UAB
Universitat Autònoma de Barcelona

Mapping the Unknown
New techniques for verification and identification of quantum states and processes

Santiago Llorens Fernández

Mapping the Unknown

New techniques for verification and identification of quantum states and processes

by
Santiago Llorens Fernández

This page is intentionally left blank

Universitat Autònoma de Barcelona

Mapping the Unknown

New techniques for verification and identification of quantum states and processes

by
Santiago Llorens Fernández

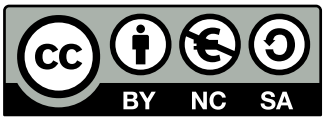
under the supervision of

Prof. Ramon Muñoz Tapia
&
Dr. Gael Sentís Herrera

A thesis submitted in partial fulfillment for the degree of
Doctor of Philosophy in Physics

Física Teòrica: Informació i Fenòmens Quàntics
Departament de Física, Facultat de Ciències

Bellaterra, 15/07/2025



A mis padres.
A mi hermana.
A todos los que me respaldan.

This page is intentionally left blank

Resum

Aquesta tesi estudia la verificació de sistemes quàntics en escenaris on l'accés a informació crucial és limitat. Mentre que en el control de qualitat clàssic els defectes poden, en principi, detectar-se de manera determinista, la naturalesa probabilística de la mecànica quàntica imposa limitacions inherents. Aquest fet planteja la pregunta: fins a quin punt és possible identificar de manera fiable defectes en estats i processos quàntics? Per respondre aquesta pregunta s'estudia la discriminació d'estats i canals quàntics en contextos universals, és a dir, quan no es coneix la forma específica de les hipòtesis. L'objectiu és establir rendiments de referència i dissenyar protocols òptims que continuïn sent vàlids sota condicions d'ignorància parcial.

La manera d'abordar l'objectiu és mitjançant la combinació d'eines d'optimització convexa amb mètodes de teoria de representacions. La programació semidefinida ofereix el marc per formular els problemes de discriminació com a problemes d'optimització sobre operadors. Per la seva banda, la dualitat de Schur–Weyl i els esquemes d'associació permeten captar la simetria dels problemes estudiats, així com reduir de manera significativa la seva complexitat. Més enllà de la discriminació d'estats, l'isomorfisme de Choi–Jamiołkowski i el formalisme de xarxes quàntiques estenen aquestes tècniques a l'anàlisi de canals quàntics.

La primera contribució se centra en el problema anomenat *edge detection*. Una seqüència de sistemes es divideix en dos dominis, cadascun preparat en un estat desconegut però diferent, i l'objectiu consisteix a identificar la frontera que els separa. Es presenta un protocol òptim per a aquest problema, juntament amb un càlcul eficient de la seva probabilitat d'èxit basat en la dualitat de Schur–Weyl i el formalisme de matrius de Gram. A més, s'explora un escenari mixt, en què els estats passen de ser coneguts a desconeguts, cosa que pot resultar útil per detectar defectes en processos de producció d'estats quàntics.

La segona contribució introdueix la *detecció quàntica d'anomalies múltiples*, en què un conjunt de fonts ha de produir un estat de referència, encara que algunes poden fallar i generar estats anòmals. Es consideren dos casos. En el primer, tant l'estat de referència com els estats anòmals són coneguts; aquest problema es resol analíticament aprofitant l'estructura algebraica dels esquemes d'associació de John-

son, la qual cosa permet obtenir estratègies òptimes tant en el règim d'error mínim com en el d'identificació sense error (inequívoca). En el segon cas, ambdues classes d'estats són desconegudes, i es desenvolupa un protocol universal la optimalitat del qual s'estableix a través de la dualitat de Schur–Weyl.

Finalment, la tercera contribució estén la detecció d'anomalies a l'àmbit dels canals quàntics. S'assumeix que els dispositius implementen una operació unitària donada, encara que un subconjunt d'aquests aplica en el seu lloc unitàries desconegudes. Es deriven protocols òptims sense error per identificar aquests dispositius defectuosos, emprant l'isomorfisme de Choi–Jamiołkowski i el formalisme de xarxes quàntiques. L'anàlisi mitjançant teoria de representació i la dualitat mixta de Schur–Weyl demostra que el rendiment òptim pot assolir-se amb estats d'entrada entrellaçats, sistemes auxiliars i estratègies paral·leles.

Resumen

Esta tesis aborda la verificación de sistemas cuánticos en escenarios donde el acceso a información crucial es limitado. Mientras que en el control de calidad clásico los defectos pueden, en principio, detectarse de manera determinista, la naturaleza probabilística de la mecánica cuántica impone limitaciones inherentes. Este hecho plantea la pregunta: ¿hasta qué punto es posible identificar de forma fiable defectos en estados y procesos cuánticos? Para responder a esta pregunta se estudia la discriminación de estados y canales cuánticos en contextos universales, es decir, cuando no se conoce la forma específica de las hipótesis. El objetivo es establecer rendimientos de referencia y diseñar protocolos óptimos que sigan siendo válidos bajo condiciones de ignorancia parcial.

La manera de abordar el objetivo es mediante la combinación de herramientas de optimización convexa con métodos de teoría de representaciones. La programación semidefinida ofrece el marco para formular los problemas de discriminación como problemas de optimización sobre operadores. Por su parte, la dualidad de Schur–Weyl y los esquemas de asociación permiten capturar la simetría de los problemas estudiados, así como reducir de manera significativa su complejidad. Más allá de la discriminación de estados, el isomorfismo de Choi–Jamiołkowski y el formalismo de redes cuánticas extienden estas técnicas al análisis de canales cuánticos.

La primera contribución se centra en el problema denominado *edge detection*. Una secuencia de sistemas se divide en dos dominios, cada uno preparado en un estado desconocido pero distinto, y el objetivo consiste en identificar la frontera que los separa. Se presenta un protocolo óptimo para este problema, junto con un cálculo eficiente de su probabilidad de éxito basado en la dualidad de Schur–Weyl y el formalismo de matrices de Gram. Además, se explora un escenario mixto, en el que los estados pasan de ser conocidos a desconocidos, lo que puede resultar útil para detectar fallos en procesos de producción de estados cuánticos.

La segunda contribución introduce la *detección cuántica de anomalías múltiples*, en la que un conjunto de fuentes debe producir un estado de referencia, aunque algunas pueden fallar y generar estados anómalos. Se consideran dos casos: en el primero, tanto el estado de referencia como los estados anómalos son conocidos;

este problema se resuelve analíticamente aprovechando la estructura algebraica de los esquemas de asociación de Johnson, lo que permite obtener estrategias óptimas tanto en el régimen de error mínimo como en el de identificación sin error (inequívoca). En el segundo caso, ambas clases de estados son desconocidas, y se desarrolla un protocolo universal cuya caracterización como óptimo se establece a través de la dualidad de Schur–Weyl.

Finalmente, la tercera contribución extiende la detección de anomalías al ámbito de los canales cuánticos. Se asume que los dispositivos implementan una operación unitaria dada, aunque un subconjunto de ellos aplica en su lugar unitarias desconocidas. Se derivan protocolos óptimos sin error para identificar estos dispositivos defectuosos, empleando el isomorfismo de Choi–Jamiołkowski y el formalismo de redes cuánticas. El análisis mediante teoría de representaciones y la dualidad mixta de Schur–Weyl demuestra que el rendimiento óptimo puede alcanzarse con estados de entrada entrelazados, sistemas auxiliares y estrategias paralelas.

Abstract

This thesis investigates the verification of quantum systems in scenarios where access to crucial information is limited. While in classical quality control faults can, in principle, be detected deterministically, the probabilistic nature of quantum mechanics imposes inherent limitations. This raises the question of how reliably faults in quantum states and processes can be identified. The challenge is addressed through the study of quantum state and channel discrimination in universal settings, where the specific form of the hypotheses is not known. The objective is to establish benchmarks of performance and to construct optimal protocols that are valid under partial ignorance.

The approach combines convex optimization with representation theory methods. Semidefinite programming provides the framework for formulating discrimination tasks as operator optimization problems. Schur–Weyl duality and association schemes capture the symmetric structure of the tasks studied and reduce their complexity. Beyond state discrimination, the Choi–Jamiołkowski isomorphism and quantum network formalism extend these techniques to the analysis of quantum channels.

The first contribution concerns the problem termed as *edge detection.* A sequence of systems is divided into two domains, each prepared in an unknown but distinct state, and the goal is to identify the boundary separating them. The optimal protocol for this task is derived, and its success probability is efficiently computed through Schur–Weyl duality and the Gram matrix formalism. A mixed scenario, where the states transition from known to unknown, is also explored, which may find applications in detecting malfunctions in quantum-state production processes.

The second contribution introduces *quantum multi-anomaly detection*, where a set of sources are intended to produce a reference state but some may malfunction and produce anomalous states instead. Two cases are studied. In the first, both the reference and anomalous states are known. This problem is studied analytically by using the algebraic structure of Johnson association schemes, yielding optimal minimum-error and unambiguous strategies. In the second, both classes of states are unknown, and a universal protocol is developed whose optimality is established through Schur–Weyl duality methods.

The final contribution extends anomaly detection to quantum channels. Devices are expected to implement a given unitary operation, but a subset of them instead applies unknown unitaries. I derive optimal zero-error protocols for identifying such faulty devices using the Choi–Jamiołkowski isomorphism and quantum network formalism. A representation-theoretic analysis of mixed Schur–Weyl duality shows that entangled input states with ancillary probes and parallel strategies suffice to achieve optimal performance.

This page is intentionally left blank

Declaration

I declare that this thesis has been composed by myself and that this work has not been submitted for any other degree or professional qualification. I confirm that the work submitted is my own, except where work which has formed part of jointly-authored publications has been included. My contribution and those of the other authors to this work have been explicitly indicated below. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

List of Publications

Publications included in the thesis

[LGS+25] S. Llorens, W. González, G. Sentís, J. Calsamiglia, R. Muñoz-Tapia, and E. Bagan, "Quantum Edge Detection", Quantum **9**, 1687 (2025).

[LSM24] S. Llorens, G. Sentís, and R. Munoz-Tapia, "Quantum multi-anomaly detection", Quantum **8**, 1452 (2024).

[SLH+24] M. Skotiniotis, S. Llorens, R. Hotz, J. Calsamiglia, and R. Muñoz-Tapia, "Identification of malfunctioning quantum devices", Physical Review Research **6**, 033329 (2024).

In preparation:

S. Llorens, A. Diebra, M. Sedlák and R. Muñoz-Tapia, "Exact identification of unknown unitary processes".

Additional publications

[DLB+25] A. Diebra, S. Llorens, E. Bagan, G. Sentís, and R. Muñoz-Tapia, "Quantum state exclusion for group-generated ensembles of pure states", arXiv:2503.02568 (2025).

[DLG+25] A. Diebra, S. Llorens, D. González-Lociga, A. Rico, J. Calsamiglia, M. Hillery, and E. Bagan, "Quantum Advantage in Identifying the Parity of Permutations with Certainty", arXiv:2508.04310 (2025).

[MLS24] S. Morelli, S. Llorens, and J. Siewert, "Exploring Imaginary Coordinates: Disparity in the Shape of Quantum State Space in Even and Odd Dimensions", arXiv:2404.15179 (2024).

Acknowledgments

This work marks a journey—a journey that has come and pass, leaving behind memories. In the cycle of research, of what has been and what is yet to come, this work arose and reached its end.

Mis más sinceros agradecimientos se los debo a mis dos tutores, Ramon y Gael. A Ramon li dec la meva sort —que per desgràcia sovint no és gaire bona— i per recordar-me que no tot el que dic "es del todo estúpido". A Gael le debo el haberme enseñado que todos en el GiQ son —y somos— unos "persons". Y a los dos les agradezco el haberme incluido en una dinámica de trabajo con tanta complicidad como la suya. Aparte de la relación personal, a Ramon li agraeixo haver-me ensenyat gran part de la meva intuïció a l'hora d'afrontar problemes, i també com ser rigorós. A Gael le agradezco mostrarme cómo se puede ser extremadamente fino —y "picky"— a la hora de escribir ciencia. A Arnau le agradezco el aire fresco que ha traído al pequeño equipo: con su entusiasmo y su calidad solo puedo desearle suerte —mucha, que la necesitará para contrarrestar. Juntament amb Ramon y Gael, agraeixo al John i a l'Emili, haver compartit amb mi el meu primer projecte d'investigació i totes les posteriors discussions i colaboracions.

Agraeixo també a tots els membres, passats i actuals, amb qui he tingut el plaer de coincidir. Thanks to the new ones, Ion, Jofre, Hari, Some, Arpan, and José. També a l'Arnau i al Jordi, per totes les converses sobre llibres de fantasia; *mil gràcies*, Jordi, per haver-me introduït en aquest món. Thanks to the ones I started this journey with, Jennifer, Naga, and my first Ph.D. partner Niklas, and also Mani and Joe, with whom I had the adventure of organizing a winter school.

Agraeixo als meus companys de despatx, Albert, Zuzana i especialment al Toni i a l'Elisabet, cadascun per motius diferents. Al Toni, per totes les pauses quan no podíem treballar més davant de l'ordinador. A l'Elisabet, per haver compartit amb mi l'aventura docent a la UAB. També vull agrair al Matt totes les vegades en què ens hem desfogat parlant dels comportaments i del rendiment dels alumnes —no sempre ha estat fàcil.

Thanks to the old ones, Maria, Phillip, and Matías; and Marco, Giulio, and Carlo (guapo), with whom, together with Walther, I shared my first conference —how crazy the "biribí"! Special mention here to Tenerife, thanks for the most fun

conference ever!

Thanks to the other old seniors, Michalis, Anna and Andreas, for the warm environment you created at GiQ. Thanks to the new seniors, Gabrielle and Martí, you will add enormously to the group. And huge thanks to Michal, for hosting me in Bratislava, for teaching me everything I know about quantum combs, and for being such a wonderful mentor.

Dentro de la universidad también ha habido compañeros de viaje: Aleix, Mario y Teo, gràcies per tots els dinars a la UAB. Pero especialmente, gracias, Ivan. Gracias por haber sido un compañero tan cercano desde el día de la matrícula de la carrera. Quién nos iba a decir que aquel encuentro casual iba a iniciar diez (¡10!) años de fantástica amistad, con tantos grandes momentos en el piso de estudiantes junto a Pujol, Fer y Gerard.

Dejando de lado la universidad, quiero agradecer a los de siempre, a los de mi pueblo. Aunque hace tiempo que nos separamos, parece que nunca lo hubiéramos hecho: Luís, Alex, Javi, Martín y Carlos (Saez). Pero, sobre todo, Carlos (Esteve), gracias por una amistad y un apoyo que ya superan los veinte años. Y en especial, Arabia, quien me ha acompañado durante estos cuatro años, me ha entendido, apoyado y soportado como nadie; junto a Coco 🐈‍⬛, claro.

Gracias a mis padres, Reyes y Juan, por vuestro apoyo incondicional en toda mi vida, mis estudios y mi trabajo. Y gracias a mi hermana, Inés, a quien además agradezco enormemente haber diseñado la portada.

This page is intentionally left blank

Contents

1 Introduction **1**
1.1 Summary of results . . . 3

2 Fundamentals **7**
2.1 The quantum state . . . 7
2.2 The quantum measurement . . . 8
2.3 The quantum channel . . . 10
2.4 The quantum tester . . . 12
2.5 Quantum state and channel identification . . . 14
2.5.1 Two states . . . 15
2.5.2 Multiple states . . . 17
2.5.3 The Gram matrix . . . 18
2.5.4 Quantum channel identification . . . 19
2.6 Universal discrimination . . . 21

3 Mathematical toolkit **23**
3.1 Semidefinite programming . . . 23
3.1.1 Primal program . . . 24
3.1.2 Dual program . . . 25
3.1.3 Weak and strong duality . . . 26
3.1.4 Complementary slackness . . . 27
3.1.5 Discrimination tasks as SDP . . . 28
3.2 Schur-Weyl dualities . . . 31
3.2.1 Representations of finite groups . . . 32
3.2.2 Partitions, Young diagrams & Schur-Weyl duality . . . 33
3.2.3 Schur-Weyl duality in quantum systems . . . 35
3.2.4 Mixed Schur-Weyl duality . . . 38
3.3 Association schemes . . . 41
3.3.1 Definition and basic properties . . . 42
3.3.2 Symmetric association schemes . . . 43

3.3.3 Intersection matrices . . . 44
3.3.4 P- and Q-polynomial schemes . . . 45
3.3.5 Johnson association scheme . . . 47

4 Quantum edge detection **51**
4.1 Unknown domains . . . 53
4.1.1 Asymptotic regime . . . 60
4.2 Known state in one domain . . . 64
4.3 Discussion . . . 67

5 Quantum multi-anomaly detection **71**
5.1 Setting of the problem . . . 72
5.2 Minimum-error . . . 74
5.2.1 Asymptotic regime . . . 76
5.3 Zero-error identification . . . 78
5.4 Universal protocol . . . 80
5.5 Discussion . . . 83

6 Identification of anomalous quantum unitary evolutions **85**
6.1 Introduction . . . 85
6.2 Setting of the problem . . . 87
6.3 Optimal zero-error protocol . . . 89
6.3.1 Single anomaly . . . 89
6.3.2 Two anomalies . . . 93
6.4 Discussion . . . 97

Outlook **101**

A Asymptotic results of quantum edge detection **105**

Bibliography **113**

1

Introduction

In a fabrication process, quality control guarantees that the product of the production line meets certain standards. Deviations from expected data might signal a fault, a contamination, or a need for recalibration. In the age of quantum technologies, yet to come, we face a parallel challenge when ensuring that quantum states are being prepared as intended and that quantum processes behave as desired.

Quality control is now understood as a mathematically rigorous field grounded in statistical inference and decision theory. Techniques such as hypothesis testing, Bayesian inference, maximum likelihood estimation, and sequential analysis form the pillars of modern inspection protocols [Car72]. These tools allow manufacturers to detect outliers, to distinguish between random fluctuations and systematic errors, and to make informed decisions under uncertainty through a rigorous statistical analysis. A central goal in modern quality control is to minimize the resources required in inspections—whether time or material costs—while maintaining high confidence in the conclusions drawn. This trade-off between cost and reliability is a centerpiece in modern inspection methods: rather than inspecting every unit, statistical techniques guide how to sample, estimate, and decide.

On the other hand, quantum systems introduce an additional layer of complexity to the already intricate techniques in classical quality control. Unlike their classical counterparts, quantum systems are inherently probabilistic; even when a quantum device is working as intended, the outcomes of measurements are not deterministic in general. An inspection of the whole production in the classical scenario may, in principle, be able to detect faulty products without errors, whereas in the quantum setup, there are cases where this is just not possible. This intrinsic randomness induces challenges for verification, as it is mandatory to account for additional errors nonexistent in the classical scenario.

These challenges motivate the study of verification of quantum systems from a statistical analysis standpoint, embracing and expanding classical methods to quantum scenarios. A central question naturally arises: *Given full access to measurement capabilities, how confidently can faults in a quantum system be detected?* In quantum mechanics, the probabilistic nature of measurement outcomes implies that even when a fault is present, it may not be detected with certainty. This naturally leads to the framework of quantum discrimination, a fundamental aspect of quantum information theory, which formalizes the problem of deciding between a finite set of hypotheses about the quantum state of a system or the process controlling it from the output provided by quantum measurements. In this thesis, a key aim is to characterize optimal protocols for quantum state and channel discrimination under various assumptions, thereby establishing performance benchmarks. Once these fundamental limits are understood, the question of resources can be addressed. In classical quality control, this means reducing the number of inspections while keeping reliability high. In the quantum scenario, it translates into minimizing the number of systems measured, or the amount of entanglement needed to achieve a given level of performance. Establishing these limits of quantum discrimination is therefore a necessary first step before any discussion regarding resource efficiency.

It is also important to note that optimal quantum verification protocols depend on the considered hypotheses, and this generally implies some prior knowledge of the faults. When such knowledge is absent, the situation changes significantly. In these cases, one must resort to *universal* strategies: protocols that do not rely on the specific description of the faults, but are instead designed to work for *any* possible scenario. Universality in this sense refers to the construction of multi-purpose verification tools, applicable regardless of the particular form of the systems or evolutions under inspection. The question is then: *If one has no information about the flaws, to what extent can one design universal strategies for verifying quantum systems?* Developing and analyzing such universal protocols is therefore essential, as it reveals what can be achieved in the most agnostic setting and establishes performance benchmarks regardless of the specific description of the quantum systems.

This dissertation aims to shed some light on these two questions. In a nutshell, how good can we infer faults in quantum systems? Can we devise universal protocols for these tasks?

What follows in this chapter is a summary of the main results of my research. Chapter 2 gives the reader a basic introduction to probability and statistics, quantum information, and quantum state discrimination, followed by a discussion of the general state of the art of the field. In Chapter 3, I describe the mathematical

framework used in my research, from symmetries to optimization theory, going through graph theory. Then, Chapters 4 to 6 comprise the body of results obtained during my PhD. The thesis ends with an outlook on future work related to the topics covered and the bibliography.

1.1 Summary of results

Quantum edge detection

The main body of this thesis is quantum state discrimination, a fundamental primitive task in quantum information. This task consists of correctly identifying the state of a quantum system that is one of a set of possible states. Conventional quantum state discrimination relies on complete knowledge of the set of candidate states. However, in many scenarios, this assumption may not hold. Quantum state discrimination protocols that do not assume this prior knowledge of the state can be thought of as universal protocols, as they do not depend on the description of the states of the quantum system, and therefore are valid for any possible quantum state. Universality addresses scenarios of partial or total ignorance about the system. Studying such protocols highlights what can be rigorously achieved in quantum inference even when not all the information is available. A particularly interesting instance of this task is what I refer to as quantum edge detection, which can be described as follows: given a string of states consisting of two domains, each of which is formed by quantum systems in the same state within the domain but different from the other, one aims to correctly identify the frontier between the two domains. With no further information, one should consider *every* possible pair of states describing each of the regions, and therefore devise a *universal* protocol. In Chapter 4, I formalize mathematically this task and derive the optimal performance of a universal protocol for large quantum systems. Having addressed the general scenario, I switch the perspective, considering a more structured problem where one of the regions is known, as it comes from a source production of states. Allegedly, at some point in time (or space), the source failed to produce the desired state, and the aim is to identify correctly when (or where) it did. The results regarding the problem of quantum edge detection are reported in

[LGS+25] S. Llorens, W. González, G. Sentís, J. Calsamiglia, R. Muñoz-Tapia, and E. Bagan, “Quantum Edge Detection”, Quantum **9**, 1687 (2025).

Quantum multi-anomaly detection

In Chapter 5, I introduce a new verification task: the multi-anomaly problem. While the quantum edge detection task aims to identify a single, abrupt change in a sequence of states, the multi-anomaly problem considers a parallel production of quantum states in which each source is intended to prepare a certain state. Some of these sources, however, may malfunction, producing sporadically anomalous outputs instead.

I study two variants of this problem. In the first case, both the expected states and the anomalies are known. The second is a universal scenario where no prior knowledge about any of the states is assumed. For each case, I derive optimal protocols and performances, and compare the results between the different variations. Moreover, for the instances where the target state is known, I introduce an unambiguous state discrimination protocol, which guarantees an error-free decision, yet adds the cost of possible inconclusive answers. Notably, the optimal protocol presents specific features that allow for resource optimization, thereby reducing inspection costs in verification controls.

The transition from Chapter 4 to this chapter reflects a natural progression: from identifying a single fault in a line to detecting multiple, possibly simultaneous, anomalies across a parallel production network. These results are presented in

[LSM24] S. Llorens, G. Sentís, and R. Munoz-Tapia, "Quantum multi-anomaly detection", Quantum **8**, 1452 (2024).

[SLH+24] M. Skotiniotis, S. Llorens, R. Hotz, J. Calsamiglia, and R. Muñoz-Tapia, "Identification of malfunctioning quantum devices", Physical Review Research **6**, 033329 (2024).

Identification of anomalous quantum unitary evolutions

Chapter 6 moves beyond detecting anomalous states and addresses the problem of verifying quantum processes. The setting is richer operationally and more aligned with quantum architectures, where each device is expected to implement a specific transformation—typically unitary—and failures manifest as unintended operations.

This problem expands the multi-anomaly detection framework from Chapter 5. In here, the anomalies are no longer produced states, but the anomalous evolution applied to a quantum system. This introduces new challenges, and entanglement plays a central role when extracting all the available information.

I study a universal variant of the problem in which the intended unitary transformation is known, but the faulty devices apply unknown unitary operations.

This setting describes an agnostic scenario, where the identification of faults must be achieved without any prior knowledge of the specific form of the anomalous evolutions. In this context, I construct universal protocols that optimize the success probability of anomaly identification, relying only on the symmetry properties inherent to the problem.

This task extends the concept of universal discrimination for quantum processes, showing that even under information constraints, anomaly identification is possible. The results complete the conceptual arc from known-state to unknown-process verification, and also highlight the power of quantum inference in uncertain regimes. The results from this part will appear in a manuscript currently in preparation

> S. Llorens, A. Diebra, M. Sedlák and R. Muñoz-Tapia, "Exact identification of unknown unitary processes". In preparation.

2

Fundamentals

The theory of probability is the only mathematical tool available to help map the unknown. It is fortunate that this tool, while tricky, is extraordinarily powerful and convenient.

Benoît Mandelbrot

2.1 The quantum state

Any attempt to verify or discriminate between quantum systems must begin with a precise formulation of how these systems are described. In quantum mechanics, this description takes the form of a mathematical object: the *quantum state*.

From a Bayesian perspective [FS13]—frequently adopted in quantum state identification—the quantum state reflects the information an agent has about a physical system. It encodes not intrinsic properties, but degrees of belief—expectations about the outcomes of future measurements. These beliefs are captured by the so-called *density matrix* ρ, an operator that acts in a complex Hilbert space $\mathcal{H}$ associated with the system. The state ρ summarizes all knowledge relevant to answering operational questions about the system.

Formally, a quantum state is a positive semidefinite operator of unit trace:

$$\rho \geq 0\,, \tag{2.1}$$

$$\mathrm{tr}\rho = 1\,. \tag{2.2}$$

A quantum state is said to be *pure* if it has rank one, meaning it can be written

as $\rho = |\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle \in \mathcal{H}$. Otherwise, the state is called *mixed*. From an informational point of view, pure states represent maximal knowledge about the system: they are extremal elements in the convex set of quantum states. Any mixed state can be expressed as a convex combination of pure states

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \, , \tag{2.3}$$

with $\sum_i p_i = 1$.

In this sense, pure states yield the sharpest predictions, while mixed states reflect residual uncertainty about the system's preparation. This scenario, in its simplest version, is usually described as a *source* that produces a quantum state $|\psi_1\rangle\langle\psi_1|$ with prior probability p, or with probability $1 - p$ the state $|\psi_2\rangle\langle\psi_2|$. The agent's belief about the quantum system is then encoded in the density matrix

$$\rho = p|\psi_1\rangle\langle\psi_1| + (1-p)|\psi_2\rangle\langle\psi_2| \, . \tag{2.4}$$

An alternative perspective on mixed states considers the entanglement of a quantum system with an external environment, whose information lies beyond the agent's knowledge. A quantum pure state belongs to a higher Hilbert space $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$, with

$$|\psi\rangle = \sum_{i,j} c_{i,j} |u_i\rangle|v_j\rangle \, , \tag{2.5}$$

where $\{|u_i\rangle\}_i$ and $\{|v_j\rangle\}_j$ are orthonormal basis for $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. If an agent has only access to the first subsystem, her partial knowledge is encoded in the quantum state

$$\rho = \mathrm{tr}_2 |\psi\rangle\langle\psi| = \sum_i p_i |u_i\rangle\langle u_i| \, , \tag{2.6}$$

which is a mixed state [NC10], with $p_i = \sum_j |c_{i,j}|^2$, and $\sum_i p_i = 1$.

2.2 The quantum measurement

If the quantum state captures what one knows about a quantum system, a quantum measurement is the act of posing a question to it. More formally, it corresponds to an interaction with the system whose outcome yields classical information.

In standard formulations of quantum mechanics, measurements are introduced through a composite postulate.

Postulate II.a

Every physical quantity corresponds to an operator A acting on a Hilbert space $\mathcal{H}$. This operator is an observable: a Hermitian operator, i.e., with real eigenvalues; whose associated eigenvectors form an orthogonal basis of the Hilbert space,

$$A = \sum_i \alpha_i \, |a_i\rangle\langle a_i| \, . \tag{2.7}$$

The result of a measurement is one of A's eigenvalues α_i, occurring with probability $p(i|\rho) = \mathrm{tr}(|a_i\rangle\langle a_i|\rho)$, according to the Born rule [NC10]. The expected value of A is then

$$\langle A \rangle = \sum_i \alpha_i \, \mathrm{tr}(|a_i\rangle\langle a_i|\rho) \, . \tag{2.8}$$

The measurement associated with the observable A can be characterized by the set of projectors $\{\Pi_i = |a_i\rangle\langle a_i|\}_i$, corresponding to a *projective measurement* with outcomes labeled by i.

Postulate II.b

If the measurement gives the result labeled by i, then the state of the system immediately after the measurement is given by

$$\rho_i = \frac{\Pi_i \rho \Pi_i}{\mathrm{tr}(\Pi_i \rho)} = |a_i\rangle\langle a_i| \, . \tag{2.9}$$

This rule is historically known as the wave function collapse postulate or Lüders-von Neumann rule [Neu18; Lüd06]. The pre-measurement state ρ instantly collapses to the post-measurement state $\rho_i = |a_i\rangle\langle a_i|$. This instantaneous action gave rise to conceptual and philosophical debate, which continues to this day. In modern interpretations such as quantum Bayesianism, this collapse is understood not as a physical process, but as an update in the agent's degrees of belief given the new information [FS13; FMS14].

A projective measurement is also known as a *projection-valued measure* (PVM) [Neu18]. However, the mathematical formalism of quantum mechanics allows for more general types of measurement. These receive the name of *positive operator-valued measure* (POVM) [Hol73], which consist of sets of *positive semidefinite* operators

$$E_i \geq 0 \, , \tag{2.10}$$

that fulfill the completeness relation

$$\sum_i E_i = \mathbb{1}\,. \tag{2.11}$$

Unlike projective measurements, POVMs do not require that measurement outcomes be associated with eigenvalues of an observable. The outcomes are simply labels corresponding to the elements $\{E_i\}_i$. The probability of obtaining outcome i is again given by the Born rule

$$p(i|\rho) = \mathrm{tr}(E_i\,\rho)\,. \tag{2.12}$$

Conditions (2.10) and (2.11) guarantee that $p(i|\rho)$ is a proper probability distribution, i.e., $p(i|\rho) \geq 0$ for all i and $\sum_i p(i|\rho) = 1$, respectively.

2.3 The quantum channel

Another ingredient in the foundational description of quantum theory is the *quantum operation*, which captures the most general physical transformation a quantum state may undergo. While early quantum mechanics focused on evolving closed systems, modern formulations recognize that realistic processes require a broader class of transformations. The appropriate framework to describe such processes is that of *quantum channels*.

The initial description of quantum mechanics [Sch26; NC10] considered the evolution of a quantum system to be governed by the Schrödinger equation:

$$i\hbar\frac{\partial\rho}{\partial t} = [H, \rho]\,, \tag{2.13}$$

where H is the *Hamiltonian*, the observable associated with the system's energy. The solution to Eq. (2.13) gives rise to a unitary evolution of the state of the system:

$$\rho(t) = U(t)\rho\,U^\dagger(t)\,, \tag{2.14}$$

with $U(t) := e^{-iHt/\hbar}$.

However, just as PVMs and pure states represent only restricted classes of measurements and states, unitary evolutions account for only a subset of all physically allowed transformations. In general, a quantum operation is represented as a linear *map*, $\Phi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$, where $\mathcal{L}(\mathcal{H})$ denotes the space of linear operators acting

on the Hilbert space $\mathcal{H}$. To ensure that Φ corresponds to a physical transformation, it must satisfy two conditions. It must be a *trace-non-increasing* map:

$$\mathrm{tr}\big(\Phi\left(\rho\right)\big) \leq 1\,, \tag{2.15}$$

for any density matrix ρ, and *completely positive*:

$$\left(\mathrm{id}_n \otimes \Phi\right)\left(Y\right) \geq 0\,, \tag{2.16}$$

for all $Y \geq 0$ and $n \in \mathbb{N}$, with id_n the identity map on a system of dimension n, that is, the action of a quantum map on a subsystem of a quantum state of a larger Hilbert space must yield a valid physical state on the total space. A *quantum channel* is a subtype of quantum operation that is *trace-preserving*, i.e., condition (2.15) is saturated.

Quantum measurements can also be understood as a special class of quantum channels, where classical information is extracted and the post-measurement state—the agent's knowledge—is updated. More generally, quantum channels represent all physical transformations compatible with the probabilistic structure of quantum theory.

A foundational result due to Kraus [KBD+83] shows that any quantum channel admits a representation in terms of a set of operators $\{K_i\}_i$ not necessarily Hermitian:

$$\Phi(\rho) = \sum_i K_i \rho K_i^\dagger\,, \tag{2.17}$$

with $\sum_i K_i^\dagger K_i = \mathbb{1}$.

Just as mixed states may arise from partial access to a larger entangled system, this decomposition can be seen either as classical uncertainty over operations or as the effective result of a unitary process on a larger—partially inaccessible—system. The latter interpretation is formalized by the Stinespring dilation theorem [Sti55], which shows that any quantum channel can be modeled as a unitary operation on a larger system, which is not entirely accessible to the agent. Analogously to quantum states, where pure states reflect maximal agent knowledge, unitary channels represent quantum operations under maximal information.

A particularly useful representation of quantum channels is given by the Choi–Jamiołkowski isomorphism. It establishes a one-to-one correspondence between *completely positive trace-preserving* (CPTP) maps, $\Phi : \mathcal{H}_A \to \mathcal{H}_B$, and positive semidefinite operators J_Φ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, given by

$$J_\Phi = \left(\mathrm{id}_A \otimes \Phi\right)|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|\,, \tag{2.18}$$

where id_A is the identity operator in subsystem A, which encodes that the A subsystem is unaffected by the channel; and $|\mathbb{1}\rangle\rangle$ is the unnormalized maximally entangled state $\sum_i |i\rangle_A |i\rangle_B$. This isomorphism allows the tools developed for quantum states to be extended to the analysis of processes in more general frameworks—such as quantum networks [CDP09]—that describe how transformations can be verified.

