UAB
Universitat Autònoma de Barcelona

New techniques to simulate electrified solid-liquid interfaces better and faster

Non-Equilibrium Green's Functions
and Equivariant Graph Neural Networks

A PhD thesis by:

Pol Febrer Calabozo

Under the supervision of:
José Miguel Alonso Pruneda
Pablo Ordejón Rontomé

For the Materials Science program at:
Universitat Autònoma de Barcelona

Institut Català de Nanociència i Nanotecnologia
2024

Per la mama, que està molt preocupada perquè el pica-pica de la defensa surti bé.

Pel papa, que segurament abans de la defensa m'informarà de la temperatura que fa a Yakutsk.

Per l'avi, que encara que no parlava anglès hagués parlat amb tothom del tribunal.

Per la iaia, que li hagués donat petons i abraçades a tothom.

Para la yaya, que en la defensa hubiese dicho "ay que guapo estás".

Para el yayo, que siempre ha estado tranquilo porque Pol no tendrá problemas para encontrar trabajo.

Bueno va, i per l'Aina també, encara que no apagués el llum només per molestar quan jo volia dormir, arriscant-se a que li tirés un Hot-Wheel a la cara.

Contents

Introduction

Electrochemistry is at the heart of many processes that are key for the present and the foreseeable future of humanity. Batteries rely on electrochemical reactions to store and release energy, and are currently key for the transition to renewable energy sources. Electrocatalysis uses electrochemical reactions to drive chemical changes that synthesize valuable products or destroy hazardous ones (e.g. in water treatment). Corrosion is often an electrochemical process that needs to be minimized to increase the lifetime of materials and infrastructure. Seeing the importance of these processes and the significant success that society has had in controlling them, one might think that we have a good understanding of the underlying physics. However, most advances in these fields have been based on empirical observations that we don't fully understand and despite that they result in very succesful applications. One of the most representative examples of this fact are Li-ion batteries. They are ubiquitous in our daily lives, powering most of our portable electronic devices. Nevertheless, their chemical stability depends heavily on the Solid Electrolyte Interface (SEI), a layer that forms on the electrodes during the first charge/discharge cycles and that we are far from fully understanding [1], [2].

One of the main reasons behind this lack of understanding is that the processes are hard to simulate computationally. Computational physics and chemistry methods have been very succesful in the last decades in supporting and enhancing experimental research by providing valuable insights of the science involved, often by accessing quantities beyond experimental observables. When it comes to describing electrons at scales of up to a few nanometers, Density Functional Theory (DFT) is the method that has been most succesful due to its balance between accuracy and computational cost. However, electrochemistry has managed to escape the reach of DFT for a long time. There are mainly two reasons for this evasion, one of them is a scale issue and the other one is conceptual.

Regarding the most trivial issue, scale, electrochemical processes often involve the interaction of a solid, typically the substrate that acts as electrode or catalyzer, with a liquid, which typically is an electrolyte that hosts reactive molecules. The fact that we need to simulate two different phases and the corresponding interface makes the simulation cells already very large. But the problem gets worse when we consider that one of the phases is a liquid. To properly characterize a liquid phase, large dimensions are needed to avoid finite size effects. Unlike with solids, small cells with periodic boundaries can not be used to reduce the computational cost while keeping a good description. Furthermore, liquids have relevant thermal oscillations that are much slower than solids and diffusion is often part of the process behind interesting reactions. This means that not only we need big simulation cells, but we also need to simulate them for long times. Problems that share these two characteristics are often tackled using parametrizations of the electron behavior to create force fields that can be used to run molecular dynamics simulations much cheaper than considering electrons *ab initio*. However, the parametrization of electrons in systems where reactions are expected has traditionally proved evasive. With the advent of machine learning (ML) force fields, the perspective is changing. Machine learning force fields have shown the potential to parametrize electron behavior in systems that we thought were out of reach for force fields. Therefore, machine learning is bound to be the solution for the scale issue in computational electrochemistry. However, machines need some source of truth to learn from, and this is where the conceptual issue comes into play.

The conceptual issue is that electrochemical processes usually happen under non-equilibrium conditions. In particular, in situations where there is a potential difference between two or more metals that are in contact with an electrolyte. This potential difference can be either due to the application of voltage by humans or due to natural processes that create potential differences. In any case, the problem to solve is the same: how do we simulate the behavior of electrons in a system where the potential is not constant? DFT, the most common method to simulate electrons, is based on the assumption that the system has a single electron potential. Bound by this limitation, computational simulations of electrochemical processes have been forced to either (1) simulate only half cell systems [3] where there is only

one metal in contact with the electrolyte and therefore the single potential approximation is reasonable or (2) incorporate artificially the effects that the potential difference has on the system, e.g. by artificially adding surface charges [4]. However, there are methods that can simulate electrons in systems with a potential difference that have been used for a long time in other fields. In particular, the Non-Equilibrium Green's Functions (NEGF) method has been used for decades to simulate electronic transport in systems with a potential difference. The NEGF method is also valid for the systems that the electrochemistry community is interested in, but there has been no adoption of the method by electrochemists.

In this thesis, I report my attempt to contribute to the electrochemistry community by tackling both problems, one at a time. By the end, I also try to combine the solution to both problems into a single application. The thesis has three main chapters. In the first chapter, I introduce TranSIESTA [5], [6], a hybrid NEGF-DFT method, as a tool for electrochemistry simulations by first exploring the most fundamental results that come out of it and then showcasing molecular dynamics of liquid water between two electrodes at applied voltage. After that, I focus on the development of a machine learning method and its application to accelerate simulations that involve NEGF-DFT. In particular, the second chapter describes Graph2Mat, a new method developed to predict electronic structure matrices, which I use to predict electron densities in generic DFT systems. The third chapter closes the loop by using Graph2Mat to predict the electron density and accelerate molecular dynamics within TranSIESTA. I wrote each chapter from a different perspective, adapting to the kind of impact I expect the work to have. Since TranSIESTA is still not used in the electrochemistry community and the aim is to present it as a valid method, the first chapter is focused on describing in detail the fundamental results that can be obtained from it in order to build confidence in the electrochemistry community. The second and third chapters are written from a more pragmatic point of view. Since machine learning for atomistic modelling is already well established I focus more on the practical results that I got using my method and how the method can be used to make an impact in the field.

Chapter 1

NEGF-DFT for electrochemistry

Non-Equilibrium Green's Functions (NEGF) are a well stablished tool in the electronic transport community. The NEGF formalism allows the simulation of atomic systems in out-of-equilibrium situations. In particular for electronic transport, the equilibrium is broken by the connection to the system of $N > 1$ electrodes with different electronic occupations (different potential or electronic temperature). A very common example of such setup is that of a system where two metallic electrodes whose Fermi levels are misaligned by a quantity V are connected to the ends of the system, as depicted in Figure 1.1. This case corresponds to applying a voltage drop V across the system. The ab initio computation of electronic structure in this non-equilibrium situation is not trivial. The typical approach of diagonalizing the Hamiltonian and filling the states according to some occupation function fails to handle these situations because of the lack of a unique Fermi level. In contrast, NEGF can solve the electron density of the system under these boundary conditions.

NEGF can not only compute the out-of-equilibrium electron density but also other properties like the density of states (DOS), both local (LDOS) or projected on orbitals (PDOS). However, NEGF is not restricted to properties that have their analogous in equilibrium situations. It can also compute properties that have an inherent out-of-equilibrium nature. Examples of such properties are electron transmission, conductivity or current flowing through the system at steady state. This is what makes NEGF a key formalism to investigate electronic transport at the nanoscale. The community has been succesfully using NEGF to investigate things like molecular electronics [7] or electronic transport in 2D devices [8], [9] for more than twenty years.

In a full electrochemical cell, the situation is somewhat analogous to that of an electronic transport setup. There can be an applied voltage and therefore a misalignment between the Fermi levels of the two electrodes. Since the NEGF formalism's purpose is to obtain electronic structure under non-equilibrium conditions, it should be no surprise that NEGF can model electrochemical setups under applied voltage. However, the computational electrochemistry community is generally unaware of the potential of NEGF for these systems. There is a general mistaken belief that NEGF are a tool specific to electronic transport and therefore using them for electrochemistry is somewhat of a "hack". There is also a part of the community that understands the generic scope of NEGF, but in that case NEGF is often mentioned in papers as a methodology that is too expensive to run realistic molecular dynamics (MD).

To break these two beliefs, I invite the reader to follow the same journey that I went on when I started exploring the idea from Mariví Fernández-Serra and my co-supervisor Pablo Ordejón of using NEGF on electrochemical setups. By writing this chapter, I hope to set the basis for future researchers to contribute to the electrochemistry community with highly relevant NEGF results. The chapter has three distinct sections that serve different purposes:

- First, I try to give a explanation of why NEGF work for electrochemistry that is deep enough to get people convinced but hopefully not too deep to scare them away (section 1.1).

- Then, I try to build confidence on the TranSIESTA method (a particular implementation of NEGF) by showing the results in simple systems and comparing them to the analytical solutions based on historical electrochemistry knowledge (section 1.2).

- Lastly, I show the results of using TranSIESTA for a complex system where a fully analytical solution is just not feasible: ab initio molecular dynamics of liquid water between gold electrodes at an applied voltage (section 1.3).

1.1 Methodology

The novelty of the work in this thesis regarding NEGF is about the use of the formalism for a new application. Although I have contributed with small bug fixes and subtle modifications along the way, I used the code already available in SIESTA under the name of TranSIESTA. There is however a significant software and conceptual stack supporting the simulations, and some of it is not generally known by the computational physicis/chemistry community. That is why I devote this section to explain the different parts involved, keeping always the focus on the target application: electrochemistry. First, I will discuss the NEGF equations, which are the core of the methodology. Then, I will briefly explain how NEGF can couple with DFT. Finally, I will introduce TranSIESTA (which implements NEGF/DFT) and its terminology to set the ground for discussing the simulations in following sections. With this, I aim to provide the reader with the necessary background to understand the results and discussion in the following sections, as well as demistifying the use of NEGF for electrochemistry for those who see NEGF as a scary black box.

