidates the importance of AI-related methods to enable automation as well as the extraction of new knowledge in electron microscopy and materials science. In fact, some of the knowledge exploited in the present work comes from other fields. The most evident example is the U-Net model used, which was originally developed for biomedical segmentation. Alternatively, the idea of segmentation and acquisition automation is mainly inspired by cryo-electron microscopy. This highlights how the AI cross-fertilisation of electron microscopy for materials science, to and from other fields can become a differentiating factor when addressing particular challenges (fig. 6.1). [6] Thus, if a similar problem has already been solved in another field, it is smart to imitate this knowledge as would be normally done within the same field. The typical access barrier preventing it is field-specific knowledge. However, AI unites these domains by providing a common nexus from which to start the bibliographic research. Thus, AI-related methods for high-energy physics or astronomy are currently reaching higher complexity than in electron microscopy imaging. On the other hand, for example, Earth sciences are still in an even earlier stage. Therefore, the multidisciplinary approach that AI unlocks can be seen as a unifying force, bringing knowledge together as a path towards great achievements.

Importantly, though, these potentially great achievements must come with responsibility. AI poses new ethical challenges which need to be carefully addressed. [8] In fact, special measures must be taken if, like us, newcomers step into the field without formal or specific training. In that sense, the following key aspects arise.

On the first hand, the established AI model is, overall, majorly based on classical programming. Thus, for both Chapters 4 and 5, it uses multiple ML/DL tools. However, their training is closed, and constrained by physical models we wanted to agglutinate into a complex unit fitting perfectly in the realm of neural networks. Thus, although we expect, for instance, that the peak finding CNN extrapolates well in FFTs unseen in its training, it won't even work in data which is not a 2D matrix. Their fixed architecture and training data tightly limit its reach and allow us to keep it thoroughly under control. Not even online training approaches would be harmful as long as basic constraints like data type or dimensionality are fixed. In fact, we

could say that the workflow developed in Chapter 4 is generative, as it can find and create models with no specific programming or training. However, there are no generative AI modules in its conception. Consequently, the generalisation capabilities are a success of the implementation and interconnection of the different modules, rather than the extensive (and also expensive and climate non-neutral) training of a generative AI. [9]

Interestingly, as indicated above, generative AI would be welcome, for instance, in cases where simulations need to be optimised based on bibliographic research. In this case, limiting the transformer architecture to the wanted scope would demand special caution. For instance, online learning would need to be controlled by an ethical committee that supervises the content fed into the generative model. First, to avoid misinformation and maintain scientific robustness. Secondly, to avoid losing the scientific scope due to irresponsible content management. Again, limiting the possible inputs could be a valid strategy, but the safest is to build upon a controlled training base.

Moreover, the whole workflow from Chapters 4 and 5 could be fit into a single generative model. Intriguingly, it could predict crystal phases not seen in the databases but fitting the diffraction pattern for solving structures. Interestingly, it could even flip the cycle: from a property or desired performance characteristic, to conceive the material and crystal phases, device morphology and configuration, and so on. Similar models based on DL have already been created. [10] Once again, specific considerations must be taken into account to ensure that the primary objectives of the task are overseen by a regulatory body, preventing any deviations from the intended focus. In summary, no real danger of losing control over the software must be considered at the current and reported stage.

Additionally, there are complementary ways to add further human control into complex workflows like the proposed one. In our case, the manual intervention consists of defining one out of four possible settings to produce results. Once defined, no human is involved at all. This *a priori* advantage could be limited in purpose to require a human cross-check, the benefits of which would be twofold. On the one hand, certain steps such as the post-processing of the segmentation or the peak identification could be validated by a trained human to enrich the analysis. On the other hand, we raise a wall to the AI on its decision tree. As discussed, in its current state, this concern is overkill, as the AI is closed and by construction cannot learn. However, reaching that point in a future version of the workflow could potentially entail these ethical issues. As a result, establishing strict control points or security checks from the very beginning is beneficial and does not hinder the original aim.