2.4 The quantum tester

The last element in our foundational framework is the *quantum tester*, which extends the role of POVMs from quantum states to quantum operations. Just as POVMs provide a mathematical description of measurements by assigning probabilities to quantum states, testers play the analogous role for processes. Testers are a particular case of *quantum networks*, and to describe them it is useful to introduce the general structure of a network first. Intuitively, a quantum network is a sequence of devices, arranged in time, through which quantum systems go. Each interaction may have an input and an output space—state preparations have only output spaces, measurements have only input spaces and channels have both input and output spaces, see Figs. 2.1a–c, and the whole network can be seen as a higher-order process that transforms incoming systems into outgoing ones, see Fig. 2.1d.

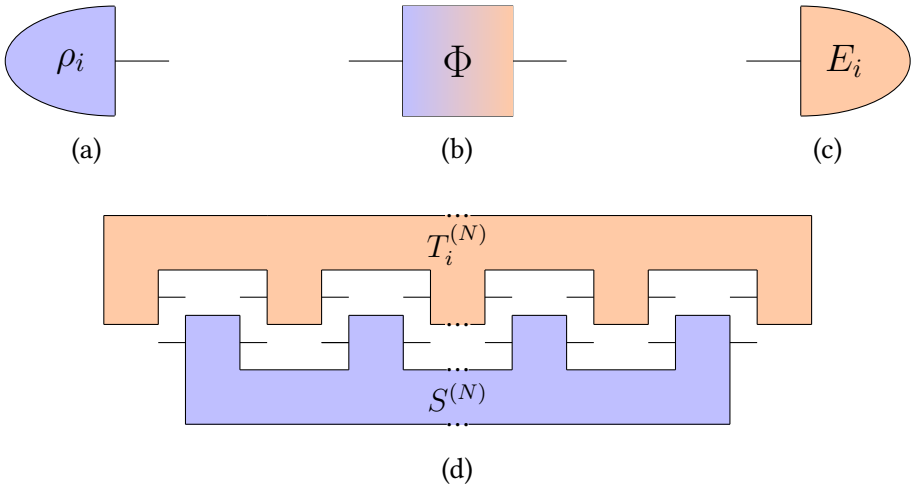


Figure 2.1: Diagrammatic representation of: (a) a state preparation, (b) a quantum channel, (c) a measurement, and (d) a quantum N-tester acting on a N-quantum network $S^{(N)}$.

What makes this framework powerful is that it provides a single mathematical object encoding the action of the entire sequence, and testers represent the most general strategies available to probe unknown quantum processes. Formally, a quantum network with N vertices, i.e., N sequential quantum interactions, can be described by a positive semidefinite operator $R^{(N)} \in \mathcal{L}\left(\bigotimes_{j=0}^{2N-1} \mathcal{H}_j\right)$, with the following properties:

$$R^{(N)} \geq 0 \tag{2.19}$$

$$\mathrm{tr}_{2n-1} R^{(n)} = \mathbb{1}_{2n-2} \otimes R^{(n-1)}\,, \qquad 2 \leq n \leq N \tag{2.20}$$

$$\mathrm{tr}_1 R^{(1)} = \mathbb{1}_0 \tag{2.21}$$

where the even (odd) indices correspond to input (output) spaces. These recursive constraints ensure that the network is *causally ordered*, i.e., the operations it encodes cannot signal from the future to the past. The operator $R^{(N)}$ is the Choi representation of the entire network and encodes the input-output behavior.

In particular, a quantum tester is a family of operators $\{T_i\}_i$ acting on a quantum network with N vertices, satisfying

$$T_i \geq 0\,, \tag{2.22}$$

$$\sum_i T_i = \mathbb{1}_{2N} \otimes R^{(N)}\,. \tag{2.23}$$

This mirrors the structure of a POVM, where the measurement operators sum to the identity, Eq. (2.11). Here, however, the tester operators sum to the Choi representation of a $N+1$-causal quantum strategy.

The probability of obtaining the i-th outcome with the tester formalism is given by the generalized Born rule

$$p(i|S) = \mathrm{tr}\left(S^{(N)}\, T_i^T\right) \tag{2.24}$$

where $S^{(N)}$ is the Choi matrix of a N-vertices quantum network, and T_i^T is the transpose of T_i. This probability is a particular instance of the rule for composing quantum networks, that is, the composition of general quantum processes: the link product [CDP08].

The link product provides a general formalism for combining quantum networks based on their shared subsystems. For two quantum channels $R \in \mathcal{L}\left(\bigotimes_{r\in\mathcal{R}} \mathcal{H}_r\right)$ and $S \in \mathcal{L}\left(\bigotimes_{s\in\mathcal{S}} \mathcal{H}_s\right)$, the composed Choi matrix is $S * R \in \mathcal{L}\left(\mathcal{H}_{\mathcal{R}\setminus\mathcal{S}} \otimes \mathcal{H}_{\mathcal{S}\setminus\mathcal{R}}\right)$ is given by

$$S * R = \mathrm{tr}_{\mathcal{R}\cap\mathcal{S}}\left[\left(\mathbb{1}_{\mathcal{S}\setminus\mathcal{R}} \otimes R^{T_{\mathcal{R}\cap\mathcal{S}}}\right)\left(S \otimes \mathbb{1}_{\mathcal{R}\setminus\mathcal{S}}\right)\right]\,, \tag{2.25}$$

where $\mathcal{R} \cap \mathcal{S}$ is the intersection of the two sets of systems $\mathcal{R}$ and $\mathcal{S}$, and $\mathcal{S} \setminus \mathcal{R}$ is the set whose elements belong to $\mathcal{S}$ but not to $\mathcal{R}$, i.e., $\mathcal{S} \setminus \mathcal{R} = \{i \in \mathcal{S},\ i \notin \mathcal{R}\}$. Two extreme scenarios arise from this product. If $\mathcal{S} \cap \mathcal{R} = \emptyset$, that is, the two quantum networks act on different sets of systems, the link product becomes the tensor product:

$$S * R = R \otimes S\,. \tag{2.26}$$

On the contrary, if the sets where they act are, indeed, the same: the outputs of R are the inputs of S; it becomes the generalized Born rule in Eq. (2.24), for example in Fig. 2.1d. This formalism will play a central role in Chapter 6.

2.5 Quantum state and channel identification

Quantum state and channel discrimination are fundamental tasks in quantum information theory. In state discrimination tasks, a system is prepared in one of several known states, each occurring with a given prior probability. The goal is to design a measurement strategy that identifies the state, optimizing performance according to a chosen objective function. A closely related task is that of quantum channel identification, which follows the same general principles applied to quantum operations.

The choice of figure of merit—that is, the operational quantity one seeks to optimize—depends crucially on the specific discrimination task. This thesis focuses on two natural criteria. The first criterion aims to minimize the probability of error, namely the likelihood that the measurement strategy yields an incorrect guess. This defines the framework of minimum-error discrimination. The second imposes a stricter requirement: no incorrect decisions are allowed. Instead, the protocol accepts inconclusive outcomes, and the goal becomes minimizing their occurrence. The protocols that fulfill no-error conditions are termed zero-error (or unambiguous). Although this thesis explores discrimination tasks under these instances, the reader can find other approaches as maximum confidence in [CAB+06] or error margins in [HHH08; SBC+13].

The remainder of this section provides a guided introduction to quantum state discrimination. It begins with the binary case—discrimination between two quantum states—under both minimum-error and zero-error criteria. The discussion then extends to the more general setting of multiple hypotheses and arbitrary figures of merit. In the case of pure states, I show how the problem can be recast in terms of the Gram matrix associated with the ensemble. Finally, I turn to the less explored domain of quantum channel identification, comparing the role of sequential versus

parallel discrimination paradigms.

2.5.1 Two states

The pioneering Helstrom's work [Hel69], established the mathematical framework for minimum-error discrimination and derived the optimal measurement in the binary case: for two known density operators ρ_0, ρ_1 with priors η_0, η_1, the minimum-error probability is

$$P_e = \frac{1}{2}\Big(1 - \big|\big|\eta_0\rho_0 - \eta_1\rho_1\big|\big|_1\Big)\,, \tag{2.27}$$

where $||\cdot||_1$ denotes the trace norm. The proof for this result is elegant and constructive, meaning that the optimal protocol arises from the proof itself.

The error probability for a fixed measurement $\{E_0, E_1\}$ is given by

$$P_e = \eta_0\,\mathrm{tr}(\rho_0\,E_1) + \eta_1\,\mathrm{tr}(\rho_1\,E_0)\,. \tag{2.28}$$

Because this problem is a binary discrimination task, we can use the completeness relation (2.11) to write all in terms of E_0—analogously, it can be done for E_1, simplifying the error probability as

$$P_e = \eta_1 + \mathrm{tr}[(\eta_0\rho_0 - \eta_1\rho_1)\,E_0]\,. \tag{2.29}$$

The matrix $Q := \eta_0\rho_0 - \eta_1\rho_1$ diagonalizes as

$$Q = \sum_i q_i\,|q_i\rangle\langle q_i| = Q_+ + Q_-\,, \tag{2.30}$$

where Q_+ and Q_- are the matrices obtained by restricting Q to the subspaces spanned by its non-negative and negative eigenvectors, respectively. Then, it is easy to see from Eq. (2.29) that the optimal protocol is given by the measurement

$$E_0^* = \Pi_-\,, \qquad E_1^* = \Pi_+\,, \tag{2.31}$$

where $\Pi_\pm$ are the projectors onto the subspaces of $Q_\pm$.

By substituting these optimal projectors into Eq. (2.29) (the optimal measurement elements), the optimal (minimal) error probability reads

$$P_e = \eta_1 + \mathrm{tr}(Q_-) = \eta_0 - \mathrm{tr}(Q_+)\,, \tag{2.32}$$

which combined read

$$P_e = \frac{\eta_0 + \eta_2}{2} - \frac{1}{2}\mathrm{tr}\big(Q_+ - Q_-\big) = \frac{1}{2}\Big(1 - \big|\big|Q\big|\big|_1\Big) \tag{2.33}$$

which is the well-known Helstrom formula in Eq. (2.27). For the particular case of pure states $|\psi_0\rangle$ and $|\psi_1\rangle$, the expression simplifies to

$$P_e = \frac{1}{2}\left(1 - \sqrt{1 - 4\eta_0\eta_1|\langle\psi_0|\psi_1\rangle|^2}\right) . \tag{2.34}$$

Throughout this thesis, I will usually denote the overlap between two states simply as $|\langle\psi_0|\psi_1\rangle| = c$. Notice that for $c = 0$, the states are orthogonal and the error probability is $P_e = 0$. In this case, one can always choose a measurement in which each state has support only on a single—different from the other—outcome, ensuring they are identified with certainty.

Another special scenario arises when the matrix $Q \geq 0$. In here, the optimal measurement elements are $E_0^* = \mathbb{1}$ and $E_1^* = 0$. In such cases—for example when $c = 1$—the optimal strategy reduces to a blind guess: selecting the most likely hypothesis based on its prior probability (e.g. for $\eta_0 \geq \eta_1$ one chooses ρ_0), without performing any measurement.

However, there might be some scenarios in which the possibility of committing an error is strictly unacceptable. For those specific discrimination tasks, the observer should work under the umbrella of *zero-error* discrimination. A zero-error POVM has to satisfy the no-error criteria, that is, whenever a decision is taken, the certainty that it is correct must be absolute,

$$\mathrm{tr}(\rho_i E_j) = 0 , \qquad \forall\, i \neq j . \tag{2.35}$$

To fulfill the completeness relation of a POVM, Eq. (2.11), an additional element is added, $E_?$. This is necessary because, when the states are not orthogonal, the zero-error constraints highly restrict the conclusive POVM elements, and they alone to not sum up to the identity, see Fig. 2.2. This element yields an inconclusive outcome, that is, the observer is not certain about which state is measuring. Zero-error tasks are impossible when at least one of the states in the prior ensemble is full-rank. To illustrate this scenario, I focus on two qubit pure states $|\psi_0\rangle$ and $|\psi_1\rangle$ [Iva87; Die88; Per88]. The POVM is constituted by $E_i = |\tilde{\varphi}_i\rangle\langle\tilde{\varphi}_i|$, with $i = 0, 1$, and $|\tilde{\varphi}_i\rangle$ a non-normalized state such that $\langle\tilde{\varphi}_i|\psi_j\rangle = \gamma_j\delta_{ij}$, and $E_? = \mathbb{1} - E_0 - E_1$, illustrated in 2.2. This POVM construction ensures that whenever E_0 clicks, the observer is certain that the prepared state was $|\psi_0\rangle$, analogously for E_1, and $E_?$ yields an inconclusive answer. This inconclusive probability is given by

$$P_? = \eta_0\mathrm{tr}(E_?|\psi_0\rangle\langle\psi_0|) + \eta_1\mathrm{tr}(E_?|\psi_1\rangle\langle\psi_1|) = 1 - \eta_0\gamma_0^2 - \eta_1\gamma_1^2 . \tag{2.36}$$

For qubits, since there is only one orthogonal projector to each state $\Pi_i^\perp = |\psi_i^\perp\rangle\langle\psi_i^\perp|$, the optimization consists of finding the maximal γ_i such that $E_? \geq 0$.

Solving the optimization task provides the optimal inconclusive probability

$$P_? = \begin{cases} 2\sqrt{\eta_0\eta_1}\, c & \text{if} \quad c \leq \sqrt{\frac{\eta_0}{\eta_1}} \leq \frac{1}{c}\,, \\ c^2 & \text{otherwise}\,, \end{cases} \tag{2.37}$$

where c is again the overlap between both states. If both states are equally probable, Eq. (2.37) reduces to $P_? = c$. In Section 3.1, I will relate this simple result to a more general result on zero-error discrimination for symmetric ensembles.

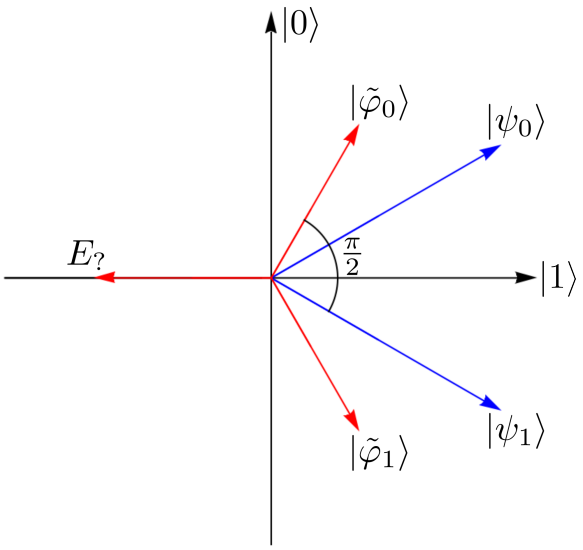


Figure 2.2: Optimal POVM for zero-error discrimination of two qubit states.

2.5.2 Multiple states

When the ensemble contains more than two states, the structure of optimal discrimination strategies becomes more complex. No closed formula exists for the optimal measurement for general ensembles, only for highly symmetric scenarios [Bar01; CDP+04]. Nevertheless, there exist optimality conditions for the POVMs. These optimality conditions, known as the Holevo-Yuen-Kennedy-Lax conditions, characterize the POVM that minimizes the average cost function, and were independently proposed by Holevo [Hol73] and by Yuen, Kennedy, and Lax [YKL75]. In the latter, they obtained the optimality conditions as a result of casting the discrimination task as a convex optimization problem. This approach will be extensively discussed in Section 3.1.

These conditions state that if a POVM with elements $\{E_i\}_i$ is optimal it must satisfy

$$\Gamma - W_i \geq 0\,, \tag{2.38a}$$

$$(W_i - \Gamma)E_i = E_i(W_i - \Gamma) = 0\,, \tag{2.38b}$$

where $\Gamma = \sum_i W_i\, E_i = \sum_i E_i\, W_i$, and $W_i = \sum_j f_{ij}\, \eta_j\, \rho_j$, with f_{ij} the cost function associated to infer hypotheses i when the actual state was ρ_j. For minimum-error, the objective function is $f_{ij} = \delta_{ij}$—it penalizes equally all erroneous answers.

However, solving the optimization task exactly is usually a hard problem. This is why it is interesting to seek protocols that are simple and close to optimality. One particularly studied protocol is the so-called square root measurement (SRM), or pretty good measurement [HW94; HJS+96]. This measurement has been proved to be optimal in various instances [EMV04; DP15], and is nearly optimal in many others, especially in multicopy settings when the number of copies of states available for measuring is large [SBC+16].

The POVM elements are given in terms of the elements of the ensemble as

$$E_i = \Lambda^{-1/2}\, \eta_i\, \rho_i\, \Lambda^{-1/2}\,, \tag{2.39}$$

where $\Lambda = \sum_i \eta_i \rho_i$.

In Chapter 5, I prove the SRM to be optimal for the discrimination task at hand, being one particular instance of the known optimal setting for the SRM. On the other hand, in Chapter 4, I show how the SRM performs near-optimally in a discrimination problem where it is not the exact optimal protocol. For both cases, the analysis was carried out using the *Gram matrix* tool that I next describe.

2.5.3 The Gram matrix

Many discrimination tasks involve ensembles of pure states. For such problems, a powerful mathematical tool is the *Gram matrix*. It encodes all the relevant information for a distinguishability task, as it captures how distinguishable the pure states in the ensemble are from each other. Given an ensemble consisting of pure states $\mathcal{E} = \{\eta_i, |\psi_i\rangle\}_i$, the Gram matrix entries are

$$G_{ij} = \sqrt{\eta_i\, \eta_j}\, \langle\psi_i|\psi_j\rangle\,. \tag{2.40}$$

Note how, by construction, this matrix is positive semidefinite.

It can be thought of as an operator acting on a Hilbert space, and being positive semidefinite, it can be decomposed as $G = X^\dagger X$, with

$$X = \sum_{i,j} \sqrt{\eta_j}\, \langle i|\psi_j\rangle |i\rangle\langle j|\,, \tag{2.41}$$

where X is a matrix whose column entries are just the states written in some basis $\{|j\rangle\}$. This orthonormal basis defines a POVM whose span is the same as that of the ensemble $\mathcal{E}$. With this Gram matrix decomposition, in the minimum-error scenario, the success probability reads

$$P_{\mathrm{s}} = \sum_j \eta_j \, |\langle \psi_j | j \rangle|^2 = \sum_j |X_{kk}|^2 \, . \tag{2.42}$$

Notice, however, that this decomposition is not unique—there are infinitely many, as

$$G = X^\dagger X = X^\dagger U^\dagger U X \, , \tag{2.43}$$

for any unitary U. The maximization of the success probability passes through finding the optimal unitary.

There is a particular decomposition of G in which $X = X^\dagger = S$, this is the so-called *square root* decomposition. The success probability for this decomposition corresponds to that of the SRM protocol, already discussed in Section 2.5.2. The performance of this protocol is close to optimality in several scenarios, and in 2015, Dalla Pozza and Pierobon [DP15] proved that this measurement is optimal if and only if the diagonal entries of the square root of the Gram matrix are all equal $S_{ii} = S_{jj}$ for all i and j. To relate the SRM in the Gram-matrix formalism with the form of the POVM elements given by Eq. (2.39), consider the operators

$$M = \sum_k |m_k\rangle\langle k| \, , \qquad \Psi = \sum_k \sqrt{\eta_k} \, |\psi_k\rangle\langle k| \, . \tag{2.44}$$

The SRM elements satisfy $|m_k\rangle = \Lambda^{-1/2}|\psi_k\rangle$, which implies $\Psi = \Lambda^{1/2} M$. Choosing the specific Gram decomposition $X := M^\dagger \Psi$, one obtains

$$X = M^\dagger \Psi = M^\dagger \Lambda^{1/2} M = X^\dagger = S \, , \tag{2.45}$$

where the second equality is obtained by substituting Ψ, and the third equality follows from the Hermiticity of $\Lambda^{1/2}$. In this way, the SRM elements in Eq. (2.39) uniquely specify the square root of the Gram matrix, S.

2.5.4 Quantum channel identification

The notion of discrimination extends beyond quantum states to arbitrary quantum channels. In quantum channel discrimination, one is given an unknown quantum operation drawn from a known set and aims to identify it. A unifying formalism for process discrimination is provided by quantum testers, which act on Choi

representations of channels. The ultimate success probability of distinguishing two channels Φ_0, Φ_1 with prior probabilities η_0, η_1 is given by an analog of the Helstrom formula:

$$P_e = \frac{1}{2}\left(1 - \left|\left|\eta_0\Phi_0 - \eta_1\Phi_1\right|\right|_\diamond\right), \tag{2.46}$$

where $|\cdot|_\diamond$ is the diamond norm, also known as the completely bounded trace norm, introduced in [AKN98]. This norm captures the maximum distinguishability between quantum channels, accounting for the fact that an observer may apply the channel to a subsystem of a larger entangled input state. Including an ancillary system—entangled with the input but untouched by the channel—can significantly enhance discrimination performance, although in some specific tasks, the inclusion of ancillary systems is not necessary to achieve optimal performance [Ací01].

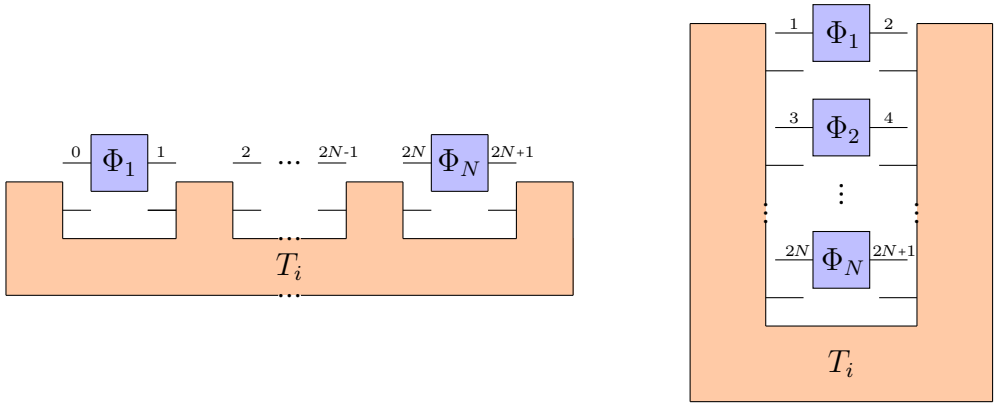
For a more general scenario in which multiple hypotheses exist—as with state discrimination—no general solution is known [PLL+19]. In analogy with the state scenario, one needs a quantum protocol that optimizes an objective cost function. In this setting, instead of using POVMs, one should account for more general quantum strategies. As discussed in Section 2.4, the objects to discriminate between quantum operations are quantum testers: for an ensemble of possible quantum processes $\{\eta_i, \Phi_i\}_i$ and a valid set of testers $\{T_i\}_i$, the success probability of correctly identifying the state reads

$$P_\mathrm{s} = \sum_i \eta_i \operatorname{tr}\left(J_{\Phi_i} T_i^T\right), \tag{2.47}$$

where C_{Φ_i} is the Choi matrices of the channel Φ_i, and the testers $\{T_i\}_i$ are constrained to Eqs. (2.19-2.23).

While Eq. (2.46) only accounts for a single use of the channel to identify, the optimization through testers is used for more convoluted scenarios, for example in multi-use settings. An example might be when several uses of the unknown channel are allowed, or when sequences of channels form the quantum operation to identify. In these scenarios, the identification protocols split into two different kinds: parallel and sequential.

Parallel identification can be thought of as the preparation of a composite state that evolves through the channels, and it has the same number of input systems as independent channels (see Fig. 2.3b). After the evolution, the whole quantum state is measured. On the other hand, sequential strategies (depicted in Fig. 2.3a) admit more involved operations, and contain the subclass of parallel strategies [HHL+10; BMQ21]. There exist other strategies that have no well-defined causal order. However, such protocols fall outside the scope of deterministic quantum circuits [CDP+13] and are not considered in this thesis.



(a) (b)

Figure 2.3: Diagrammatic representation of (a) sequential and (b) parallel quantum testers.

2.6 Universal discrimination

I finish this chapter by presenting the concept of *universal* discrimination, which plays a central role in a sizeable part of this thesis.

Whether for state or channel discrimination, the vast majority of current literature focuses on optimizing identification protocols when the set of hypotheses is fully characterized. Although this approach is suitable for certain cases and offers a benchmark for more agnostic scenarios, general tasks may not have this prior knowledge about the hypotheses.

In contrast, universal discrimination protocols aim to work for unknown states or processes. Of course, in the complete absence of prior knowledge, performance cannot exceed random guessing. To remain operationally meaningful, universal discrimination tasks assume some underlying structure or symmetry in the hypotheses.

What distinguishes universal protocols is that they are not tailored to any specific description of the states or channels under inspection. Instead, they rely solely on the structural features, such as repetitions among multiple subsystems. This makes them fundamentally different from the usual protocols: rather than optimizing over a known set of hypotheses, universal discrimination addresses the situation where the actual form of the states or channels is unknown, focusing only on their collective properties. In this sense, universal protocols provide a natural framework for verification under ignorance. By focusing on the aspects

that remain invariant for any specific form of the hypotheses, it is possible to obtain protocols that are robust under ignorance. This concept highlights the central role of symmetries and structural assumptions in nontrivial discrimination tasks and forms a cornerstone of the methods developed in this thesis.

In the following chapters, I will present three settings that all fall within the framework of multi-hypothesis discrimination, together with optimal protocols—designed to operate in a fully agnostic manner. In Chapter 4, two distinct domains composed of unknown states are considered. The goal is to identify the boundary between these domains using a protocol that is not tailored to the specific states defining them, but rather uses the differences across the two regions. In Chapters 5 and 6, a collection of quantum devices is assumed to operate according to a certain expected behavior, but some faults may occur. The protocol is designed to detect these deviations by leveraging the structural information of the entire system and the changes induced by such faults.

3

Mathematical toolkit

Symmetry, as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty and perfection.

Hermann Weyl

In this chapter, I gather most of the relevant mathematical tools used in the rest of the thesis. I begin with semidefinite programming (SDP), the natural way to write discrimination tasks as optimizations over positive operators, and I outline the primal and dual formulations in relevant quantum discrimination problems.

Next, I review representation theory, with a focus on the symmetry methods that support universal protocols, specifically Schur–Weyl and mixed Schur–Weyl dualities. These tools, together with Haar averages, reveal the block structure arising in Hilbert spaces when treating universal discrimination tasks.

Finally, I introduce association schemes—focusing on the Johnson scheme—as the mathematical objects controlling the structure of the Gram matrices in certain symmetric scenarios. In such problems, association schemes offer a complete characterization of optimal protocols and performances of POVMs.

3.1 Semidefinite programming

Quantum information theory heavily relies on linear algebra, with positive semidefinite matrices playing a central role. Quantum states, measurements, channels, and

testers are represented by positive semidefinite operators. Discrimination tasks consist of optimizing a cost function over these operators. It is natural—and very convenient—to use *SDP* [VB96] to approach the problems presented in this thesis.

In this section, I introduce the formalism of SDP, establish its connection to quantum discrimination tasks, and provide insights on its application in identification techniques.

3.1.1 Primal program

A SDP is an optimization problem over Hermitian operators acting on a complex vector space. The goal is to maximize—or minimize—a linear function of an operator variable X, subject to linear constraints. The form of the *primal* SDP used throughout this work is:

$$\max_{X} \quad \operatorname{tr}(AX) \tag{3.1}$$

$$\text{subject to} \quad \Phi(X) = B\,, \tag{3.2}$$

$$\Gamma(X) \leq C\,. \tag{3.3}$$

where A, B and C are Hermitian operators, and $\Phi(\cdot)$ and $\Gamma(\cdot)$ are Hermiticity-preserving linear maps. These maps encapsulate the linear constraints that the entries of X must satisfy. This form is widely used in quantum information applications [SC23]. However, alternative but equivalent formulations also exist. One replaces inequalities with equality constraints by introducing *slack* variables [Wat18]. Another, proposed in [VB96], differs substantially from the one adopted in this work and will not be discussed further.

One can easily recast the program and eliminate the inequalities by introducing a new variable $X' = X \oplus Z$, where Z is a slack variable, and convert the SDP to

$$\max_{X} \quad \operatorname{tr}(A'X') \tag{3.4}$$

$$\text{subject to} \quad \Phi'(X') = B\,, \tag{3.5}$$

$$\Gamma'(X') = C\,, \tag{3.6}$$

with $A' = A \oplus 0$, where 0 is the zero-matrix in the space spanned by Z. The new maps $\Phi'(X') = \Phi(X)$ and $\Gamma'(X') = \Gamma(X) + Z$ act only on the subspace spanned by A.

An operator variable X that satisfies the constraints is called *feasible*, and the set of such X is the *feasible set*, denoted by $\mathcal{A}$. If the constraints are strictly satisfied—that is, $\Phi(X) = B$ and $\Gamma(X) < C$—for some operator, the program is said to be *strictly feasible*.

The solution of the optimization problem is denoted by α and receives the name of *optimal primal value*

$$\alpha = \sup_{X \in \mathcal{A}} \operatorname{tr}(AX) \ . \tag{3.7}$$

Notice the distinction between the optimization *problem* and the *optimal value*. In the program, "max" indicates the search for a feasible X that attains the optimum. If the supremum is attained by some feasible X^*, then the maximum exists and $\alpha = \operatorname{tr}(AX^*)$; otherwise, no feasible point attains α. Such non-attainment can occur, for example, when the feasible set is not compact. An explicit instance appears in Chapter 6, where the discrimination task has no maximizer, yet its value can still be characterized, and standard SDP duality tools remain applicable.

3.1.2 Dual program

Every SDP has an associated *dual program*, which can be derived via the Lagrangian method, where new variables—called *dual variables*—play the role of Lagrange multipliers associated with each constraint. Consider the primal program of Eqs. (3.1–3.3), its associated Lagrangian takes the form:

$$\mathcal{L}(X, Y, Z) = \operatorname{tr}(AX) + \operatorname{tr}\Big(Y\big(B - \Phi(X)\big)\Big) + \operatorname{tr}\Big(Z\big(C - \Gamma(X)\big)\Big) , \tag{3.8}$$

where Y and $Z \geq 0$ are Hermitian operators acting as the dual variables. For any primal feasible X (so that $B - \Phi(X) = 0$ and $C - \Gamma(X) \geq 0$), the second term in Eq. (3.8) vanishes and the third is non-negative because $Z \geq 0$. Therefore

$$\mathcal{L}(X, Y, Z) \geq \operatorname{tr}(AX) , \tag{3.9}$$

that is, the Lagrangian upper-bounds the primal objective for *every* feasible X.

Collecting the terms involving X in Eq. (3.8), one has

$$\mathcal{L}(X, Y, Z) = \operatorname{tr}(YB) + \operatorname{tr}(ZC) + \operatorname{tr}\Big(X\big(A - \Phi^\dagger(Y) - \Gamma^\dagger(Z)\big)\Big) , \tag{3.10}$$

where $\Phi^\dagger(\cdot)$ and $\Gamma^\dagger(\cdot)$ are the dual maps of $\Phi(\cdot)$ and $\Gamma(\cdot)$, respectively, defined such that

$$\operatorname{tr}(\Phi(X)Y) = \operatorname{tr}(X\Phi^\dagger(Y)) , \quad \text{for all } X, Y. \tag{3.11}$$

Since the optimal primal value is $\alpha = \sup_X \operatorname{tr}(AX)$, any useful bound must upper bound the primal objective function $\operatorname{tr}(AX)$ for every feasible X. This motivates

defining the dual function in terms of the supremum of the Lagrangian

$$g(Y,Z) := \sup_X \mathcal{L}(X,Y,Z) = \begin{cases} \mathrm{tr}(YB) + \mathrm{tr}(ZC) & \text{if} \quad \Phi^\dagger(Y) + \Gamma^\dagger(Z) = A \\ +\infty & \text{otherwise.} \end{cases} \tag{3.12}$$

One seeks to minimize the dual function to characterize the optimal value α, which leads to a minimization over the pair of dual variables (Y, Z). The dual function is finite if and only if

$$\Phi^\dagger(Y) + \Gamma^\dagger(Z) = A\,, \tag{3.13}$$

it defines the constraint of the dual program. Putting these elements together, the dual program reads

$$\min_{Y,Z} \quad \mathrm{tr}\,(BY) + \mathrm{tr}\,(CZ) \tag{3.14}$$

$$\text{subject to} \quad \Phi^\dagger(Y) + \Gamma^\dagger(Z) = A\,, \tag{3.15}$$

$$Z \geq 0\,. \tag{3.16}$$

The set of feasible solutions of the dual program is denoted by $\mathcal{B}$, and the *optimal dual value* is given by

$$\beta = \inf_{\{Y,Z\}\in\mathcal{B}} \mathrm{tr}\,(BY) + \mathrm{tr}\,(CZ)\ . \tag{3.17}$$

3.1.3 Weak and strong duality

A fundamental property of SDP is weak duality: for any primal feasible $X \in \mathcal{A}$ and dual feasible $(Y, Z) \in \mathcal{B}$, it holds that

$$\mathrm{tr}(AX) \tag{3.18}$$

$$= \mathrm{tr}\Big(\big(\Phi^\dagger(Y) + \Gamma^\dagger(Z)\big)X\Big) \tag{3.19}$$

$$= \mathrm{tr}\big(Y\Phi(X) + Z\,\Gamma(X)\big) \tag{3.20}$$

$$\leq \mathrm{tr}\big(YB + ZC\big)\,, \tag{3.21}$$

where the first relation is obtained by the equality constraint of the dual program Eq. (3.15), and the last inequality is given by the constraints of the primal program, Eqs. (3.2–3.3).

Taking the supremum over all $X \in \mathcal{A}$ and the infimum over all $(Y, Z) \in \mathcal{B}$, establishes that

$$\alpha \leq \beta\,. \tag{3.22}$$

A direct implication of weak duality is that every dual feasible operators $(Y, Z) \in \mathcal{B}$ gives an upper bound for α. Likewise, any primal feasible operator $X \in \mathcal{A}$ establishes a lower bound for β.

If it happens that for $X \in \mathcal{A}$ and $(Y, Z) \in \mathcal{B}$, one finds $\mathrm{tr}(AX) = \mathrm{tr}(BY + CZ)$, then

$$\alpha = \beta\,, \tag{3.23}$$

and both X and (Y, Z) are optimal for the primal and dual, respectively. This relation is called *strong duality*.

Unlike weak duality, strong duality is not always guaranteed to hold. Nevertheless, the following theorem establishes a sufficient condition for strong duality to hold.