Note that the only subsection required to follow further discussions is subsection 1.1.3, where Tran-SIESTA terminology is discussed. The rest is aimed only to those who wish to have a high level understanding of what is going on under the hood, and might be revisited after reading the results of the simulations.

1.1.1 Non-Equilibrium Green's Functions

Let's begin by visualizing the kind of situation that we would like to simulate. Figure 1.1 shows a sketch of such a situation. The system of interest, which we call the Center (C) region, is coupled to two electrodes, Left (L) and Right (R). When voltage is applied, the L and R electrodes are forced to stay at different potentials. What this means in terms of electronic structure is that the Fermi levels (E_f) of L and R are misaligned by the voltage drop (V):

$$E_f^L - E_f^R = eV \tag{1.1}$$

with e being the charge of the electron (which is negative). This situation is out of equilibrium, as there are some electrons in one electrode that would rather be in the other electrode. Let's take the case where $V > 0$ and therefore $E_f^R > E_f^L$. If there was no external force, electrons on the R electrode with an energy higher than E_f^L would eventually move to the L electrode until $E_f^R = E_f^L$, i.e. equilibrium is reached. However, if there is an external agent such as a battery, the non-equilibrium situation can be maintained indefinitely. It is this non-equilibrium case that we are interested in simulating, as it is what happens for example when we are charging a battery or when we are trying to catalyze a reaction with the help of an external voltage. We need a methodology that can give us the electronic structure of the Center region under these conditions.

We will focus on computing the electron density, but NEGF can be used to compute many other properties. If there is one way of computing electron density (ρ) in a system that scientists are most familiar with, it is probably not through NEGF. The most common way of solving electronic structure is by interpreting the Schrödinger equation

$$\hat{H}\Psi = E\Psi \tag{1.2}$$

as an eigenvalue problem. The Hamiltonian is often defined in a certain basis (planewaves, Slater orbitals, Gaussians, wavelets, atomic orbitals, etc..) and is represented as a matrix H. Diagonalizing the H matrix gives us the eigenvectors Ψ, which are interpreted as the one-electron wavefunctions of the system, and their corresponding eigenvalues E, which are the energies. We then fill the states according to some occupation function to get the electron density:

$$\rho = \sum_i |\Psi_i|^2 n_f(E_i) \tag{1.3}$$

where the occupation function $n_f(E)$ is tipically the Fermi-Dirac distribution function:

$$n_f(E) = \frac{1}{e^{\frac{E - E_f}{kT}} + 1} \tag{1.4}$$

that depends on the Fermi level E_f and the electronic temperature T. By looking at our situation in Figure 1.1 it quickly becomes clear that our system does not have a single E_f and therefore is not easy to define a $n_f(E)$. The easiest way out of this issue is to take a completely different approach.
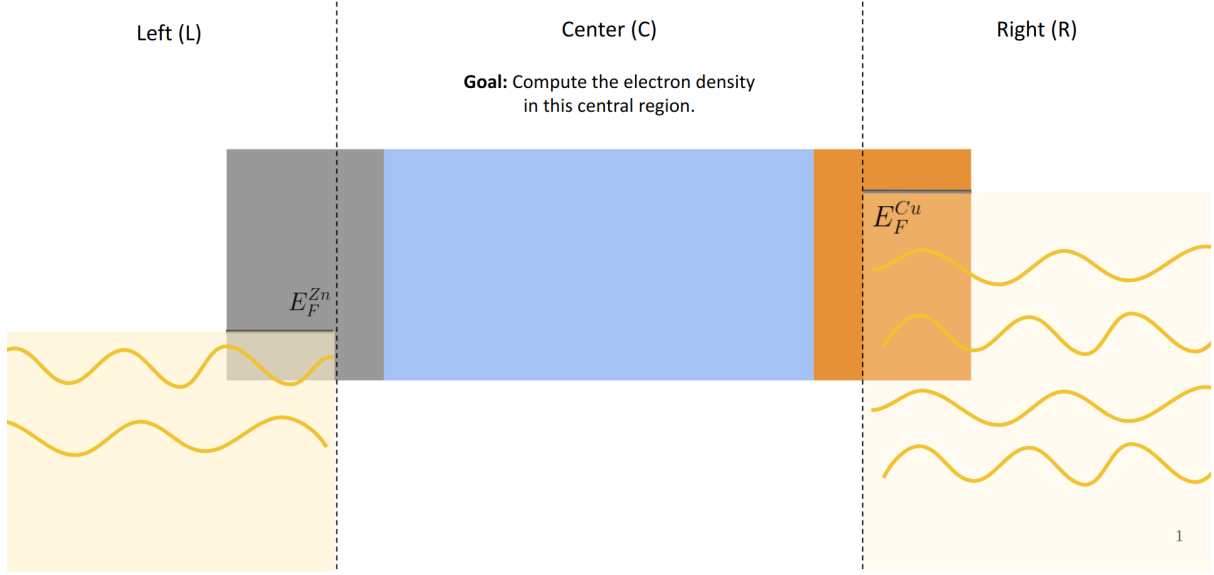
Figure 1.1: Sketch of a system with applied voltage. Our system of interest is the Center region, which is connected to two electrodes (Left and Right). The electrodes Fermi levels are misaligned by an energy of V eV, with V being the voltage drop that we want to apply.

Let's think of the system from another perspective: our Center region is now filled with states that are extensions of the states from the L and R electrodes, as depicted in Figure 1.2. This is called the scattering states approach, because the Center region is seen as a scattering center for the otherwise unperturbed bulk electrode states.

We can now split the density into two contributions, one for each electrode:

$$\rho = \rho^L + \rho^R = \sum_i |\Psi_i^L|^2 n_f^L(E_i^L) + \sum_i |\Psi_i^R|^2 n_f^R(E_i^R) \tag{1.5}$$

The form of each contribution is the same as in Equation 1.3. However, we now have well defined occupation functions, $n_f^L(E)$ and $n_f^R(E)$, which use the fermi levels E_f^L and E_f^R of the L and R electrodes. By solving the occupation functions problem, however, we have created another one: how do we get the states Ψ^L and Ψ^R and their energies? This might seem easy, but it is not. And this is the point where Green's functions come into play.

Figure 1.3 shows a sketch of the situation for the L electrode, which I will discuss, but the same discussion applies to the R electrode. The L electrode has its own states Ψ_L^o, where the o superscript stands for "unperturbed". These states Ψ_L^o can be easily computed by creating a unit cell of the electrode and running a "normal" (diagonalizing the Hamiltonian) calculation with periodic boundaries. However, we are not interested in Ψ_L^o. What we really need is Ψ_L, i.e. the scattered states, which come out once you connect the electrode to the Center region.

For that purpose, let's think from the perspective of the Center region. We can build the Schrödinger equation for the Center region:

$$H_C \Psi = E \Psi \tag{1.6}$$

where H_C is the Hamiltonian of the Center region. When the L electrode is connected, this can be seen as a perturbation from the perspective of the Center region. I am now going to make a derivation of the Green's Function that is not formally correct but it is useful to easily create some intuition on the reason behind its expression. Following, we add a generic perturbation v to the Schrödinger equation:

$$H_C \Psi = E \Psi + v \tag{1.7}$$

We are interested in the states Ψ once we have added this perturbation, so we solve for Ψ:

$$\Psi = (H_C - E)^{-1} v \tag{1.8}$$

In this scenario, $(H_C - E)^{-1}$ is the operator that we need to apply to the perturbation in order to get the wavefunctions of the system. $(H_C - E)^{-1}$ can then be interpreted as some kind of response of the
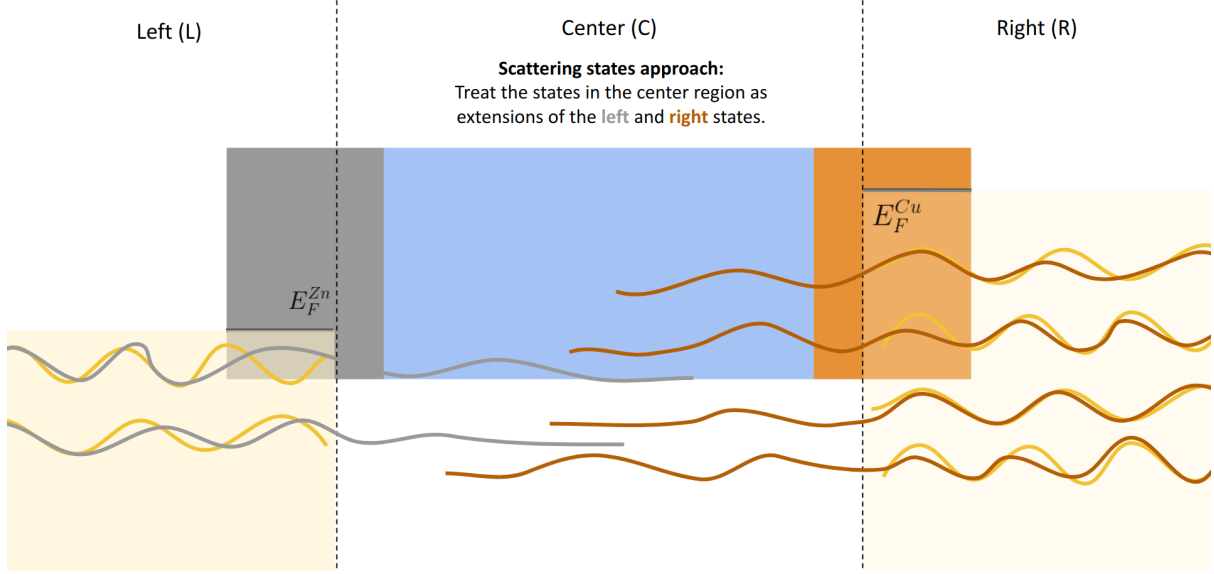
Figure 1.2: Sketch of the scattering states approach. Gold colored waves represent the original states of the L and R electrodes. In gray (L) and brown (R), I represent the same states once the electrodes are connected to the Center region. The states penetrate into the Center region, probably decaying as they go deeper and deeper into the system. At the bulk of the electrodes, the states remain the same. That is to say that the perturbation is localized at the interface between the electrodes and the Center region.