Finally, it is worth addressing the understandable major and typical concern paired with AI advancements. The automation we propose must not be seen as a way to erase job positions, but to make the current ones more profitable. Thus, although

now the job of many can indeed be done by just an individual, the idea is that these many can do more tasks in the same working time. Hypothetically, if we could only ensure a comprehensive analysis of one single sample per week, with these tools, we could analyse one sample per day or just within hours. We then skyrocket the progress and accelerate the research loop extensively mentioned in the Introduction. Of course, the data analysis progress must be accompanied by advances in the acquisition, manufacturing, and beyond, to avoid a research cycle where the players endlessly pass the bottlenecks back and forth.

Therefore, the results we reported and their implications revolutionise our initial research objectives and open a new paradigm of accelerated research discovery in the framework of electron microscopy and materials science.

References

[1] E. Luna et al. "Indium distribution at the interfaces of (Ga,In)(N,As)GaAs quantum wells". In: *Applied Physics Letters* 92.14 (Apr. 2008), p. 141913. ISSN: 0003-6951. DOI: 10.1063/1.2907508. eprint: https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/1.2907508/13695069/141913_1_online.pdf. URL: https://doi.org/10.1063/1.2907508.

[2] Alexander Kirillov et al. *Segment Anything*. 2023. arXiv: 2304.02643 [cs.CV].

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4.

[4] J. Yi, Z. Yuan, and J. Peng. "Adversarial-Prediction Guided Multi-Task Adaptation for Semantic Segmentation of Electron Microscopy Images". In: *Proc. - Int. Symp. Biomed. Imaging*. 2020, pp. 1205–1208.

[5] T. M. Quan et al. "Removing imaging artifacts in electron microscopy using an asymmetrically cyclic adversarial network without paired training data". In: *Proc. - 2019 Int. Conf. Comput. Vis. Work. ICCVW 2019*. 2019, pp. 3804–3813. DOI: 10.1109/ICCVW.2019.00473.

[6] Marc Botifoll, Ivan Pinto-Huguet, and Jordi Arbiol. "Machine learning in electron microscopy for advanced nanocharacterization: current developments, available tools and future outlook". In: *Nanoscale Horiz.* 7 (12 2022), pp. 1427–1477. DOI: 10.1039/D2NH00377E. URL: http://dx.doi.org/10.1039/D2NH00377E.

[7] M. Stöger-Pollach. "Optical properties and bandgaps from low loss EELS: Pitfalls and solutions". In: *Micron* 39.8 (2008), pp. 1092–1110. ISSN: 0968-4328. DOI: https://doi.org/10.1016/j.micron.2008.01.023. URL: https://www.sciencedirect.com/science/article/pii/S0968432808000334.

[8] Charles Radclyffe, Mafalda Ribeiro, and Robert H. Wortham. "The assessment list for trustworthy artificial intelligence: A review and recommendations". In: *Frontiers in Artificial Intelligence* 6 (2023). DOI: 10.3389/frai.2023.1020592. URL: https://www.frontiersin.org/articles/10.3389/frai.2023.1020592.

[9] *The Environmental Impact of ChatGPT*. https://earth.org/environmental-impact-chatgpt/. Accessed: 2024-04-23.

[10] Amil Merchant et al. "Scaling deep learning for materials discovery". In: *Nature* 624.7990 (Dec. 2023), pp. 80–85. ISSN: 1476-4687. DOI: 10.1038/s41586-023-06735-9. URL: https://doi.org/10.1038/s41586-023-06735-9.

Appendix A

Supplementary data

A.1 Chapter 4

A.1.1 Complementary performance metrics of the peak finding models

	Precision	Recall	F1
Counts	1000	1000	1000
Mean	78.33	59.13	55.16
St. dev.	32.63	31.61	28.38
Min.	0	0	0
25%	64.12	33.33	37.5
50%	100	57.89	58.2
75%	100	92.45	76
Max.	100	100	100

TABLE A.1: Performance metrics of the machine learning-based 2D Gaussian fitting (2DGML) on the full labelled dataset.

	Precision	Recall	F1
Counts	912	912	912
Mean	85.81	57.24	60.33
St. dev.	23.05	30.08	24.05
Min.	10.81	3.13	6.06
25%	77.78	33.33	43.98
50%	100	54.69	60
75%	100	85.71	77.78
Max.	100	100	100

TABLE A.2: Performance metrics of the machine learning-based 2D Gaussian fitting (2DGML) on the labelled dataset elements for which precision is higher than 10%.