Theorem 3.1. (Slater's condition [Sla14]) *For every SDP the following statements hold*

- *If $\mathcal{A} \neq \emptyset$ and there exists a set of dual operators that are strictly feasible, i.e. $Z > 0$ and $\Phi^\dagger(Y) + \Gamma^\dagger(Z) = A$, then $\alpha = \beta$ and there exists a primal optimal operator X such that $\mathrm{tr}(AX) = \alpha$.*
- *If $\mathcal{B} \neq \emptyset$ and there exists a strictly feasible primal operator, i.e. $\Phi(X) = B$ and $\Gamma(X) < C$, then $\alpha = \beta$ and there exists a dual optimal set of operators (Y, Z) such that $\mathrm{tr}(BY + CZ) = \beta$.*

Notice how the existence of a primal (dual) optimal does not imply the existence of a dual (primal) optimal operator, but still strong duality holds. As discussed earlier, this is due to the definition of α and β with the supremum and infimum, respectively, as strong duality is more than just the existence of optimal operators. Strong duality states that the optimal values of both primal and dual programs are equal, independently of whether their solutions are attainable or not.

A direct corollary follows from Theorem 3.1

Corollary 3.1.1. *If both the primal and dual programs are strictly feasible, then strong duality holds $\alpha = \beta$, and both solutions are attainable.*

3.1.4 Complementary slackness

When strong duality holds and the optimum is attained, a key condition known as *complementary slackness* characterizes the relationship between primal and dual solutions.

Suppose X^* is an optimal solution to the primal problem, and strong duality holds. Then the Lagrangian becomes:

$$\mathcal{L} = \mathrm{tr}(AX^*) + \mathrm{tr}\Big(Z\big(C - \Gamma(X^*)\big)\Big)\,. \tag{3.24}$$

Since both Z and $C - \Gamma(X)$ are positive semidefinite, the Lagrangian is equal to the optimal primal value α if and only if

$$\mathrm{tr}\Big(Z\big(C - \Gamma(X^*)\big)\Big) = 0\,. \tag{3.25}$$

Which implies that Z and $C - \Gamma(X^*)$ have orthogonal supports. This orthogonality relation gives practical insight: once a candidate X is found, the slackness condition helps to construct a matching dual variable Z; and from Eq. (3.15)

$$\Phi^\dagger(Y) + \Gamma^\dagger(Z) = A\,, \tag{3.26}$$

and Z one can construct a matching Y. If the dual duple (Y, Z) outputs the same value as the primal objective, the problem is solved, as strong duality holds ($\alpha = \beta$) and the solutions for both versions of the problem are optimal.

3.1.5 Discrimination tasks as SDP

Having introduced the general framework of SDP, I now turn to its application in quantum discrimination tasks. These problems can be naturally written as SDPs, as they often involve optimizing over positive semidefinite operators subject to linear constraints.

First, for quantum state discrimination, one has an ensemble of quantum states $\{\rho_i, \eta_i\}_i$ with prior probabilities η_i, and aims to construct a measurement $\{E_i\}_i$ that maximizes a general cost-function associated with the identification of the states. This optimization can be written as an SDP in the following form

$$\max_{\{E_i\}} \quad \sum_i \mathrm{tr}(W_i E_i) \tag{3.27}$$

$$\text{subject to} \quad E_i \geq 0, \quad \text{for all} \quad i \tag{3.28}$$

$$\sum_i E_i = \mathbb{1} \tag{3.29}$$

where the E_i are the POVM elements, and recall that $W_i = \sum_j \eta_j f_{ij} \rho_j$ plays the role of the objective operator (A) in the standard form of an SDP. The associated

Lagrangian for this SDP is

$$\mathcal{L} = \sum_i \mathrm{tr}(W_i E_i) + \sum_i \mathrm{tr}(E_i Z_i) - \mathrm{tr}\Big(Y\Big(\sum_i E_i - \mathbb{1}\Big)\Big) \tag{3.30}$$

$$= \mathrm{tr}\, Y + \sum_i \mathrm{tr}(E_i(W_i + Z_i - Y))\,. \tag{3.31}$$

Therefore, the dual version of the SDP is of the form

$$\min_Y \quad \mathrm{tr}\, Y \tag{3.32}$$

$$\text{subject to} \quad Y - W_i \geq 0 \quad \text{for all} \quad i\,. \tag{3.33}$$

Note that the variables Z_i drop out, as they only enforce semidefinite positivity. Notice also how the constraint of the dual SDP corresponds to the first Holevo-Yuen-Kennedy-Lax condition, Eq. (2.38a). The second condition, Eq. (2.38b), is obtained via complementary slackness: for the Lagrangian in Eq. (3.31) to be equal to the optimal value of the primal $\mathrm{tr}(E_i Z_i) = 0$, substituting $Z_i = Y - W_i$ one has $\mathrm{tr}(E_i(Y - W_i)) = 0$ and therefore

$$E_i(Y - W_i) = (Y - W_i)E_i = 0\,. \tag{3.34}$$

A particularly relevant instance of discrimination tasks arises when the figure of merit is minimum-error. For this protocol, primal and dual SDPs read

Minimum-error

Primal program	Dual program
$\max_{\{E_i\}} \quad \sum_i \eta_i \mathrm{tr}(\rho_i E_i)$	$\min_Y \quad \mathrm{tr}\, Y$
$\text{s.t.} \quad \sum_i E_i = \mathbb{1}$	$\text{s.t.} \quad Y - \eta_i \rho_i \geq 0,\ \forall i$
$E_i \geq 0,\ \forall i$	$Y = Y^\dagger$

Another important instance of quantum state discrimination is the unambiguous protocol. This corresponds to designing a POVM $\{E_i\}$ with one element for each state and an additional element for the inconclusive outcome, such that

$$\mathrm{tr}(E_i \rho_j) = 0\,, \quad \text{for all} \quad i \neq j\,. \tag{3.35}$$

This condition ensures that no erroneous decisions are drawn from the measurement. The two versions of the problem read

Zero-error

Primal program		Dual program	
$\max_{\{E_i\}}$	$\sum_i \eta_i \mathrm{tr}(\rho_i E_i)$	$\min_{Y}$	$\mathrm{tr}\, Y$
s.t.	$\sum_i E_i \leq \mathbb{1}$	s.t.	$Y - \eta_i \rho_i + \sum_{j \neq i} \nu_{ij} \rho_j \geq 0,\ \forall i$
	$E_i \geq 0\,,\ \forall i$		$Y = Y^\dagger$
	$\mathrm{tr}(\rho_i E_j) = 0\,,\ i \neq j$		$\nu_{ij} \in \mathbb{R}$

As described in Section 2.5, for the special case of an ensemble consisting of pure states, the POVM elements admit the form of unnormalized pure states $E_i = |\tilde{\varphi}_i\rangle\langle\tilde{\varphi}_i|$ such that $|\langle\psi_i|\tilde{\varphi}_j\rangle|^2 = \gamma_i \delta_{ij}$. By imposing this specific structure of the measurement in the SDP, the program simplifies as follows.

For the first constraint, by multiplying both terms of the inequality by $X^\dagger$ on the left, and by X on the right, one obtains $\Gamma_D \leq G$, where recall that $X = \sum_{ij} \sqrt{\eta_i} |\psi_i\rangle\langle j|$, and $G = X^\dagger X$. The matrix Γ_D is the diagonal matrix with $\{\gamma_i\}$ as its elements. The second constraint is just that all γ_i are positive, or just that $\Gamma_D \geq 0$. Lastly, the zero-error constraint is already guaranteed by choosing the POVM elements of this form. With this transformation, the objective function is a maximization over coefficients $\{\gamma_i\}_i$: $\max_{\{\gamma_i\}} \sum_i \gamma_i$, and the two versions of the SDP read

Zero-error (pure states)

Primal program		Dual program	
$\max_{\Gamma}$	$\mathrm{tr}\, \Gamma$	$\min_{X}$	$\mathrm{tr}(XG)$
s.t.	$\Gamma_D \leq G$	s.t.	$X_D \geq \mathbb{1}$
	$\Gamma \geq 0$		$X \geq 0$

This formulation has two main advantages. The first is that it reduces the dimensionality of the task, only focusing on the subspace of the Hilbert space spanned by the ensemble. This allows for running efficient SDP solvers. The second advantage is even more significant: while the dual formulation of the SDP for unambiguous discrimination is not attainable (see Chapter 6), it becomes attainable in the simplified case of pure states.

To discriminate between quantum channels, similar principles apply. These tasks can be cast as SDP, but the constraints on the variables differ from those

of POVMs, recall Eqs. (2.19-2.23). For the case of minimum-error, the analog for channel discrimination reads

Minimum-error

Primal program		Dual program	
$\max_{\{T_i\}}$	$\sum_i \eta_i \mathrm{tr}(C_i T_i^T)$	$\min_{\{y,\, Y^{(N)}\}}$	y
s.t.	$\sum_i T_i = \mathbb{1}_{2N} \otimes R^{(N)}$	s.t.	$yY^{(N)} - \eta_i C_i^T \geq 0,\ \forall i$
	$\mathrm{tr}_{2n-1} R^{(n)} = \mathbb{1}_{2n-2} \otimes R^{(n-1)}$		$\mathrm{tr}_{2n} Y^{(n)} = \mathbb{1}_{2n-1} \otimes Y^{(n-1)}$
	$\mathrm{tr}_1 R^{(1)} = 1$		$\mathrm{tr}_2 Y^{(1)} = \mathbb{1}_1$
	$T_i \geq 0,\ \forall i$		$Y^{(N)} = Y^{(N)\dagger}$

The elegant formulation of the dual version shows that the dual variable $Y^{(N)}$ is a quantum comb that shares the same structure as the hypotheses $\{C_i\}_i$ to be distinguished and stands as the counterpart of the testers $\{T_i\}_i$, showing the interplay between both mathematical structures, shown in Fig. 2.1. Needless to say that one can recover the simpler state discrimination task from this formulation, as quantum networks constitute the most general mathematical object to describe quantum operations.

3.2 Schur-Weyl dualities

The optimality of certain quantum discrimination protocols often depends not on the specific states involved, but on their symmetry properties. This motivates a study of representation-theoretic tools and algebraic structures that involve such symmetries.

In this section, I formalize the concept of universal discrimination, a central theme throughout the thesis. Universal protocols are those that perform effectively for any set of quantum states that follow a certain underlying structure, regardless of their specific form. To study such scenarios, one requires tools that are basis-independent. In this thesis, I use Schur-Weyl duality to devise such protocols for universal discrimination tasks.

The section is organized as follows. I begin by reviewing basic concepts from representation theory following [FH13; Sag01], focusing on group actions on tensor product spaces. I then introduce Schur-Weyl duality and the decomposition of Hilbert spaces into irreducible components. The relationship between this decomposition and quantum state discrimination is highlighted next, in which the

integration with respect to the Haar measure plays a central role. For a mathematical detailed work on Haar integration, see [CŚ06], for a more quantum-information view, see [Mel24], and for a symbolic implementation, see [PM17]. Finally, I sketch how Schur-Weyl duality extends to quantum channels, via the so-called Mixed Schur-Weyl duality, which will have its own introduction and will be later discussed in Chapter 6.

3.2.1 Representations of finite groups

Let $\mathcal{G}$ be a finite group and V a complex vector space. A representation of $\mathcal{G}$ in V is a group homomorphism $R : \mathcal{G} \to \mathrm{GL}(V)$, where $\mathrm{GL}(V)$ denotes the group of invertible linear operators on V.

A *subrepresentation* W of a representation is a vector subspace of V which is invariant under the action of $\mathcal{G}$. A representation is called *irreducible* if it has no trivial invariant subspaces, i.e., its only invariant subspaces are $\{0\}$ and itself. Hereafter, the term *irrep* will be used as shorthand for irreducible representation.

Given two representations V and W, both their direct sum $V \oplus W$ and tensor product $V \otimes W$ are also representations.

A fundamental principle in representation theory is the Schur's lemma, which states

Theorem 3.2. (Schur's Lemma [Sch05]) *Let $\varphi : V \to W$ be a linear map between irreps of a group $\mathcal{G}$, respecting its action. Then:*

- *If $V \ncong W$, then $\varphi = 0$.*
- *If $V \cong W$, then $\varphi = \kappa \mathbb{1}$, for some $\kappa \in \mathbb{C}$.*

Here $V \cong W$ denotes that the irreps V and W are isomorphic.

This theorem will later explain the block-diagonal structure induced by Schur-Weyl duality.

When the representation R acts on $V^{\otimes n}$, it induces the natural action

$$R(g)(v_1 \otimes v_2 \otimes \ldots \otimes v_n) = R(g)(v_1) \otimes R(g)(v_2) \otimes \ldots \otimes R(g)(v_n)\,. \tag{3.36}$$

Independently, a second group—the symmetric group S_n—acts naturally on $V^{\otimes n}$ by permuting the tensor products. For each element $\sigma \in S_n$, the associated linear operator P_σ acts as

$$P_\sigma(v_1 \otimes v_2 \otimes \ldots \otimes v_n) = v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \ldots \otimes v_{\sigma^{-1}(n)}\,. \tag{3.37}$$

Crucially, these two actions commute

$$P_\sigma \, R(g) = R(g) \, P_\sigma \,, \quad \text{for all } g \in \mathcal{G} \,, \sigma \in S_n \,. \tag{3.38}$$

The special case in which $\mathcal{G} = \mathrm{U}(d)$ acts on $V = \mathbb{C}^d$ defines the setting for *Schur-Weyl duality*. The fact that the actions of $\mathrm{U}(d)$ and S_n commute implies that each group lies in the commutant of the other. Schur-Weyl duality then states that the space $(\mathbb{C}^d)^{\otimes n}$ admits a decomposition into a direct sum of irreps of both groups simultaneously [FH13]. To make this structure explicit, it is necessary first to introduce the concepts of partitions and Young diagrams [You00; Ful97].

3.2.2 Partitions, Young diagrams & Schur-Weyl duality

A partition $\lambda \vdash n$ is a sequence of non-increasing integer numbers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ such that $\sum_{i=1}^{k} \lambda_i = n$. Each partition can be graphically represented by a Young diagram, which is a collection of n boxes organized in left-justified rows with row i containing λ_i boxes, as in Fig. 3.1.

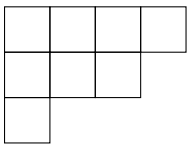


Figure 3.1: Young diagram corresponding to the partition $\lambda = (4, 3, 1)$.

These objects simultaneously label the irreps of both $\mathrm{U}(d)$ and S_n, provided that the number of non-zero elements of λ does not exceed d, i.e., the Hamming weight $|\lambda|$ satisfies $|\lambda| \leq d$. Specifically, the irrep of S_n labeled by λ is denoted by S^λ, and the one corresponding to $\mathrm{U}(d)$ is denoted by U_λ. Schur-Weyl duality [Wey46] then gives the decomposition:

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq d}} \mathrm{U}_\lambda \otimes \mathrm{S}^\lambda \,. \tag{3.39}$$

Throughout this thesis, the convention used when working in the Schur basis is that the first component in the tensor product corresponds to the action of the $\mathrm{U}(d)$ group, labeled with a subscript denoting the irrep. In contrast, the second corresponds to the action of S_n, labeled with a superscript.

The dimensions and multiplicity of the irreps U_λ and S^λ are also dual: the dimension of the irrep U_λ, denoted by d_λ is equal to the multiplicity of the irrep

S^λ, and viceversa, the multiplicity of U_λ, denoted by m_λ is equal to the dimension of S^λ. The convention in this thesis is that, whenever the terms "dimension" and "multiplicity" appear in the context of Schur-Weyl duality, they refer by default to the irrep U_λ.

To compute these dimensions, two combinatorial objects are particularly relevant. A semi-standard Young tableau (SSYT) of shape λ is a filling with numbers in $\{0, 1, \ldots, d-1\}$, in a non-decreasing way across rows and strictly increasing down columns. The number of such tableaux gives $d_\lambda = |\mathrm{SSYT}(\lambda, d)|$, examples of such diagrams can be seen in Fig. 3.2.

0	0	0
1	1	

0	0	1
1	1	

Figure 3.2: The two SSYT for partition $\lambda = (3, 2)$. The dimension of the irrep $d_{(3,2)} = 2$.

On the other hand, a standard Young tableau (SYT) of shape λ is a filling of the boxes of the Young diagram with the integers $\{1, 2, \ldots, n\}$, such that entries increase strictly across each row and down each column. The number of SYTs corresponds to $m_\lambda = |\mathrm{STY}(\lambda)|$, depicted in Fig. 3.3.

1	2	3
4	5	

1	2	4
3	5	

1	2	5
3	4	

1	3	4
2	5	

1	3	5
2	4	

Figure 3.3: All five SYT for partition $\lambda = (3, 2)$. The multiplicity of the irrep $m_{(3,2)} = 5$.

These counts can also be obtained via Bratteli diagrams [Bra72]—also known as Young lattices—which provide a graphical way to organize all partitions $\lambda \vdash n$. In these diagrams, each node represents a partition, and directed edges connect partitions that differ by a single added box. The number of distinct paths from the empty partition to a given λ in this diagram equals $m_\lambda = |\mathrm{Paths}(\lambda)|$ (see Fig. 3.4).

Finally, explicit expressions for both m_λ and d_λ are given by the hook-length formulas [FRT54]. For any box (i, j)—that is in row i and column j—the hook length is denoted by $h_\lambda(i, j)$, defined as the number of boxes to the right in the

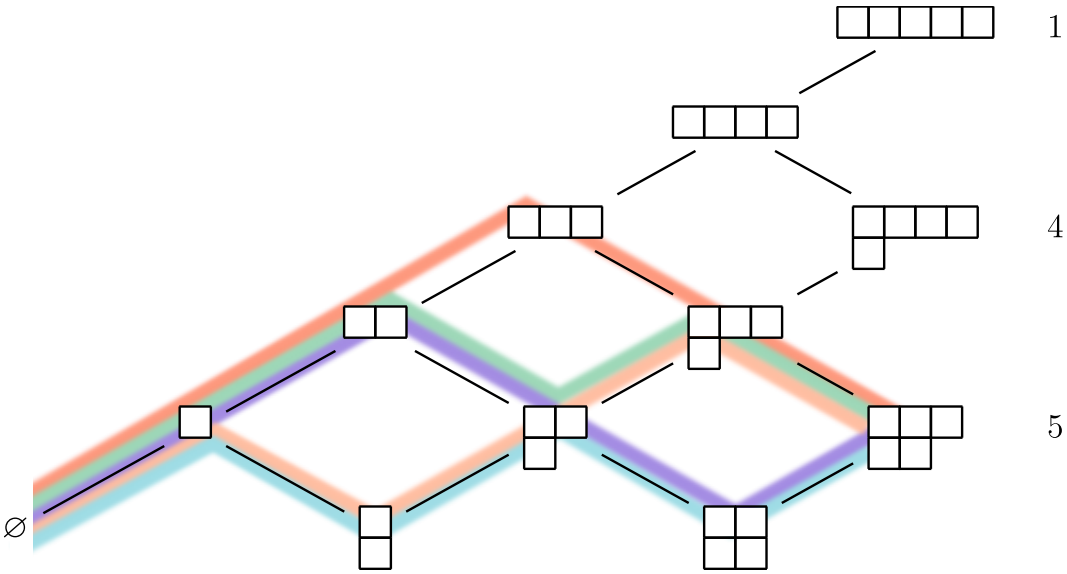


Figure 3.4: Bratteli diagram for $n = 5$ and $d = 2$. The multiplicities are shown next to each final Young tableaux. The example $\lambda = (3, 2)$, has multiplicity $m_{(3,2)} = 5$, as already shown in Fig. 3.3, depicted by color lines, representing each of the possible paths to arrive at partition $(3, 2)$.

same row and below in the same column, including the (i, j) box itself. Then, the multiplicity of the irrep labeled by λ is given by

$$m_\lambda = \frac{n!}{\prod_{(i,j)\in\lambda} h_\lambda(i,j)}\,, \tag{3.40}$$

and the dimension of the irrep, with at most d rows, by

$$d_\lambda = \prod_{(i,j)\in\lambda} \frac{d+j-i}{h_\lambda(i,j)}\,. \tag{3.41}$$

3.2.3 Schur-Weyl duality in quantum systems

Universal discrimination protocols are intended to work without access to the classical description of the states defining the hypotheses. However, some underlying structure must be present; in its absence, complete ignorance would make discrimination no better than random guessing. In this thesis, the ensembles considered

consist of strings formed by multiple copies of two pure states, arranged in a certain order,

$$\rho_{k,\sigma} = U_\sigma |\psi_0\rangle\langle\psi_0|^{\otimes(n-k)} \otimes |\psi_1\rangle\langle\psi_1|^{\otimes k} U_\sigma^\dagger \,, \tag{3.42}$$

that is, $n-k$ copies of state $|\psi_0\rangle$ and $k \leq n$ copies of $|\psi_1\rangle$. Here, in contrast with the previous section, the permutations are represented by unitary matrices U_σ, which act on the Hilbert space by reordering the tensor factors into a certain order. These states arise naturally in production processes of quantum states, as discussed in Chapter 1. Typically, $|\psi_0\rangle$ corresponds to the intended state, while $|\psi_1\rangle$ represents those generated by malfunctions in the production. In this context, the most general ensemble of hypotheses is given by $\{\rho_{k,\sigma}\}_{k,\sigma}$, where both the number of faults k and their position in the string, σ, may vary.

In the scenario where the forms of $|\psi_0\rangle$ and $|\psi_1\rangle$ are unknown to the observer, one could consider different figures of merit to quantify the performance of a universal protocol designed to identify the number and position of $|\psi_1\rangle$ states. In this thesis, the focus is on the average success probability of correctly identifying the hypotheses—where the average is taken over all possible choices of $|\psi_0\rangle$ and $|\psi_1\rangle$. The objective is then to construct a POVM $\{E_{k,\sigma}\}_{k,\sigma}$ that performs optimally on average, i.e., one POVM that maximizes

$$\int \mathrm{d}\psi_0 \mathrm{d}\psi_1 \mathrm{tr}\left(E_{k,\sigma}\rho_{k,\sigma}\right) \,, \tag{3.43}$$

where the outcome (k,σ) is associated with the state $\rho_{k,\sigma}$. Since the measurement element $E_{k,\sigma}$ is independent of the particular choice of ψ_0 and ψ_1, it remains unchanged under the integral. In addition, in the discrimination tasks considered in this thesis, the two states can be assumed to be independent, allowing the averages to be taken separately for each state.

It is at this point that representation theory and Schur-Weyl duality become relevant. All global states $\rho_{k,\sigma}$ are composed of pure states, and the average is taken with respect to the Haar measure. Then the goal is to identify what is the object

$$\int \mathrm{d}U \, U^{\otimes k} |\psi\rangle\langle\psi|^{\otimes k} U^{\dagger\otimes k} \,. \tag{3.44}$$

Here, the operator $U^{\otimes k}$, can be decomposed into irreps as $\bigoplus_\lambda U_\lambda$, and a corollary of Schur's lemma applies:

Corollary 3.2.1. *Let A be a linear map between irreps V and W, and consider*

$$A' = \frac{1}{|\mathcal{G}|} \sum_{g\in\mathcal{G}} R_W(g)^{-1} A \, R_V(g) \,. \tag{3.45}$$

Then,

- *If* $V \ncong W$, *then* $A' = 0$.
- *If* $V \cong W$, *then* $A' = \mathbb{1}_V/\dim(V)$.

This corollary extends to the continuous group $\mathrm{U}(d)$ by replacing the sum by the integration over the Haar measure.

The state $|\psi\rangle\langle\psi|^{\otimes k}$ has support only in the symmetric subspace of k parties, that is $\lambda = (k, 0)$, also denoted by the label "sym". Then, Schur's Lemma implies

$$\int \mathrm{d}U \, U^{\otimes k}|\psi\rangle\langle\psi|^{\otimes k}U^{\dagger\otimes k} = \frac{\mathbb{1}_k^{\mathrm{sym}}}{d_k^{\mathrm{sym}}} \,. \tag{3.46}$$

As a consequence, the average over the states $\rho_{k,\sigma}$ yields

$$\rho_{k,\sigma} = c_k U_\sigma \, \mathbb{1}_{n-k}^{\mathrm{sym}} \otimes \mathbb{1}_k^{\mathrm{sym}} \, U_\sigma^\dagger \,, \tag{3.47}$$

where c_k is a normalization constant that depends on the dimensions of the symmetric subspaces

$$c_k = \frac{1}{d_{n-k}^{\mathrm{sym}} \, d_k^{\mathrm{sym}}} \,, \tag{3.48}$$

with the dimensions given by the Hook-length formula

$$d_k^{\mathrm{sym}} = \binom{k+d-1}{d-1} \,. \tag{3.49}$$

In this particular case, the computation of the dimension of the symmetric subspace admits a simpler approach than applying the Hook-length formula in Eq. (3.41). Since the corresponding Young tableau consists of a single row, the rules for SSYTs require the entries to be non-decreasing along the row. Consequently, determining the number of SSYTs reduces to counting the number of ways to distribute k indistinguishable balls into d labeled boxes, that is, the "balls into boxes" or "stars and bars" problem, see [Sta11].

Notice that multiplicities do not appear in Eq. (3.48), and the normalization constant c_k depends only on the dimensions, as the multiplicities m_{n-k}, and m_k are equal to unity. This is easily seen by counting paths in Young lattices: there is only a single way to reach the fully symmetric Young diagram, as illustrated in Fig. 3.4.

The state $\rho_{k,\sigma}$ in Eq. (3.47) can be further simplified. The subspaces $\mathbb{1}_{n-k}^{\mathrm{sym}} \otimes \mathbb{1}_k^{\mathrm{sym}}$ are *reducible* representations in $(\mathbb{C}^d)^{\otimes n}$. When writing the state in the Schur basis—the basis in which the space has block-diagonal form—it takes the form

$$\rho_{k,\sigma} = c_k \bigoplus_{\substack{\lambda \vdash n \\ |\lambda| \leq 2 \\ \lambda_2 \leq k}} \mathbb{1}_\lambda \otimes \Omega_{k,\sigma}^\lambda \,, \tag{3.50}$$

where $\mathbb{1}_\lambda$ is the identity on the $\mathrm{U}(d)$ irrep labeled by partition λ, and Ω_k^λ encodes the information of the order of the tensor factors, that is, the corresponding representation of S_n.

Notice that the sum goes only over bipartitions—partitions $\lambda = (\lambda_1, \lambda_2)$ with at most two rows—regardless of the local dimension d. This restriction arises from the composition rules of Young tableaux: the subspaces in Eq. (3.48), $\mathbb{1}_{n-k}^{\mathrm{sym}}$ and $\mathbb{1}_k^{\mathrm{sym}}$, correspond to Young diagrams $(n-k, 0)$ and $(k, 0)$, respectively. These subspaces are isomorphic to the spaces of total spin $j = \frac{n-k}{2}$ and $j' = \frac{k}{2}$. When coupled, their tensor product cannot generate a component corresponding to three fully antisymmetric particles, namely $\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix}$, because the composition effectively involves only two symmetric systems.

Note also that the representation of the state in the part corresponding to $\mathrm{U}(d)$ is the identity, that is because by construction it commutes with every $U^{\otimes n} = \bigoplus_\lambda U_\lambda \otimes \mathbb{1}^\lambda$—or simply, because all information about this group has been erased by the integration. Because of this, the protocol is completely independent of the information of the specific states ψ_0 and ψ_1.

It is also worth noting that the operators Ω_k^λ are rank-1 projectors. That is, in the composition at hand, each irrep labeled by λ appears with only once. Equivalently, in the spin picture, when two spin components are combined, there is a unique way of obtaining each total spin, $J = \frac{\lambda_1 - \lambda_2}{2}$, from j and j'. This uniqueness ensures that Ω_k^λ projects onto a one-dimensional subspace, and is therefore rank one.

As a final comment, without loss of generality, if all states in an ensemble have the form in Eq. (3.50), the optimal POVM can also be chosen to respect the same block structure

$$E_{k,\sigma} = \bigoplus_\lambda \mathbb{1}_\lambda \otimes \mathcal{E}_{k,\sigma}^\lambda \,, \tag{3.51}$$

as for any POVM $\{E'_{k,\sigma}\}_{k,\sigma}$ one can find another POVM $\{E_{k,\sigma}\}_{k,\sigma}$ with

$$E_{k,\sigma} = \int \mathrm{d}U \; U^{\otimes n} E'_{k,\sigma} \, U^{\dagger \otimes n} \,, \tag{3.52}$$

that yields the same success probability, due to the cyclic property of the trace, and the fact that all states $\rho_{k,\sigma}$ of the form of Eq. (3.50) commute with the action of $U^{\otimes n}$.

3.2.4 Mixed Schur-Weyl duality

The classical Schur-Weyl duality relates the actions of the symmetric group and the general linear group on tensor powers of a vector space. This duality can be extended to more general settings. While I will not go into these mathematical

generalizations in full detail, I briefly introduce one such extension: *mixed* Schur-Weyl duality following [CDD+08; GBO23]. Mixed Schur-Weyl duality has gathered attention in recent years in the field of quantum information, see [BSH24; GBO23; GO24; Gri25; SMM+25]. In this thesis, this duality appears naturally in universal discrimination tasks of quantum channels, where both input and output spaces are involved, leading to a mixed tensor structure of vector spaces that I now describe. This will be explored in detail in Chapter 6.

Mixed Schur-Weyl duality involves a different algebraic structure known as the walled Brauer algebra [Bra37; Koi89; Tur90], denoted by $\mathfrak{B}_{r,s}$, which acts together with the general linear group $\mathrm{GL}(V)$ on a space of the form

$$V^{\otimes r} \otimes (V^*)^{\otimes s}\,, \tag{3.53}$$

where V is a complex vector space and V^* its dual.

The general linear group acts naturally on this mixed tensor product: on the V components via the usual representation $R(g)$, and on the V^* components via the dual representation $R^*(g)$, defined by $R(g^{-1})^T = R^*(g)$. In the context of quantum information, this dual action corresponds to $U^* = (U^\dagger)^T$, i.e., the complex conjugate of the unitary operator.

The second algebra appearing in the duality, the walled Brauer algebra, plays a role analogous to that of the symmetric group in the standard case. However, a distinction arises: while the representation of S_n is isomorphic to the group itself, the case for the walled Brauer algebra is different; when $d < r + s$, the algebra $\mathfrak{B}_{r,s}$ is not isomorphic to the algebra generated by the matrix representation of its elements $\mathcal{A}_{r,s}$, see [CDD+08; GBO23]. The elements of $\mathcal{A}_{r,s}$ consist of permutations $\sigma \in S_{r+s}$ that have been partially transposed on the V^* components. In some quantum information works, this has been referred to as the algebra of partially transposed permutations. An example of such an element,

$$A = \sum_{\substack{i,j,k=0 \\ i',j',k'=0}}^{d-1} |i',j',i',k'\rangle\langle i,j,j,k| \in \mathcal{A}^d_{2,2}\,, \tag{3.54}$$

is illustrated in Fig. 3.5.

Although I will not go into the full details of the walled Brauer algebra, a few key aspects relevant to its application are worth highlighting.

The composition rules for elements of $\mathfrak{B}_{r,s}$ closely resemble those of the symmetric group S_m. Diagrammatically, one can compose two elements by stacking them vertically and identifying the bottom row of the first diagram with the top row

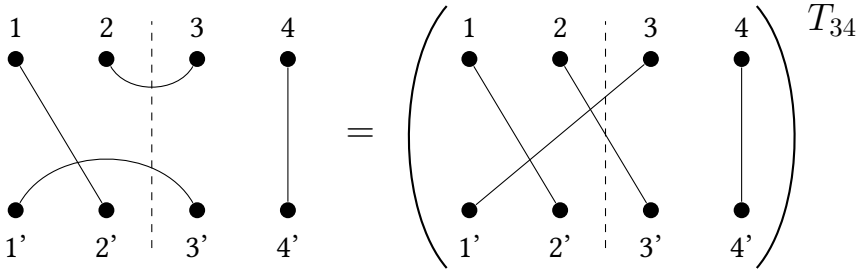


Figure 3.5: An example of an element A [c.f. Eq. (3.54)] of the walled Brauer algebra $\mathfrak{B}^d_{2,2}$, acting on $V^{\otimes 2} \otimes (V^*)^{\otimes 2}$. The dashed line separates the usual vector spaces from the dual vector spaces.

of the second diagram. Then, by concatenating these diagrams and following the corresponding lines, one obtains the resulting composition. That is, a first element of the algebra acts on a vector, and the second element acts on the result of the previous transformation (analogously to element composition for the symmetric group). However, an important distinction arises: whenever a closed loop is formed in the composition, the resulting diagram must be multiplied by a scalar factor equal to d, the dimension of the underlying vector space V. An example of such a composition is illustrated in Fig. 3.6.

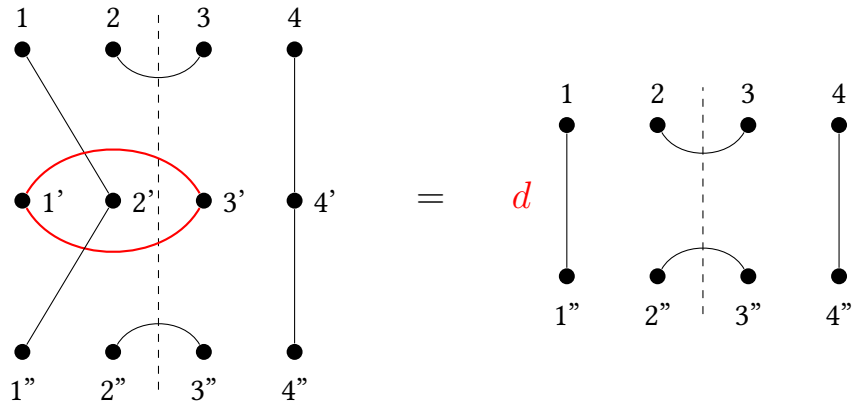


Figure 3.6: An example of an element of the walled Brauer algebra, acting on $V^{\otimes 2} \otimes (V^*)^{\otimes 2}$. The dashed line separates the usual vector spaces from dual vector spaces.

A contraction between two tensor factors corresponds to an element of $\mathfrak{B}_{r,s}$ in which both the top and bottom legs of two indices are connected—top-to-top

and bottom-to-bottom—as shown on the right-hand side of Fig. 3.6. In matrix representation, such a contraction is the operator $\sum_{i,j} |ii\rangle\langle jj|$. From the perspective of quantum information, this operator coincides with the Choi matrix of the identity channel:

$$J_{\mathrm{id}} = \sum_{i,j} |ii\rangle\langle jj| \,. \tag{3.55}$$

The fact that an element of this algebra coincides with the Choi representation of the identity channel already hints that the use of this tool is appropriate for working with quantum networks in symmetric scenarios, as discussed in Chapter 6.