system to an external perturbation. In fact, the Green's function is the response of the system to an external perturbation, and it has this exact expression:

$$G_C(E) = (H_C - E)^{-1} \qquad (1.9)$$

In our case, as I said before, the perturbation is going to be the connection of the L electrode with unperturbed states ψ_L^o. It turns out that the final expression for the scattered states of the left electrode in the central region (Ψ_L) is as simple as:

$$(\Psi_L)_i = G_C(E_i)(\Psi_L^o)_i \qquad (1.10)$$

where we can make one final comment about G. The response of the system is energy dependent, and therefore we need to apply the response at the energy of the electrode's unperturbed state. In practical terms, it is more convenient to compute the response to the density of states. The density of states (DOS) is defined as:

$$DOS_L(E) = \sum_i (\Psi_L)_i (\Psi_L)_i^* \delta(E - E_i) \qquad (1.11)$$

and using Equation 1.10 we can rewrite it in terms of the unperturbed states Ψ_L^o:

$$DOS_L(E) = G_C(E) \left[\sum_i (\Psi_L^o)_i (\Psi_L^o)_i^* \delta(E - E_i) \right] G_C^\dagger(E) \qquad (1.12)$$

where the central term is simply the unperturbed density of states $DOS_L^o(E)$, leaving:

$$DOS_L(E) = G_C(E) DOS_L^o(E) G_C^\dagger(E) \qquad (1.13)$$

To compute the electron density from the density of states we just need to apply the occupation function $n_L(E)$, and integrate over all energies. If we sum over all electrodes connected to our system (subindex e), we end up getting the final equation to compute the electron density through NEGF:

$$\rho_C = \sum_e \int G_C(E) DOS_e^o(E) G_C^\dagger(E) n_e(E) dE \qquad (1.14)$$

The reader should notice that in deriving the equation we have not used any concept that is specific to electronic transport. We have used Green's functions (GF) because they allow us to define the states
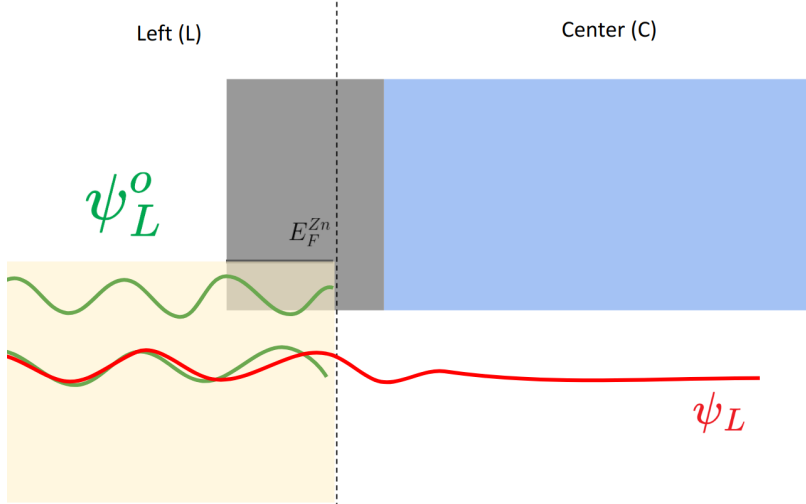
Figure 1.3: Sketch of the left electrode's scattering states. In green, Ψ_L^o, the unperturbed states of the L electrode. In red, Ψ_L, one of the scattering states penetrating into the Center region.

of our system in terms of the coupled electrodes' states. And the fact that our states are defined in terms of the electrodes' states has came in very handy to treat the non-equilibrium (NE) situation, because in contrast with the Center region, each electrode still has a well defined occupation function when the system is out of equilibrium. There is therefore nothing magic or exotic about NEGF, and they are a natural tool to use for any simulation that involves a voltage drop through the system. In other words, using NEGF for electrified interfaces in full electrochemical cells is not a stretch, but rather a natural application of the method.

1.1.2 Interaction between NEGF and DFT

We have seen that NEGF give us a way to compute the electron density of a system from its Hamiltonian. DFT allows us to do the reverse: given the electron density, we can compute the Hamiltonian using well defined energy contributions together with exchange correlation functionals. Although the connection between NEGF and DFT is not formally trivial, in practice codes usually implement the loop depicted in Figure 1.4. This loop is no different than the typical self-consistency loop that DFT-based codes use to solve electronic structure, except that NEGF is now in the position where usually one finds the diagonalization of the Hamiltonian. By implementing this loop, it is possible to compute the electronic structure of out-of-equilibrium systems ab initio with the accuracy of DFT.
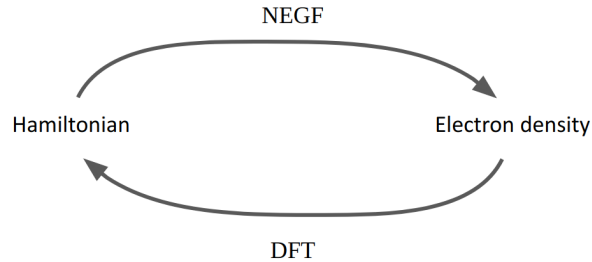
Figure 1.4: Self consistent loop using NEGF to compute the electron density from the Hamiltonian and DFT to compute the Hamiltonian from the electron density. This is the loop that TranSIESTA uses to solve the electronic structure of systems out-of-equilibrium.

1.1.3 TranSIESTA

TranSIESTA is a particular implementation of the NEGF/DFT methodology. The name TranSIESTA stands for "Transport SIESTA" as a reference to the original application of the code. Over time, the

name has remained the same, which certainly does not help communicating to the community the point that this whole section is trying to make: NEGF (and TranSIESTA) are not just for transport.

TranSIESTA is currently integrated within the SIESTA code, an open source DFT code that uses localized basis sets. The usage of TranSIESTA involves instructing SIESTA to perform the loop depicted in Figure 1.4. That is, do everything as normal SIESTA does, but use NEGF as the method to compute the electron density from the Hamiltonian. One can therefore use TranSIESTA together with the rest of the functionality in SIESTA to simulate systems out of equilibrium. For example, our end goal is to run ab initio molecular dynamics of solid-liquid electrified interfaces with TranSIESTA. Although dynamics for this kind of systems are routinely run by computational chemists under equilibrium conditions, there is still no literature on using NEGF (and TranSIESTA) to capture their non-equilibrium behavior. This makes the demonstration of TranSIESTA's capabilities for electrochemistry a much needed steping stone not only for SIESTA's users but also for the wider computational electrochemistry community.

Figure 1.5 shows a typical setup for the simulation of an electrochemical system with applied voltage. The region of interest for our simulation is sandwiched between two metallic electrodes. Our region of interest (in the picture "Simulation cell") contains the phase that we are interested in (electrolyte, liquid, gas...) and the interfaces of such phase with the electrodes. The two electrodes extend infinitely to the sides, while the rest of directions contain periodic images of the full system. In a TranSIESTA calculation, the user fixes the potential (Fermi level) of the electrodes so that the difference between them is the desired voltage drop. TranSIESTA then solves the electron density in the simulation cell to find the steady state of the system under the applied voltage conditions.
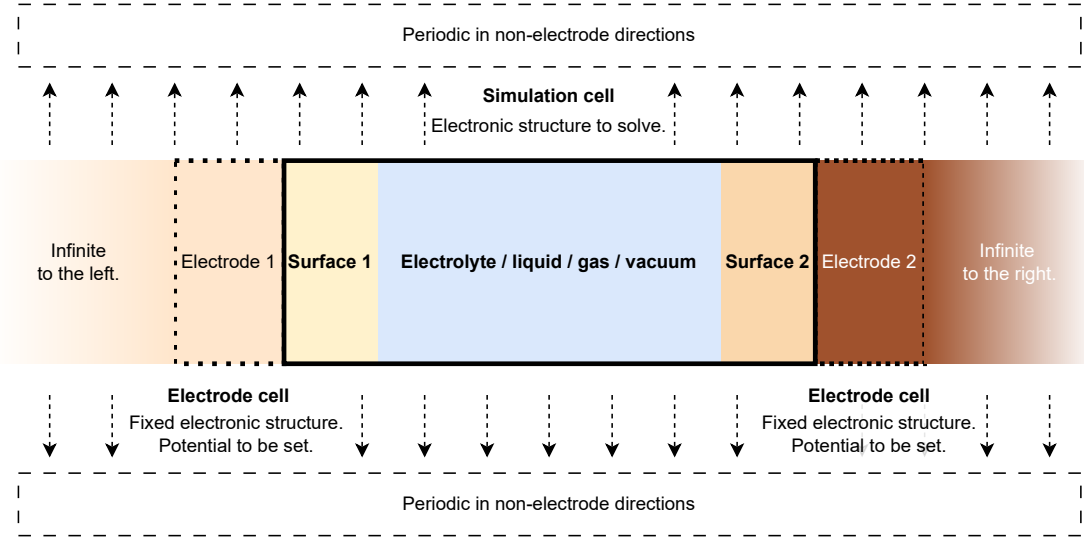
Figure 1.5: Sketch of a typical transiesta setup to simulate an electrochemical cell. The system is divided into three regions: two electrodes and one central region. The two electrodes, whose electronic structure is computed before running the TranSIESTA calculation, are coupled to the central region. Electrodes extend to infinity on their respective sides, and the rest of directions contain periodic images of both the electrodes and the central region. The central region (indicated as "Simulation cell") is really where we are interested in computing quantities. In the pictured setup, it contains an electrolyte, liquid, gas or vacuum together with the interphase of such system with the electrode.