	Precision	Recall	F1
Counts	1000	1000	1000
Mean	69.03	48.55	46.42
St. dev.	31.98	32.19	26.85
Min.	0	0	0
25%	46.67	21.88	28.57
50%	76.92	36.84	42.86
75%	100	73.48	63.64
Max.	100	100	100

TABLE A.3: Performance metrics of the encoder-decoder convolutional neural network (CNN) on the full labelled dataset.

	Precision	Recall	F1
Counts	929	929	929
Mean	74.21	46.19	49.82
St. dev.	26.88	30.39	24.73
Min.	11.11	3.7	5.88
25%	55.56	21.74	31.82
50%	80	35.29	45.45
75%	100	69.23	66.67
Max.	100	100	100

TABLE A.4: Performance metrics of the encoder-decoder convolutional neural network (CNN) on the labelled dataset elements for which precision is higher than 10%.

	Precision	Recall	F1
Counts	1000	1000	1000
Mean	62.33	73.87	56.25
St. dev.	37.06	24.42	31.55
Min.	0	0	0
25%	33.33	56.42	38.78
50%	73.91	76.47	64.71
75%	100	100	78.95
Max.	100	100	100

TABLE A.5: Performance metrics of the 1D Scanner (1DScan) on the full labelled dataset.

	Precision	Recall	F1
Counts	807	807	807
Mean	76.92	71	69.17
St. dev.	24.4	22	19.08
Min.	10.14	4.76	7.14
25%	63.64	54.69	57.89
50%	81.82	71.43	70
75%	100	89.47	82.35
Max.	100	100	100

TABLE A.6: Performance metrics of the 1D Scanner (1DScan) on the labelled dataset elements for which precision is higher than 10%.

	Precision	**Recall**	**F1**
Counts	1000	1000	1000
Mean	69.78	70.89	61.87
St. dev.	32.5	25.84	27.51
Min.	0	0	0
25%	51.29	53.85	50
50%	78.52	73.68	67.86
75%	100	100	81.25
Max.	100	100	100

TABLE A.7: Performance metrics of the global peak finding method on the full labelled dataset.

	Precision	**Recall**	**F1**
Counts	909	909	909
Mean	76.64	70.26	67.84
St. dev.	25.36	23.9	20.93
Min.	10.14	4.35	8.33
25%	62.5	53.85	55.56
50%	81.82	71.43	69.23
75%	100	90.48	82.61
Max.	100	100	100

TABLE A.8: Performance metrics of the global peak finding method on the labelled dataset elements for which precision is higher than 10%.

A.2 Chapter 5

A.2.1 Significance tests

The following statistical tests have been performed to prove the statistical veracity of the statements we make throughout the article, in particular when comparing band gap values between regions, or when comparing different fit models. The values tabulated for the t distribution, which can be treated as a Gaussian given the number of samples used, are one-sided for the sake of our test purpose.

Test 1: ZnTe interface pixels. Indirect fit vs Direct fit

Let us test if the indirect fit is better than the direct one in ZnTe interfacial pixels.

Null hypothesis: Both direct and indirect fits respond the same way: $H_0 \rightarrow \bar{r^2}_i = \bar{r^2}_d$

Alternative hypothesis: The indirect model fits the curve in the ZnTe interface pixels better than the direct model: $H_1 \rightarrow \bar{r^2}_i > \bar{r^2}_d$

- Total number of ZnTe interface pixels = 904
- Mean value of the r^2 coefficients for their indirect fit = 0.97191584
- Standard deviation of the r^2 coefficients for indirect fit = 0.021585
- Mean value of the r^2 coefficients for their direct fit = 0.9277318
- Standard deviation of the r^2 coefficients for direct fit = 0.047758

$$t = \frac{\bar{r^2}_i - \bar{r^2}_d}{\sqrt{\frac{{s_{r^2_i}}^2 + {s_{r^2_d}}^2}{n^2}}} = 762.12 \tag{A.1}$$

The null hypothesis is rejected, and we can accept the alternative hypothesis with a confidence higher than 99.9995%: i.e., the indirect fit is significantly better than the direct one in ZnTe interface pixels.

Test 2: ZnTe 'bulk' pixels. Indirect fit vs Direct fit (Control)

Let us test if the indirect fit is better than the direct one in ZnTe bulk pixels.