Just as in standard Schur-Weyl duality, the tensor space $V^{\otimes r} \otimes (V^*)^{\otimes s}$ decomposes into a direct sum of irreps, one for each equivalence class under the joint action of $\mathrm{GL}(V)$ and the walled Brauer algebra.

$$(\mathbb{C}^d)^{\otimes r} \otimes (\mathbb{C}^{d^*})^{\otimes s} = \bigoplus_{\hat{\lambda} \in \hat{\mathfrak{B}}_{r,s}} \mathrm{U}_{\hat{\lambda}} \otimes \mathrm{B}^{\hat{\lambda}} \tag{3.56}$$

However, the labels for these irreducible components are no longer single partitions, but rather pairs of partitions

$$\hat{\lambda} = \{\lambda_L, \lambda_R\} \,, \tag{3.57}$$

where λ_L corresponds to the structure in the V tensor components and λ_R to the dual ones. The set of all such admissible pairs is denoted by $\hat{\mathcal{A}}_{r,s}$, and each pair satisfies

$$\sum_i (\lambda_L)_i = r \tag{3.58}$$

$$\sum_j (\lambda_R)_j = s \tag{3.59}$$

$$|\lambda_L| + |\lambda_R| \leq d \tag{3.60}$$

I conclude by noting that the mixed Schur-Weyl duality appears naturally in the universal discrimination of quantum channels, where both input and output spaces are involved, leading to the mixed tensor structure $V^{\otimes n} \otimes (V^*)^{\otimes n}$. This will be explored in detail in Chapter 6.

3.3 Association schemes

This final section introduces a class of algebraic-combinatorial structures known as association schemes. These schemes provide an abstract framework for describing

regular patterns of overlaps and symmetries between quantum states, and are particularly well suited to capturing the structure of Gram matrices arising in Chapter 5. In this thesis, association schemes are used to characterize the eigenspaces of such Gram matrices and to identify conditions under which the SRM is optimal. The focus is placed on commutative and symmetric schemes, which admit a rich spectral theory through their Bose–Mesner algebra.

Here, following [BBI+21], I introduce the basic definitions and algebraic tools needed in the later discussion.

3.3.1 Definition and basic properties

An *association scheme* is a pair $(X, \{R_i\}_{i=0}^n)$, where X is a finite set and $R_i \subseteq X \times X$ is a binary relation on the direct product. These relations satisfy the following properties

(1) $R_0 = \{(x, x) \mid x \in X\}$.

(2) $X \times X = \bigcup_{i=0}^k R_i$, and $R_i \cap R_j = \emptyset$ if $i \neq j$.

(3) Define the transpose $R_i^T = \{(x, y) \mid (y, x) \in R_i\}$. Then there exists $i' \in \{0, 1, \ldots, n\}$ such that $R_i^T = R_{i'}$.

(4) For i, j, k, the number $p_{ij}^k = |\{z \in X \mid (x, z) \in R_i\,, (z, y) \in R_j\}|$ is independent of the choice of x and y.

The numbers p_{ij}^k are the structural constants of the association scheme, and receive the name of intersection numbers.

An association scheme is called commutative if

(5) For any i, j, k, $p_{ij}^k = p_{ji}^k$,

and symmetric if it holds

(6) For all i, $R_i^T = R_i$.

In this thesis, I focus exclusively on symmetric schemes. These automatically satisfy the commutativity condition.

An alternative—and often more convenient—description of association schemes uses adjacency matrices of graphs. An adjacency matrix of a graph with $|X|$ vertices is a $(0, 1)$-matrix of size $|X| \times |X|$ whose entries are defined as

$$A_{x,y} = \begin{cases} 1 & \text{if vertices } x \text{ and } y \text{ are connected} \\ 0 & \text{otherwise.} \end{cases} \tag{3.61}$$

Each relation R_i can be represented by an adjacency matrix whose entries are equal to one if the pair $(x, y) \in R_i$.

The set of adjacency matrices $\{A_i\}_{i=0}^n$ fulfill

(1') $A_0 = \mathbb{1}$.

(2') $\sum_{i=0}^n A_i = J$, the all-ones matrix.

(3') For each i, it exists i' such that $A_i^T = A_{i'}$.

(4') For each i, j there exist non-negative integers p_{ij}^k such that

$$A_i A_j = \sum_{k=0}^{n} p_{ij}^k \, A_k \, . \tag{3.62}$$

Conditions (5) and (6) translate into

(5') $A_i A_j = A_j A_i$ for all i, j,

(6') $A_i^T = A_i$ for all i,

respectively.

3.3.2 Symmetric association schemes

Given a symmetric association scheme $(X, \{R_i\}_{i=0}^n)$, the set of adjacency matrices $\{A_i\}_{i=0}^n$ generates a matrix algebra known as the Bose–Mesner algebra [BM59], denoted by $\mathcal{A}$

$$\mathcal{A} = \mathrm{span}\{A_0, A_1, \ldots, A_n\} \, . \tag{3.63}$$

This algebra is of dimension $n + 1$, and it is closed under both the usual matrix product and the Hadamard product—or entrywise multiplication: $(A \circ B)_{ij} = A_{ij} B_{ij}$. Moreover, since all matrices in $\mathcal{A}$ commute, they can be diagonalized simultaneously in a certain set of eigenspaces $\{E_i\}_{i=0}^n$, called the central orthogonal. These eigenspaces constitute another basis for the Bose-Mesner algebra, and fulfill some relations similar to those of the set $\{A_i\}$

(1") $E_0 = \frac{1}{|X|} J$.

(2") $\sum_{i=0}^n E_i = \mathbb{1}$.

(3") $E_i^T = E_{i'}$ for some i and i'.

(4") $E_i \circ E_j = \frac{1}{|X|} \sum_{i=0}^{n} q_{ij}^k E_k$.

(5") $E_i \circ E_j = E_j \circ E_i$.

The structural constants q_{ij}^k of the algebra with respect to the Hadamard product receive the name of Krein numbers [Del73; Kre73].

The two bases $\{A_i\}_{i=0}^n$ and $\{E_i\}_{i=0}^n$ are connected via

$$A_i = \sum_{j=0}^{n} p_i(j) E_j \,, \tag{3.64}$$

and

$$E_i = \sum_{j=0}^{n} q_i(j) A_j \,. \tag{3.65}$$

The matrices P with $(P)_{ij} = p_i(j)$ and Q with $(Q)_{ij} = q_i(j)$ are the so-called *first* and *second eigenmatrix*, respectively. The name is not naive, as $p_i(j)$ are the eigenvalues of A_i in the subspace of the projector E_i, Eq. (3.64), and conversely for $q_i(j)$, but with respect to the Hadamard product. These matrices satisfy the relation

$$PQ = QP = |X| \mathbb{1} \,. \tag{3.66}$$

The dimensions of such subspaces spanned by E_i are the multiplicities of the association scheme, denoted by m_i. Another important quantity to be defined is k_i, defined by an eigenvalue equation $A_i J = k_i J = k_i |X| E_0$, and it corresponds to structural constants of regular graphs, see [BBI+21] for the connection of regular graphs and symmetric association schemes. This will be used later when a specific association scheme is presented.

3.3.3 Intersection matrices

There is still another important object that helps reveal the internal structure of the association schemes used in this thesis: the *intersection matrices*. These matrices encode the structural constants of the scheme in an algebraic form as

$$(B_i)_{jk} = p_{ij}^k \,. \tag{3.67}$$

The algebra generated by the set $\{B_0, B_1, \ldots, B_n\}$ under usual matrix multiplication is isomorphic to the Bose-Mesner algebra $\mathcal{A}$ [BBI+21]—with the standard matrix product.

A dual version of this construction arises when considering the Krein parameters. The corresponding matrices are called the *dual* intersection matrices, and are defined as

$$(B_i^*)_{jk} = q_{ij}^k \,. \tag{3.68}$$

In this case, the algebra generated by the matrices $\{B_0, B_1, \ldots, B_n\}$, with respect to the usual matrix product, is isomorphic to the Bose-Mesner algebra generated by the scaled idempotents $\{|X|E_i\}_{i=0}^n$ under the Hadamard product [BBI+21].

For the purpose of this thesis, it will be sufficient to know that these isomorphisms exist and that they are preserved under the respective products. In particular, in some important cases where the matrices B_i and B_i^* take a specific structured form, the corresponding association schemes are known to be related to a family of orthogonal polynomials.

3.3.4 P- and Q-polynomial schemes

Some association schemes admit a structure known as P-polynomial (or distance-regular) or Q-polynomial. These properties give rise to recurrence relations among the adjacency matrices or the idempotents, and correspond to schemes whose eigenmatrices are generated by orthogonal polynomials.

A P-polynomial scheme is an association scheme for which the intersection matrix $B := B_1$ is tridiagonal

$$B = \begin{Bmatrix} b_0 & b_1 & \cdots & b_{n-1} & * \\ a_0 & a_1 & \cdots & a_{n-1} & a_n \\ * & c_1 & \cdots & c_{n-1} & c_n \end{Bmatrix} . \tag{3.69}$$

This is equivalent to the intersection numbers to satisfy

$$p_{1j}^k \begin{cases} = 0 & \text{if} \quad |k-j| > 1 \,, \\ \neq 0 & \text{if} \quad |k-j| \leq 1 \,. \end{cases} \tag{3.70}$$

so that multiplication by A_1 only connects adjacent adjacency matrices in the algebra. Then, from Eq. (3.62), the relation

$$A_1 A_j = b_{j-1} A_{j-1} + a_j A_j + c_{j+1} A_{j+1} \,, \tag{3.71}$$

holds for each j. This recursive structure allows all the adjacency matrices to be written as polynomials evaluated at A_1. Define a polynomial with the recurrence relation

$$x v_j(x) = b_{j-1} v_{j-1}(x) + a_j v_j(x) + c_{j+1} v_{j+1}(x) \,. \tag{3.72}$$

where $v_{-1}(x) = 0$, and $v_1(x) = 1$. Then, one has

$$A_j = v_j(A_1)\,, \tag{3.73}$$

every matrix of $\mathcal{A}$ can be written as a certain polynomial of degree j evaluated at A_1.

This property has an important spectral consequence. Let θ_i with $i \in \{0, 1, \ldots, n\}$ denote the $n+1$ distinct eigenvalues of A_1. Since $A_i = \sum_{j=0}^{n} p_i(j)E_j$, one has

$$P_{ij} = v_i(\theta_j)\,, \tag{3.74}$$

meaning the entries of the first eigenmatrix P are simply evaluations of the polynomials $v_i(x)$ at the eigenvalues θ_j. Thus, computing the spectrum of a single adjacency matrix A_1 determines the eigenstructure of the entire algebra. This fact will be crucial in Chapter 5, where eigenvalues of the Gram matrix are obtained by this recurrence.

Similarly, a scheme is said to be Q-polynomial if the Krein parameters satisfy an analogous tridiagonal condition

$$q_{1j}^{k} \begin{cases} = 0 & \text{if} \quad |k-j| > 1\,, \\ \neq 0 & \text{if} \quad |k-j| \leq 1\,, \end{cases} \tag{3.75}$$

which implies that the corresponding dual intersection matrix $B^* := B_1^*$ has tridiagonal form

$$B^* = \begin{Bmatrix} b_0^* & b_1^* & \cdots & b_{n-1}^* & * \\ a_0^* & a_1^* & \cdots & a_{n-1}^* & a_n^* \\ * & c_1^* & \cdots & c_{n-1}^* & c_n^* \end{Bmatrix}. \tag{3.76}$$

In this way, one can define a polynomial with a recurrence relation

$$x v_j^*(x) = b_{j-1}^* v_{j-1}^*(x) + a_j^* v_j^*(x) + c_{j+1}^* v_{j+1}^*(x)\,. \tag{3.77}$$

so that

$$|X|E_j = v_j^*(|X|E_1)\,. \tag{3.78}$$

Here, the polynomial structure is taken with respect to the Hadamard product, not the usual matrix product.

This dual recurrence provides the spectral data of the idempotents in much the same way that the P-polynomial structure governs the adjacency matrices. In schemes where both P- and Q-polynomial structures are present, the eigenmatrices P and Q are generated by classical families of orthogonal polynomials.

3.3.5 Johnson association scheme

I conclude this chapter by introducing the association scheme that governs the structure of a particular set of matrices arising in the multi-anomaly problem of Chapter 5. The hypotheses in that Chapter are associated with subsets of k elements from a larger set of n. These combinatorial objects give rise to the *Johnson association scheme*, whose structure is closely tied to the *Johnson graphs*. The overlaps between the states considered there mirror this combinatorial framework. Solving the problem requires the set of tools provided by the theory of association schemes. In what follows, I introduce the basic definitions and results of the Johnson scheme that will play a central role in Chapter 5.

Definition 3.1. *A Johnson graph $J(n,k)$ is a graph whose vertices are the k-element subsets of a n-element set. Any pair of vertices r, s are connected by an edge if the intersection of the subsets $|r \cap s| = 1$* (see Figure 3.7).

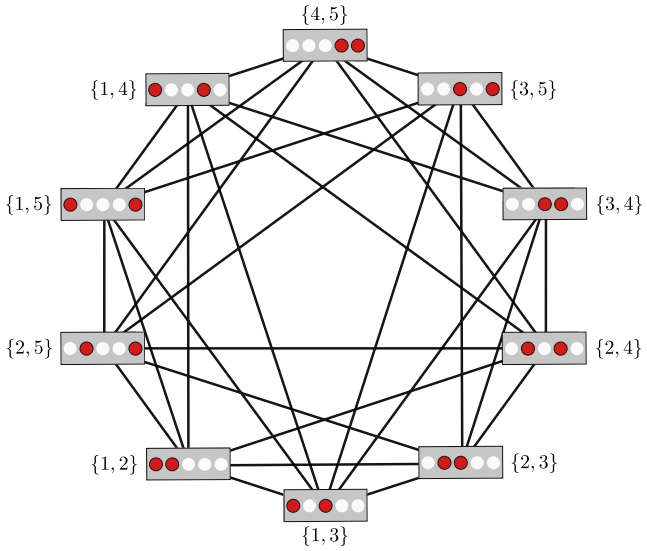


Figure 3.7: Jonhson graph $J(5,2)$ for a set of $n = 5$ objects with subsets of cardinality $k = 2$, where two vertices r and s are connected by an edge if $|r \cap s| = 1$.

As discussed earlier, this graph can be represented by an adjacency matrix, A, whose $k+1$ eigenvalues and corresponding multiplicities read

$$\lambda_j(A) = (k-j)(n+j+1) - k\,, \tag{3.79a}$$

$$m_j(A) = \binom{n}{j} - \binom{n}{j-1}\,, \tag{3.79b}$$

for $j = 0, \ldots, k$. Johnson graphs can be generalized to cover different intersection relations between the k-element subsets. These generalized graphs are called *generalized* Johnson graphs.

Definition 3.2. *A generalized Johnson graph $J(n, k, i)$ is a graph whose vertices are the k-element subsets of a n-element set. Any pair of vertices r, s are connected by an edge if the intersection of the subsets $|r \cap s| = i$.*

By fixing n and k, one obtains a family of $k + 1$ generalized Johnson graphs. Their adjacency matrices, $\{A_i\}_{i=0}^{k}$, satisfy conditions (1')–(6') and together span the Bose–Mesner algebra of the Johnson association scheme. To describe this algebra fully, it is useful to move between two complementary bases: the adjacency basis $\{A_i\}_{i=0}^{k}$ and the idempotent basis $\{E_i\}_{i=0}^{k}$, which satisfies conditions (1")–(6"). The connection between these bases is given in the eigenmatrices P and Q (Eqs. 3.64–3.65), whose entries specify how adjacency relations translate into projection operators and vice versa.

A central feature of symmetric association schemes is that these eigenmatrices have a highly structured form, described by families of classical orthogonal polynomials. In the case of the Johnson scheme, the relevant polynomials are the Hahn, Q_j, and dual Hahn, R_i, polynomials [Hah49; KS96]. They encode the spectral data of the scheme, as they determine the eigenvalues of the adjacency matrices, and thereby capture how the combinatorial structure of the graphs is reflected in their spectra. The entries of the eigenmatrices P and Q are given by

$$p_i(j) = k_i \, Q_j(i; -n + k - 1, -k + 1, k) \,, \tag{3.80a}$$

$$q_j(i) = m_j \, Q_j(i; -n + k - 1, -k + 1, k) \,, \tag{3.80b}$$

where $k_i = \binom{k}{i}\binom{n-k}{i}$ is the valency (the number of neighbors of each vertex) in the Johnson graph $J(n, k, i)$, and m_j is the multiplicity of the j-th eigenvalue in Eq. (3.79b).

While the Hahn polynomials describe the entries of the eigenmatrices, their dual counterpart—the dual Hahn polynomials—are the orthogonal polynomials associated with the scheme, which provides the eigenvalues and recurrence relations of the adjacency matrices [c.f. Eq.(3.73)]. This duality between two polynomial families arises from the duality between the two bases of the algebra, and it provides a complete characterization of the scheme.

To see this more explicitly, recall that any adjacency matrix A_i can be written in terms of some polynomial evaluated at A—in this specific case, the dual Hahn polynomials—as

$$A_i = (-1)^i \binom{k}{i}, R_i(A + k; 0, n - 2k, k) \,, \tag{3.81}$$

where the dual Hahn polynomial of degree i (R_i) is defined by

$$R_i(\lambda(x);\gamma,\delta,M) = {}_3F_2\left(\begin{matrix}-i, & -x, & x+\gamma+\delta+1 \\ & \gamma+1, & -M\end{matrix}; 1\right), \tag{3.82}$$

with $\lambda(x) = x(x+\gamma+\delta+1)$, and ${}_3F_2$ is the generalized hypergeometric function

$${}_pF_q\left(\begin{matrix}a_1, & \ldots, & a_p \\ b_1, & \ldots, & b_q\end{matrix}; z\right) = \sum_{m=0}^{\infty} \frac{(a_1)_m \cdots (a_p)_m}{(b_1)_m \cdots (b_q)_m} \frac{z^m}{m!}. \tag{3.83}$$

Here, $(a)_m$ is the rising Pochhammer symbol (or rising factorial) [AAR99]: $(a)_m := a(a+1)\ldots(a+m)$ and $(a)_0 := 1$.

For the Johnson association scheme, the relevant parameters of the dual Hahn polynomials are $x = j = 0, \ldots, k$, $\gamma = 0$, $\delta = n - 2k$, and $M = k$.

Finally, it is worth noting the explicit duality visible in Eqs. (3.80a–3.80b). While the eigenmatrix entries $p_i(j)$ are expressed in terms of Hahn polynomials, they could equivalently be written using dual Hahn polynomials, since the two families satisfy the relation

$$R_i(\lambda(j);\gamma,\delta,M) = Q_j(i;\gamma,\delta,M),, \tag{3.84}$$

which encapsulates the orthogonality relation of the eigenmatrices [c.f. Eq. (3.66)].

Quantum Edge Detection

4

Quantum edge detection

Life is a travelling to the Edge *of knowledge, then a leap taken.*

David Herbert Lawrence

The study of sudden changes is fundamental to understand countless natural phenomena, from phase transitions in material science [Ful20] to tipping points in climate systems [RCC+21]. However, these changes do not appear exclusively in nature—they can also emerge in controlled, human-designed processes. Consider, for instance, the production of quantum systems in a specific state. At some point, the source may unexpectedly fail, preparing a state different from the intended one, see Fig. 4.1 below. Could we reliably detect such a change in production? Sentís and collaborators raised this question in their work on quantum change point (QCP) [SBC+16]. They proved that identifying such change is indeed possible, achieving non-zero success probabilities even for infinitely long preparation series. However, their approach relies on prior knowledge of the states involved—specifically, the classical information describing these states. This leads to the question of whether it is possible to devise a universal protocol capable of detecting sudden changes without depending on the specific pair of states. We termed this problem as *quantum edge detection*.

The name quantum edge detection is inspired by its classical counterpart, a fundamental technique in image processing. In that context, an *edge* refers to a sudden change in intensity or color between neighboring pixels—an abrupt variation that typically signals a boundary between distinct regions in an image [Can86; PP93]. Detecting these transitions allows us to locate where a uniform structure gives rise to a different one.

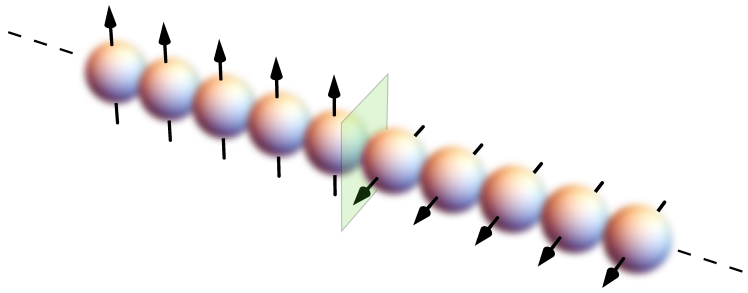


Figure 4.1: Illustration of a qubit array divided into two domains. Quantum edge detection aims to identify the boundary (highlighted in green) separating them.

In our case, the data is no longer a two-dimensional array of pixels but a one-dimensional string of quantum systems. The classical description—intensity or color—of each pixel is replaced by a quantum state. There are several differences, yet, the conceptual analogy remains; we aim to identify a sharp transition between two otherwise uniform regions, based solely on structural features. Quantum edge detection inherits the motivation of its classical counterpart, but confronts entirely different challenges.

At a deeper level, this question touches on a broader principle: that of learning from data. Learning is the process by which a system extracts structure from available (or input) data. Learning does not always require explicit instructions or supervision; often, it arises simply from the ability to recognize regularities or to distinguish the typical pattern. In classical settings, learning is often associated with prediction or classification, tasks where labeled data are available, and models are trained to reproduce a desired output. But not all learning takes place in such clearly defined environments. Sometimes, learning consists of simply detecting when something has changed. And when no prior knowledge is available, such detection becomes a form of *unsupervised* inference, that is, the system must rely entirely on the internal features of the data to reach a decision.

This perspective proves particularly relevant for certain quantum tasks, where underlying structures, such as symmetries and correlation patterns, can be revealed. In the case of quantum edge detection, the goal is to locate a boundary between two domains composed of unknown states. The task does not rely on identifying the domains themselves, but rather on recognizing the differences between them.

The protocol must find the point where the string changes its structure, and a new one arises.

The protocol responds to the structure present in the data, using the redundancy of the string to draw a statistical distinction between its two domains. As the length of the input increases, the distinction becomes sharper. While seemingly counterintuitive at first (the problem should become increasingly harder), the protocol benefits from longer sequences, as more data enhances the underlying structure, and with it, the success probability of correctly identifying the edge increases.

In this chapter, we explore this idea in depth. We formalize the edge detection task, develop optimal discrimination strategies in the universal setting, and analyze their performance both numerically and asymptotically. In doing so, we uncover a learning-like behavior embedded in quantum inference, able to recognize the change without ever knowing what that change consists of.

4.1 Unknown domains

Let $|\phi_0\rangle,\ |\phi_1\rangle$ denote the states of the qudits in each domain and $k = 1, 2, \ldots, N$ the position of the edge we aim to detect. The global state of the string can be expressed as:

$$|\Phi_k\rangle = \underbrace{|\phi_0\rangle \otimes \cdots \otimes |\phi_0\rangle}_{N-k} \otimes \overbrace{|\phi_1\rangle \otimes \cdots \otimes |\phi_1\rangle}^{k} . \tag{4.1}$$

The detection protocol, including post-processing, is defined by a POVM, whose elements, E_k, are associated with the outcomes of a measurement, indicating the possible locations of the edge along the string. Under the assumption that we have no classical description of the states $|\phi_0\rangle$ and $|\phi_1\rangle$, we aim to find the optimal protocol that maximizes the average success probability

$$P_{\mathrm{s}} = \frac{1}{N} \int \mathrm{d}\phi_0 \, \mathrm{d}\phi_1 \sum_k \mathrm{tr}\left(|\Phi_k\rangle\langle\Phi_k| E_k\right) . \tag{4.2}$$

This approach considers a uniform prior distribution of the edge between the two domains, reflecting a total lack of information about its location. The universality of the detection protocol is ensured by assuming the positive operators $\{E_k\}_k$ to be independent from the states of the particles in the string. This independence allows $\{E_k\}_k$ to be moved outside the integral in Eq. (4.2), thereby enabling the

definition of effective states ρ_k for the string. Specifically,

$$\rho_k = \int |\phi_0\rangle\langle\phi_0|^{\otimes(N-k)} \otimes |\phi_1\rangle\langle\phi_1|^{\otimes k} \mathrm{d}\phi_0 \, \mathrm{d}\phi_1 \,, \tag{4.3}$$

where the average over the states $|\phi_0\rangle$ and $|\phi_1\rangle$ acknowledges that they are unknown to the observer. A straightforward application of the Schur lemma (see Section 3.2.3) yields a block-diagonal structure for the states ρ_k in the Schur basis [FSC+22; SMM+19]

$$\rho_k = c_k \bigoplus_\lambda \mathbb{1}_\lambda \otimes \Omega_k^\lambda \,, \tag{4.4}$$

where, recall from Eq. (3.48) that $c_k = 1/(d_{n-k}^{\mathrm{sym}} \, d_k^{\mathrm{sym}})$, and λ labels the irreps arising from the joint action of the unitary group $\mathrm{SU}(d)$ and the symmetric group S_N on the vector space $(\mathbb{C}^d)^{\otimes N}$ of the entire string, where the symmetric group operates by permuting the particles within the string.

As already discussed in Section 3.2.2, the labels λ are usually identified with partitions of N, or equivalently with Young Diagrams of N boxes, and in this scenario, only two-row Young diagrams arise because, as shown in Eq. (3.47), the density matrices ρ_k act on the tensor product of two symmetric irreducible subspaces (each represented by one-row Young diagram).

Consequently, all required labels have the simple form of bipartitions $\lambda = (N - \lambda_2, \lambda_2)$ with $\lambda_2 = 0, 1, 2, \ldots, \lfloor N/2 \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. Hereafter, we introduce a slight abuse of notation, and we will refer to λ_2 simply as λ, as any bipartition λ is univocally defined by its second component. Each projector Ω_k^λ acts on the S_N irreducible subspace λ and is rank-1, (see Section 3.2.3), thus, it can be written as $\Omega_k^\lambda = |\Omega_k^\lambda\rangle\langle\Omega_k^\lambda|$. Note that the state ρ_k has support in irreducibles with $\lambda = 0, 1, 2, \ldots, \min\{N - k, k\}$.

Given the structure of ρ_k as shown in Eq. (4.4), one can choose, without loss of generality, universal POVM elements of the form

$$E_k = \bigoplus_\lambda \mathbb{1}_\lambda \otimes \mathcal{E}_k^\lambda \,, \tag{4.5}$$

which implies that it suffices to consider two-step protocols, starting with a projective measurement onto the irreducible subspaces followed by a restricted—to the irrep—POVM on the posterior outcome. The first measurement, referred to as weak Schur sampling, is represented by the projectors $\{\mathbb{1}_\lambda \otimes \mathbb{1}^\lambda\}_{\lambda=0}^{\lfloor N/2 \rfloor}$ and outputs a value, λ^*, for the label λ. Subsequently, a measurement with POVM elements

$\{\mathbb{1}_{\lambda^*} \otimes \mathcal{E}_k^{\lambda^*}\}_{k\in K_{\lambda^*}}$ is executed on the posterior state belonging to the λ^* subspace. For a generic value of λ, K_λ is the range of values of k for which ρ_k has support in the λ subspace, $\lambda \leq k \leq N-\lambda$.

The second step of the detection protocol involves a discrimination strategy among the set of pure states $\{|\Omega_k^{\lambda^*}\rangle\}_{k\in K_{\lambda^*}}$ with a prior probability mass function given by $\eta_k^{\lambda^*} / \sum_{k'} \eta_{k'}^{\lambda^*}$. In this equation, $\eta_k^{\lambda^*}$ represents the joint probability of the edge being located at k and obtaining the outcome λ^* from the weak Schur sampling. This probability is expressed as:

$$\eta_k^{\lambda^*} = \begin{cases} \dfrac{d_{\lambda^*}}{N d_{N-k}^{\mathrm{sym}}\, d_k^{\mathrm{sym}}}\,, & \text{if } k \in K_{\lambda^*} \\ 0\,, & \text{otherwise,} \end{cases} \tag{4.6}$$

where d_λ is explicitly given by the Hook-length formula in Eq. (3.41), which for bipartitions simplifies to:

$$d_\lambda = \frac{N-2\lambda+1}{N-\lambda+1}\binom{d+\lambda-2}{d-2}\binom{d+N-\lambda-1}{d-1}. \tag{4.7}$$

Combining the above results, the average success probability, which we aim to maximize, reads

$$P_{\mathrm{s}} = \sum_{\lambda=0}^{\lfloor N/2\rfloor} \sum_{k\in K_\lambda} \eta_k^\lambda \operatorname{tr}\left(\Omega_k^\lambda \mathcal{E}_k^\lambda\right). \tag{4.8}$$

The dimension of the irreducible sub-spaces to which the states $|\Omega_k^\lambda\rangle$ belong grows exponentially with N [Gon24], making optimization over $\{\mathcal{E}_k^\lambda\}_{k,\lambda}$ via direct computation infeasible even for small values of N.

To address this challenge, we rely on the Gram matrix G^λ of the set of pure states $\{|\Omega_k^\lambda\rangle\}_{k\in K_\lambda}$ for each value of λ. As discussed in detail in Section 2.5.3, this matrix provides a complete description of the quantum state ensemble relevant for the discrimination task. Its entries are given by

$$G_{kk'}^\lambda = \sqrt{\eta_k^\lambda \eta_{k'}^\lambda}\ \langle\Omega_k^\lambda|\Omega_{k'}^\lambda\rangle. \tag{4.9}$$

Since G^λ is an $|K_\lambda|\times|K_\lambda|$ matrix, the computational complexity of the optimization problem becomes polynomial in N. It takes the simple form:

$$G_{kk'}^\lambda = v_k^\lambda u_{k'}^\lambda, \quad k \leq k' \tag{4.10}$$

(for $k' \leq k$, exchange k and k'), where:

$$u_k^\lambda = \sqrt{\eta_k^\lambda \frac{\binom{N-k}{\lambda}}{\binom{k}{\lambda}}}, \quad v_k^\lambda = \sqrt{\eta_k^\lambda \frac{\binom{k}{\lambda}}{\binom{N-k}{\lambda}}}. \tag{4.11}$$

Note that the Gram matrix G^λ can be represented as $G^\lambda = \mathrm{triu}(\mathsf{u}\mathsf{v}^T) + \mathrm{tril}(\mathsf{v}\mathsf{u}^T)$, where tril (triu) denotes lower (upper, including the diagonal) triangular part. Here, u and v are column vectors with components given by (4.11). This property ensures that G^λ is semiseparable [VVM08], offering significant computational advantages that complement the complexity reduction achieved by using the Gram matrix formulation. In particular, we will exploit the fact that the inverse of a semiseparable matrix is tridiagonal, resulting in sparsity.

Another way of checking that G^λ is semiseparable from a representation theory approach arises from the following theorem

Theorem 4.1. (Theorem 2.15 in [VVM08]) *A real symmetric matrix M of size n is semiseparable if there exist numbers $d_1, \ldots, d_n, s_1, \ldots, s_n, c_1, \ldots, c_n$ such that*

$$M_{ij} = \begin{cases} d_j s_j s_{j+1} \ldots s_{i-1} c_i & \text{if} \quad i > j \\ d_i s_i s_{i+1} \ldots s_{j-1} c_j & \text{if} \quad i \leq j \end{cases} \tag{4.12}$$

assuming $s_i s_{i+1} \ldots s_{j-1} = 1$ if $i = j$.

For our matrices G^λ, the coefficients read $d_i = c_i = \sqrt{\eta_i^\lambda}$ and $s_i = \langle \Omega_i^\lambda | \Omega_{i+1}^\lambda \rangle$, for $i \in K_\lambda$. The only missing step is to prove that the product sequence

$$\langle \Omega_i^\lambda | \Omega_{i+1}^\lambda \rangle \langle \Omega_{i+1}^\lambda | \Omega_{i+2}^\lambda \rangle \ldots \langle \Omega_{j-1}^\lambda | \Omega_j^\lambda \rangle = \langle \Omega_i^\lambda | \Omega_j^\lambda \rangle \,, \tag{4.13}$$

which reduces to proving the simpler relation

$$\langle \Omega_k^\lambda | \Omega_{k'}^\lambda \rangle \langle \Omega_{k'}^\lambda | \Omega_{k''}^\lambda \rangle = \langle \Omega_k^\lambda | \Omega_{k''}^\lambda \rangle \,, \tag{4.14}$$

for $k \leq k' \leq k''$. This relation is easily seen from the operator structure of the hypotheses, in Eq. (3.47), and the relation

$$(\mathbb{1}_{n-k}^{\mathrm{sym}} \otimes \mathbb{1}_k^{\mathrm{sym}})(\mathbb{1}_{n-k'}^{\mathrm{sym}} \otimes \mathbb{1}_{k'}^{\mathrm{sym}})(\mathbb{1}_{n-k''}^{\mathrm{sym}} \otimes \mathbb{1}_{k''}^{\mathrm{sym}}) = (\mathbb{1}_{n-k}^{\mathrm{sym}} \otimes \mathbb{1}_k^{\mathrm{sym}})(\mathbb{1}_{n-k''}^{\mathrm{sym}} \otimes \mathbb{1}_{k''}^{\mathrm{sym}}) \,, \tag{4.15}$$

again with $k \leq k' \leq k''$. Focusing on the middle term in the left-hand side of the relation, the first tensor term acts trivially to the left, as $(n-k') \leq (n-k)$. Similarly, the second tensor term acts trivially to the right. Writing Eq. (4.15) in

the Schur basis, one obtains Eq. (4.14) for each λ, and thus G^λ is proven to be semiseparable.