In practice, running a TranSIESTA calculation is not a simple one step process. Figure 1.6 shows the typical workflow to be followed in order to perform a TranSIESTA simulation. The first step is to get the electronic structure of the electrodes with SIESTA (to get the DOS_e^o in equation Equation 1.14). Then, one must create a simulation cell coupling the atomic coordinates of the region of interest with the electrodes. Notice that electrodes can be reused for different calculations. Once the full system is constructed we can run the TranSIESTA calculation, indicating the Fermi level of each electrode. For example, if we want to run a calculation with an applied voltage of 1V, we set the Fermi level of the left electrode to +0.5 eV and the Fermi level of the right electrode to -0.5 eV. Setting the Fermi level of the electrodes will just induce a rigid shift of their energy levels, while keeping the same electronic structure. TranSIESTA will then solve the electronic structure and find the self-consistent Hamiltonian

and electron densities of the system. Displaying the results is then just a matter of postprocessing the output files. In particular, the core of the analysis in this thesis has been done with the *sisl* python package, a tool to postprocess DFT output.
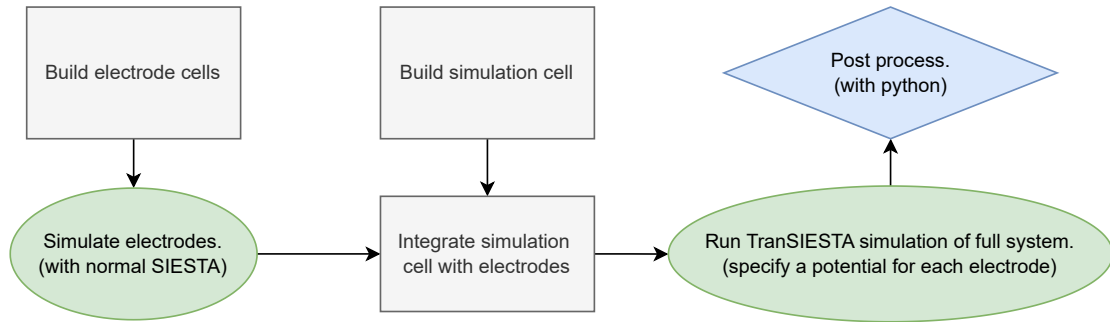

Figure 1.6: Sketch of the typical workflow in TranSIESTA, which is followed for the simulations presented in this thesis. Rectangular forms represent structure manipulation steps, ellipses represent DFT calculations, and the rhombus represents postprocessing of results.

1.2 Electrochemical potentials and TranSIESTA

The goal of chapter 1 is to explore the usage of TranSIESTA for electrochemistry. Before moving into simulating solid-liquid interfaces, where the behavior is the result of a non-trivial interplay between many factors, it makes sense to use simpler systems to probe the most basic electrochemistry concepts. Answering the question of whether TranSIESTA reproduces the most basic concepts of electrochemistry is crucial to make it a useful tool. Not only that, but it is also important to establish a clear relation between the quantities calculated by TranSIESTA and the quantities in electrochemistry textbooks. It is only by addressing these two points that electrochemistry practitioners will be able to discuss deeper electrochemical phenomena that come out of TranSIESTA simulations.

To fulfill this goal, in this section I explore the behavior of (apparently) very simple systems: capacitors with a fixed medium. That is, two electrodes separated by some medium that contains either no atoms or atoms that do not move. Considering a fixed medium is conceptually much simpler because (1) it avoids thermodynamics and (2) the dielectric response involves only the behavior of electrons. In particular, I will discuss two types of fixed mediums as illustrated in Figure 1.7. First I will discuss the simplest capacitor, which is a vacuum capacitor. Then, as an intermediate step between the vacuum capacitor and the solid-liquid interface, I will discuss a capacitor with a medium of rigid water. That is, water molecules that are all oriented in the same direction. In both cases, I first discuss the analytical expected results in terms of electrochemical potentials and then I compare the analytical solution to the results of TranSIESTA simulations.
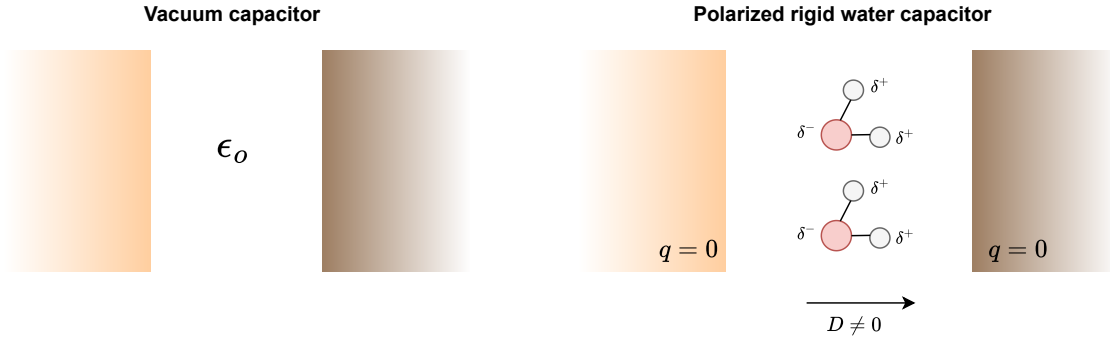
Figure 1.7: Illustration of the two types of fixed medium used to assess TranSIESTA capabilities in this section. First, we discuss vacuum capacitors, and then a capacitor with a medium of water molecules all oriented in the same direction (rigid water). This second medium can have a net polarization due to the orientation of the molecules even if the external field that they see is 0 (electrodes are not charged).

1.2.1 Conventions for electrochemical potentials

The definition of applied voltage drop between two metallic electrodes L (left) and R (right) has already been given in Equation 1.1 in terms of the Fermi levels of the two electrodes. In electrochemistry, however, it is more common to talk about potentials. The Fermi level happens to be the same as what is known as the electrochemical potential for electrons ($\bar{\mu}_{e^-}$). Therefore Equation 1.1 can be rewritten as:

$$V = \bar{\mu}_{e^-}^{L} - \bar{\mu}_{e^-}^{R} \tag{1.15}$$

This is merely a change of terminology, but it allows us to think in a different way and split the electrochemical potential into several contributions. The electrochemical potential for a particle type i in a material M is the total work that must be done to bring such a particle from infinitely far in vacuum to the material's bulk. One can arbitrarily split the electrochemical potential into multiple contributions in the form of potentials which can be observable or just a useful theoretical construct. I have chosen to follow the splitting proposed in the book Modern Electrochemistry 2A by J.O. Bockris et al. [10]:

$$\bar{\mu}_i = \mu_i + z_i F(\chi + \psi) \tag{1.16}$$

because it is easy to relate the different contributions to TranSIESTA quantities/parts of the system. In this equation:

- μ_i is the chemical potential of the particle i in the bulk of the material. It is purely a bulk property which represents the chemical stability of the particle i inside the material.

- z_i is the charge of the particle i.

- F is the Faraday constant.

- ψ is the Volta potential or outer potential. It is the electrostatic potential that a charged particle must overcome to arrive at the surface of the material from infinitely far away due to the charge excess of the surface.

- χ is the dipole potential or inner potential. It is the electrostatic potential that a charged particle must overcome to travel across the dipoles in the surface of the material coming from outside.

I will now ask the reader to agree on a convention that will make the discussions simpler. These potentials are defined for any kind of particle, but for now we will keep atoms fixed and electrons are the only particle that we will discuss. For that reason, I will adopt a convention tipically used in solid state physics which is to talk about potentials in terms of the potential energy of one electron. Figure 1.8 shows two example potential profiles using this convention, where lower potential regions are more favorable for electrons. Using this convention, the electrochemical potential of an electron is:

$$\bar{\mu}_{e^-} = \mu_{e^-} + \chi + \psi \tag{1.17}$$

where symbols mean exactly the same, but the electrostatic potentials χ and ψ are now multiplied by $-F$ with respect to the previous definition.
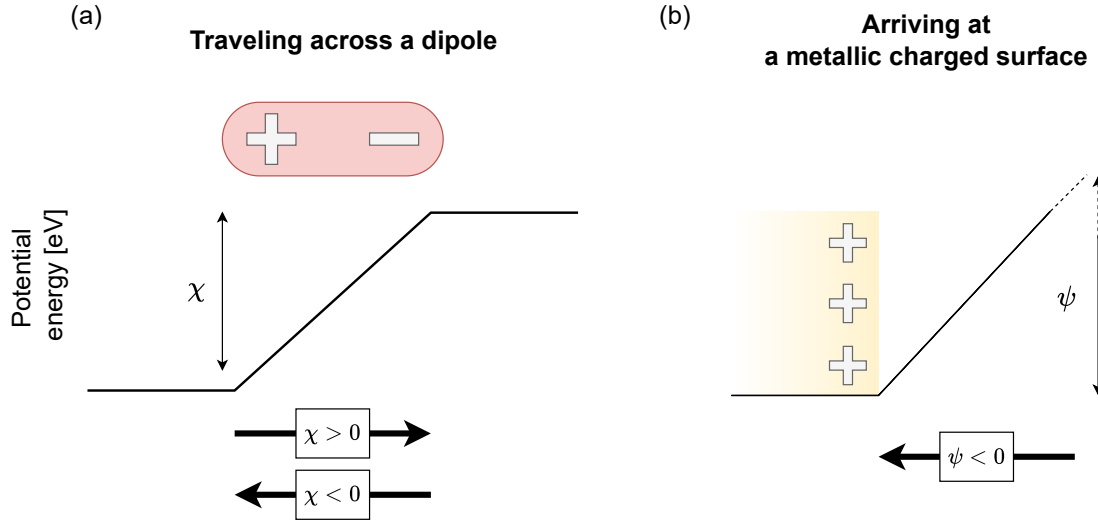
Figure 1.8: Conventions for electrostatic potentials in this thesis. Potentials are defined in terms of the potential energy of one electron. Lower potential regions are therefore regions that are more favorable for electrons. The figure displays examples of potential profiles for two cases: (a) An electron crossing a dipole. This kind of potentials will be referred to using the letter χ (b) An electron arriving at a metallic charged surface. I will use the letter ψ to refer to these potentials.