Null hypothesis: Both direct and indirect fits respond the same way: $H_0 \rightarrow \bar{r^2}_i = \bar{r^2}_d$

Alternative hypothesis: The indirect model fits the curve in the ZnTe bulk pixels better than the direct does: $H_1 \rightarrow \bar{r^2}_i > \bar{r^2}_d$

- Total number of ZnTe bulk pixels = 22755
- Mean value of the r^2 coefficients for their indirect fit = 0.951974713
- Standard deviation of the r^2 coefficients for indirect fit = 0.05486646

- Mean value of the r^2 coefficients for their direct fit = 0.85257474
- Standard deviation of the r^2 coefficients for direct fit = 0.1427154

$$t = \frac{\bar{r^2}_i - \bar{r^2}_d}{\sqrt{\frac{{s_{r^2_i}}^2 + {s_{r^2_d}}^2}{n^2}}} = 14793.1 \tag{A.2}$$

The null hypothesis is rejected, and we can accept the alternative hypothesis with a confidence higher than 99.9995%: i.e., the indirect fit is significantly better than the direct one in ZnTe bulk pixels.

Test 3: E_g ZnTe interface pixels (indirect) vs E_g ZnTe 'bulk' pixels (direct)

Let us test if the band gap values obtained at the bulk pixels for ZnTe are bigger than the ones obtained for the interfacieal pixels of ZnTe (i.e., band gap reduction at the interface)

Null hypothesis: Both band gaps are the same statistically: $H_0 \rightarrow \bar{E}_{g_b} = \bar{E}_{g_{int}}$

Alternative hypothesis: The band gap in the bulk is statistically bigger than the interface band gap: $H_1 \rightarrow \bar{E}_{g_b} > \bar{E}_{g_{int}}$

- Total number of ZnTe bulk pixels = 22755
- Total number of ZnTe interfacial pixels = 904
- Mean value of the band gap for ZnTe's bulk pixels = 2.316184 eV
- Standard deviation of the band gap for ZnTe's bulk pixels = 0.1865171 eV
- Mean value of the band gap for ZnTe's interfacial pixels = 2.281633 eV
- Standard deviation of the band gap for ZnTe's interfacial pixels = 0.291775 eV

$$t = \frac{\bar{E}_{g_b} - \bar{E}_{g_{int}}}{\sqrt{\frac{{s_{E_{g_b}}}^2}{{n_b}^2} + \frac{{s_{E_{g_{int}}}}^2}{{n_{int}}^2}}} = 107.014 \tag{A.3}$$

The null hypothesis is rejected, and we can accept the alternative hypothesis with a confidence higher than 99.9995%: i.e., there is a significant decrease in the band gap energy of the ZnTe interface compared to the bulk ZnTe.

Test 4: E_g ZnSe interface pixels vs E_g ZnSe 'bulk' pixels

Let us test if the band gap values obtained at the bulk pixels for ZnSe are bigger than the obtained in the interface pixels of ZnSe (i.e. band gap reduction at the interface)

Null hypothesis: Both band gaps are the same statistically: $H_0 \rightarrow \bar{E}_{g_b} = \bar{E}_{g_{int}}$

Alternative hypothesis: The band gap in the bulk is statistically bigger than the interface band gap: $H_1 \rightarrow \bar{E_{g_b}} > \bar{E_{g_{int}}}$

- Total number of ZnSe bulk pixels = 14772
- Total number of ZnSe interfacial pixels = 549
- Mean value of the band gap for ZnSe's bulk pixels = 2.773333 eV
- Standard deviation of the band gap for ZnSe's bulk pixels = 0.14892 eV
- Mean value of the band gap for ZnSe's interfacial pixels = 2.7545514 eV
- Standard deviation of the band gap for ZnSe's interfacial pixels = 0.25151 eV

$$t = \frac{\bar{E_{g_b}} - \bar{E_{g_{int}}}}{\sqrt{\frac{{s_{E_{g_b}}}^2}{{n_b}^2} + \frac{{s_{E_{g_{int}}}}^2}{{n_{int}}^2}}} = 40.987 \tag{A.4}$$

The null hypothesis is rejected, and we can accept the alternative hypothesis with a confidence higher than 99.9995%: i.e., there is a significant decrease in the band gap energy of the ZnSe interface compared to the bulk ZnSe.