The inherent symmetries of 1D quantum edge detection enable us to simplify our optimization task into a series of more manageable pure-state discrimination subproblems, each framed within an irrep subspace λ. Even though we significantly reduce the complexity via Schur-Weyl duality, this is insufficient to completely solve the problem analytically. In general, quantum state discrimination problems lack known analytical solutions, except for two states [Hel69] or for highly symmetric sets of states [Bar01; CDP+04; SMM22]. Since our sub-problems do not fit into any of these simple cases, we must resort to numerical optimization. In practice, however, without relying on Schur-Weyl duality, a direct numerical method would encounter exponential complexity due to the high dimensionality (d^N) of the effective mixed states ρ_k. Yet Schur-Weyl duality alone cannot sufficiently tame the problem without invoking the Gram matrix formulation. As discussed earlier, using G^λ, with a size at most N^2, provides an exponential advantage, enabling us to efficiently employ SDP optimization, whose complexity scales only polynomially with the size of G^λ.

The key step is to view the N columns of the symmetric and positive definite matrix $\sqrt{G^\lambda}$, each denoted as $\left[\sqrt{G^\lambda}\right]_k$, as representations of the N unnormalized states $\sqrt{\eta_k^\lambda}|\Omega_k^\lambda\rangle$. The consistency of this representation can be verified by noting that

$$\left[\sqrt{G^\lambda}\right]_k^T \left[\sqrt{G^\lambda}\right]_{k'} = \left[(\sqrt{G^\lambda})^2\right]_{kk'} = G^\lambda_{kk'}\,. \tag{4.16}$$

The success probability of the optimal discrimination strategy on a given irrep λ can then be computed as

$$P_\lambda = \max_{\{\mathcal{E}_k\}} \sum_{k\in K_\lambda} \mathrm{tr}\left(\mathcal{E}_k^\lambda \left[\sqrt{G^\lambda}\right]_k \left[\sqrt{G^\lambda}\right]_k^T\right) \tag{4.17}$$

$$\text{s.t.} \quad \mathcal{E}_k^\lambda \geq 0\,, \tag{4.18}$$

$$\sum_{k\in K_\lambda} \mathcal{E}_k^\lambda = \mathbb{1}^\lambda, \tag{4.19}$$

where $\mathcal{E}_k^k$ is a $|K_\lambda| \times |K_\lambda|$ matrix representation of the POVM operator $\mathcal{E}_k^\lambda$. We observe that the optimization problem in Eq. (4.17) takes the form of an SDP. The average success probability of the optimal detection protocol is then given by

$$P_\mathrm{s} = \sum_{\lambda=0}^{\lfloor N/2\rfloor} P_\lambda. \tag{4.20}$$

Here P_λ denotes the (joint) probability of obtaining the outcome λ from the weak Schur sampling and successfully detecting the edge.

The green points in Fig. 4.2 illustrate the results of our calculation, obtained using Eqs. (4.17) and (4.20) for $d = 2$ (qubits). Similar trends are observed for other values of d. These points represent the ultimate limit to the achievable success probability in detecting the edge of a domain along a string of quantum particles.

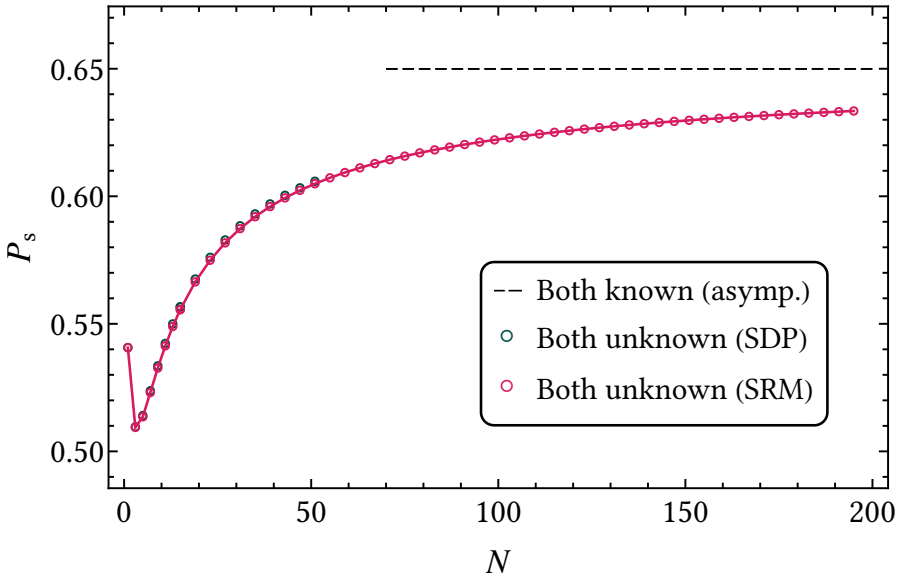


Figure 4.2: Success probability as a function of the string length N for the scenarios considered in this Chapter: when all states are unknown (green and pink circles). The maximum success probability data points, obtained through SDP optimization, are shown in green, while the SRM success probability points are in pink. The black dashed line represents the average asymptotic success probability derived from the QCP, assuming knowledge of all states. All data points correspond to N ranging from 2 to 18 in steps of 2, and from 22 to 198 in steps of 4.

While not always optimal [DP15], the SRM (see Section 2.5) often yields accurate lower bounds on success probability across numerous discrimination problems. It has been shown to offer an asymptotically optimal quantum change point detection protocol in [SBC+16], a similar discrimination task that will be relevant in the asymptotic regime of edge detection. That work also presents general conditions for the SRM's optimality and provides bounds on the success probability it achieves. However, in the present case, the SRM does not meet the required conditions for these results to hold for finite-sized problems. In spite of this, we will verify its asymptotic optimality for edge detection.

The POVM elements of the SRM are defined as

$$\mathcal{E}_k^\lambda = \Lambda_\lambda^{-1/2} \left(\eta_k^\lambda \, \Omega_k^\lambda\right) \Lambda_\lambda^{-1/2}, \tag{4.21}$$

where $\Lambda_\lambda = \sum_k \eta_k^\lambda \, \Omega_k^\lambda$, resembling Eq. (2.39), but restricted to the irreducible λ. In our lower-dimensional representation, this fixes the definition of the matrix elements of E_k in Eq. (4.17) to $[E_k]_{ii'} = \delta_{ki}\delta_{ki'}$ independently of λ. With this choice, the expression for P_λ^{SRM} simplifies to

$$P_\lambda^{\mathrm{SRM}} = \sum_{k \in K_\lambda} \left([\sqrt{G^\lambda}]_{kk} \right)^2. \tag{4.22}$$

In simple terms, the joint probability of obtaining λ and succeeding amounts to the sum of the squares of the diagonal entries of $\sqrt{G^\lambda}$. The success probability $P_{\mathrm{s}}^{\mathrm{SRM}} = \sum_\lambda P_\lambda^{\mathrm{SRM}}$ gives a lower bound to P_{s} since it results from a particular choice for E_k in Eq. (4.17).

Since no optimization is required to compute P_λ^{SRM}, it becomes feasible to evaluate the success probability P^{SRM} for significantly larger values of N. Additionally, the SRM equips us with the tools to derive analytical results for the asymptotic behavior of P_{s} as N grows large, a topic we discuss in the next subsection.

In Fig. 4.2, the pink points, connected by lines, depict the results of our numerical evaluation of P^{SRM} for $d = 2$ (similar patterns hold for other values of d). It is noteworthy that the pink points differ almost imperceptibly from the green ones, which represent the optimal protocol, indicating that the SRM is nearly optimal. Both sets of points fall underneath the dashed line but gradually approach it. This line represents a scenario where the quantum states of the domains are known, and this information is used for edge detection. More precisely, it depicts the average success probability over $|\phi_0\rangle$ and $|\phi_1\rangle$ in Eq. (4.2) when the measurement is allowed to depend on these states and $N \to \infty$. This average, that we denote as $P_{\mathrm{s}}^{\mathrm{known}}$, can be easily computed using the results from the QCP derived in [SBC+16]. The asymptotic limit of the success probability of the QCP depends solely on the overlap between the two states ϕ_0 and ϕ_1

$$P_{\mathrm{s}}^{\mathrm{QCP}}(\phi_0, \phi_1) = \frac{4(1-c^2)}{\pi^2} K^2(c^2)\,, \tag{4.23}$$

where $c = |\langle\phi_0|\phi_1\rangle|$ denotes the overlap of the states in each domain, and $K(x)$ is the complete elliptic function of the first kind [AS48]. Moreover, the average over ϕ_0 and ϕ_1 can be replaced by the average over their overlap,

$$\mathrm{d}\phi_0 \, \mathrm{d}\phi_1 = \mu_d(c^2) \, \mathrm{d}c^2\,, \tag{4.24}$$

where $\mu_d(c^2)\,\mathrm{d}c^2$ represents the uniform (probability) measure for qudits [SMM+19], given by $\mu_d(c^2) = (d-1)(1-c^2)^{d-2}$. The average, $P_{\mathrm{s}}^{\mathrm{known}}$, can thus be expressed as

$$P_{\mathrm{s}}^{\mathrm{known}}(d) = \int \mathrm{d}\phi_0\,\mathrm{d}\phi_1\, P_{\mathrm{s}}^{\mathrm{QCP}}(\phi_0,\phi_1) = \int_0^1 \frac{4(1-c^2)}{\pi^2} K^2(c^2)\,\mu_d(c^2)\,\mathrm{d}c^2\,. \quad (4.25)$$

This success probability serves as an upper bound to P_{s} in the scenario considered in this work. The convergence of the pink points to this line also showcases the asymptotic optimality of the SRM.

Intuition might suggest that the success probability should monotonically decrease with N, as one could argue that with a longer string of particles, there would be more possible locations for the edge, making the detection task more challenging. This is indeed the behavior of the success probability for the QCP, akin to the scenario where the states of the domains are known. However, Fig. 4.2 disproves this intuition, except for very short strings of fewer than eight particles. For longer strings ($N \geq 8$), the trend is reversed: edge detection becomes more successful as the number of particles grows larger. This suggests that the detection protocol *learns* from the unknown states of the domains and having more copies helps learning more about them, similarly to what happens in a state estimation problem [MP95], overcoming the increasing difficulty of discriminating possible edge locations.

To test this idea, we simulated an estimate-and-discriminate strategy where a small fraction (of order $\sqrt{N}$) of particles from each end of the string is used to estimate the states of the domains. This information is then fed into the QCP protocol to precisely locate the edge. While we can only perform this simulation for modest string lengths, the results align with those shown in Fig. 4.2, with the success probability approaching the value achieved by the optimal edge detection protocol.

4.1.1 Asymptotic regime

We assess the performance of the optimal detection protocol in the asymptotic regime of large N by computing the limiting value of the success probability. The calculation is performed indirectly, by verifying that the lower bound $P_{\mathrm{s}}^{\mathrm{SRM}}$ provided by the SRM strategy matches the upper bound $P_{\mathrm{s}}^{\mathrm{known}}$ resulting from averaging the success probability under the assumption of known domain states. This alignment, that we present later in Table 4.1, not only reveals the value of the maximum success probability but also confirms the asymptotic optimality of the SRM strategy.

To this end, we will show that the success probability has an asymptotic expansion of the form

$$P_{\mathrm{s}}^{\mathrm{SRM}} \sim p_0(d) + \frac{p_1(d)}{N} + O(N^{-2}) \tag{4.26}$$

and present a method for computing the coefficients $p_n(d)$, $n = 0, 1, 2, \ldots$, to arbitrary accuracy. The method shares similarities with the application of perturbation theory in fields like high-energy physics, particularly relying on techniques to expedite the convergence of asymptotic series through the utilization of Padé approximants [Bak64].

Here, we only provide an overview of the method, focusing on computing $p_0(2)$, the leading coefficient for qubits ($d = 2$), with 0.03% uncertainty. We will see that the computed value coincides with the upper bound $P_{\mathrm{s}}^{\mathrm{known}}$, within the specified accuracy. Results for larger dimensions ($d = 3, 4, 8$) are also given, but technical details are deferred to Appendix A.

The essence of our calculations lies in a 'perturbative expansion' of the inverses of the Gram matrices G^λ, rescaled appropriately. We recall that these inverses are tridiagonal matrices [c.f. Eqs. (4.10–4.11)]. The sparsity of these matrices, with their off-diagonal entries treated as perturbations, is crucial for efficiently keeping track of the orders of the expansion.

The rescaling of the Gram matrices is performed as follows:

$$\tilde{G}^\lambda = \frac{(N/2)^2}{N - 2\lambda + 1} G^\lambda, \tag{4.27}$$

ensuring that $\tilde{G}^\lambda = \mathbb{1} + O(1/N)$ for large N. Thus, $(\tilde{G}^\lambda)^{-1} = \mathbb{1} + \Delta_\lambda$, where Δ_λ contains non-zero terms only in the super- and sub-diagonals, each of order $O(1/N)$, remember that the inverse of a semiseparable matrix is tridiagonal, and all G^λ are semiseparable. Using the binomial series, we obtain

$$\sqrt{\tilde{G}^\lambda} = (\mathbb{1} + \Delta_\lambda)^{-1/2} = \mathbb{1} + \sum_{r=1}^{\infty} \binom{-1/2}{r} (\Delta_\lambda)^r. \tag{4.28}$$

In practical calculations, the series is truncated to the desired order of approximation, rescaled back, and then substituted into Eq. (4.28). This procedure yields an asymptotic expression for the joint probability P_λ^{SRM}, recall Eq. (4.22), whose marginal probability (over λ) provides the desired asymptotic series (4.26) for the success probability. The initial terms can be expressed as:

$$P_\lambda^{\mathrm{SRM}} = \frac{4(2j+1)^2}{N^3} \left[1 - \frac{4}{N} + \frac{12}{N^2} - \frac{8(j^2+j+12)}{3N^3} - \frac{4j^4 + 8j^3 + \ldots}{15N^4} + \ldots \right]. \tag{4.29}$$

Here, $j = N/2 - \lambda$ denotes the total spin quantum number of the string of particles (qubits), which serves as an alternative and commonly used label for the irreps of $\mathrm{SU}(2)$.

The calculation of the marginal, $\sum_j P^{\mathrm{SRM}}_{\lambda(j)}$, is involved due to the scaling of the range of values of j with N, resulting in infinitely many terms in the expansion (4.29) that contribute to the coefficient $p_0(2)$. This can be addressed by introducing a scaled version of j as $x := j/(N/2)$. In the limit $N \to \infty$, the variable x can be thought of as a real variable whose range is $[0, 1]$. In terms of x, Eq. (4.29) takes the form

$$\frac{N}{2} P^{\mathrm{SRM}}_{\lambda(x)} = \sum_{r=1}^{\infty} a_r x^{2r}, \tag{4.30}$$

and the leading-order term $p_0(2)$ can be computed as

$$p_0(2) = \int_0^1 \frac{N}{2} P^{\mathrm{SRM}}_{\lambda(x)} \, \mathrm{d}x. \tag{4.31}$$

However, a difficulty arises as we have not identified a closed expression for the general term of the sequence $\{a_r\}_{r\in\mathbb{N}}$, and the truncated series, which we managed to compute up to $O(x^{30})$, is unable to capture the singular behavior of the integrand near $x = 1$ with sufficient accuracy (dashed line in Fig. 4.3). Padé approximants can be used to speed up the convergence of the series as they are known to accurately describe the behavior of asymptotic expansions near singular points [BO13].

In brief, for a given function $f(x)$, its Padé approximant of order $[n/m]$ is the rational function

$$[n/m]_f(x) := \frac{\sum_{r=0}^{n} A_r x^r}{1 + \sum_{r=1}^{m} B_r x^r}, \tag{4.32}$$

whose Maclaurin series expansion matches that of $f(x)$ up to order $n + m$. This condition determines the coefficients A_r and B_r [Bak64].

The solid line in Fig. 4.3 represents the best Padé approximant, of order $[14/14]$, to the expansion in Eq. (4.30), truncated at $O(x^{30})$ and represented by the dashed line. Additionally, the red crosses and open blue circles depict numerical values of $P^{\mathrm{SRM}}_{\lambda(j)}$ for a few choices of j (equivalently x) and for $N = 500$ and $N = 4600$ respectively. The figure readily shows the significant improvement provided by the Padé approximant near $x = 1$ (see Appendix A for more details).

By replacing the integrand in Eq. (4.31) with its Padé approximant and performing the integral, we obtain the value $p_0(2) \approx 0.6499$. This leads us to conclude that the success probability for qubits approaches a finite value as $N \to \infty$, given by

$$P^{\mathrm{SRM}}_{\mathrm{s}} \approx 0.6499 + O(1/N). \tag{4.33}$$

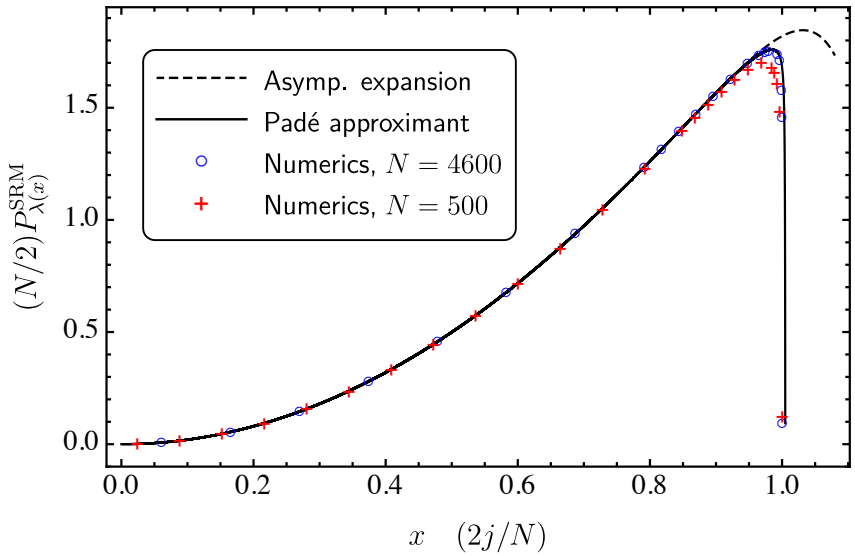


Figure 4.3: Joint probability $P^{\mathrm{SRM}}_{\lambda(x)}$ plotted against the scaled angular momentum $x = 2j/N$. The dashed curve represents the asymptotic power series truncated at $O(x^{30})$, while the solid curve depicts its $[14/14]$ Padé approximant. These curves closely match each other, deviating only near $x = 1$. The range of x extends beyond its physical domain (the unit interval) to show this deviation more clearly. The red crosses and blue circles denote numerical evaluations of the joint probability at selected points, corresponding to $N = 500$ and $N = 4600$ respectively. The solid line is seen to fit the numerical points corresponding to the larger value of N almost perfectly.

Using the same procedure, one can compute $p_0(d)$ for any value of the local dimension d. Results, along with their estimated error margins, for some selected values can be found in the second column of Table 4.1.

For the same values of d, the third column of Table 4.1 presents the averaged success probability assuming measurements can depend on the domain states. Since our focus is solely on the limiting values as $N \to \infty$, we can rely on the asymptotic result provided in [SBC+16]. The numerical integration has been carried out to high precision, ensuring that all digits shown in the third column of the table are significant. As anticipated earlier, the agreement observed with the values in the second column tells us that they are optimal and serves as a numerical verification of the asymptotic optimality of the SRM strategy.

d	$p_0(d)$	$P_{\mathrm{s}}^{\mathrm{known}}(d)$
2	$0.6499 \pm 2 \times 10^{-4}$	0.64991
3	$0.792308 \pm 3 \times 10^{-6}$	0.792311
4	$0.852860 \pm 1 \times 10^{-6}$	0.8528600
8	$0.9323011 \pm 2 \times 10^{-7}$	0.9323011

Table 4.1: Limiting success probability as $N \to \infty$ for the SRM protocol (second column) and average success probability assuming known domain states, computed from the QCP results in [SBC+16] (third column). The error intervals in the second column reflect two different procedures used to compute the Padé approximant, see Appendix A.

4.2 Known state in one domain

In this section, we consider a situation where the state of the particles in one domain is known, specifically set to $|0\rangle$ without loss of generality, while the state of the particles in the other domain remains unknown. Consequently, measurements may depend solely on the known state $|0\rangle$. This setup is particularly relevant in the context of time series data or manufacturing processes. Consider a scenario where a particle source is designed to produce particles in the state $|0\rangle$. However, due to a malfunction, the source begins producing particles in an unspecified state after a certain period. Given a string of N particles produced by this source, our objective is to develop a detection protocol that optimally identifies where along the string the state of the particles changed, and determine its success probability.

In this scenario, the effective states of the string read [cf. Eq. (3.46)]

$$\rho_k = \frac{|0\rangle\langle 0|^{\otimes(N-k)} \otimes \mathbb{1}_k^{\mathrm{sym}}}{d_k^{\mathrm{sym}}}, \tag{4.34}$$

where, as in the previous section 4.1 , we assume that the location of the edge, or state change, is uniformly distributed along the string. The specific form (4.34) of the effective states implies several adjustments in the detection strategy developed for both domains unknown.

We observe that the density matrices ρ_k are diagonal in the basis $\{\, |w_k^{\mathbf{n}}\rangle \,\}$, where $\mathbf{n} = (n_0, n_1, \ldots, n_{d-1})$, with n_i the number of particles in the i-th level of the system $|i\rangle$, that fulfill $\sum_{i=0}^{d-1} n_i = k$, and $k = 0, 1, \ldots, N$; which is defined as

follows:

$$|w_k^{\mathbf{n}}\rangle = |0\rangle^{\otimes N-k} \otimes \frac{\sum_\sigma P_\sigma \, |0\rangle^{\otimes n_0} \otimes |1\rangle^{\otimes n_1} \otimes \cdots \otimes |d-1\rangle^{\otimes n_{d-1}}}{\sqrt{\binom{k}{\mathbf{n}}}}. \tag{4.35}$$

Here, $\binom{k}{\mathbf{n}}$ is the multinomial factor defined as

$$\binom{k}{\mathbf{n}} := \frac{k!}{\prod_{l=0}^{d-1} n_l!}\,, \tag{4.36}$$

and the set $\{P_\sigma\}_\sigma$ stands for the unitary representation of S_k on $(\mathbb{C}^d)^{\otimes k}$ that permutes the k particles. The sum over σ extends only to those *effective* permutations σ that do not leave the state $|0\rangle^{\otimes n_0} \otimes |1\rangle^{\otimes n_1} \otimes \cdots \otimes |d-1\rangle^{\otimes n_{d-1}}$ invariant. Clearly, only states with the same numbers $n_1, n_2, \ldots, n_{d-1}$ may exhibit a non-vanishing overlap. This class of bases, known as Jordan bases [BBF+06], play a key role in programmable discriminators [SBC+10; HB07].

The optimal protocol, due to the block-diagonal structure of the hypotheses, also involves two sequential measurements, as in the case of both domains unknown. First we measure the total number of different excitations (number of states $|1\rangle, |2\rangle$, etc) to obtain a multi-valued label $(n_1, n_2, \ldots, n_{d-1})$, thereby projecting each state ρ_k into the posterior state $|w_k^{\mathbf{n}}\rangle$. Subsequently, a second measurement, akin to the corresponding step in the protocol for unknown states, is performed to discriminate among the states $|w_k^{\mathbf{n}}\rangle$, each occurring with respective prior probabilities given by:

$$\eta_k^{\mathbf{n}} = \frac{1}{N d_k^{\mathrm{sym}}}\,. \tag{4.37}$$

It is useful (and meaningful) to assign a new label $\tilde{\mathbf{n}}$ to each label $\mathbf{n}$, according to

$$\mathbf{n} \to \tilde{\mathbf{n}} = (N - k + n_0, n_1, \ldots, n_{d-1})\,, \tag{4.38}$$

as this label captures both the result of the first projective measurement and the total number of states in the known state $|0\rangle$. Notice that two states will have non-vanishing overlap only if all entries of the label except $\tilde{n}_0$ are equal.

Following a similar approach as in Section 4.1, we compute the probability of successful discrimination using the Gram matrices $G^{\tilde{\mathbf{n}}}$ of the sets $\{|w_k^{\mathbf{n}}\rangle\}$, defined as the $(N-k+n_0)\times(N-k+n_0)$ Hermitian matrices with entries $G^{\mathbf{n}}_{kk'} = \sqrt{\eta_k^{\mathbf{n}}\eta_{k'}^{\mathbf{n}'}}\langle w_k^{\mathbf{n}}|w_{k'}^{\mathbf{n}'}\rangle$. From Eq. (4.35), we readily find

$$G^{\tilde{\mathbf{n}}}_{k\,k'} = v_k^{\mathbf{n}} u_{k'}^{\mathbf{n}} := \sqrt{\eta_k^{\mathbf{n}}\eta_{k'}^{\mathbf{n}} \frac{\binom{k}{\mathbf{n}}}{\binom{k'}{\mathbf{n}}}}, \quad k \le k', \tag{4.39}$$

with $G^{\tilde{\mathbf{n}}}_{k\,k'} = G^{\tilde{\mathbf{n}}}_{k'\,k}$ if $k' < k$. The form of $G^{\tilde{\mathbf{n}}}$ indicates that they are semiseparable matrices, as observed with the Gram matrices in Section 4.1. All the properties discussed there for G^{λ} also apply to $G^{\tilde{\mathbf{n}}}$. This parallelism between Eqs. (4.39) and (4.11) is noteworthy, where $\tilde{\mathbf{n}}$ takes the role of λ, and only one multinomial is present in Eq. (4.39), reflecting that in this case only one domain of unknown states exists.

Yet another difference can be depicted, as in Section 4.1, the number of states in the first domain uniquely defines a label λ, while for this problem, the number of excitations $N - \tilde{n}_0$ gives rise to several distinct $\tilde{\mathbf{n}}$. However, the matrices $G^{\tilde{\mathbf{n}}}$ are independent of all entries $\tilde{n}_l$ with $l > 0$. The multinomials in Eq. (4.39) simplify to only depend on $N - \tilde{n}_0$. This allows us to construct effective Gram matrices of the form

$$G^{\tilde{n}_0}_{k\,k'} = \sum_{n_1,n_2,\ldots} G^{\tilde{\mathbf{n}}}_{k\,k'} = \binom{N - \tilde{n}_0 + d - 2}{d-2} \sqrt{\eta_k^{\tilde{n}_0} \eta_{k'}^{\tilde{n}_0} \frac{\binom{k}{N-\tilde{n}_0}}{\binom{k'}{N-\tilde{n}_0}}}\,, \tag{4.40}$$

to further reduce the complexity of the problem for the following success probability computations.

The calculations of the maximum and SRM success probabilities, the former obtained via SDP optimization, follow the methodology described earlier in the Chapter. Numerical results are depicted in Fig. 4.2, facilitating comparison with results obtained for unknown domain states. The blue circles represent the SDP results, calculated for N up to 50, while the orange circles represent the numerical evaluation of the SRM success probability, extended to much larger values of N.

Consistently, these points lie between those corresponding to the scenario where both domain states are unknown (green and pink) and the asymptotic dashed line representing the opposite scenario, in which both states are known. Remarkably, the blue and orange circles closely align, indicating that the SRM protocol remains nearly optimal in this setting as well.

Although we do not perform a detailed asymptotic analysis for this scenario, the success probability is bounded from below by that of the fully unknown case. This guarantees that the optimal asymptotic performance reaches the same value as in the known QCP scenario. Notably, the increase in success probability with growing system size is also observed here, highlighting once again the protocol's learning behavior—its ability to compensate for the increased complexity of the task caused by the larger number of hypotheses.

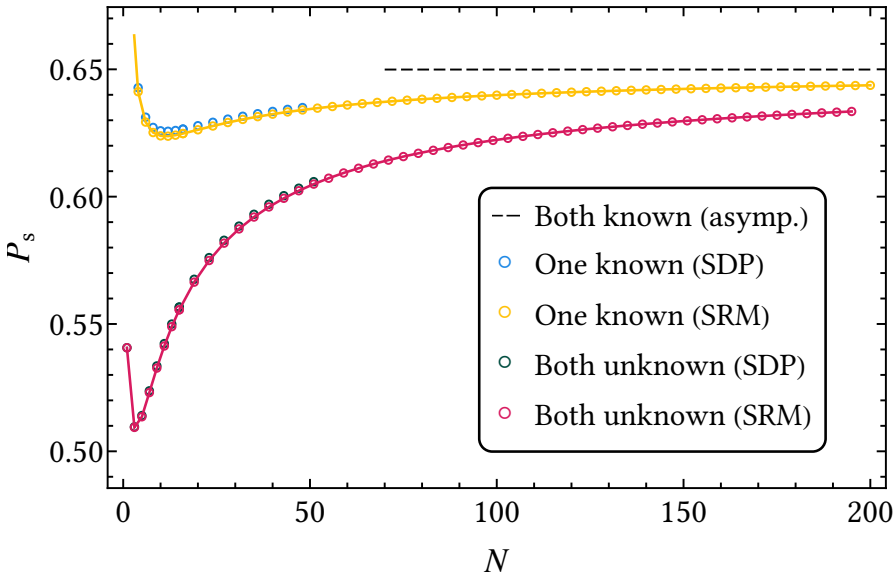


Figure 4.4: Success probability as a function of the string length N for the scenarios considered in this Chapter: when all states are unknown (green and pink circles) and when the states of one domain are known (blue and orange circles). The maximum success probability data points, obtained in both scenarios through SDP optimization, are shown in green and blue, respectively, while the SRM success probability points are in pink and orange, respectively. The black dashed line represents the average asymptotic success probability derived from the QCP, assuming knowledge of all states. All data points correspond to N ranging from 2 to 18 in steps of 2, and from 22 to 198 in steps of 4.

4.3 Discussion

In this Chapter, we have analyzed the problem of quantum edge detection. In the first part of the Chapter, we have focused on a one-dimensional string composed of two different regions, each of which is formed by unknown—yet different—quantum states. We have restructured the highly complex discrimination task into smaller discrimination sub-problems of pure states. This reduction in complexity has allowed us to perform an exhaustive numerical analysis. On one hand, we utilized SDP to obtain the optimal average success probability of correctly identifying the edge between both domains for values of N up to 50. On the other hand, we studied the performance of an extensively used protocol, the SRM. The numerical analysis of this protocol does not depend on an optimization task; thus, its complexity is

less than that of SDP, reaching values of N up to 200.

This numerical advantage is due to the use of representation-theoretic tools introduced in Chapter 3, particularly Schur-Weyl duality. By taking into account the symmetries of the problem, we were able to decompose the large Hilbert space into smaller subspaces, each associated with a pure-state discrimination task. Direct brute-force optimization becomes infeasible even for moderate values of N, as the Hilbert space grows exponentially with the system size, scaling as d^N. While one might hope to use a Gram matrix approach to reduce complexity, this is not directly applicable here, as the states to be distinguished are mixed, and no known Gram matrix formulation exists for this class of states. The block structure obtained through Schur-Weyl duality not only reduces computational cost but also provides a clearer analytical understanding of the problem. This decomposition allowed us to evaluate the SRM protocol's performance for significantly larger values of N, and to establish its asymptotic optimality in this universal setting.

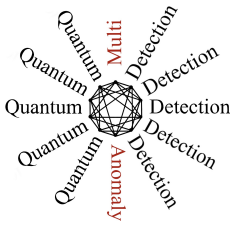
The numerical study suggests a convergence—in the asymptotic limit of large strings—between the success probability of quantum edge detection and that of the known-state scenario. To investigate this behavior, we performed an asymptotic expansion of the SRM protocol's success probability. However, no closed-form expression could be derived for the coefficients of this expansion, and its truncated form failed to capture the singular behavior observed for small values of λ. To overcome this limitation, we employed Padé approximants to accelerate the convergence of the series. This method produced strong agreement with the numerical results, accurately reproducing the singular behavior at small λ, and allowed us to estimate the asymptotic success probability for quantum edge detection across several values of the local dimension d. These results highlight the practical utility of Padé approximants in capturing nontrivial features of the asymptotic behavior, especially in instances where standard expansions fail capturing the observed data.

In the second part of this Chapter, we investigated an intermediate scenario in which one of the domains is composed of known states, while the other remains unknown. This setting also falls within the universal framework, but it captures a more practical situation where the observer has access to information about one domain—possibly through prior state preparation—while a sudden change leaves the second domain unspecified. Despite this asymmetry, the protocol remains universal in the sense that the POVM is not tailored to a particular quantum statem except to the reference one. As in the unknown scenario, we found that the problem could be decomposed into smaller—more tractable—sub-problems. A two-step protocol naturally arises: the first involves a sampling operation—with different structure from that of the unknown case—followed by pure-state discrimination tasks. The

Gram matrices characterizing this second step were also found to be semiseparable, maintaining the same structural properties observed in the fully unknown scenario. Numerical computations for both the optimal and SRM protocols showed strong agreement, indicating that the SRM remains nearly optimal even for small values of N.

From a technical perspective, our analysis establishes the asymptotic optimality of the SRM for two distinct instances of quantum edge detection. In both the fully universal setting and the intermediate case with one known domain, the SRM closely approaches the optimal success probability, even for relatively small system sizes. This is particularly significant because there are no known results regarding the optimality of the SRM for the type of discrimination problem encountered in this Chapter. Beyond these specific conclusions, our results showcase the effectiveness of our technique, which uses weak Schur sampling, the Gram matrix formulation of discrimination, reducing the complexity of the optimization, and acceleration of perturbative expansions through Padé analysis. The latter, though not commonly used in quantum information, proved particularly valuable in deriving asymptotic limits for the problem at hand. Altogether, our study shows that precise edge detection in one-dimensional strings of quantum systems is achievable, and it can reach remarkable levels of performance. In particular, we have shown that for large systems of qubits, the edge can be identified with success probability of approximately 65%, while for higher-dimensional systems, this value can approach unity as the local dimension increases. The intuition behind this fact is that, as the local dimension grows, two arbitrary quantum states—as the ones conforming the regions—tend to be more orthogonal as they live in a larger Hilbert space, which makes them easier to distinguish and thus enhances the performance of edge detection.

Throughout this Chapter, we have considered a scenario in which two different quantum states form separated domains within a one-dimensional string. Another relevant instance of universal discrimination arises in the context of anomaly detection. In that problem, the outliers are not confined to a defined region but instead appear at arbitrary positions within a domain composed by a different state. In Chapter 5, we investigate this primitive problem in detail, and compare its behavior with the edge detection scenario analyzed here.



5

Quantum multi-anomaly detection

giQ

It's in the Anomalies *that nature reveals its secrets.*

Johann Wolfgang von Goethe

The identification of rare events that deviate from expected behavior is known as *anomaly detection*. It is a fundamental primitive in classical data analysis and signal processing [TNB14], with a wide range of applications—from detecting intrusions in network security [SP04; TJ03], to identifying fraudulent financial activity [AMI16], or diagnosing faults in industrial systems and manufacturing processes [ZVF+23].