Sometimes, it will also become handy to talk in terms of the work function. The work function (Φ) of a material is the minimum work that an electron must do to get from the bulk of the material to just outside the surface. It is therefore the opposite of the electrochemical potential, but it doesn't include the Volta potential because the electron is kept just outside the surface.

$$\Phi^M = -\mu_{e^-}^M - \chi^M \tag{1.18}$$

The electrochemical potential of an electron can then be expressed in terms of the work function and the Volta potential as:

$$\bar{\mu}_{e^-}^{M} = -\Phi^{M} + \psi^{M} \tag{1.19}$$

Symbol	Name	Description
$\mu_{e^-}^M$	Chemical potential	It represents the stability of an electron inside a material M. It is a quantity inherent to a given bulk configuration. If the material didn't have a surface, this would be the electrochemical potential or Fermi level. It is different for different bulk materials. Even different crystalline phases of the same chemical composition have different μ_{e^-}!
χ	Dipole potential	Potential difference that the electron must overcome to travel across a dipole. It might refer to a surface dipole of a material M, which is then denoted as χ^M. The surface dipoles depend on the geometry of the surface and therefore χ^M is different for different surfaces of the same material. E.g. a $Au_{fcc}[111]$ surface has a different χ^M than a $Au_{fcc}[110]$ surface.
ψ^M	Volta potential	Potential difference that the electron must overcome to arrive at the surface of the material from infinitely far away in vacuum due to the charge excess of the surface. This is 0 for a charge neutral surface.
$\bar{\mu}_{e^-}^M$	Electrochemical potential	Total work that has to be made to bring the electron from infinitely far in vacuum to the material's bulk. Computed as $\bar{\mu}_i^M = \mu_i^M + z_i F(\chi^M + \psi^M)$.
Φ^M	Work function	Work that an electron inside the bulk must do to get just outside of the material. It is work in the opposite direction as the electrochemical potential, hence the negative signs. However it doesn't include the Volta potential because the electron is not moved far from the surface. Computed as $\Phi^M = -\mu_e^M - \chi^M$.

Table 1.1: Summary of electron potentials discussed in this thesis. Potentials are defined in terms of the potential energy of one electron.

1.2.2 Vacuum capacitors

Once the conventions are clear, it is now time to see how all these quantities are related to a TranSIESTA simulation. As discussed previously, in this section I will consider a system of two electrodes (L and R) separated by vacuum. Figure 1.9 shows a sketch of such a system together with an example of its potential profile. The main point of that figure is to observe how the different TranSIESTA parts participate in the potential profile. The chemical potential μ_{e^-} is determined by the bulk electrode (e.g. fcc Au) and the dipole potential χ is determined by the surface of the electrode (e.g. the [111] surface). Together, these two potentials define the work function Φ of the electrode. Additionally, the surface of the electrodes can adquire a charge excess, denoted as a charge density σ (density/area), which determines the Volta potential ψ.

We can now make an electron go from the bulk of the left electrode to the bulk of the right electrode, following the profile in Figure 1.9b. If we do so by going through the imaginary battery outside the system, the only contribution is the applied voltage V. If we take the hard path and go through the simulation cell, we will have to sum each potential difference that we encounter throughout the way. In either case, the potential difference obtained should be the same, which leaves us with:

$$V = -\chi^L - \mu_{e^-}^L + \Delta\psi + \mu_{e^-}^R + \chi^R \tag{1.20}$$

Note that here I have used $\Delta\psi$ instead of ψ^L and ψ^R. Being the system periodic, the surface of the electrodes is infinite. The infinite surface creates a constant field that extends infinitely far away. Therefore, ψ would be infinite and impossible to compute. However, we can split the Volta potential into the work to bring the electron up to some arbitrary point P where the potential is well defined (e.g. the center of the capacitor) and the work to bring the electron from there to the surface of the electrode:

$$\psi^M = \psi^{\infty \to P} + \psi^{P \to M} \tag{1.21}$$

If we then take the difference between the left and right electrodes, the first contribution vanishes and we are left only with the second contribution, which is well defined:

$$\Delta\psi = \psi^{P \to R} - \psi^{P \to L} \tag{1.22}$$

In fact, the quantity $\Delta\psi$ is simply the potential drop across the vacuum region of the capacitor, and it therefore follows the textbook formula for a capacitor:

$$\Delta\psi = Ed = \frac{\sigma}{\epsilon}d \tag{1.23}$$

where ϵ is the permittivity of the medium between plates and we have considered that both electrodes have the same charge with oposite sign ($\sigma^R = -\sigma^L$). This latter condition is not explicitly imposed by the TranSIESTA simulation. However, if there was a net charge in the system the field would extend infinitely into the bulk of the electrode. This is not physical, since the potential deep into the metallic electrode must be flat. Charge neutrality is in fact a requirement for the NEGF formalism to be correct, therefore in practice it is a constraint in our system.

Going back to the equality defined by the potential profile, we can simplify it by packing the chemical potential and the dipole potential into the work function, as done in the simplified profile of Figure 1.9c:

$$V = \Phi_L + \Delta\psi - \Phi_R \tag{1.24}$$

It feels natural to do this if both μ_{e^-} and χ are properties that are not affected by the changes in electronic structure induced by non-equilibrium. The chemical potential is certainly not dependent on the applied voltage, but there is no obvious reason for surface dipoles to stay constant as the applied voltage changes. This is something that will be explored in following sections. Plugging Equation 1.23 into Equation 1.24, we can get an expression for the charge excess on the electrodes:

$$\sigma_L = (V + \Phi_R - \Phi_L)\frac{\epsilon}{d} \tag{1.25}$$

This equation invites the reader to believe that everything is known already without running the TranSIESTA calculation. The applied voltage V is a known constraint that we impose on the system, work functions Ψ for common surfaces are tabulated and ϵ and d are simple parameters of the capacitor. However, this is not quite so. There are some subtleties that are not trivially known:
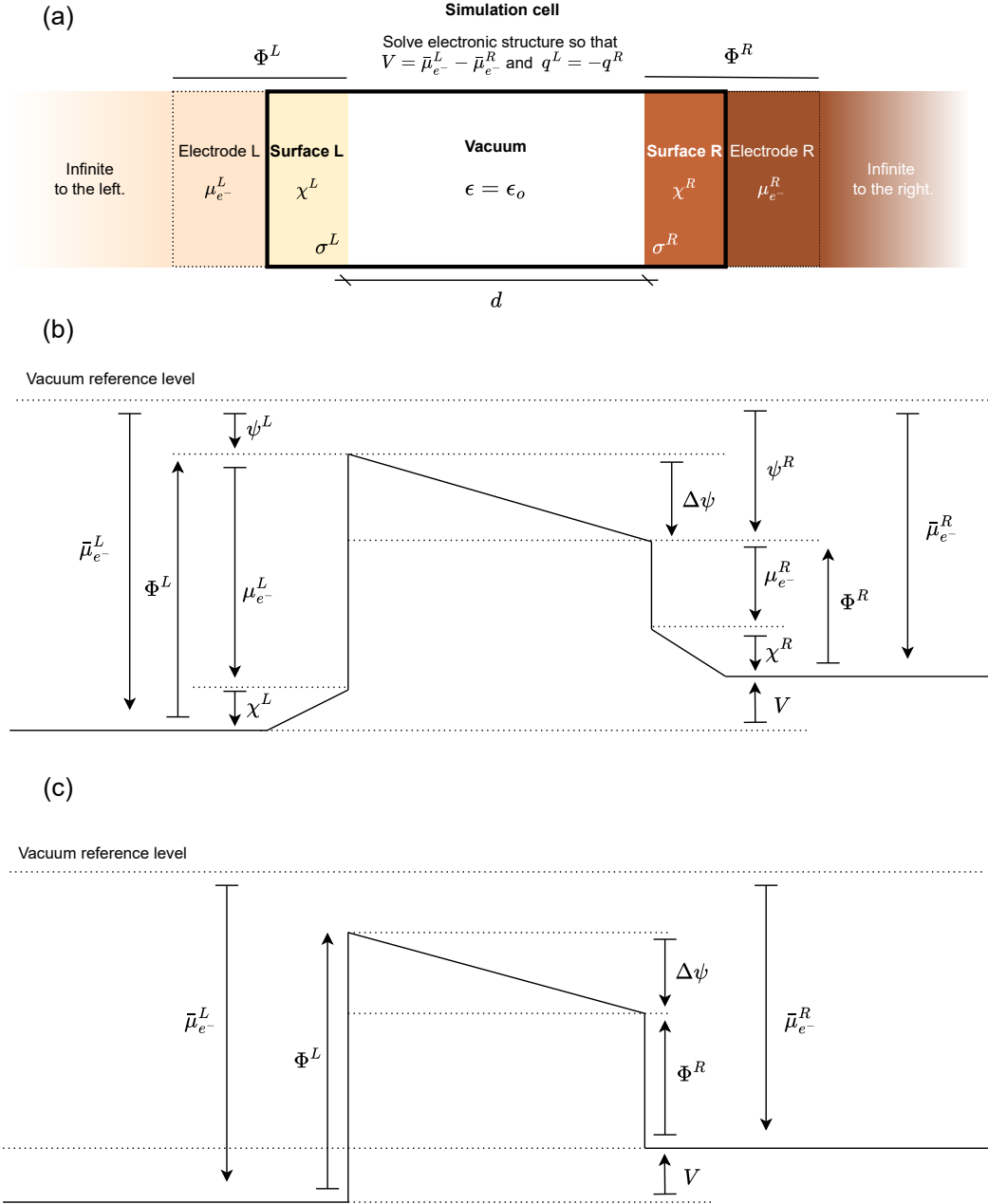
Figure 1.9: TranSIESTA setup for a vacuum capacitor. Each region contributes in some way to the electrochemical potentials.