Appendix B

List of publications

For Chapter 3

1. AUTHORS: Kushagra Aggarwal, Andrea Hofmann, Daniel Jirovec , Ivan Prieto, Amir Sammak, **Marc Botifoll**, Sara Martí-Sánchez, Menno Veldhorst, Jordi Arbiol, Giordano Scappucci, Jeroen Danon, and Georgios Katsaros

   TITLE: Enhancement of proximity-induced superconductivity in a planar Ge hole gas

   JOURNAL: *Physical Review Research*, **3**, L022005 (2021)

   IMPACT FACTOR: 4.2

2. AUTHORS: D. Jirovec, A. Hofmann, A. Ballabio, P. M. Mutter, G. Tavani, **M. Botifoll**, A. Crippa, J. Kukucka, O. Sagi, F. Martins, J. Saez-Mollejo, I. Prieto, M. Borovkov, J. Arbiol, D. Chrastina, G. Isella, and G. Katsaros

   TITLE: A singlet-triplet hole spin qubit in planar Ge

   JOURNAL: *Nature Materials*, **20**, 1106–1112 (2021)

   IMPACT FACTOR: 43.841

3. AUTHORS: Marco Valentini, Maksim Borovkov, Elsa Prada, Sara Martí-Sánchez, **Marc Botifoll**, Andrea Hofmann, Jordi Arbiol, Ramón Aguado, Pablo San-Jose, and Georgios Katsaros

   TITLE: Majorana-like Coulomb spectroscopy in the absence of zero-bias peaks

   JOURNAL: *Nature*, **612**, 442–447 (2022)

   IMPACT FACTOR: 69.504

4. AUTHORS: Brian Paquelet Wuetz, Davide Degli Esposti, Anne-Marije J Zwerver, Sergey V Amitonov, **Marc Botifoll**, Jordi Arbiol, Lieven MK Vandersypen, Maximilian Russ, and Giordano Scappucci

   TITLE: Reducing charge noise in quantum dots by using thin silicon quantum wells

JOURNAL: *Nature Communications*, **14**, 1385 (2023)

IMPACT FACTOR: 17.690

5. AUTHORS: Alberto Tosato, Vukan Levajac, Ji-Yin Wang, Casper J Boor, Francesco Borsoi, **Marc Botifoll**, Carla N Borja, Sara Martí-Sánchez, Jordi Arbiol, Amir Sammak, Menno Veldhorst, and Giordano Scappucci

   TITLE: Hard superconducting gap in germanium

   JOURNAL: *Communications Materials*, **4**, 23 (2023).
   https://doi.org/10.1038/s43246-023-00351-w

   IMPACT FACTOR: 7.8

6. AUTHORS: Davide Degli Esposti, Lucas EA Stehouwer, Önder Gül, Nodar Samkharadze, Corentin Déprez, Marcel Meyer, Ilja N Meijer, Larysa Tryputen, Saurabh Karwal, **Marc Botifoll**, Jordi Arbiol, Sergey V Amitonov, Lieven MK Vandersypen, Amir Sammak, Menno Veldhorst, and Giordano Scappucci

   TITLE: Low disorder and high valley splitting in silicon

   JOURNAL: *npj Quantum Information* **10**, 32 (2024).
   https://doi.org/10.1038/s41534-024-00826-9

   IMPACT FACTOR: 7.6

7. AUTHORS: Marián Janík, Kevin Roux , Carla Borja Espinosa , Oliver Sagi, Abdulhamid Baghdadi, Thomas Adletzberger, Andrea Ballabio, **Marc Botifoll**, Alba Garzón Manjón, Jordi Arbiol, Daniel Chrastina, Giovanni Isella, Ioan M. Pop, and Georgios Katsaros

   TITLE: In situ Resistance Control of Granular Aluminium Superinductors for Hybrid Circuit Quantum Electrodynamics

   JOURNAL: In preparation

For Chapter 4

8. AUTHORS: **Marc Botifoll**, Ivan Pinto-Huguet, and Jordi Arbiol

   TITLE: Machine learning in electron microscopy for advanced nanocharacterization: current developments, available tools and future outlook

   JOURNAL: *Nanoscale Horizons*, **7**, 1427-1477 (2022)