As a field, anomaly detection has been influenced by diverse approaches, including statistics, machine learning, spectral theory, and classical information theory [CBK09]. In this Chapter, we tackle the problem from a quantum information perspective. In this context, the data consists of a string of quantum systems described by specific quantum states: the expected behavior is modeled by a *reference* state, while the *anomalies* (or outliers) are individual quantum systems in a different quantum state.

Our goal is to detect the position of these anomalous quantum systems within a larger system. To this end, we formulate the task as a quantum state discrimination problem, where each possible configuration of the anomalies corresponds to a hypothesis to be tested. We investigate two distinct scenarios. In the first, we assume that both the reference and anomalous states are known to the observer. In this case, we characterize optimal measurements for both the minimum-error and zero-error settings. In the second, we consider a fully universal scenario, in which no prior information about the quantum states is available. Here, we show that it

is still possible to construct an effective, state-independent protocol that identifies the anomalies based solely on structural features of the data.

In the known-state scenario, the task exhibits a rich underlying symmetry. The assumption that anomalies are equally likely to occur at any position along the string leads to an algebraic structure governed by association schemes [BBI+21]. Using standard results from this framework, we present the optimal protocols for the minimum-error and zero-error scenarios and characterize their performances analytically. Moreover, for the minimum-error protocol, we further analyze the asymptotic behavior of the success probability in the large-string limit. We found that it converges to the success probability of the zero-error protocol, which remains constant for any string length and depends only on the number of outliers and the overlap between the reference and anomalous states.

In the universal scenario, where both states are unknown, the protocol exhibits a learning behavior similar to that discussed in Chapter 4 for the quantum edge detection problem [LGS+25a]. Here, the *pattern* is determined by the reference states, which provide the structure necessary for the protocol to learn. As the number of reference states increases, the protocol gathers more information about the expected behavior, thereby improving its ability to detect deviations from this pattern without any prior knowledge.

This Chapter approaches quantum anomaly detection from both known and universal perspectives. First, we analyze the known-state scenario, providing optimal measurements and performance bounds for both the minimum- and zero-error protocols. We then turn to the fully universal case, where no prior information about the states is assumed, and derive the optimal protocol and its performance. Altogether, this work offers a clear benchmark for the performance of quantum anomaly detection protocols and serves as a reference point for future studies, as the one covered in Chapter 6, where the anomalies are no longer prepared states but quantum processes.

5.1 Setting of the problem

Let us start by describing the key elements of the multi-anomaly detection problem. We denote the reference state of the string by $|\phi_0\rangle$, and the anomalous state by $|\phi_1\rangle$. For a fixed string length n and a given number k of anomalies, the total number of hypotheses to consider is $N = \binom{n}{k}$. Without loss of generality, we can consider the overlap between the reference and an anomalous state to be real and denote it as $c = |\langle\phi_0|\phi_1\rangle|$. We further assume that all hypotheses occur with equal prior probability. Each hypothesis corresponds to a global quantum state $|\Psi_r\rangle$, where

$r \subset \{1, \ldots, n\}$ is a subset of positions with cardinality $|r| = k$ that labels the position of the anomalies (see Fig. 5.1 for illustrative examples).

We now construct the Gram matrix of the set of hypotheses, which contains all the relevant information for a discrimination task. It is defined as the matrix of overlaps of the states, which in our case reads

$$G_{rs} = \langle \Psi_r | \Psi_s \rangle = (c^2)^{\delta(r,s)} , \tag{5.1}$$

where $\delta(r, s) := k - |r \cap s|$ quantifies the distance between two subsets r and s, and $|r \cap s|$ is the number of elements belonging to the intersection of the subsets r and s, and corresponds to half the Hamming distance [Ham50]. In Fig. 5.1, we illustrate the cases of one and two anomalies for $n = 4$ systems. In the first case, the distance between different hypotheses is always $\delta = 1$, whereas in the second case, both $\delta = 1$ and $\delta = 2$ distances appear.

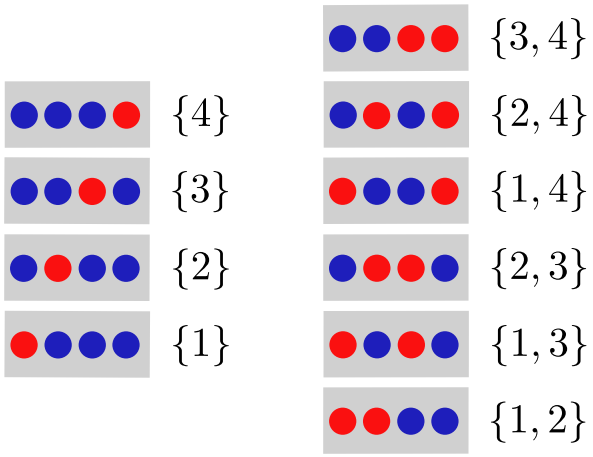


Figure 5.1: Set of hypotheses for the cases of $n = 4$. Subfigure (a) corresponds to $k = 1$, notice that the distance between any two hypotheses is $\delta = 1$. Subfigure (b) corresponds to $k = 2$, where distances can take values of $\delta = 1$, for example hypotheses $\{3, 4\}$ and $\{2, 4\}$ or $\delta = 2$, for instance hypotheses $\{3, 4\}$ and $\{1, 2\}$.

Decomposing the Gram matrix G in powers of c^2, we obtain a collection of symmetric matrices $\{A_i\}_{i=0}^{k}$, whose entries are 0's and 1's. This decomposition reads

$$G = \sum_{i=0}^{k} (c^2)^i A_i \, . \tag{5.2}$$

Each matrix A_i corresponds to the adjacency matrix of a generalized Johnson graph, and together they form the Bose–Mesner algebra $\mathcal{A}$ of the Johnson association scheme [BBI+21]. Consequently, the Gram matrix G belongs to this algebra as well. The elements of $\mathcal{A}$ commute and have zeros in their diagonal entries, except for the identity matrix $\mathbb{1} = A_0$. As we see below, the standard results of this association scheme—reviewed in Section 3.3—allow us to provide proofs of optimality as well as an elegant method to obtain analytical expressions for the success probability.

5.2 Minimum-error

As discussed in Chapter 2, the minimum-error protocol aims to minimize the error probability of guessing incorrectly. To each POVM element Π_r, we associate the hypothesis $|\Psi_r\rangle$. The conditional probability of obtaining outcome Π_s given the state $|\Psi_r\rangle$ is given by the Born rule $\mathrm{tr}\left(|\Psi_r\rangle\langle\Psi_r|\Pi_s\right)$ (here, we slightly abuse notation by labeling the measurement outcome with the same index as the corresponding POVM element). Success occurs when, upon measuring $|\Psi_r\rangle$, one obtains outcome Π_r, and then the success probability reads

$$P_{\mathrm{s}} = \sum_r \eta_r \mathrm{Tr}\left(|\Psi_r\rangle\langle\Psi_r|\Pi_r\right), \tag{5.3}$$

where η_r denotes the prior probability of the hypothesis $|\Psi_r\rangle$. From this point onward, we assume that all hypotheses are equally likely, i.e., $\eta_r = \eta_s = 1/N = \binom{n}{k}^{-1} \forall\, r, s$, reflecting the complete lack of knowledge of the observer regarding the position of the anomalies.

A particularly relevant protocol in this context is the SRM. The matrix $\sqrt{G} = S$ defines the POVM through its elements S_{rs}, as they are projections into a measurement basis $\{|m_k\rangle\}$, i.e., $S_{rs} = \langle m_r|\Psi_s\rangle$ (see Chapter 2 for a detailed discussion). Therefore, under equal priors, the average success probability of the SRM protocol reads

$$P_{\mathrm{s}} = \frac{1}{N}\sum_r |S_{rr}|^2 . \tag{5.4}$$

Now, since the Gram matrix Eq. (5.2) belongs to the Bose-Mesner algebra $\mathcal{A}$, any function of G also belongs to it, and in particular its square root $S = \sqrt{G}$. Therefore, the matrix S can be written in terms of the basis $\{A_i\}_{i=0}^k$ as

$$S = \sum_{i=0}^{k} s_i A_i . \tag{5.5}$$

Thus, the diagonal entries of S are all equal, i.e., $S_{ll} = S_{hh} = \forall l, h$, since only the identity element $A_0 = \mathbb{1}$ contributes to the diagonal. Optimality then follows directly: as reviewed in Chapter 2, the SRM is optimal if and only if all diagonal entries of S are equal [DP15].

The algebraic structure of the Johnson association scheme allows us to further characterize the success probability in terms of the eigenvalues of the Gram matrix. To this end, we first express the success probability in a more convenient form. Since all diagonal entries of S are equal, they can be written as $S_{ll} = (\operatorname{tr} S)/N$, with $\operatorname{tr} S = \sum_l \sqrt{\lambda_l}$, where λ_l are the eigenvalues of the Gram matrix. This leads to the compact expression [SBC+16]:

$$P_{\mathrm{s}} = \left(\frac{\sum_{l=1}^{N} \sqrt{\lambda_l}}{N} \right)^2 . \tag{5.6}$$

Hence, the task reduces to determining the eigenvalues λ_l. Since all adjacency matrices A_i commute, they are all simultaneously diagonalizable in a common eigenbasis, $\{E_i\}$, which corresponds to the second basis of the Bose-Mesner algebra introduced in Section 3.3. Any matrix belonging to $\mathcal{A}$ diagonalizes in this basis. Therefore, the eigenvalues of G are given by the weighted sum of the eigenvalues of the adjacency matrices A_i

$$\lambda_j(G) = \sum_{i=0}^{k} (c^2)^i \lambda_j(A_i) \, . \tag{5.7}$$

In the remainder of this Chapter, we will omit the dependence on G and simply write $\lambda_j := \lambda_j(G)$. For the eigenvalues of the adjacency matrices, we will, however, write the explicit dependence, $\lambda_j(A_i)$.

Recall that every P- and Q-polynomial admits a description of its adjacency matrices—and their corresponding eigenvalues—in terms of a family of orthogonal polynomials evaluated at the first adjacency matrix $A := A_1$. In the case of the Johnson association scheme, the eigenvalues $\lambda_j(A_i)$ can be expressed as the evaluation of the associated orthogonal polynomial at the eigenvalue $\lambda_j(A)$ [BBI+21]

$$\lambda_j(A_i) = (-1)^i \binom{k}{i} R_i(\lambda_j(A) + k; 0, n - 2k, k) \, , \tag{5.8}$$

where R_i is the dual Hahn polynomial of degree i [KS96], which are the orthogonal polynomials arising from the Johnson scheme [c.f. Eq. (3.82) for its explicit combinatorial structure].

Introducing Eq. (5.8) in Eq. (5.7), we observe that the eigenvalues of G, expressed as sums of orthogonal polynomials weighted by powers of c^2, correspond to the generating functions of the dual Hahn polynomials. These take the form

$$\lambda_j = (1-c^2)^j \, {}_2F_1\left(\begin{matrix} j-k, & -n+k+j \\ & 1 \end{matrix}; c^2\right), \tag{5.9}$$

with corresponding multiplicities

$$m_j = \binom{n}{j} - \binom{n}{j-1}, \tag{5.10}$$

where $j = 0, 1, \ldots, k$, with the convention that $\binom{n}{-1} = 0$, and ${}_2F_1$ is the hypergeometric function [given in Eq. (3.83)]. Note that the number of distinct eigenvalues is $k+1$, i.e., just the number of anomalies plus one, and these are monotonically decreasing with j. For instance, in the simplest case of one anomaly, covered in [SLH+24], there are only two different eigenvalues, $\lambda_0 = 1 + (n-1)c^2 > \lambda_1 = (1-c^2)$ with multiplicities 1 and $n-1$, respectively.

Taking into account the degeneracy of the eigenvalues, the success probability of the minimum-error protocol reads

$$P_{\mathrm{s}} = \left(\sum_{j=0}^{k} \frac{m_j}{N} \sqrt{\lambda_j}\right)^2, \tag{5.11}$$

where $\{\lambda_j\}_j$ are given in Eq. (5.9), and $\{m_j\}_j$ in Eq. (5.10).

5.2.1 Asymptotic regime

For a fixed number of anomalies and an increasing number of total systems n, it could be expected that the success probability vanishes as the number of possible locations for the anomalies—and hypotheses—goes to infinity. However, as we next show, the asymptotic limit of P_{s} for $n \to \infty$ remains finite.

We begin by noting that the eigenvalues of the Gram matrix are polynomials in n of degree $k-j$ for $j = 0, 1, \ldots, k$, that is, $\lambda_j \sim O(n^{k-j})$. On the other hand, the ratio between the multiplicities of each eigenvalue and the total number of hypotheses scales as $m_j/N \sim O\left(n^{j-k}\right)$. Then, it follows that, in the large n limit, the leading contributions to Eq. (5.11) arise from λ_k and $\sqrt{\lambda_k \lambda_{k-1}}$, i.e., the two smallest eigenvalues.

By substituting the values of the multiplicities from Eq. (5.10), and the explicit expression of the eigenvalues λ_k and λ_{k-1} from Eq. (5.8), we obtain the asymptotic expansion

$$P_{\mathrm{s}} = (1-c^2)^k + \frac{2kc(1-c^2)^{k-1/2}}{\sqrt{n}} + O\left(\frac{1}{n}\right) . \tag{5.12}$$

Thus, for large n, the success probability approaches $(1-c^2)^k$ from above, converging slowly [$O\left(1/\sqrt{n}\right)$ as noted]. Note that this expansion is valid in the regime $k/n \ll 1$.

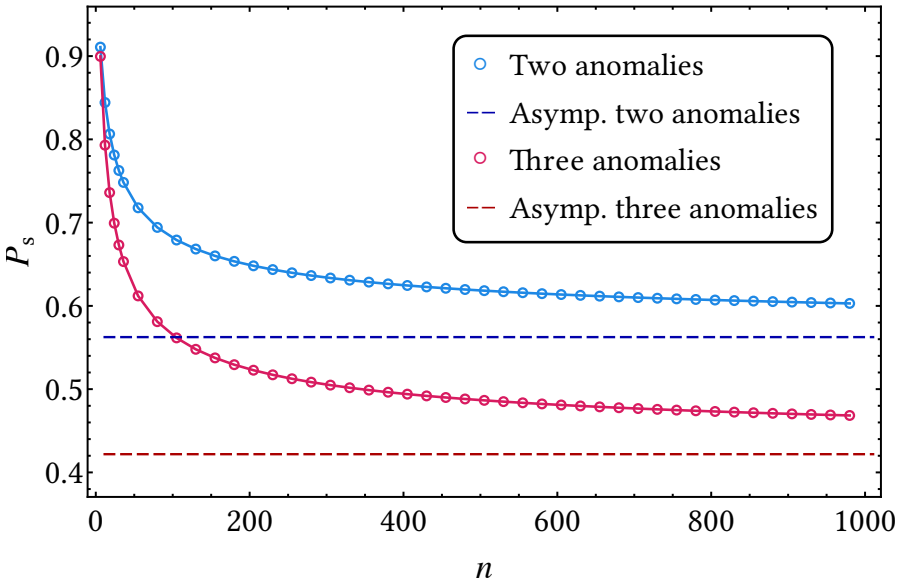


Figure 5.2: Comparison of the success probability of two anomalies, the upper curve (blue), and the case of three anomalies, the lower curve (red), both for a fixed overlap $c = \frac{1}{2}$. The dashed, upper and lower lines correspond to the asymptotic limits of the success probabilities of two and three anomalies, respectively.

In Fig. 5.2, we plot the success probability for the cases of two and three anomalies as a function of the string length n. As shown, the convergence to the asymptotic value is rather slow, as the leading correction scales with $O(1/\sqrt{n})$. The gap between the finite-n success probability and its asymptotic limit remains sizable even for relatively large values of n. This gap is more pronounced for higher values of k, due to the factor k appearing in the second term of Eq. (5.12).

5.3 Zero-error identification

The scenario of zero-error identification can also be addressed naturally within this framework. Here, the goal is to identify the position of the anomalies with certainty, i.e., without allowing any probability of error. Naturally, this can only be achieved at the expense of allowing for inconclusive outcomes.

The objective is to design a protocol that maximizes the rate of conclusive outcomes, or equivalently, that minimizes the rate of inconclusive ones. The corresponding POVM consists of $N+1$ elements: $\{\Pi_r\}_r$ with $r \subset \{1,\dots,n\}$ and $|r| = k$, and an extra $\Pi_?$ element associated with the inconclusive outcome. Each Π_r element identifies with zero-error the state $|\Psi_r\rangle$. Accordingly, the structure of the POVM operators is tightly constrained by the zero-error conditions $\mathrm{tr}\,(\Pi_r|\Psi_s\rangle\langle\Psi_s|) = 0$ for $r \neq s$.

The optimization problem can be formulated as a SDP, as discussed in Section 3.1.5:

$$P_{\mathrm{s}} = \frac{1}{N} \max_{\Gamma} \ \mathrm{tr}\,\Gamma \tag{5.13a}$$

$$\text{s.t.} \quad G - \Gamma_D \geq 0 \tag{5.13b}$$

$$\Gamma \geq 0 \tag{5.13c}$$

Here, Γ_D is the diagonal part of matrix Γ i.e., $[\Gamma_D]_{rs} = \delta_{rs}\Gamma_{rr}$. The entries Γ_{rr} represent the conditional probabilities of correctly identifying $|\Psi_r\rangle$ without error.

This formulation arises from the structure of the POVM elements that satisfy the zero-error conditions. Such elements must be of the form $E_k = \gamma_k|\tilde{\varphi}_k\rangle\langle\tilde{\varphi}_k|$, where the (unnormalized) states satisfy $\langle\tilde{\varphi}_k|\Psi_l\rangle = \delta_{kl}$, i.e., these states are the rows of the (pseudo) inverse of the matrix $X = \sum_k |\Psi_k\rangle\langle k|$, with $\{|k\rangle\}$ an orthonormal basis of dimension N. The constraint $G - \Gamma_D \geq 0$ then follows from the positivity condition of the inconclusive POVM element, $\mathbb{1} - \sum_k \gamma_k|\tilde{\varphi}_k\rangle\langle\tilde{\varphi}_k| \geq 0$, by multiplying it from the left and right by $X^\dagger$ and X, respectively (see Section 3.1.5 or [SCM17] for further details).

As already discussed in Section 3.1, SDPs are more than a numerical tool; all the theory behind this mathematical field allows us to find analytical solutions to discrimination tasks. In this scenario, we consider an ansatz for the primal program in Eqs. (5.13a–5.13c), and then, via the Lagrangian method and complementary slackness, find a dual instance that attains the same value as the primal program. The ansatz we consider is $\Gamma = (1-c^2)^k\mathbb{1}$, which yields a success probability

$$P_{\mathrm{s}} = (1-c^2)^k\,. \tag{5.14}$$

The operator Γ belongs to the feasible set of the primal program in Eqs. (5.13a–5.13c), as it is a positive definite matrix and the smallest eigenvalue of the Gram matrix, $\lambda_{\min} = (1-c^2)^k$. If the primal variable Γ is optimal, the Lagrangian

$$\mathcal{L} = \operatorname{tr}\Gamma^* + \operatorname{tr}\Big(Y^*\big(G - \Gamma_D^*\big)\Big) + \operatorname{tr}\left(Z^*\,\Gamma^*\right), \tag{5.15}$$

should output the same value as the primal program, P_{s}. A standard practice is to set the slack variable $Z = 0$, with this, we note that $\mathcal{L} = P_{\text{s}}$ if the quantity

$$\operatorname{tr}\Big(Y^*\big(G - \Gamma_D^*\big)\Big) = 0\,, \tag{5.16}$$

that is, if the dual variable Y^*, and $G - \lambda_{\min}\mathbb{1}$ have orthogonal supports. Therefore, the dual variable must be proportional to the projector associated with the smallest eigenvalue of the Gram matrix, $Y^* \propto E_k$. To find the explicit dual variable, we look at the dual instance of the SDP

$$P_{\text{s}} = \min_{Y}\ \operatorname{tr} G\,Y\,, \tag{5.17a}$$

$$\text{s.t.}\ \ Y_D - \frac{1}{N}\mathbb{1} \geq\ 0 \tag{5.17b}$$

$$Y \geq 0\,. \tag{5.17c}$$

The dual instance outputs the same value as the primal program for the dual variable

$$Y = \frac{N}{m_k}E_k\,. \tag{5.18}$$

From this choice, it follows that $P_s = \operatorname{tr} G\,Y$ coincides with the value of the primal SDP. So, if Y is a feasible solution to Eqs. (5.17a–5.17c), it is an optimal solution as the values of both primal and dual objective functions coincide. We just have to prove that Y_D, the matrix of the diagonal terms of Y, satisfies $Y_D \geq \mathbb{1}$ (note that Y is non-negative by construction as E_k is an orthogonal projection). Let us decompose E_k as a linear combination of adjacency matrices, $E_k = (1/N)\sum_{j=0}^{k} q_k(j)A_j$, where $q_k(j)$ are given by the Hahn polynomials (see Section 3.3.5) and note again that the only matrix with diagonal elements different from 0 is $A_0 = \mathbb{1}$, i.e.,

$$Y_D = \frac{q_k(0)}{m_k}\mathbb{1}\,. \tag{5.19}$$

But $q_k(0) = m_k$, due to the duality of the association scheme and their associated polynomials [BBI+21]. Hence $Y_D = \mathbb{1}$, which completes the proof. This result could

have been readily anticipated, as the symmetry of the problem already requires that the optimal Γ_D satisfies $\Gamma_D \propto \mathbb{1}$ and then Eqs. (5.13a–5.13c) are just the program defining the lowest eigenvalue of G.

We also see that Eq. (5.14) coincides with the leading term of the minimum-error success probability in Eq. (5.12), thus we conclude that the unambiguous measurement is also optimal for the minimum-error protocol for large n. This measurement can be realized by a local protocol that checks if each particle has a projection or not in the complementary subspace of the reference state. This result can be extended to the case where the anomalies are mixed states since the optimal protocol is fixed by the reference state.

5.4 Universal protocol

We finally tackle the anomaly identification task when the reference and anomalous states are unknown to the observer. That is, we aim at finding a universal protocol that does not use any information about the state—only that they are different—and hence works for any arbitrary pair of reference and anomalous states.

In this scenario, all hypotheses are generated by applying a permutation to the fiducial state, in which all k anomalies are located at the end of the string

$$\rho_\sigma = \int U_\sigma |\phi_0\rangle\langle\phi_0|^{\otimes(n-k)} \otimes |\phi_1\rangle\langle\phi_1|^{\otimes k} U_\sigma^\dagger \, \mathrm{d}\phi_0 \, \mathrm{d}\phi_1 \,, \tag{5.20}$$

where ϕ_0 (ϕ_1) corresponds to the reference (anomalous) state, U_σ stands for the unitary transformation that acts on the tensor-product space as a permutation $\sigma \in \mathcal{S}_n \subset S_n$, in the subset of relevant permutations for k anomalies of S_n—the symmetric group—that is, permutations yielding the same state ρ_σ are accounted only once. Here, $\mathrm{d}\phi_i$ is the measure of the uniform distribution of $|\phi_i\rangle$.

It is convenient to index each hypothesis by the permutation σ that takes a fiducial state where all anomalies occur in the last positions of the hypothesis in question. Using Schur lemma (see Section 3.2), the hypotheses, in the Schur basis, read

$$\rho_\sigma = c_k \bigoplus_\lambda \mathbb{1}_\lambda \otimes \Omega_\sigma^\lambda \,, \tag{5.21}$$

where the normalization constant is $c_k = 1/(d_k^{\mathrm{sym}} d_{n-k}^{\mathrm{sym}})$, and d_k^{sym} is the dimension of the symmetric subspace of k parties, given by the hook length formula (see Section 3.2.3 for further details).

As already discussed in Section 3.2, the parameter λ labels the irrep of the joint action of the groups $\mathrm{SU}(d)$ and the symmetric group S_n over the whole state space

$(\mathbb{C}^d)^{\otimes n}$, and it is usually identified with bipartitions $(n-\lambda,\lambda)$ of n elements. The first tensor element, $\mathbb{1}_\lambda$, is defined over the subspace associated with the irreducible label λ of $\mathrm{SU}(d)$, and the second factor, Ω_σ^λ, stands for an operator that acts in the irrep of S_n labeled by λ, and it is a rank-1 projector (see Section 3.2.3, for a detailed discussion). Notice that, at variance with Chapter 4, we have an extra index σ that stands for the ordering of the two different states within the array, while in the quantum edge detection problem, all sequences were ordered.

We can now derive the explicit expression of the success probability for the minimum-error protocol. The optimal measurement, fortunately, turns out to be again the SRM. For mixed states, its elements are $\Pi_\sigma = \Lambda^{-1/2}\rho_\sigma\Lambda^{-1/2}$, where

$$\Lambda = \sum_{\sigma\in\mathcal{S}_n}\rho_\sigma = \bigoplus_\lambda C_\lambda \mathbb{1}_\lambda \otimes \mathbb{1}^\lambda\,. \tag{5.22}$$

where we have used Schur lemma, since Λ commutes with any element $\sigma\in S_n$ as well as with any element $U\in\mathrm{SU}(d)$, and C_λ is a proportionality constant for each partition λ.

As in the previous section, we write the dual SDP for the success probability:

$$\min_Y \operatorname{tr} Y \tag{5.23a}$$

$$\text{s.t.}\quad Y-\rho_\sigma\geq 0\quad \text{for all}\quad \rho_\sigma \tag{5.23b}$$

If we consider $Y=\sum_\sigma \Pi_\sigma\rho_\sigma$, primal and dual programs give the same value. To prove optimality, one only has to check that Y satisfies the feasibility conditions. Naturally, the constraints $Y-\rho_\sigma\geq 0$, with $\sigma\in\mathcal{S}_n$, are just the well-known Holevo conditions [Hol73], as described in Section 2.5.2.

It is easy to compute that

$$Y = c_k \bigoplus_\lambda \mathbb{1}_\lambda\otimes\mathbb{1}^\lambda\,, \tag{5.24}$$

and to check that $Y-\rho_\sigma\geq 0$ holds, since $\mathbb{1}^\lambda\geq\Omega_\sigma^\lambda$.

Hence, the optimal success probability of the universal protocol is just $(\operatorname{tr} Y)/N$, which reads

$$P_{\mathrm{s}} = \sum_{\lambda=0}^{k}\left(\frac{(n-2\lambda+1)^2}{(n-\lambda+1)^2}\frac{\binom{n-\lambda+d-1}{d-1}}{\binom{n-k+d-1}{d-1}}\frac{\binom{n}{\lambda}}{\binom{n}{k}}\frac{\binom{\lambda+d-2}{d-2}}{\binom{k+d-1}{d-1}}\right), \tag{5.25}$$

where we have used the explicit hook length formulae for the dimensions of the irreps of $\mathrm{SU}(d)$ and S_n for bipartitions $(n-\lambda,\lambda)$ [see Eqs. (3.40–3.41)].

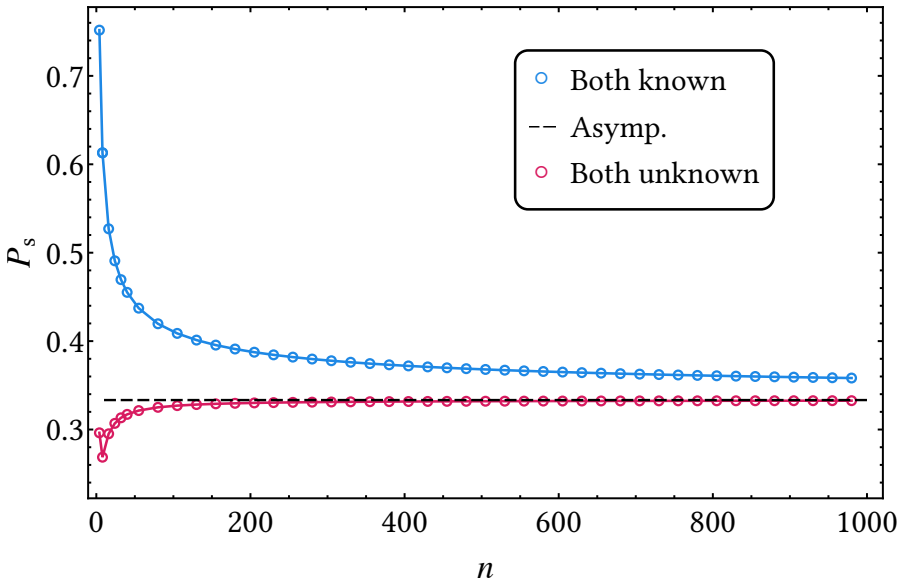


Figure 5.3: Comparison between the average (over c) success probability for the case of known reference and anomalous states of dimension $d = 2$, the upper curve (blue online), and the success probability for the universal protocol, the lower curve (red online). The dashed black line corresponds to the asymptotic limit of both success probabilities $P_s = 1/(k+1)$

The asymptotic behavior of the success probability in the regime of large n is dominated by the partition $\lambda = k$, as the most important factor in Eq. (5.25) is the ratio $\binom{n}{\lambda}/\binom{n}{k}$ and the contributions of the rest of partitions $\lambda = l < k$ vanish as $O(1/n^{k-l})$. The asymptotic expansion of the success probability then reads

$$P_s = \frac{d-1}{d-1+k} + \mathcal{O}\left(\frac{1}{n}\right) , \tag{5.26}$$

Interestingly, we observe for this problem the same learning behavior as for quantum edge detection [LGS+25a]. The success probability for the universal protocol increases with the string length, reaching the asymptotic value of the average success probability for known states and anomalies (see Fig. 5.3),

$$P_s^{\text{known}} = \int_0^1 (1-c^2)^k (1-c^2)^{(d-2)} (d-1)\, \mathrm{d}c^2 = \frac{d-1}{d-1+k} \tag{5.27}$$

where $(d-1)(1-c^2)^{(d-2)}\mathrm{d}c^2 =: \mathrm{d}\mu(c^2)$ is the uniform measure of the square of the overlap [SMM+19], and we have dropped the vanishing terms in Eq. (5.12).

The reasons behind these similar behaviors stem from the fact that, for known states and large n, the optimal measurement approaches an unambiguous protocol that detects anomalies by projecting on the orthogonal space of the reference state. Therefore, the knowledge of the anomalous state is not required, for it would be impossible to determine this state with a fixed (and small) number of instances distributed at random places. In this asymptotic limit, only the reference state needs to be known to devise the protocol, and this can be done by using a vanishing number of systems (e.g. $\propto \sqrt{n}$). The success probability Eq. (5.25) is depicted in Fig. 5.3. We observe that, as n grows, counterintuitively at first sight, P_{s} increases. The logic that explains this behavior is that increasing n allows the protocol to learn about the reference state, which gives an advantage that exceeds the detrimental effect of having a larger set of possible hypotheses. It is also evident that, although the universal protocol reaches the value Eq. (5.27) rather fast, the average of the success probability for known reference states and anomalies is notably above this value even for significantly large string lengths. Note also that, as expected, $P_{\mathrm{s}}^{\mathrm{known}}$ in Eq. (5.27) tends to one as $d \to \infty$, which reflects the fact that random states tend to be more orthogonal as d grows.

5.5 Discussion

We have presented two distinct instances of multiple-anomaly identification. In the case where both the reference and anomalous states are known, we proved that the SRM protocol is optimal by exploiting the algebraic structure underlying the problem. Specifically, by noting that the Gram matrix arising from the ensemble of quantum states belongs to the Bose–Mesner algebra associated with the Johnson association scheme.

This result extends naturally to other quantum discrimination tasks in which the Gram matrix is structured by association schemes. For instance, in [SMM+19], the classification of known states is linked to the Hamming scheme, while in certain extensions of the QCP problem [SBC+16], the relevant algebraic structures correspond to cyclic graphs and their associated schemes. Furthermore, our proof applies to broader algebraic settings: the only requirement is that the diagonal elements of the square root of the Gram matrix depend solely on the identity matrix. Similar structures appear in other problems, such as quantum state exclusion [DLB+25], although this property was not explicitly used there.

We characterized the asymptotic performance of the SRM protocol, showing that the success probability is non-vanishing in the limit of large strings. Its asymptotic value depends only on the overlap between the reference and anomalous states,

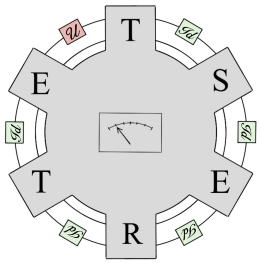
and the number of anomalies. The decay towards this value is of order $O(n^{-1/2})$, indicating a relatively slow convergence.

We also addressed the problem in the zero-error discrimination setting, where no incorrect guesses are allowed, at the expense of inconclusive answers. In this context, we identified the optimal protocol and fully characterized its performance. Our approach combined SDP complementary slackness and theoretical tools from association schemes to derive the analytical solution to the problem. Interestingly, the optimal protocol turns out to be local: it tests each party individually to check if it has support on the reference state or not.

This protocol exhibits two particularly relevant features. First, it allows for online implementation, that is, parties can be measured sequentially, and if the k anomalies are found, the process can be halted early, thereby reducing the number of measurements required. Second, it is non-destructive on the reference states, which are typically the intended output of the state source—another significant advantage in terms of resource preservation. While the minimum-error protocol achieves a higher overall success probability, the zero-error strategy offers stronger guarantees and resource-saving advantages. In the asymptotic regime, their performances converge, indicating that unambiguous detection can be achieved with minimal sacrifice in the success probability. Moreover, this protocol remains optimal even when the anomalous state is unknown, i.e., in a universal setting. In this intermediate scenario, the measurement is not tailored to any specific anomalous state, yet it remains local and non-destructive.

The final part of the Chapter explored a fully universal scenario, in which neither the reference state nor the anomalous states are known to the observer. To address this task, we once again employed SDP techniques to derive the analytical solution. We found that the SRM remains optimal in this setting as well. By using representation-theoretic tools, we were able to characterize the performance of the protocol and derive a closed expression valid for arbitrary string length, number of anomalies, and local system dimension.

In this setting, the success probability increases with the size of the string, asymptotically approaching the average success probability obtained in the known-state case, analogous to that observed in Chapter 4. In here, the protocol effectively learns from the reference state, as when the string length increases, there is more available information to be extracted. In contrast, learning from the anomalies is unfeasible, since they form only a vanishing fraction of the string. This asymmetry is captured in the success probability, which coincides with that of the local strategy based on projection onto the reference state.