- As discussed, the work functions might not be constant with applied voltage because there is no general reason for dipole potentials χ to remain unchanged.

- The permittivity ϵ is trivial for vacuum ($\epsilon = \epsilon_o$). Any other medium makes the determination of ϵ not so trivial. For a given snapshot (fixed atomic positions) of a medium, the reaction to external fields is only due to the response of electrons. Therefore ϵ here is the electronic part of the permittivity, or in other words the permittivity at high frequencies. Furthermore, usually the polarizability of the medium is not uniform and there is also a small vacuum region between the medium and the electrode. These two factors make the total ϵ an effective permittivity that depends on many factors. For all these reasons, the only way of determining ϵ for a given snapshot of a medium is to compute the electronic self consistent response ab initio.

- The capacitor width d is trivial at the macroscale or even at the mesoscale. However, at the microscale the error that one can make by defining it "by eye" is not negligible. Do we choose d to be the distance between the atom centers at the surface, or do we consider the surface to end some Ångström away from the last atom?

- The charge density σ is the total charge of the system divided by the surface area of the electrodes. However, it does not indicate how charge is distributed in the surface. Is it a uniform charge distribution? Does charge accumulate only on the first layer or does it extend to the second layer or outside the surface? The answer to this second question will affect the value of d.

The task of TranSIESTA is to find the electron density that makes all these variables consistent while satisfying the NEGF equations. Following, I will show how TranSIESTA successfully manages the situation of symmetric and asymmetric vacuum capacitors.

1.2.2.1 A symmetric capacitor: Au — vacuum — Au

The first TranSIESTA simulation that I will discuss is an utterly simple system. Yet, this simple system will get us started in the path to determine the fitness of TranSIESTA for electrochemistry. The system is depicted in Figure 1.10 and it is a vacuum capacitor made of two fcc gold electrodes terminated by a [111] surface. There is a vacuum region of 30 Å between both surfaces. I will use this system not only to understand its physics, but also to introduce the analysis techniques that will be used throughout the thesis.
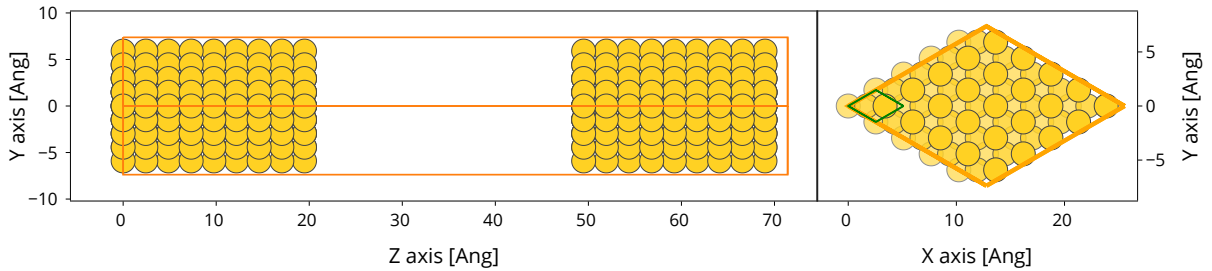
Figure 1.10: Illustration of the symmetric Au — vacuum — Au capacitor. The plot in the left shows the view across the capacitor, while the plot on the right shows a view of the electrode surface from the top (both surfaces are equal). The system is periodic in the x and y directions and gold electrodes will be connected on the -Z (left) and +Z (right) directions. There is a vacuum region of 30 Å between the electrodes. The surface is a $Au^{fcc}_{[111]}$ surface. The surface's unit cell is a 5x5 supercell (drawn in orange) of the minimal cell (drawn in green), so that comparison with the asymmetric capacitor in the next section is easier.

This is a symmetric capacitor, therefore we should expect that the difference between work functions of both surfaces is 0. This is surely true at 0V, but once the applied voltage introduces asymmetry into the system there is no guarantee that the equality will hold. Let's assume for now that the work functions are equal for all V. The symmetry reduces equation Equation 1.24 to:

$$V = \Delta\psi \tag{1.26}$$

which predicts that the potential drop across the vacuum region should be exactly equal to the applied voltage. The expected profiles at 0V and finite voltage are shown in Figure 1.11. The profiles also depict the charge excess that the electrodes must acquire to generate the required field in the vacuum region. The rest of this section is devoted to analyze the results coming out of TranSIESTA calculations and to compare them to the expected results.
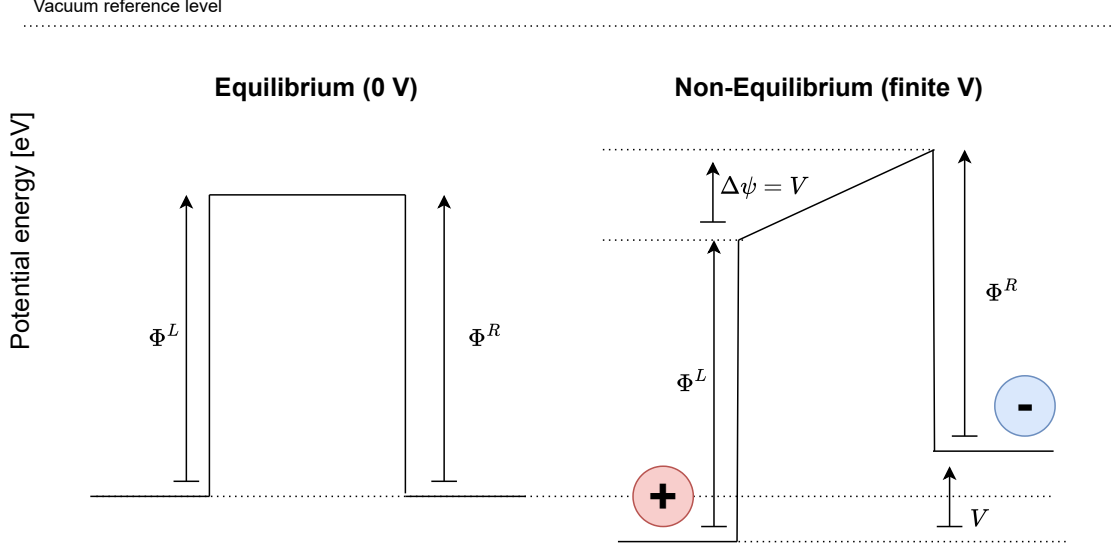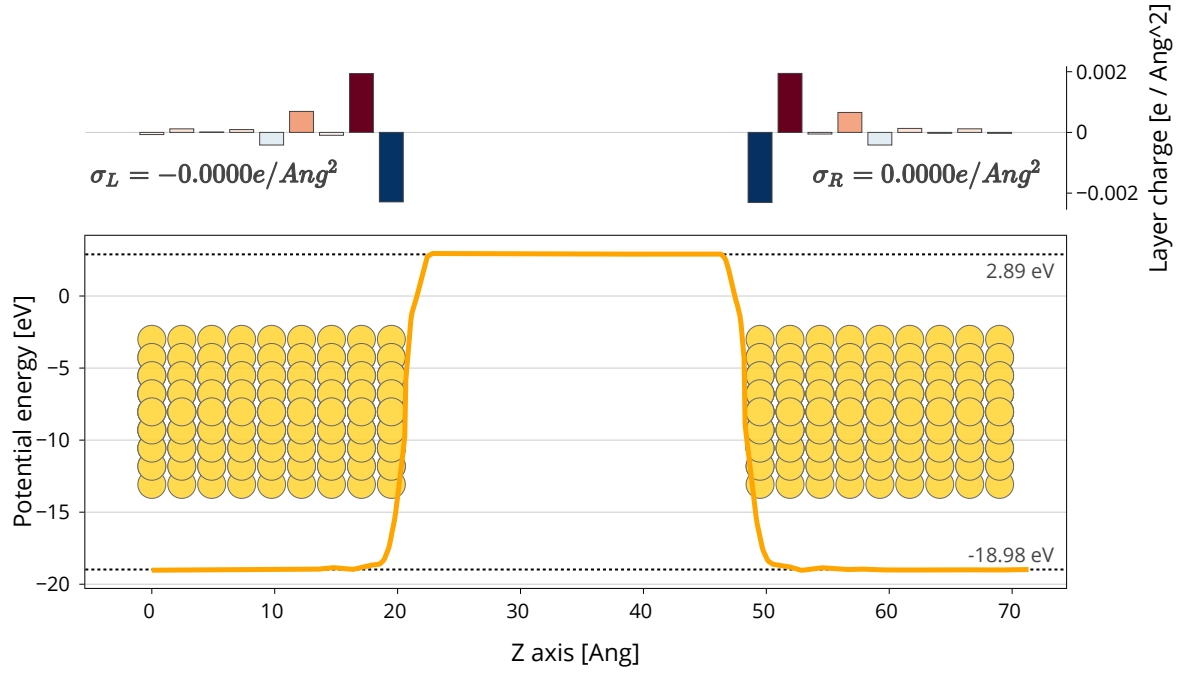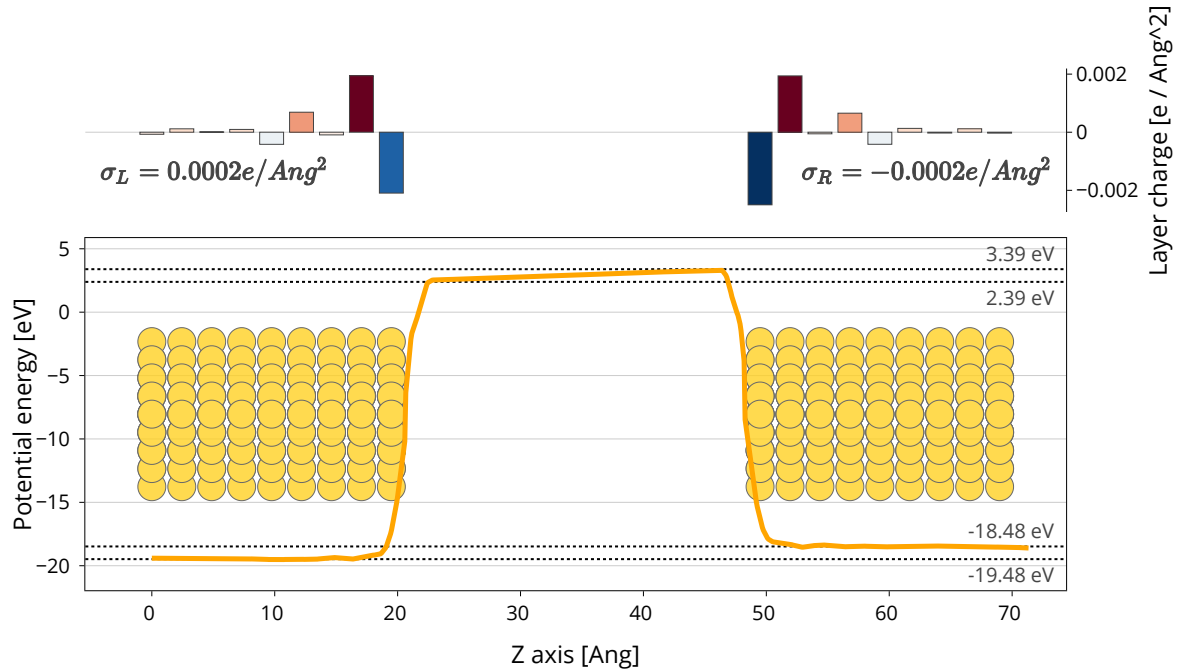