   IMPACT FACTOR: 10.989

9. AUTHORS: **Marc Botifoll***, Ivan Pinto-Huguet, Enzo Rotunno, Thomas Galvani, Catalina Coll, Maria Chiara Spadaro, Martin Børstad Eriksen, Yann-Michel Niquet, Georgios Katsaros, Gonzalo Merino, Pablo Ordejón, Stephan Roche, Vincenzo Grillo, and Jordi Arbiol* (* Corresponding author)

TITLE: End-to-end Artificial Intelligence Workflow for the Autonomous Electron Microscopy Unveiling of Physical Knowledge

JOURNAL: To be submitted

For Chapter 5

10. AUTHORS: **Marc Botifoll**$^{\gamma}$, Sara Martí-Sánchez$^{\gamma}$, Eitan Oksenberg, Christian Koch, Carla Borja, Maria Chiara Spadaro, Valerio Di Giulio, Quentin Ramasse, Javier García de Abajo, Ernesto Joselevich, and Jordi Arbiol ($^{\gamma}$ Equal contribution)

    TITLE: Sub-nanometer mapping of strain-induced band structure variations in planar nanowire core-shell heterostructures

    JOURNAL: *Nature Communications*, **13**, 4089 (2022)

    IMPACT FACTOR: 17.690

11. AUTHORS: Nicholas Stephen, Demie Kepaptsoglou, Robert Lawrence, Ivan Pinto-Huguet, **Marc Botifoll**, Jordi Arbiol, Agnieszka Gocalinska, Enrica Mura, Emanuele Pelucchi and Miryam Arredondo

    TITLE: Evaluating the local bandgap across InGaAs multi quantum wells in a metamorphic laser via EELS

    JOURNAL: To be submitted

Complementary publications:

12. AUTHORS: Pavel Aseev, Guanzhong Wang, Luca Binci, Amrita Singh, Sara Martí-Sánchez, **Marc Botifoll**, Lieuwe J. Stek, Alberto Bordin, John D. Watsin, Frenk Boeakhout, Daniel Abel, John Gamble, Kevin Van Hoogdale, Jordi Arbiol, Leo P. Kouwenhoven, Gijs de Lange and Philippe Caroff

    TITLE: Ballistic InSb Nanowires and Networks via Metal-Sown Selective Area Growth

    JOURNAL: *Nano Letters*, **19**, 9102-9111 (2019)

    IMPACT FACTOR: 12.437

13. AUTHORS: Dawei Yang, Zhifu Liang, Chaoqi Zhang, Jordi Jacas Biendicho, **Marc Botifoll**, Maria Chiara Spadaro, Qiulin Chen, Mengyao Li, Alberto Ramon, Ahmad Ostovari Moghaddam, Jordi Llorca, Jiaao Wang, Joan Ramon Morante, Jordi Arbiol, Shu-Lei Chou, and Andreu Cabot

    TITLE: $NbSe_2$ Meets C_2N: A 2D-2D Heterostructure Catalysts as Multifunctional Polysulfide Mediator in Ultra-Long-Life Lithium–Sulfur Batteries

    JOURNAL: *Advanced Energy Materials*, **11**, 2101250 (2021)

    IMPACT FACTOR: 29.368

14. AUTHORS: María Isabel Díez-García, Andrés Alberto García Blanco, Sebastian Murcia-López, **Marc Botifoll**, Jordi Arbiol, Mohammad Qamar, and Joan Ramon Morante

    TITLE: Flexible and Binder-Free Iron Phosphide Electrodes Using a Three-Dimensional Support for High Hydrogen Productivity

    JOURNAL: *ChemElectroChem* 2023, e202201152

    IMPACT FACTOR: 4.782

15. AUTHORS: Qing Sun, Guifang Zeng, Jing Li, Shang Wang, **Marc Botifoll**, Hao Wang, Deping Li, Fengjun Ji, Jun Cheng, Huaiyu Shao, Yanhong Tian, Jordi Arbiol, Andreu Cabot, and Lijie Ci

    TITLE: Is Soft Carbon a More Suitable Match for SiO_x in Li-Ion Battery Anodes?