6

Identification of anomalous quantum unitary evolutions

Life is a process of becoming, a combination of states we have to go through. Where people fail is that they wish to elect a state and remain in it. Evolution *is the movement through states.*

Anaïs Nin

6.1 Introduction

In Chapter 5 we introduced the task of anomaly detection for quantum states. The same principles can be naturally extended to quantum processes, and in the present Chapter, we analyze this more general setting. We consider a set of devices that are assumed to apply a fixed and known unitary transformation, which we take as the reference behavior. Among them, a subset is anomalous in the sense that they implement a different unitary transformation, unknown to the observer. Our goal is to identify, with zero-error, which of the devices are anomalous. Although one could also address the problem under the minimum-error criterion, zero-error protocols are particularly natural in this setting: the devices that behave as intended should remain untouched, while only the faulty ones need to be singled out with certainty.

Furthermore, in the universal scenario of the previous chapter, the first correction to perfect identification was shown to scale as the inverse of the number of devices, and a similar behavior is expected here. In fact, for the case of a single anomaly, the optimal minimum-error protocol coincides with the zero-error strategy followed by classical post-processing, supporting the choice of the zero-error formulation throughout this Chapter.

As in previous chapters of this thesis, we frame the task as a discrimination problem. The difference is that the objects we aim to discriminate are now quantum channels rather than quantum states. To treat them, we use Choi–Jamiołkowski isomorphism, which maps quantum channels to positive operators. This allows us to apply the quantum networks formalism, where the role of POVMs is played by quantum testers. In the same line of Chapters 4 and 5, as the anomalies are unknown to the observer, we devise universal protocols that work for any possible unitary transformation. The figure of merit is then the average—over all possible unitaries—success probability of correctly identifying the position of the anomalies.

The complexity of the task depends on the number of anomalies. We begin by studying the case in which only one of the devices is anomalous. In this regime, the Choi matrices present local symmetries that enable us to restrict the search for optimal testers to a reduced family of operators. These operators can be described in terms of the algebra of partially transposed permutations. Working within this mathematical framework, we derive an analytic expression for the optimal success probability, valid for arbitrary local dimension and any number of devices. The corresponding protocol is shown to be implementable in a local and parallel manner, using maximally entangled states as probes.

The situation changes dramatically when turning to the case of two anomalies. The local symmetries are no longer present; instead, a global symmetry of the whole set of parties appears. This difference enlarges the space of feasible testers—even after reducing the search to the subspaces that respect the symmetry—and the optimization problem becomes substantially more demanding. Despite this increase in complexity, we establish that for qubits and parallel implementations, the same local protocol introduced for the single-anomaly scenario remains optimal. This is a nontrivial finding, as the lack of local symmetries makes the proof considerably more involved. Numerical evidence for small system sizes further indicates that sequential strategies offer no advantage over parallel ones, lending additional support to the result.

The proof of optimality combines SDP techniques with representation-theoretic tools. In particular, the algebra of partially transposed permutations and its block-diagonal decomposition via mixed Schur–Weyl duality play a central role, as they

reduce the analysis of otherwise high-dimensional operators into tractable components. These methods allow us to demonstrate that the optimal success probability in the two-anomaly qubit case is non-vanishing independently of the number of devices.

6.2 Setting of the problem

We begin by introducing the mathematical framework and notation for the task considered in this Chapter. Each device $\mathcal{D}_i$, with $i = 1, 2, \ldots, n$, implements a unitary channel acting on a single system. A subset of k devices is faulty, and apply an *unknown* anomalous unitary transformation W, while the remaining devices apply the expected *known* unitary V.

Since the action of the known unitary can be inverted, without loss of generality, we reduce the problem to distinguishing the identity channel $\mathcal{I}(\rho) := \rho$, from the action of the unitary

$$\mathcal{U}(\rho) := U\rho U^\dagger\,, \quad \text{with } U = V^\dagger W\,. \tag{6.1}$$

For a given subset $r \subset \{1, \ldots, n\}$ of size $|r| = k$, we define the corresponding hypothesis as the channel

$$\tilde{\mathcal{F}}_r(\cdot) := \mathcal{I}_{\bar{r}}(\cdot) \otimes \mathcal{U}_r^{\otimes k}(\cdot)\,, \tag{6.2}$$

where $\bar{r}$ denotes the complementary set of r. That is, every non-faulty device acts trivially, while faulty devices apply the same unknown unitary U.

The anomaly-identification problem thus consists of distinguishing among the set of channels

$$\{\tilde{\mathcal{F}}_r : \; r \subset \{1, \ldots, n\}\,, \; |r| = k\}\,, \tag{6.3}$$

corresponding to the $N = \binom{n}{k}$ possible locations of the faulty devices. We model our complete lack of knowledge about the location of the anomalies by setting all prior probabilities $\eta_r = 1/N$.

Throughout this Chapter, we use the Choi-Jamiołkowski isomorphism and the tester formalism. For a channel $\Phi : \mathcal{H}_{\text{in}} \to \mathcal{H}_{\text{out}}$, its Choi matrix is

$$J_\Phi = \sum_{i,j=0}^{d-1} |i\rangle\langle j| \otimes \Phi(|i\rangle\langle j|) \in \mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{out}}\,, \tag{6.4}$$

where $|i\rangle$ is the computational basis. For notational clarity, we use the subindices "in" and "out" (note the difference with Section 2.4, in which odd and even numbers denote in/out subspaces).

The unambiguous discrimination task is carried out by quantum testers, which play for channels an analogous role to that of POVMs for states: they are collections of positive operators $\{T_r\}_r$, associated with possible outcomes of the discrimination protocol, which satisfy constraints ensuring complete positivity and causality of the overall network (see Secs. 2.4 and 3.1.5). In this setting, the average success probability reads

$$P_{\mathrm{s}} = \int \mathrm{d}U \sum_r \frac{1}{N} \operatorname{tr}\left(T_r^T \tilde{F}_r\right) , \tag{6.5}$$

where T_r denotes the tester associated with identifying the anomaly at positions labeled by r, and the operator $\tilde{F}_r$ denotes the Choi matrix of channel $\tilde{\mathcal{F}}_r$. The average is taken with respect to the Haar measure [CŚ06; Mel24], since the anomalous unitaries U are unknown and thus can be assumed to be uniformly distributed.

As in Chapters 4 and 5, the universality of the detection protocol is ensured by the independence of the testers $\{T_r\}_r$ from the Choi matrices. This independence allows us to define the effective Choi matrices F_r as

$$F_r = \int \mathrm{d}U \tilde{F}_r = |\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|^{\otimes(n-k)}_{\bar{r}_{\mathrm{in}}\bar{r}_{\mathrm{out}}} \otimes C^{(k)}_{r_{\mathrm{in}}r_{\mathrm{out}}} , \tag{6.6}$$

where $|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|$ denotes the Choi matrix of the identity channel and the subscript indicates the corresponding bipartite systems. The matrix $C^{(k)}$ captures the action of the anomalous unitaries and is given by

$$C^{(k)}_{r_{\mathrm{in}}r_{\mathrm{out}}} = \int (\mathbb{1}^{\otimes k}_{r_{\mathrm{in}}} \otimes U^{\otimes k}_{r_{\mathrm{out}}}) \, |\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|^{\otimes k} \, (\mathbb{1}^{\otimes k}_{r_{\mathrm{in}}} \otimes U^{\otimes k}_{r_{\mathrm{out}}})^\dagger \, \mathrm{d}U \, . \tag{6.7}$$

This integral admits a closed form when $k \leq d$

$$C^{(k)}_{r_{\mathrm{in}}r_{\mathrm{out}}} = \sum_{\sigma,\pi\in S_k} \mathrm{Wg}(\sigma\pi^{-1}, d) \, U_{r_{\mathrm{in}}}(\sigma) \otimes U^*_{r_{\mathrm{out}}}(\pi) \, , \tag{6.8}$$

where S_k is the symmetric group on k elements, $U(\sigma)$ is the unitary representation of the permutation $\sigma \in S_k$ in a Hilbert space of local dimension d, and $\mathrm{Wg}(\sigma, d)$ is the Weingarten function [Wei78; CMN22]. The full derivation of this closed form can be found in [CŚ06] and lies outside the scope of this Chapter. Still, some intuition can be given: integrating over the unitary group effectively projects the problem onto the algebra generated by permutations. The outcome is thus a convex combination of permutation operators from the symmetric group, with the Weingarten function providing the corresponding coefficients.

6.3 Optimal zero-error protocol

As discussed previously, our goal is to unambiguously identify the anomalous devices. The optimal success probability in channel discrimination is achieved using testers. The most general implementation of a quantum tester is a sequential (or adaptive) strategy, in which a prepared state is transformed by a quantum channel and the output of this action can be subsequently transformed by another channel, instead of being measured. The SDP formulation for the zero-error discrimination of quantum channels via sequential implementation of testers reads

$$P_{\mathrm{s}} = \max_{\{T_r\}} \sum_{|r|=k} \mathrm{tr}(T_r^T F_r) \tag{6.9a}$$

$$\text{s.t.} \quad \mathrm{tr}(T_r^T F_s) = 0 \quad \forall r \neq s \tag{6.9b}$$

$$T_r \geq 0 \tag{6.9c}$$

$$T_? + \sum_r T_r = \mathbb{1}_n \otimes R^{(n)} \tag{6.9d}$$

$$\mathrm{tr}_{i_{\mathrm{in}}}(R^{(i)}) = \mathbb{1}_{i_{\mathrm{out}}} \otimes R^{(i-1)} \quad i = 2, \ldots, n \tag{6.9e}$$

$$\mathrm{tr}_{1_{\mathrm{in}}}(R^{(1)}) = 1\,. \tag{6.9f}$$

Here, each operator T_r is the tester corresponding to the hypothesis in which the anomaly is located at positions labeled by r, while $T_?$ represents the inconclusive outcome. The constraint in Eq. (6.9b) is the zero-error condition for channel identification and the rest arise from the causal relations of quantum networks discussed in Section 2.4 [c.f. Eqs. (2.19–2.23)].

6.3.1 Single anomaly

We begin by considering the case of a single anomaly, $k = 1$. In this setting, the Choi matrices $\{F_r\}$ exhibit a very convenient symmetry: each F_r is invariant under the action of local unitaries $U \otimes U^*$ acting on the in–out pair of every bipartite system $j \in \{1, \ldots, n\}$. Explicitly, they satisfy

$$F_r = \int \left(\bigotimes_{j=1}^{N} (U_{j_{\mathrm{in}}} \otimes U^*_{j_{\mathrm{out}}}) \right) F_r \left(\bigotimes_{j=1}^{N} (U_{j_{\mathrm{in}}} \otimes U^*_{j_{\mathrm{out}}})^\dagger \right) \prod_{j=1}^{N} \mathrm{d}U_j\,. \tag{6.10}$$

This symmetry follows from the fact that $|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|$ lies in the algebra of partially transposed permutations, and the Haar averaging in Eq. (6.8) for $k = 1$ produces a Choi matrix proportional to the identity operator, which itself belongs to the same algebra (see Fig. 6.1 for a diagrammatic representation).

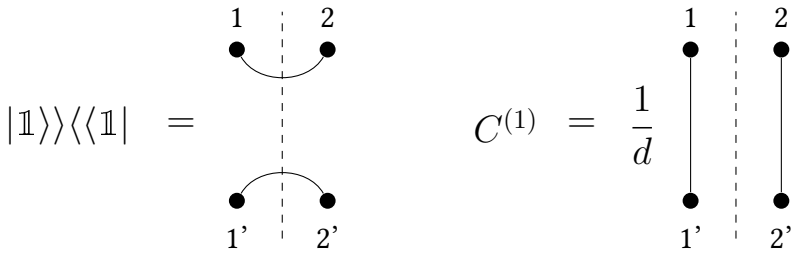


Figure 6.1: Diagrammatic representation of the two invariant—under the group action of $\mathrm{SU}(d)$—elements appearing in the Choi matrices F_r in Eq. (6.10). The two elements define the two invariant orthogonal subspaces in $\mathcal{A}^d_{1,1}$, i.e., Π_0 and Π_1.

Without loss of generality, we can restrict our search for optimal testers to those that respect the same symmetry. We define the projectors $\Pi_0 = \frac{1}{d}|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|$ and $\Pi_1 = \mathbb{1}\otimes\mathbb{1} - \Pi_0$, and consider operators of the form

$$E_s = \bigotimes_{j=1}^{n} \left(\mathit{1}_{\bar{s}}(j)\,\Pi_0 + \mathit{1}_s(j)\Pi_1\right)_{j_{\mathrm{in}}j_{\mathrm{out}}}, \tag{6.11}$$

where $\mathit{1}_s(j)$ is the indicator function

$$\mathit{1}_s(j) = \begin{cases} 1 & \text{if } \ j \in s\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{6.12}$$

That is, the operator E_s is a tensor product with Π_1 in the parties that belong to the set s and Π_0 in the remainder of the systems. These operators form an orthogonal basis for $(\mathcal{A}^d_{1,1})^{\otimes n}$, the tensor product of algebras of partially transposed permutations. The testers can thus be written as

$$T_r = \sum_{s\subseteq\{1,\ldots,n\}} a^r_s\, E_s\,, \tag{6.13}$$

with a^r_s a set of positive coefficients. The same basis also allows us to expand the Choi matrices $\{F_r\}_r$ as

$$F_r = d^{n-2}(E_\emptyset + E_r)\,, \tag{6.14}$$

where $\emptyset$ denotes the empty set.

In this basis, the success probability in Eq. (6.9a) reduces to the simple expression

$$P_{\mathrm{s}} = d^n \frac{d^2-1}{d^2} \sum_{|r|=1} a_r^r \,, \tag{6.15}$$

which we aim to maximize subject to the constraints in Eqs. (6.9b–6.9f). Defining the coefficients $\alpha_s = a_s^? + \sum_{|r|=k} a_s^r$, the completeness condition becomes

$$T_? + \sum_{|r|=1} T_r = \mathbb{1}_n \otimes R^{(n)} = \sum_{s\subseteq\{1,\ldots,n\}} \alpha_s \, E_s \,, \tag{6.16}$$

while the constraint $\mathrm{tr}_{n_{\mathrm{in}}} T^{(n)} = \mathbb{1}_{n_{\mathrm{out}}} \otimes T^{(n-1)}$ implies

$$\sum_{s\subseteq\{1,\ldots,n-1\}} \alpha_{s\cup\{n\}} = \sum_{s\subseteq\{1,\ldots,n-1\}} \alpha_s \,, \tag{6.17}$$

Iterating this relation for all constraints in Eqs. (6.9e–6.9f) we obtain that all α_s must be equal, so Eq. (6.16) becomes

$$\alpha \sum_r E_r = \alpha \mathbb{1}^{\otimes 2n} = \mathbb{1}_n \otimes R^{(n)} \,, \tag{6.18}$$

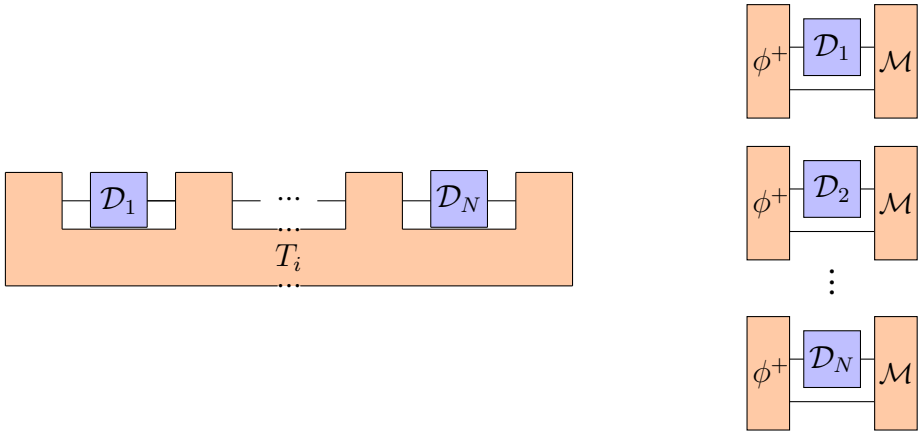
which fixes $\alpha = d^{-n}$.

Since $a_r^r \leq \alpha = d^{-n}$, the optimal success probability is bounded by

$$P_{\mathrm{s}} \leq 1 - \frac{1}{d^2} \,. \tag{6.19}$$

This bound is achievable by taking $T_r = d^{-n} E_r$ for each $|r| = 1$, and $T_? = d^{-n}\mathbb{1}^{\otimes 2n} - \sum_{|r|=1} T_r$.

A realization achieving this bound proceeds as follows. One prepares the product state $|\Psi\rangle = |\phi^+\rangle^{\otimes n}$, where $|\phi^+\rangle = \frac{1}{\sqrt{d}} \sum_i |ii\rangle$ denotes the maximally entangled state of two d-dimensional systems. The first subsystem of each copy of $|\phi^+\rangle$ is sent through one of the devices $\mathcal{D}_j$, while the second subsystem is left untouched. A local—in each bipartite system—measurement is then performed to check whether the resulting state has projection onto $|\phi^+\rangle$ or not, corresponding to the application of the two-outcome measurement $\mathcal{M} = \{\Pi_0, \Pi_1\}$. If the outcome "1" is obtained, the corresponding device is identified with certainty as anomalous; if the outcome is instead "0", the remaining devices are tested in the same way until



(a) (b)

Figure 6.2: Diagrammatic comparison of (a) a generic sequential strategy and (b) our local protocol.

the anomaly is located. If none of the local measurements clicks "1", the protocol output is inconclusive. This realization is depicted in Fig 6.2b.

This protocol parallels the unambiguous multi-anomaly state discrimination protocol presented in Chapter 5, but with two important distinctions: the procedure addresses quantum processes rather than quantum states, and ancillary systems are required here to implement the protocol. In both scenarios, the method is local, in the sense that measurements act independently on each system—in here, bipartite systems—and it can be applied in an online fashion, allowing the anomaly to be detected without the need to test all devices. Furthermore, it is non-destructive: when a device operates as intended, the output state remains undisturbed by the measurement. Also, in the limit of large local dimension d, the success probability approaches one. This is because the outputs of random unitaries tend to become more orthogonal as the Hilbert space dimension increases, making the anomalies easier to detect. The scaling behavior presents an important distinction with respect to the universal state-discrimination protocol discussed in Chapter 5: while in the state case the leading correction scales as $1/d$, here it decreases as $1/d^2$. This difference is natural, since the discrimination now effectively takes place in a bipartite subsystem of dimension d^2. Finally, it is noteworthy that we have considered the most general causal strategy for the protocol [c.f. Eqs. (6.23a–6.9f)]; however, in the end, it turns out that the optimal implementation corresponds to a *parallel* strategy. To the best of our knowledge, it is not known in which general tasks involving unknown unitaries parallel implementations are sufficient and can be assumed

beforehand. For instance, in [QDS+19], a parallel strategy suffices for the probabilistic inversion of a single use of an unknown unitary, while for multiple uses, adaptive strategies perform better. Similarly, in [SBZ19], parallel implementations are shown to be optimal for the storage of m unitaries. In contrast, in the context of unitary discrimination, ancillary systems are unnecessary when the unitaries are known [Ací01; SLH+24]. In contrast, in the setting of unknown unitaries studied here, we prove that the optimal protocol is parallel and requires ancillary systems (see Fig. 6.2b for a diagrammatic comparison with the most general causal strategy).

6.3.2 Two anomalies

We now turn to the more challenging scenario of two anomalies. In the single-anomaly scenario of the previous Section, the symmetry of the Choi matrices allowed us to restrict the search for optimal testers to a small subspace. That symmetry was local and factorized across all parties, which greatly simplified both the analysis and the construction of optimal protocols for arbitrary local dimension d and for the most general causal strategies.

This situation changes when considering two anomalies, $k = 2$. Here, the Choi matrix of the two anomalous channels takes the form

$$C^{(2)} = \frac{1}{d^2-1}\left(\mathbb{1}_{\text{in}} \otimes \mathbb{1}_{\text{out}} - \frac{1}{d}\,\mathbb{1}_{\text{in}} \otimes \mathrm{P}_{\text{out}} - \frac{1}{d}\,\mathrm{P}_{\text{in}} \otimes \mathbb{1}_{\text{out}} + \mathrm{P}_{\text{in}} \otimes \mathrm{P}_{\text{out}}\right), \quad (6.20)$$

where $\mathrm{P} = (|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|)^{T_1}_{12}$ denotes the unitary representation of the permutation operator: $\mathrm{P}|i\rangle|j\rangle = |j\rangle|i\rangle$ (see Fig. 6.3 for a schematic representation of the Choi matrix).

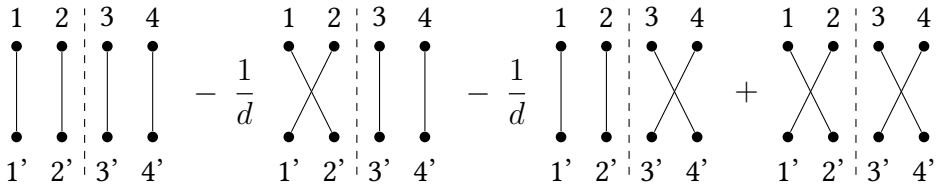


Figure 6.3: Schematic representation representation of the Choi matrix $C^{(2)}$ as elements of $\mathcal{A}^d_{2,2}$.

Unlike in the single-anomaly case, this matrix does not possess the same local symmetry on each bipartite subsystem. Instead, the symmetry of the hypotheses is

given by the action of a global tensor product unitary, namely

$$F_r = \int \left(U_{\text{in}} \otimes U_{\text{out}}^*\right)^{\otimes n} F_r \left(U_{\text{in}}^\dagger \otimes U_{\text{out}}^T\right)^{\otimes n} \mathrm{d}U \,, \tag{6.21}$$

which does not factorize across sites. As a consequence, the space of symmetric operators compatible with this invariance is larger—in fact, it corresponds to the whole algebra of partially transposed permutations $\mathcal{A}_{n,n}^d$—and the simplifications available in the single-anomaly analysis no longer apply here.

Because of this broader symmetry class, finding the optimal protocol for general d and general testers becomes substantially more challenging. In what follows, we focus on the qubit case ($d = 2$) and restrict the optimization to *parallel* quantum testers. In this setting, the symmetry constraints and feasibility conditions in the SDP take a simpler form, allowing for a fully analytical solution. More importantly, for $k \leq n/2$—a natural regime given that k represents the number of anomalies—numerical checks for small values of n ($n = 4$ and $n = 5$) show that parallel and sequential strategies yield the same optimal success probability. While a formal proof is still missing, the evidence for small cases suggests that, in this scenario, sequential strategies offer no advantage over parallel ones. This makes the parallel setting a justified choice for our analysis. With these restrictions in place, we now show that the local protocol introduced for the single-anomaly case remains optimal when $d = 2$ and $k = 2$.

The proof proceeds in two steps. First, we verify that the proposed testers form a feasible solution to the primal SDP. Then, we construct a suitable dual ansatz to demonstrate strong duality, which allows us to establish the optimality of the protocol. The explicit proof will be carried out for the case $n = 4$, and we will use this result for proving optimality for arbitrary values of n.

The protocol is defined by testers of the form $T_r = \frac{1}{d^n} E_r$, which yields a success probability

$$P_{\text{s}} = \frac{1}{d^2} \operatorname{tr}\left(C^{(2)} \Pi_1^{\otimes 2}\right) = \frac{d^4 - 2d^2 + 2}{d^4} = \frac{5}{8} \,. \tag{6.22}$$

As discussed in Section 6.3.1, these testers satisfy all constraints in Eqs. (6.9a–6.9e), as the parallel-comb setting is a special case for the adaptive implementation. To prove optimality, we construct an ansatz for the dual problem whose value matches that of the primal.

The dual formulation for parallel testers follows from the Lagrangian method

and resembles the one presented in Section 3.1.5

$$\min_{\{y,\,Y\}} \quad y \tag{6.23a}$$

$$\text{s.t.} \quad Y - \eta_r F_r^T + \sum_{s\neq r} \nu_{sr} F_s^T \geq 0 \;\forall\, r \tag{6.23b}$$

$$y\mathbb{1}_{\text{in}} - \text{tr}_{\text{out}} Y \geq 0 \tag{6.23c}$$

$$Y = Y^\dagger \tag{6.23d}$$

$$y \in \mathbb{R} \tag{6.23e}$$

$$\nu_{sr} \in \mathbb{R} \quad \forall\, r, s \tag{6.23f}$$

The constraint in Eq. (6.23c) suggests that Y should be chosen to saturate the inequality, with $\text{tr}\, Y = y\, d^n$, and $y = 5/8$. A natural starting point is

$$Y = \frac{d^n}{N} \sum_{|r|=2} \sqrt{T_r}\, F_r \sqrt{T_r}\,. \tag{6.24}$$

This choice satisfies Eqs. (6.23c–6.23e), but violates the constraint in Eq. (6.23b) for any real coefficients $\{\nu_{rs}\}$. This violation is known in the literature for the case of zero-error inference (see, e.g., [DLB+25]). As anticipated in Section 3.1, the *infimum* of the dual problem is finite but not attainable—yet strong duality still holds, since feasible dual points can be constructed arbitrarily close to it. Note that Y is invariant under bipartite system permutations; therefore, it is sufficient to consider a fixed r in Eq. (6.23b) to check feasibility for all r.

To prove strong duality, we consider the matrix $Y - \frac{1}{N} F_r + \sum_{s\neq r} \nu_{rs} F_s$, for a fixed set of anomalies r, which by construction lies in the algebra of partially transposed permutations. This property, conveniently, allows us to use Schur-Weyl duality and rewrite the operator in a block diagonal form

$$Y - \frac{1}{N} F_r + \sum_{s\neq r} \nu_{rs} F_s = \bigoplus_{\hat{\lambda}\in\hat{\mathcal{A}}^2_{4,4}} \mathbb{1}_{\hat{\lambda}} \otimes \Xi_r^{\hat{\lambda}}\,, \tag{6.25}$$

where $\hat{\lambda}$ labels the irreps of the joint action of $\mathcal{A}^2_{4,4}$ and $\text{SU}(2)$. These labels are usually identified with pairs of partitions (λ_L, λ_R) that fulfill $|\lambda_L| + |\lambda_R| \leq 2$, depicted in Fig. 6.4 (see Section 3.2.4 for a detailed discussion). However, as every element of the left-hand side of Eq. (6.25) contains a double contraction, i.e., terms of the form $|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|^{\otimes 2}$, only a restricted set of irreps contributes. In fact, non-vanishing

components appear exclusively in partitions $\big((2),(2)\big)$, $\big((1),(1)\big)$, and $(\varnothing,\varnothing)$. This follows from the way contractions affect the partitions: each contraction effectively reduces the *left* Young diagram by one box, thereby restricting which irreps can survive (see e.g. [GBO23]). As a consequence, all other irreps vanish and can be discarded. This observation is the key to the dimensionality reduction, although the operator formally lives in a Hilbert space of dimension 2^8, its relevant information is confined to the direct sum of just three components in the multiplicity space, $\Xi_r^{\hat{\lambda}}$, which makes the subsequent analysis tractable.

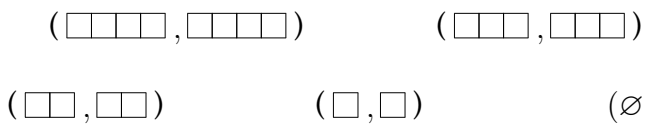


Figure 6.4: Pairs of Young diagrams associated with the irreps of the joint action of $\mathrm{SU}(2)$ and $\mathcal{A}_{4,4}^2$. The first row contains the spaces in which the elements of Eq. (6.25) have vanishing representation, while the second row contains the spaces in which the representation is non-trivial.

Another important observation is that, due to the symmetry of the constraint in Eq. (6.25), we can set $\nu_{rs}=\nu$ for every $s\neq r$. It is easy to see that for sufficiently large values of ν, the operators $\Xi_r^{((2),(2))}$ and $\Xi_r^{((1),(1))}$ are positive semidefinite. The case of $\Xi_r^{(\varnothing,\varnothing)}$ is more complex, and turns out that all its eigenvalues are non-negative, except for one, which behaves as $-\frac{k^{(1)}}{\nu}+\mathcal{O}\left(\nu^{-2}\right)$, where $k^{(1)}$ is a constant of $\mathcal{O}(1)$.

This allows us to introduce the new dual variables

$$Y_\epsilon=\frac{1}{N}\sum_{|r|=k}\sqrt{T_r}F_r\sqrt{T_r}+\frac{\epsilon}{d^n}\mathbb{1}_{\mathrm{out}}\otimes\mathbb{1}_{\mathrm{in}}\,, \tag{6.26}$$

with $\epsilon\simeq\frac{k^{(1)}}{\nu}$, and $y_\epsilon=5/8+\epsilon$. This choice of dual variables constitute a feasible instance of the SDP for sufficiently large values of ν. Taking the limit $\epsilon\to 0^+$ ($\nu\to\infty$) yields a dual value arbitrarily close to $5/8$. Since the primal optimum cannot exceed the dual infimum, this establishes strong duality and proves that, for qubits with $n=4$ and $k=2$, the optimal success probability is $P_{\mathrm{s}}=5/8$.

For an arbitrary number of devices n, the local protocol attains a constant success probability $P_{\mathrm{s}}=5/8$. To prove its optimality, we now show that this value is, in fact, an upper bound—which our protocol saturates.

Let $P_{\mathrm{s}}^*(n)$ denote the optimal success probability for n devices, and suppose that an external agent provides the additional information that the two anomalous

devices are guaranteed to be among a fixed subset $\tilde{\mathcal{D}}_4$ of four devices. We denote the optimal success probability for this scenario by $P_{\mathrm{s}}^*(n|\tilde{\mathcal{D}}_4)$. Clearly, knowing $\tilde{\mathcal{D}}_4$ can only help, so

$$P_{\mathrm{s}}^*(n) \leq P_{\mathrm{s}}^*(n|\tilde{\mathcal{D}}_4)\,. \tag{6.27}$$

Note that the hypotheses for this scenario are exactly those of $n = 4$ tensored by $|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|^{\otimes(n-4)}$, which represents the Choi matrices of the remaining devices. This extra factor carries no distinguishing information—as it is identical for all hypotheses—so both optimal success probabilities are equal: $P_{\mathrm{s}}^*(n|\tilde{\mathcal{D}}_4) = P_{\mathrm{s}}^*(4)$.

Combining these observations with the fact that the local protocol achieves $P_{\mathrm{s}}(n) = 5/8$ for any n gives

$$\frac{5}{8} \leq P_{\mathrm{s}}(n) \leq \frac{5}{8}\,, \tag{6.28}$$

which proves the optimality of the protocol for all n. □

6.4 Discussion

In this Chapter, we have studied the task of zero-error discrimination of anomalous unitary channels. To derive the optimal protocols, we have worked in the quantum networks formalism, using quantum testers to discriminate between the hypotheses, which, in contrast with previous chapters, are now defined by the Choi matrices of the quantum channels applied by the devices. By exploiting the symmetry of these matrices, we reduced the search for optimal testers to a smaller set, generated by the algebra of partially transposed permutations.

We first examined the case of a single anomalous device, $k = 1$. In this scenario, the symmetry of the hypotheses is particularly convenient, as the corresponding Choi matrices are invariant under local transformations of the form $U \otimes U^*$ on each input–output pair. This property reduces the problem to a highly symmetric setting, from which we derived an analytic expression for the optimal success probability, $P_{\mathrm{s}} = 1 - \frac{1}{d^2}$, valid for any local dimension d and any number of devices n. This bound is attained by a local protocol in which the observer prepares n copies of a maximally entangled state, sends a copy through each device, and then performs local projective measurements to determine whether a device has performed an anomalous transformation.

In the second part of the Chapter, we investigated the case of two anomalies, $k = 2$. Here, the problem becomes more challenging. The Choi matrices no longer share the same local invariance as in the single-anomaly case, but are instead

invariant only under global transformations of the form $U \otimes U^*$ acting simultaneously across all parties. This broader symmetry class within the algebra of partially transposed permutations enlarges the space of feasible testers, making the optimization considerably harder. We then restrict to the qubit case and to parallel quantum testers, as numerical checks for small system sizes indicate that sequential strategies do not outperform parallel ones. Within this framework, we proved that the local protocol introduced in the single-anomaly case remains optimal also for two anomalies. This proof depend on the reduction of the highly-dimensional operators into more manageable components with the mixed Schur-Weyl duality, making the problem both numerically and analytically tractable. Working in this representation, we established optimality through SDP, the protocol for two anomalies in the qubit case yields a success probability $P_{\mathrm{s}} = 5/8$, independently of the number of devices n.

The detection strategy presented here closely resembles the unambiguous identification of anomalous states discussed in Chapter 5. In both cases, the protocols are local, non-destructive, and can be applied in an online fashion. The main difference lies in the objects being discriminated and in the implementation: in Chapter 5, the anomalies correspond to prepared quantum states, while here they are quantum processes. This distinction required the use of the Choi–Jamiołkowski isomorphism and the quantum network formalism, which provide the appropriate framework for channel discrimination tasks.

In line with the other discrimination tasks studied in this thesis, the protocols derived here are universal in the sense that they make no assumptions about the anomalous unitary transformations. They are valid for any anomalous unitar y channel, without tailoring the protocol to a specific case. However, it is noteworthy that we do not observe the learning behavior found in Chapters 4 and 5. In those tasks, the protocol could effectively *learn* from the reference—yet unknown—systems. By contrast, in the problem addressed here, the reference systems are already known, and the unknown objects are precisely the anomalies. As a result, no learning behavior is observed, since the anomalies do not provide any structure from which the protocol can gain information. This is consistent with the universal setting analyzed in Chapter 5, where the protocol asymptotically extracts information from the reference states—representing the majority of the systems—but not from the anomalies.

Beyond the theoretical interest, these results may find application in the verification of quantum devices. In particular, anomaly-detection protocols of the kind studied here could be adapted to certify the correct functioning of quantum gates or subroutines in a quantum circuit, where faulty components correspond to

anomalous unitaries. The universal character of the protocol is especially appealing in practice, as it avoids the need for a specific tailoring with respect to the faults, whose specific description is rarely available to the experimentalist.

Outlook

In this thesis I have presented three distinct problems that together form a coherent narrative on universal discrimination of quantum states and processes. The main results have already been summarized in the Introduction 1, and each of the Chapters 4, 5, and 6 has concluded with a discussion of its specific contributions. In this final Chapter, the focus shifts to a broader perspective and I outline possible directions for future research and highlight open problems that emerge from the questions addressed throughout this thesis.