Figure 1.11: Expected potential profiles of the Au — vacuum — Au capacitor at 0V and with an applied voltage, under the assumption that the work functions remain constant. To generate the required field in the vacuum region, the surface of the electrodes must acquire a charge excess, which is represented by the + and - circles. A positive charge is deficit of electrons, and a negative charge is an excess of electrons.

Let's start with the potential profile at 0V. It sounds simple, but already there are some complications that a TranSIESTA user will face. After convergence, SIESTA outputs the total potential of the system in a real space 3D grid. The first thought is that just averaging the potential in the X and Y directions will give us the potential profile along Z. However, the potential contains huge atomic oscillations that do not allow the profile to be observed. There are many ways to solve this problem, each with their own drawbacks. I have chosen to clean the profiles following 2 steps: (1) Fourier transform the profile, mask out the frequencies of atomic oscillations, and inverse Fourier transform back to real space; (2) use a gaussian filter to smooth out the remaining oscillations. I apply this "cleaning" only to the region where there are atoms, leaving the vacuum region untouched. By doing this, I obtain the potential profile shown in Figure 1.12a for 0V. Due to the smoothing, discussing the details at the interface is meaningless. Therefore, there is no way to observe the dipole potential χ in this profile. However, we can see that the most relevant aspects of the profile are satisfied. The potentials are flat and equal in both electrodes, and the potential is also flat in the vacuum region. It is important to note that the difference between the potential at the electrodes and the potential in the vacuum region is not the work function Φ. That would require further alignment of the electrode potential with the reference levels. Nevertheless, the misalignment is the same for both electrodes so the difference between work functions ($\Phi_R - \Phi_L$, in this case 0) is correctly captured by the profile. On top of the potential profile, I illustrate the net charge of each gold layer. Although the dipole potential χ is not directly visible in the profile, the charges show that there are indeed dipoles at the surface of the electrodes. An electron coming from the bulk will have to travel across these dipoles. At the edges of our system the dipoles fade away, and the electronic structure fades into the bulk electronic structure of the Au electrode. It is a requirement of the NEGF formalism that the edges of the system have bulk-like electronic structure, and the layer charges show that I have enough gold layers to allow for the necessary screening to happen.

(a) Potential profile at V = 0 eV.

(b) Potential profile at V = 1 eV.

Figure 1.12: Potential profiles obtained from TranSIESTA calculations of the symmetric Au — vacuum — Au capacitor. On top of the profile, the hirshfeld net charge of each layer is shown with a bar graph (positive charge is deficit of electrons). The total net charge density σ of each side is also displayed.

It is nice to see that the potential profile at 0V is what we would expect, but one could get this profile without NEGF, since this is the equilibrium situation. The potential profile in Figure 1.12b is the first non-equilibrium result that I showcase in this thesis. It is the potential profile for V = 1 eV. A quick comparison with the expected non-equilibrium profile in Figure 1.11 shows that the analytic result is perfectly reproduced. The potential at the electrodes is flat, but now the left electrode is 1 eV lower than the right electrode. This 1 eV difference is of course the condition that we imposed on the system. The potential drop inside the vacuum region ($\Delta\psi$) is however not something that we directly impose. According to the profile, TranSIESTA has determined that $\Delta\psi$ is also 1 eV. This result is consistent with (but not a definitive proof of) the assumption that we made in Equation 1.26 about the work functions not depending on voltage. To generate the field in the vacuum region, TranSIESTA has charged the surfaces of the electrodes with a net positive charge in the left electrode and a net negative charge in the right electrode. This is also what was expected from the analytic solution.

Apart from plotting the raw profiles, it is sometimes useful to plot profiles of the differences between two calculations at different voltages. In Figure 1.13 I show the differences between a calculation at V = 1 eV and at V = 0 eV. Taking the difference removes the atomic oscillations completely, which obviates the need for smoothing and therefore allows us to see the sharpest details. This sharpness comes at a price, as the difference only shows the effects of the voltage change. All information of the original profiles is lost. However, we can still infer some interesting things from the profiles in Figure 1.13. The potential difference profile indicates that the calculation at V = 1 eV has (1) imposed a rigid shift on the potential of the electrodes and (2) introduced the potential drop in the vacuum region. The fact that the change in the potential is completely flat on the electrodes again suggests that the work functions are not changing with voltage, as otherwise we would see some non-flat section near the surface. The charge differences show that only the most outer layer of the electrodes accumulates the extra charge. In fact, the excess charge is accumulated on the "external" side of the outermost layer. This means that $Au_{fcc}^{[111]}$ is an excellent screener of external fields, as opposed to worse metals that might distribute the excess charge deeper within the electrode. A deeper penetration of the charge would result in a softer corner in the potential difference profile at the surface.
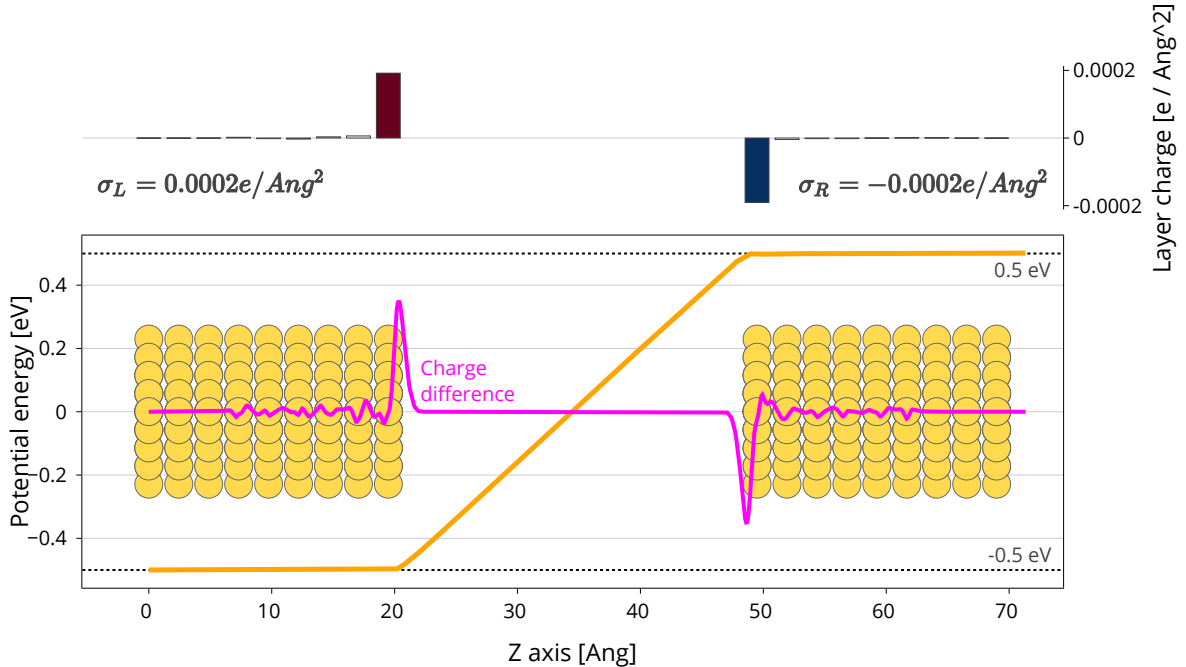
Figure 1.13: Profile of the difference between TranSIESTA calculations at V = 1 eV and at V = 0 eV (1 eV quantities minus 0 eV quantities). Along with the potential difference profile (in orange), the charge difference profile is also plotted in magenta with no corresponding axis, just to visualize the localization of the charge differences. On top of the profile, the difference in hirshfeld charges of each layer is shown with a bar graph. The total charge density difference σ of each side is also displayed. Positive charge is deficit of electrons.