    JOURNAL: *Small*, 2023, 2302644. https://doi.org/10.1002/smll.202302644

    IMPACT FACTOR: 15.153

16. AUTHORS: María Isabel Díez-García, Guillem Montaña-Mora, **Marc Botifoll**, Andreu Cabot, Jordi Arbiol, Mohammad Qamar, and Joan Morante

    TITLE: Cobalt-Iron Oxyhydroxide Obtained from the Metal Phosphide: A Highly Effective Electrocatalyst for the Oxygen Evolution Reaction at High Current Densities

    JOURNAL: *ACS Applied Energy Materials*, 2023, **6**, 11, 5690–5699

    IMPACT FACTOR: 6.959

17. AUTHORS: Gong, LI; Zhang, Chao Yue; Li, Junshan; Montaña-Mora, Guillem; **Botifoll, Marc**; Zhou, Jin Yuan; Arbiol, Jordi; Kallio, Tanja; Martinez-Alanis, Paulina; and Cabot, Andreu

    TITLE: Enhanced Electrochemical Hydrogenation of Benzaldehyde to Benzyl Alcohol on Pd@Ni-MOF by Modifying the Adsorption Configuration

    JOURNAL: *ACS Applied Materials and Interfaces*, 2024,
    https://doi.org/10.1021/acsami.3c13920

    IMPACT FACTOR: 10.38

18. AUTHORS: Xiaonan Deng, Jizhe Cui, Yu-chun Lin, Simian Zhang, Huihui Yu, **Marc Botifoll**, Shoutian Sun, Qinghong Yuan, Jordi Arbiol, Zhengcao Li, Zheng Zhang, Yue Zhang, Qihua Xiong, Rong Yu, and Chen Wang

    TITLE: Hyper-mismatch atomically coherent lateral layered semiconductor heterointerface and interfacial layer reconfigured moiré superlattice

    JOURNAL: Under review

19. AUTHORS: Quach, Vien-Duong, Spadaro, Maria Chiara, **Botifoll, Marc**, Vezin, Herve, Colbeau-Justin, Christophe, Dumeignil, F., Arbiol, Jordi, Wojcieszak, Robert, Ghazzal, and Mohamed Nawfal

    TITLE: Engineering CuO_2/TiO_2 interface for optimal charge carrier separation and photocatalytic hydrogen generation

    JOURNAL: Under review

20. AUTHORS: Linlin Yang, Ren He, **Marc Botifoll**, Yongcai Zhang, Yang Ding, Chong Di, ChuanSheng He, Ying Xu, Jordi Arbiol, Yingtang Zhou, and Andreu Cabot

    TITLE: Enhanced oxygen evolution reaction and zinc-air battery performance via electronic spin modulation at metal active sites in reconstructed $Ni/MnFe_2O_4$ heterostructures

    JOURNAL: Under review

21. AUTHORS: Jing Yu, Chao Yue Zhang, Chen Huang, Oleg Usoltsev, Ashley P. Black, Kapil Gupta, Maria Chiara Spadaro, Ivan Pinto-Huguet, **Marc Botifoll**, Canhuang Li, Javier Herrero-Martín, Jin Yuan Zhou, Alexandre Ponrouch, Ruirui Zhao, Lluís Balcells, Andreu Cabot and Jordi Arbiol

    TITLE: Vacancy-mediated electron spin engineering to promote Li-S redox reactions

    JOURNAL: Under review

22. AUTHORS: Yifei Wu, Xiaonan Deng, Simian Zhang, Xiaoze Li, Yuqi Wang, Shengxian Ke, Qinghua Zhang, Ivan Pinto-Huguet, **Marc Botifoll**, Di Bao, Zhengcao Li, Jordi Arbiol, Lin Gu, Honghua Fang, Qihua Xiong, and Chen Wang