The problem of edge detection in Chapter 4 can be extended in several ways. A very natural extension is to consider higher-dimensional arrangements of quantum systems. While this generalization may appear conceptually straightforward, its implementation poses significant challenges. In particular, the symmetry techniques that are effective in one dimension do not translate directly to two- or three-dimensional settings, and even the notions of "region" or "boundary" are not well defined for quantum systems, unlike in classical spatial domains. Another natural extension is to consider sequences with more than two states that generate multiple edges, which introduces additional complexity in the design of discrimination protocols and the definition of optimal strategies. Further progress may also be achieved by exploring alternative figures of merit, for instance, criteria that penalize outcomes according to their distance from the true edge rather than through a simple yes/no assessment, or some other criteria [VM19]. In addition, it is worth studying restrictions to local strategies—similar to those studied in Chapters 5 and 6—as means to reduce the resources required by the protocols; possibly measuring the states in an online fashion, similar to the work presented in [SMM18]. Apart from these extensions, edge detection also connects with the broader family of change point problems [SBC+16; SCM17], from which edge detection formulation originally arose. Natural generalizations of these concepts have already been explored in the literature, such as change point detection for Hamiltonian dynamics [Nak23] and unambiguous change point identification for quantum channels [Nak25], which effectively extend edge detection protocols to channel discrimination. In this context, two directions remain largely unexplored: the development of unambiguous protocols for universal edge detection in quantum

states, and the construction of universal strategies for edge detection of unknown quantum channels.

In Chapter 5 I presented the problem of identification of anomalous states. This research project also leaves open directions for further research. A first open question is the determination of an optimal unambiguous protocol in the universal scenario where both the reference and the anomalous states are unknown. Another natural extension is to move beyond state discrimination and consider quantum channels. In particular, in [SLH+24] results for the case of known unitaries, and Pauli and amplitude damping channels are presented, but restricted to the single-anomaly scenario. Extending these results to an arbitrary number of anomalies, as well as to universal settings—in which not only the anomalies but also the reference channels or unitaries are unknown—would be of significant interest.

These possible extensions are not only of theoretical interest but also connect directly to possible applications. Quantum computing relies on the precise preparation of initial states and their controlled evolution. At the start of any computation, qubits must be initialized in the same fixed state, and any imperfection at this stage can propagate through the circuit, ultimately compromising the final output. The techniques developed in Chapter 5 provide a way to certify the computational resources at this input level by identifying potential malfunctions in the preparation step. In addition to state preparation, the results of Chapter 6 extend this verification to the dynamical phase of quantum computing. Since quantum circuits are built from unitary gates, the task of detecting anomalous unitaries directly aligns with the need to ensure that circuits behave as intended and produce faithful outputs. The universality of the protocols developed here offers a distinct advantage: because they are not tailored to any specific malfunction, they can be applied in practice to inspect quantum circuits without requiring prior knowledge of the faults.

Beyond such possible applications, there are also fundamental research directions. A further line of research concerns the study of local strategies for universal protocols. Restricting discrimination protocols to local operations may allow the observer to reduce the resources required by the protocol and enable more practical implementations. It remains to be understood whether the performance of such strategies would be comparable to the global case, or even non-vanishing. Lastly, the association scheme techniques introduced in Chapter 5 could potentially be applied to other problems in quantum state discrimination. The advantage offered by the underlying symmetries of such schemes provides a powerful analytic tool, but the main challenge lies in identifying the relevant association scheme for each problem. For instance, in [SMM+19] the problem structure corresponded to the Hamming association scheme, although a different set of techniques was employed

there. This leaves the door open for association scheme techniques in other relevant scenarios in quantum state discrimination.

The universal results presented in Chapter 6 also give rise to a number of open questions. Among them, a complete analytical proof that parallel strategies achieve the same success probability as sequential ones is (still) missing, the extension of the analysis to an arbitrary number of anomalies, and the generalization of the multiple-anomaly setting to arbitrary local dimensions. Beyond these challenges, several directions appear as natural continuations of this work. One is the generalization of the universal setting to include not only the anomalous but also the reference unitaries, in line with the scenario considered in Chapter 5. Another is to move beyond unitaries and explore the case of general quantum channels in the universal setting, but this seems more involved a priori. Also, the extension of this problem to classification tasks, in which both the number and the positions of the anomalies are arbitrary could be of interest, previous results on this line are presented in [SMM+19] for unknown states. Lastly, it is worth mentioning related works that address universal protocols for unknown unitaries in contexts different from discrimination tasks, such as the inversion of an unknown unitary transformation [QDS+19] or the storage and retrieval of unknown unitary channels [SBZ19].

A

Asymptotic results of quantum edge detection

The method we employed to calculate the asymptotic success probability was previously discussed in Section 4.1.1. Here, we present some missing steps, including intermediate expressions (many of which are organized in tables), to assist interested readers in replicating our findings. We use the variables $j = N/2 - \lambda$ and $m = N/2 - w$ throughout the appendix. While denoting angular momentum and magnetic quantum number, respectively, only for $d = 2$, the use of j and m in deriving the asymptotic success probability has proven notably advantageous, resulting in simpler and more transparent expressions also for $d > 2$.

The entries of the Gram matrix, $G := G^{\lambda}$, where we drop the label λ to simplify the notation, are given in Eqs. (4.10) and (4.11). They can be expressed as follows:

$$G_{k\,k'} = \frac{2j+1}{\frac{N}{2}+j+1} \frac{\binom{d+\frac{N}{2}-j-2}{d-2}\binom{d+\frac{N}{2}+j-1}{d-1}}{\sqrt{\binom{d+k-1}{d-1}\binom{d+\bar{k}-1}{d-1}\binom{d+k'-1}{d-1}\binom{d+\bar{k}'-1}{d-1}}} \sqrt{\frac{\binom{k}{\frac{N}{2}-j}\binom{\bar{k}'}{\frac{N}{2}-j}}{\binom{k'}{\frac{N}{2}-j}\binom{\bar{k}}{\frac{N}{2}-j}}}, \tag{A.1}$$

where the last (d-independent) factor can be clearly identified with the overlap $\langle \Omega_k^{\lambda} | \Omega_{k'}^{\lambda} \rangle$, while the remaining ($d$-dependent) factors correspond to the priors η_k^{λ} and $\eta_{k'}^{\lambda}$. Here $\bar{k} = N - k$, $\bar{k}' = N - k'$, and the range of k and k' is $(N/2) - j \leq k \leq (N/2) + j$. We assume $k \leq k'$ ($\bar{k}' \leq \bar{k}$); for $k' \leq k$, we simply exchange k with k' and $\bar{k}$ with $\bar{k}'$, i.e., G is symmetric. The Gram matrix is also persymmetric, meaning it is symmetric with respect to the anti-diagonal. This implies that G is invariant under the exchange $k, k' \leftrightarrow \bar{k}, \bar{k}'$.

The scaled Gram matrix, used to derive the asymptotic success probability, is [cf. Eq. (4.27)]:

$$\tilde{G} = \frac{(N/2)^2}{(d-1)(2j+1)} G. \tag{A.2}$$

The inverse of $\tilde{G}$ is tridiagonal (because $\tilde{G}$ is semiseparable). Its diagonal and super-diagonal entries are respectively given by the sequences:

$$[\tilde{G}^{-1}]_{\text{diag}} = \mathcal{B}_m^j \frac{j(j+1)+\frac{N}{2}\left(\frac{N}{2}+1\right)-2m^2}{\left(\frac{N}{2}-j\right)\left(\frac{N}{2}+j+1\right)} \tag{A.3}$$

$(-j \leq m \leq j)$, and

$$[\tilde{G}^{-1}]_{\text{super}} = -\sqrt{\mathcal{B}_m^j \mathcal{B}_{m+1}^j} \frac{\sqrt{\left(\frac{N}{2}-m\right)\left(\frac{N}{2}+m+1\right)(j-m)(j+m+1)}}{\left(\frac{N}{2}-j\right)\left(\frac{N}{2}+j+1\right)} \tag{A.4}$$

$(-j \leq m \leq j-1)$, where

$$\mathcal{B}_m^j = \frac{1+\frac{N}{2}+j}{(N/2)\left(\frac{N}{2}-j+d-2\right)!} \frac{\left(\frac{N}{2}-j\right)!\left(\frac{N}{2}+j\right)!}{\left(\frac{N}{2}+j+d-1\right)!} \frac{\left(\frac{N}{2}-m+d-1\right)!\left(\frac{N}{2}+m+d-1\right)!}{\left(\frac{N}{2}-m\right)!\left(\frac{N}{2}+m\right)!}. \tag{A.5}$$

As explained in Section 4.1.1, we introduce a scaled version of j, defined as $x = j/(N/2)$, and compute the Maclaurin series expansion of the joint probability $P(x) := P^{\text{SRM}}_{\lambda(x)}$ (indices are dropped to avoid clutter) up to a high order, $2R$, in x. At the leading order in $1/N$, the joint probability is given by the series:

$$\frac{N}{2} P(x) = \sum_{r=1}^{\infty} a_r x^{2r}. \tag{A.6}$$

The (exact) coefficients a_r have been computed up to $r = 15$ or $r = 18$, depending on d, using *Wolfram Mathematica 13.3*, with the corresponding code available in [LGS+25b]. For $d = 2, 3, 4$, and 8, they are collected in Tables A.1–A.4.

The Padé approximants [Bak64] (referred to as Padés hereafter) to $(N/2)P(x)$ can be computed from those tables, as outlined in Section 4.1.1. Once again, we relied on *Mathematica* to perform this tedious calculation. Diagonal Padés, of the form shown in Eq. (4.32) with $n = m$, have proven to deliver lower uncertainly and better stability for the problem at hand. We express them as:

$$\mathcal{P}_{2s}(x) := \frac{\sum_{r=0}^{s} A_r x^{2r}}{1 + \sum_{r=1}^{s} B_r x^{2r}}. \tag{A.7}$$

l	a_{2l-1}	a_{2l}
1	2	0
2	$-\frac{1}{30}$	$-\frac{1}{35}$
3	$-\frac{229}{10080}$	$-\frac{101}{5544}$
4	$-\frac{5725}{384384}$	$-\frac{5111}{411840}$
5	$-\frac{554293}{52715520}$	$-\frac{2137221}{236487680}$
6	$-\frac{249919231}{31783944192}$	$-\frac{76576105}{11076222976}$
7	$-\frac{10870862389}{1772195676160}$	$-\frac{8413125001}{1533630873600}$
8	$-\frac{1506595197973}{304973453721600}$	–

Table A.1: Coefficients a_r of the MacLaurin series of $(N/2)P(x)$ defined in Eq. (A.6) for $d = 2$.

l	a_{2l-1}	a_{2l}
1	4	$-\frac{8}{3}$
2	$-\frac{1}{15}$	0
3	$\frac{3}{560}$	$\frac{3}{616}$
4	$\frac{739}{192192}$	$\frac{307}{102960}$
5	$\frac{1044347}{448081920}$	$\frac{874097}{472975360}$
6	$\frac{23645057}{15891972096}$	$\frac{5047729}{4153583616}$
7	$\frac{891227339}{886097838080}$	$\frac{168502693}{200038809600}$
8	$\frac{3749131111}{5258162995200}$	$\frac{9600936701}{15756961775616}$
9	$\frac{1398178715421307}{2662296261608079360}$	–

Table A.2: Coefficients a_r of the MacLaurin series of $(N/2)P(x)$ defined in Eq. (A.6) for $d = 3$.

l	a_{2l-1}	a_{2l}
1	6	-8
2	$\frac{31}{10}$	$\frac{3}{35}$
3	$\frac{9}{1120}$	0
4	$-\frac{125}{128128}$	$-\frac{25}{27456}$
5	$-\frac{3875}{5431296}$	$-\frac{51075}{94595072}$
6	$-\frac{393277}{963149824}$	$-\frac{3454141}{11076222976}$
7	$-\frac{125777259}{521234022400}$	$-\frac{145143689}{766815436800}$
8	$-\frac{15293179829}{101657817907200}$	$-\frac{249322259}{2059733565440}$
9	$-\frac{174838753670533}{1774864174405386240}$	$-$

Table A.3: Coefficients a_r of the MacLaurin series of $(N/2)P(x)$ defined in Eq. (A.6) for $d = 4$.

l	a_{2l-1}	a_{2l}
1	14	-56
2	$\frac{3353}{30}$	-127
3	$\frac{120347}{1440}$	$-\frac{11675}{396}$
4	$\frac{226243}{54912}$	$\frac{35777}{411840}$
5	$\frac{7310429}{896163840}$	$\frac{119175}{94595072}$
6	$\frac{1121953}{4540563456}$	$\frac{83349}{1582317568}$
7	$\frac{12326391}{1265854054400}$	0
8	$-\frac{20469449}{11295313100800}$	$-\frac{61408347}{35015470612480}$
9	$-\frac{24207914731}{17927920953589760}$	$-$

Table A.4: Coefficients a_r of the Maclaurin series of $(N/2)P(x)$ defined in Eq. (A.6) for $d = 8$.

For each d, we have analyzed the behavior of the sequence $\mathcal{P}_2(x), \mathcal{P}_4(x), \ldots$ that can be derived from Tables A.1–A.4 and concluded that it exhibits convergence

and stability across the entire range $0 \leq x \leq 1$. From our analysis, we estimated the corresponding asymptotic success probability by integrating the highest-order Padé of the sequence, $\mathcal{P}_{\text{high}}(x)$, as in Eq. (4.31):

$$P_{\text{s}} \sim p_0(d) \approx \int_0^1 \mathcal{P}_{\text{high}}(x)dx. \tag{A.8}$$

These estimated values correspond to the lower margins provided in the second column of Table 4.1. For completeness, Tables A.5–A.7 gather the coefficients of the highest-order diagonal Padés for $d = 2, 3, 4$. However, for higher dimensionality, the Padés do not seem to offer any improvement over the Maclaurin series expansion, and they are omitted here.

Alternatively, one could find the primitive of Eq. (A.6), represented by the series

$$Q(x) = \sum_{r=1}^{\infty} \frac{a_r}{2r+1} x^{2r+1}, \tag{A.9}$$

and then, to estimate the asymptotic success probability $P_{\text{s}} \sim p_0(d) = Q(1)$, apply Padé acceleration to the truncated series approximation obtainable from Tables A.1–A.4. Note that $Q(x)$ is an odd function, preventing the construction of diagonal Padés. For this reason, we examine the behavior of the Padé sequence $\{\mathcal{Q}_{2n}^{2n-1}(x), \mathcal{Q}_{2n}^{2n+1}(x)\}_n$, where $\mathcal{Q}_m^n(x) = [n/m]_Q(x)$. This sequence also exhibits rapid convergence and stability in the unit interval and, consequently, we estimate the asymptotic success probability from its highest-order Padé, $\mathcal{Q}_{\text{high}}$, as

$$P_{\text{s}} \sim p_0(d) \approx \mathcal{Q}_{\text{high}}(1). \tag{A.10}$$

The results are the upper margins shown in the second column of Table 4.1.

We estimate the accuracy of this method by computing the difference between the upper and lower margins obtained with the two described alternatives. In all cases, this yields an estimated uncertainty of less than 0.03%, which we see as a clear indication of the consistency of our Padé analysis.

For a local dimension of around 20 or higher, certain Maclaurin coefficients exhibit significant growth, alternating signs, and the series becomes unstable. While Padé approximants offer some assistance in the analysis, the estimated accuracy diminishes, requiring a different approach to obtain precise results for the success probability.

r	A_r	B_r
1	2.	-3.47961
2	-6.95921	4.68403
3	9.33473	-3.03572
4	-5.98403	0.951594
5	1.82375	-0.125282
6	-0.22237	0.00507854
7	0.00725633	-0.0000178728

Table A.5: Coefficients of the Padé approximant $\mathcal{P}_{14}(x)$ for $d = 2$.

r	A_r	B_r
1	4.	-1.56531
2	-8.9279	-0.406743
3	2.48052	1.89668
4	8.77573	-1.14072
5	-9.58824	0.240021
6	3.87205	-0.0158881
7	-0.633517	0.000108498
8	0.0319437	0.00000116197

Table A.6: Coefficients of the Padé approximant $\mathcal{P}_{16}(x)$ for $d = 3$.

r	A_r	B_r
1	6.	-2.7713
2	-24.6278	2.85673
3	42.4108	-1.33735
4	-39.3833	0.276189
5	20.9823	-0.0205167
6	-6.2558	0.000174461
7	0.928717	0.00000312601
8	-0.0502589	0.0000000704169

Table A.7: Coefficients of the Padé approximant $\mathcal{P}_{16}(x)$ for $d = 4$.

For asymptotically large local dimensions, assuming additionally that $d \gg N$, the Gram matrix becomes

$$G_{k\,k'} = \frac{(2j+1)\sqrt{\bar{k}!\bar{k}'!k!k'!}}{\left(\frac{N}{2}-j\right)!\left(\frac{N}{2}+j+1\right)!}\sqrt{\frac{\binom{k}{\frac{N}{2}-j}}{\binom{k'}{\frac{N}{2}-j}}\frac{\binom{\bar{k}'}{\frac{N}{2}-j}}{\binom{\bar{k}}{\frac{N}{2}-j}}} + O\left(1/d\right), \tag{A.11}$$

where $k \leq k'$. One can follow a similar procedure as outlined to obtain an expansion analogous to Eq. (4.29). Its first terms are given by:

$$P_\lambda = \frac{2(2j+1)^2}{N^2}\left[1 - \frac{2(2j^2+2j+3)}{3N} + \frac{4(j^2+j+3)(4j^2+4j+5)}{15N^2} + \dots\right]. \tag{A.12}$$

A glimpse at the few terms presented in this equation already reveals that in this regime, the appropriate scaling of j to determine the asymptotic behavior is $y = j/\sqrt{N/2}$. In this case, the Maclaurin series corresponding to the leading term in $1/N$ is:

$$\frac{N}{2}P(y) = \sum_{l=1}^{\infty}\frac{(-1)^{l+1}(2y^2)^l}{(2l-1)!!} = 2yF(y), \tag{A.13}$$

where $F(y)$ is identified as Dawson's integral, defined as [AS48]:

$$F(y) := \mathrm{e}^{-y^2}\int_0^y \mathrm{e}^{t^2}dt. \tag{A.14}$$

The asymptotic probability of success is computed as $P_\mathrm{s} \sim 2\sqrt{N/2}\int_0^1 xF\left(\sqrt{N/2}\,x\right)dx$ (since $y = \sqrt{N/2}\,x$), which approaches unity as N tends to infinity.

Bibliography

[AS48] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Vol. 55 (US Government printing office, 1948).

[Ací01] A. Acín, "Statistical Distinguishability between Unitary Operations", Physical Review Letters **87**, 177901 (2001).

[AKN98] D. Aharonov, A. Kitaev, and N. Nisan, "Quantum circuits with mixed states", in Proceedings of the thirtieth annual ACM symposium on Theory of computing - STOC '98 (1998), pp. 20–30.

[AMI16] M. Ahmed, A. N. Mahmood, and M. R. Islam, "A survey of anomaly detection techniques in financial domain", Future Generation Computer Systems **55**, 278–288 (2016).

[AAR99] G. E. Andrews, R. Askey, and R. Roy, *Special Functions*, Encyclopedia of Mathematics and its Applications (Cambridge University Press, Cambridge, 1999).

[Bak64] G. A. Baker Jr, *The theory and application of the Padé approximant method*, tech. rep. (Los Alamos National Lab. (LANL), Los Alamos, NM (United States), 1964).

[BSH24] M. Balanzó-Juandó, M. Studziński, and F. Huber, "Positive maps from the walled Brauer algebra", Journal of Physics A: Mathematical and Theoretical **57**, 115202 (2024).

[BBI+21] E. Bannai, E. Bannai, T. Ito, and R. Tanaka, *Algebraic Combinatorics* (Walter de Gruyter GmbH & Co KG, 2021).

[Bar01] S. M. Barnett, "Minimum-error discrimination between multiply symmetric states", Physical Review A **64**, 030303 (2001).

[BMQ21] J. Bavaresco, M. Murao, and M. T. Quintino, "Strict Hierarchy between Parallel, Sequential, and Indefinite-Causal-Order Strategies for Channel Discrimination", Physical Review Letters **127**, 200504 (2021).

[BO13] C. M. Bender and S. A. Orszag, *Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory* (Springer Science & Business Media, 2013).

[BBF+06] J. A. Bergou, V. Bužek, E. Feldman, U. Herzog, and M. Hillery, "Programmable quantum-state discriminators with simple programs", Physical Review A **73**, 062334 (2006).

[BM59] R. C. Bose and D. M. Mesner, "On Linear Associative Algebras Corresponding to Association Schemes of Partially Balanced Designs", The Annals of Mathematical Statistics **30**, 21–38 (1959).

[Bra72] O. Bratteli, "Inductive limits of finite dimensional C^*-algebras", Transactions of the American Mathematical Society **171**, 195–234 (1972).

[Bra37] R. Brauer, "On algebras which are connected with the semisimple continuous groups", Annals of Mathematics **38**, 857–872 (1937).

[Can86] J. Canny, "A Computational Approach to Edge Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-8**, 679–698 (1986).

[Car72] P. L. Carter, "A Bayesian Approach to Quality Control", Management Science (1972).

[CBK09] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey", ACM computing surveys (CSUR) **41**, 1–58 (2009).

[CDP08] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Quantum Circuit Architecture", Physical Review Letters **101**, 060401 (2008).

[CDP09] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Theoretical framework for quantum networks", Physical Review A **80**, 022339 (2009).

[CDP+04] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, "Covariant quantum measurements that maximize the likelihood", Physical Review A **70**, 062105 (2004).

[CDP+13] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "Quantum computations without definite causal structure", Physical Review A **88**, 022318 (2013).

[CMN22] B. Collins, S. Matsumoto, and J. Novak, "The weingarten calculus", Notices of the American Mathematical Society **69**, 1 (2022).

[CŚ06] B. Collins and P. Śniady, "Integration with Respect to the Haar Measure on Unitary, Orthogonal and Symplectic Group", Communications in Mathematical Physics **264**, 773–795 (2006).

[CDD+08] A. Cox, M. De Visscher, S. Doty, and P. Martin, "On the blocks of the walled Brauer algebra", Journal of Algebra **320**, 169–212 (2008).

[CAB+06] S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, "Maximum Confidence Quantum Measurements", Physical Review Letters **96**, 070401 (2006).

[DP15] N. Dalla Pozza and G. Pierobon, "Optimality of square-root measurements in quantum state discrimination", Physical Review A **91**, 042334 (2015).

[Del73] P. Delsarte, "An algebraic approach to the association schemes of coding theory", PhD thesis (Université Catholique de Louvain, 1973).

[DLB+25] A. Diebra, S. Llorens, E. Bagan, G. Sentís, and R. Muñoz-Tapia, "Quantum state exclusion for group-generated ensembles of pure states", arXiv:2503.02568 (2025).

[Die88] D. Dieks, "Overlap and distinguishability of quantum states", Physics Letters A **126**, 303–306 (1988).

[EMV04] Y. Eldar, A. Megretski, and G. Verghese, "Optimal detection of symmetric mixed quantum states", IEEE Transactions on Information Theory **50**, 1198–1207 (2004).

[FSC+22] M. Fanizza, M. Skotiniotis, J. Calsamiglia, R. Muñoz-Tapia, and G. Sentís, "Universal algorithms for quantum data learning", Europhysics Letters **140**, 28001 (2022).

[FRT54] J. S. Frame, G. d. B. Robinson, and R. M. Thrall, "The Hook Graphs of the Symmetric Group", Canadian Journal of Mathematics **6**, 316–324 (1954).

[FMS14] C. A. Fuchs, N. D. Mermin, and R. Schack, "An introduction to QBism with an application to the locality of quantum mechanics", American Journal of Physics **82**, 749–754 (2014).

[FS13] C. A. Fuchs and R. Schack, "Quantum-Bayesian coherence", Reviews of Modern Physics **85**, 1693–1715 (2013).

[Ful97] W. Fulton, *Young Tableaux: With Applications to Representation Theory and Geometry* (Cambridge University Press, 1997).

[FH13] W. Fulton and J. Harris, *Representation Theory: A First Course* (Springer Science & Business Media, 2013).

[Ful20] B. Fultz, *Phase transitions in materials* (Cambridge University Press, 2020).

[Gon24] W. L. González Olaya, "Kronecker states: a powerful source of multipartite maximally entangled states in quantum information", PhD thesis (Universidad de los Andes, 2024).

[Gri25] D. Grinko, "Mixed Schur–Weyl duality in quantum information", PhD thesis (Institute for Logic, Language and Computation, 2025).

[GBO23] D. Grinko, A. Burchardt, and M. Ozols, "Gelfand-Tsetlin basis for partially transposed permutations, with applications to quantum information", arXiv:2310.02252 (2023).

[GO24] D. Grinko and M. Ozols, "Linear Programming with Unitary-Equivariant Constraints", Communications in Mathematical Physics **405**, 278 (2024).

[Hah49] W. Hahn, "Über Orthogonalpolynome, die q-Differenzengleichungen genügen", Mathematische Nachrichten **2**, 4–34 (1949).

[Ham50] R. W. Hamming, "Error detecting and error correcting codes", The Bell System Technical Journal **29**, 147–160 (1950).

[HHL+10] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, "Adaptive versus nonadaptive strategies for quantum channel discrimination", Physical Review A **81**, 032339 (2010).

[HJS+96] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, "Classical information capacity of a quantum channel", Physical Review A **54**, 1869 (1996).

[HW94] P. Hausladen and W. K. Wootters, "A 'pretty good'measurement for distinguishing quantum states", Journal of Modern Optics **41**, 2385–2390 (1994).

[HHH08] A. Hayashi, T. Hashimoto, and M. Horibe, "State discrimination with error margin and its locality", Physical Review A **78**, 012333 (2008).

[HB07] B. He and J. A. Bergou, "Programmable unknown quantum-state discriminators with multiple copies of program and data: A Jordan-basis approach", Physical Review A **75**, 032316 (2007).

[Hel69] C. W. Helstrom, "Quantum detection and estimation theory", Journal of Statistical Physics **1**, 231–252 (1969).

[Hol73] A. S. Holevo, "Statistical decision theory for quantum systems", Journal of Multivariate Analysis **3**, 337–394 (1973).

[Iva87] I. D. Ivanovic, "How to differentiate between non-orthogonal states", Physics Letters A **123**, 257–259 (1987).

[KS96] R. Koekoek and R. F. Swarttouw, "The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue", arXiv:math/9602214 (1996).

[Koi89] K. Koike, "On the decomposition of tensor products of the representations of the classical groups: by means of the universal characters", Advances in Mathematics **74**, 57–86 (1989).

[KBD+83] "The first representation theorem", in *States, effects, and operations fundamental notions of quantum theory: lectures in mathematical physics at the university of texas at austin*, edited by K. Kraus, A. Böhm, J. D. Dollard, and W. H. Wootters (Springer Berlin Heidelberg, Berlin, Heidelberg, 1983), pp. 42–61.

[Kre73] M. G. Kreyn, *Generalization of Some Investigations of A. M. Lyapunov on Linear Differential Equations with Periodic Coefficients*, NASA Technical Translation from Doklady Akademii Nauk SSSR, Vol. 73, No. 3 (1950), pp. 445–448, TT F-14,835 (NASA, 1973).

[LGS+25a] S. Llorens, W. González, G. Sentís, J. Calsamiglia, R. Muñoz-Tapia, and E. Bagan, "Quantum Edge Detection", Quantum **9**, 1687 (2025).

[LGS+25b] S. Llorens, W. González, G. Sentís, J. Calsamiglia, R. Muñoz-Tapia, and E. Bagan, *Quantum-Edge-Detection*, Available at: https://github.com/Emili0-0/Quantum-Edge-Detection, 2025.

[Lüd06] G. Lüders, "Concerning the state-change due to the measurement process", Annalen der Physik **518**, 663–670 (2006).

[MP95] S. Massar and S. Popescu, "Optimal Extraction of Information from Finite Quantum Ensembles", Physical Review Letters **74**, 1259–1263 (1995).

[Mel24] A. A. Mele, "Introduction to Haar Measure Tools in Quantum Information: A Beginner's Tutorial", Quantum **8**, 1340 (2024).

[Nak23] K. Nakahira, "Identification of quantum change points for Hamiltonians", Physical Review Letters **131**, 210804 (2023).

[Nak25] K. Nakahira, "Unambiguous discrimination of the ¡change point for quantum channels", arXiv:2508.06785 (2025).

[Neu18] J. v. Neumann, *Mathematical Foundations of Quantum Mechanics: New Edition* (Princeton University Press, 2018).

[NC10] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, 10th anniversary edition (Cambridge university press, Cambridge, 2010).

[PP93] N. R. Pal and S. K. Pal, "A review on image segmentation techniques", Pattern Recognition **26**, 1277–1294 (1993).

[Per88] A. Peres, "How to differentiate between non-orthogonal states", Physics Letters A **128**, 19 (1988).

[PLL+19] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, "Fundamental limits to quantum channel discrimination", npj Quantum Information **5**, 1–8 (2019).

[PM17] Z. Puchała and J. A. Miszczak, "Symbolic integration with respect to the Haar measure on the unitary group", Bulletin of the Polish Academy of Sciences Technical Sciences **65**, 21–27 (2017).

[QDS+19] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, "Reversing unknown quantum transformations: Universal quantum circuit for inverting general unitary operations", Physical review letters **123**, 210502 (2019).

[RCC+21] P. D. L. Ritchie, J. J. Clarke, P. M. Cox, and C. Huntingford, "Overshooting tipping point thresholds in a changing climate", Nature **592**, 517–523 (2021).

[Sag01] B. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions* (Springer Science & Business Media, 2001).

[Sch26] E. Schrödinger, "An Undulatory Theory of the Mechanics of Atoms and Molecules", Physical Review **28**, 1049–1070 (1926).

[Sch05] I. Schur, "Neue begründung der theorie der gruppencharaktere", Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin: Jahrgang 1905; Erster Halbband Januar bis Juni, 406–432 (1905).

[SBZ19] M. Sedlák, A. Bisio, and M. Ziman, "Optimal probabilistic storage and retrieval of unitary channels", Physical Review Letters **122**, 170502 (2019).

[SBC+13] G. Sentís, E. Bagan, J. Calsamiglia, and R. Muñoz-Tapia, "Programmable discrimination with an error margin", Physical Review A **88**, 052304 (2013).

[SBC+16] G. Sentís, E. Bagan, J. Calsamiglia, G. Chiribella, and R. Muñoz-Tapia, "Quantum Change Point", Physical Review Letters **117** (2016).

[SBC+10] G. Sentís, E. Bagan, J. Calsamiglia, and R. Muñoz-Tapia, "Multicopy programmable discrimination of general qubit states", Physical Review A **82**, 042312 (2010).

[SCM17] G. Sentís, J. Calsamiglia, and R. Muñoz-Tapia, "Exact Identification of a Quantum Change Point", Physical Review Letters **119**, 140506 (2017).

[SMM22] G. Sentís, E. Martínez-Vargas, and R. Muñoz-Tapia, "Online identification of symmetric pure states", Quantum **6**, 658 (2022).

[SMM18] G. Sentís, E. Martínez-Vargas, and R. Muñoz-Tapia, "Online strategies for exactly identifying a quantum change point", Physical Review A **98**, 052305 (2018).

[SMM+19] G. Sentís, A. Monràs, R. Muñoz-Tapia, J. Calsamiglia, and E. Bagan, "Unsupervised Classification of Quantum Data", Physical Review X **9**, 041029 (2019).

[SP04] V. A. Siris and F. Papagalou, "Application of anomaly detection algorithms for detecting SYN flooding attacks", in IEEE Global Telecommunications Conference, 2004. GLOBECOM'04. Vol. 4 (2004), pp. 2050–2054.

[SLH+24] M. Skotiniotis, S. Llorens, R. Hotz, J. Calsamiglia, and R. Muñoz-Tapia, "Identification of malfunctioning quantum devices", Physical Review Research **6**, 033329 (2024).

[SC23] P. Skrzypczyk and D. Cavalcanti, *Semidefinite Programming in Quantum Information Science* (IOP Publishing, 2023).

[Sla14] M. Slater, "Lagrange Multipliers Revisited", in *Traces and Emergence of Nonlinear Programming*, edited by G. Giorgi and T. H. Kjeldsen (Springer Basel, Basel, 2014), pp. 293–306.

[Sta11] R. P. Stanley, *Enumerative combinatorics, volume 1*, 2nd, Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, 2011).

[Sti55] W. F. Stinespring, "Positive Functions on C*-Algebras", Proceedings of the American Mathematical Society **6**, 211–216 (1955).

[SMM+25] M. Studziński, T. Młynik, M. Mozrzymas, and M. Horodecki, "Irreducible matrix representations for the walled Brauer algebra", arXiv: 2501.13067 (2025).

[TNB14] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential analysis: Hypothesis testing and changepoint detection* (CRC press, 2014).

[TJ03] M. Thottan and C. Ji, "Anomaly detection in IP networks", IEEE Transactions on signal processing **51**, 2191–2204 (2003).

[Tur90] V. G. Turaev, "Operator invariants of tangles, and R-matrices", Mathematics of the USSR-Izvestiya **35**, 411 (1990).

[VVM08] R. Vandebril, M. Van Barel, and N. Mastronardi, *Matrix computations and semiseparable matrices: linear systems*, Vol. 1 (JHU Press, 2008).

[VB96] L. Vandenberghe and S. Boyd, "Semidefinite Programming", SIAM Review **38**, 49–95 (1996).

[VM19] E. M. Vargas and R. Muñoz-Tapia, "Certified answers for ordered quantum discrimination problems", Physical Review A **100**, 042331 (2019).

[Wat18] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).

[Wei78] D. Weingarten, "Asymptotic behavior of group integrals in the limit of infinite rank", Journal of Mathematical Physisc **19** (1978).

[Wey46] H. Weyl, *The Classical Groups: Their Invariants and Representations* (Princeton University Press, 1946).

[You00] A. Young, "On Quantitative Substitutional Analysis", Proceedings of the London Mathematical Society **s1-33**, 97–145 (1900).

[YKL75] H. Yuen, R. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory", IEEE Transactions on Information Theory **21**, 125–134 (1975).

[ZVF+23] J. Zipfel, F. Verworner, M. Fischer, U. Wieland, M. Kraus, and P. Zschech, "Anomaly detection for industrial quality assurance: A comparative evaluation of unsupervised deep learning models", Computers & Industrial Engineering **177**, 109045 (2023).