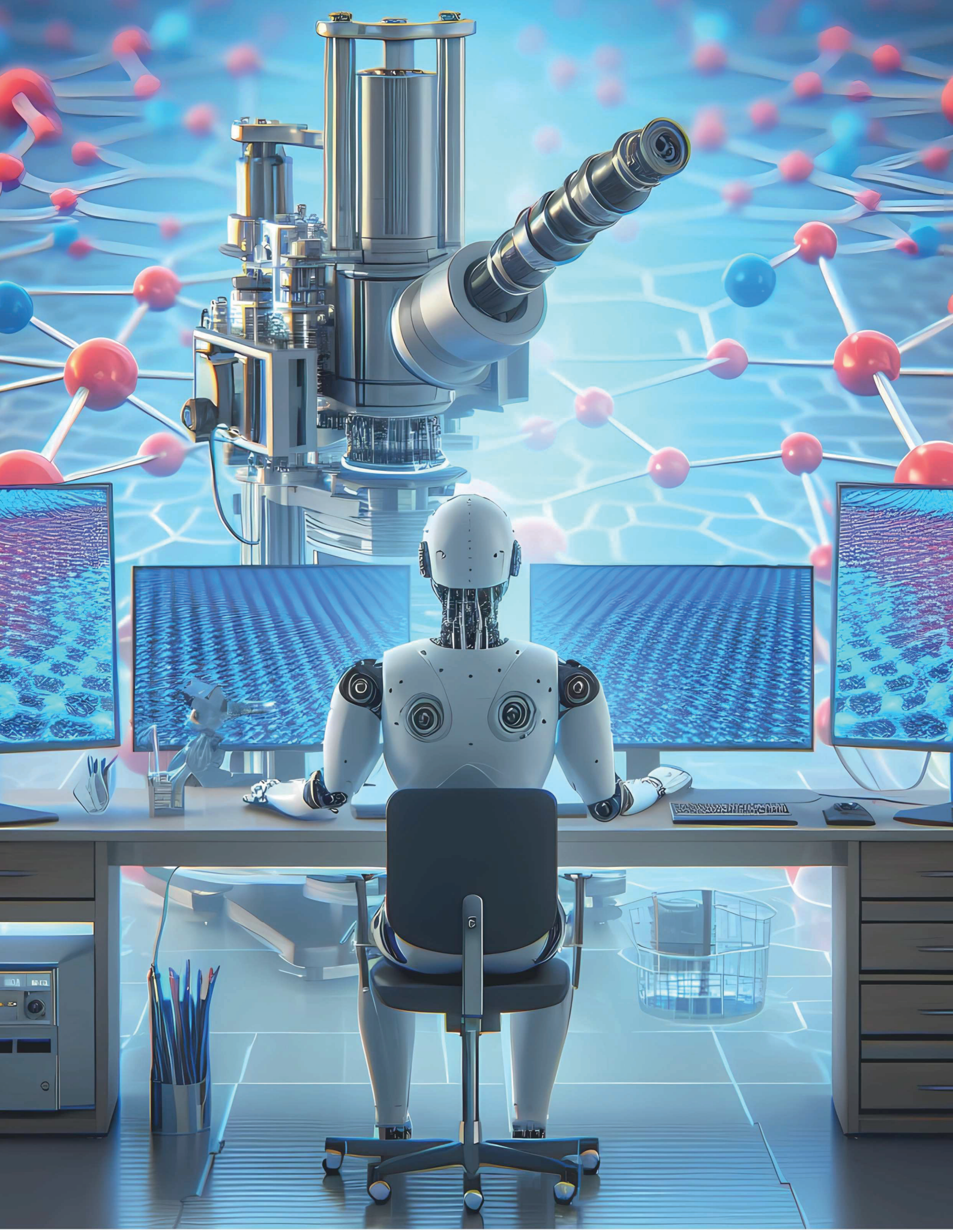
    TITLE: Ultrastrong interface modulated exciton anisotropic diffusion-based exciton diode

    JOURNAL: To be submitted

23. AUTHORS: Yuqi Wang, Xiaonan Deng, Simian Zhang, Yuan Luo, Yuxin Zhai, Jing Chen, Yifei Wu, Shengxian Ke, Ivan Pinto-Huguet, **Marc Botifoll**, Qinghua Zhang, Dezhao Wu, Di Bao, Zhengcao Li, Yong Xu, Lin Gu, Jordi Arbiol, Sheng Meng, Qihua Xiong, and Chen Wang

    TITLE: Highly precise defect-mediated lateral atomic layer semiconductor optical superlattice

    JOURNAL: To be submitted



AGAUR Agència de Gestió d'Ajuts Universitaris i de Recerca

GAeN

UAB Universitat Autònoma de Barcelona

ICN2 Institut Català de Nanociència i Nanotecnologia

CSIC Consejo Superior de Investigaciones Científicas

BIST Barcelona Institute of Science and Technology