After this thorough analysis of the V = 1 eV case, I would like to zoom out a little bit to observe the evolution of a quantity with applied voltage. In particular, I want to discuss the tendency of the charge excess on the electrodes. In Equation 1.25 I derived a simple expression for the charge excess on the electrodes. This expression is further simplified with the assumption that the work functions are equal and constant for the case of the Au — vacuum — Au capacitor. It reduces to the typical high school formula for a capacitor:

$$\sigma = V \frac{\epsilon}{d} \tag{1.27}$$

Figure 1.14 shows the tendency obtained from TranSIESTA calculations for the excess charge on the electrodes as a function of the applied voltage. The trend is clearly linear, which again supports our constant work functions assumption. This is in fact a much better proof, since we can be sure that the assumption holds at least from -2V to 2V. Now, fitting the data to a line we can get the slope of this linear relation. Simply dividing the permittivity of the medium (in this case vacuum) in atomic units by the slope gives us an estimate of the capacitor width d. By crunching the numbers, I get d = 28.02 Å. This is a good time to recall that the distance between the center of the outermost layers was 30 Å. It makes good sense that vacuum doesn't start right at the last atom center. Ultimately, the discrepancy between values proves the point that I made while discussing Equation 1.25: parameters that are obvious at the macroscale might be not so obvious at the microscale. Even for simple systems like the one I have chosen to start discussions with, a self-consistent solution of the electronic structure out of equilibrium (e.g. with NEGF/DFT) is fundamental to understand exactly what is going on.
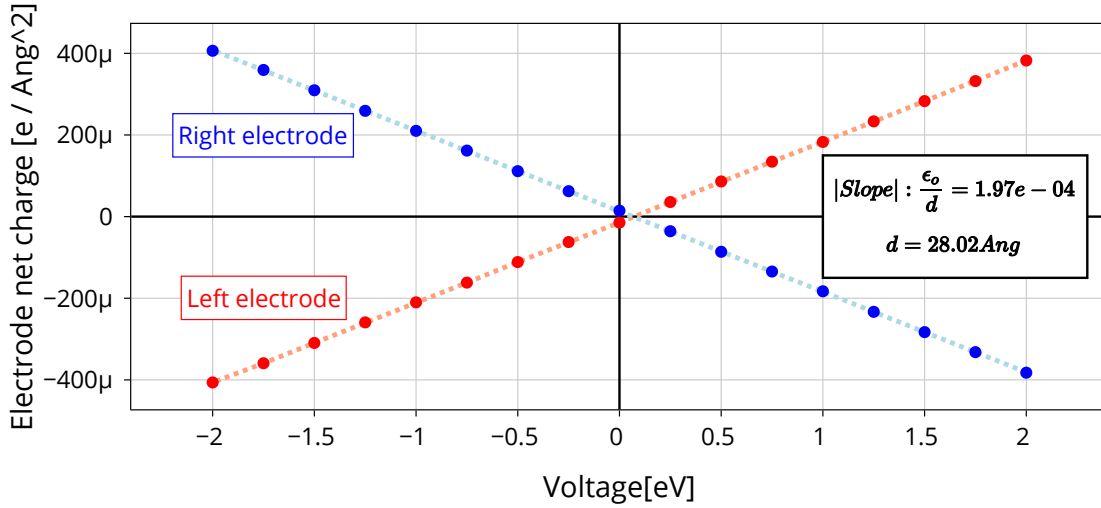
Figure 1.14: Plot of electrode excess charge vs voltage. Opposite electrodes accumulate opposite charges and cross at V = 0 eV, where they are both charge neutral. With a linear fit, we can get the slope of the line and estimate the capacitor width d.

1.2.2.2 An asymmetric capacitor: Graphene — vacuum — Au

In our path toward understanding very complex systems, let's take one small step away from the symmetric capacitor. In this section, I am taking the symmetric capacitor and adding a small asymmetry: on top of the left electrode surface, I add a graphene layer. The resulting system is depicted in Figure 1.15. Since the graphene layer is 3 Å away from the surface, the vacuum region is now 27 Å wide.

The interest of this system lies in the fact that now both surfaces are different. Therefore the surfaces should have different work functions Φ due to a difference in the dipole potential χ. In Figure 1.16, I show the result of running a "normal" (no NEGF) SIESTA calculation with the slab dipole correction. The slab dipole correction introduces a field in the center of the vacuum region to keep the vacuum potential flat. The potential drop that this field creates is precisely the difference in work functions between the two surfaces. In this case, the calculation shows that $\Delta\Phi = \Phi_R - \Phi_L = -0.52$ eV.

Figure 1.15: Asymmetric Graphene — vacuum — Au capacitor. It is the same as the symmetric capacitor in Figure 1.10, but adding a graphene layer on top of the left electrode. 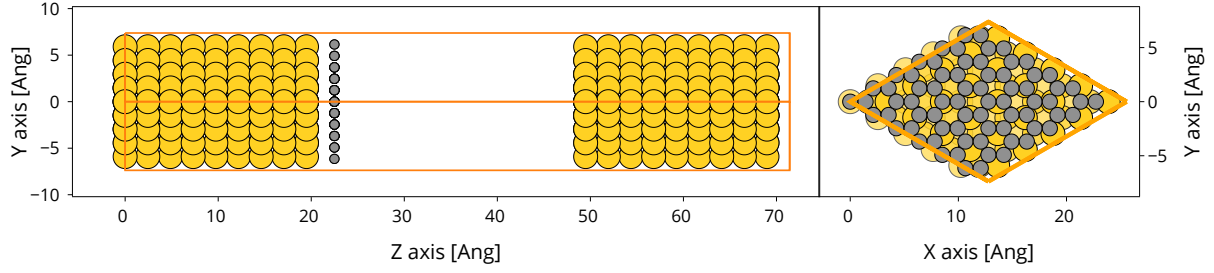The plot in the left shows the view across the capacitor, while the plot on the right shows a view of the left electrode's surface from the top. The system is periodic in the x and y directions and gold electrodes will be connected on the -Z (left) and +Z (right) directions. Before the addition of the graphene layer, there was a vacuum region of 30 Å between the electrodes. The graphene layer is 3 Å away from the surface, which leaves a vacuum region of 27 Å between the electrodes. The surface of the left electrode is an hexagonal lattice match of a 5x5 gold fcc [111] surface, and a 6x6 graphene supercell. The graphene lattice is slightly stretched and the coordinates haven't been relaxed.
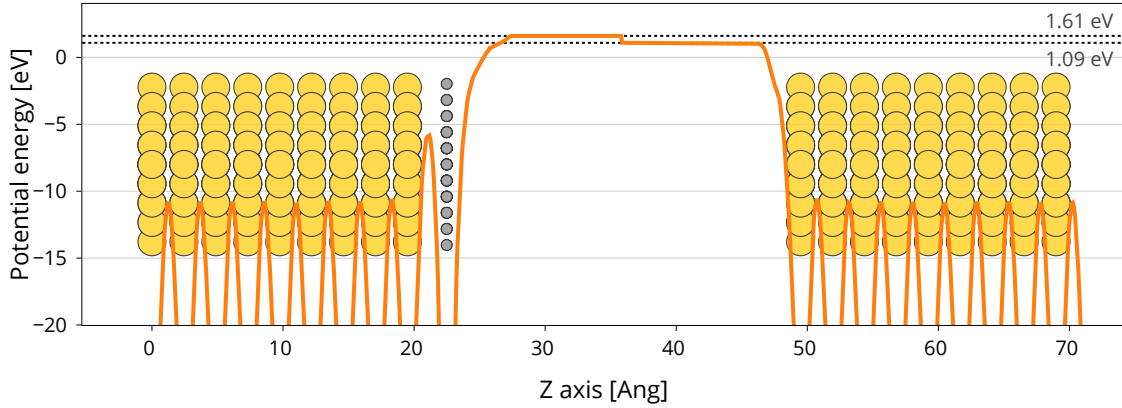
Figure 1.16: Work function difference from slab dipole corrected SIESTA simulation.

With the confirmation from a DFT calculation that there is a difference in work functions, we can now write an expression for the potential drop in the vacuum region from Equation 1.24:

$$\Delta\psi = V + \Delta\Phi \tag{1.28}$$

This equation tells us that in the equilibrium situation (V = 0 eV) there already is a field in the vacuum region. The vacuum field must be caused by an imbalance of charge in the system, in particular a non zero and opposite net charge on each electrode. This means that equilibrium involves our capacitor to be charged! This fact comes out trivially from the equations and it is well known, but it can easily be overlooked by a novice computational electrochemist. With this new situation, the expected profiles are the ones I draw in Figure 1.17. The vacuum potential drop is offset by the difference in work functions. Of particular interest is the point at which $\Delta\Phi$ exactly compensates the applied voltage. That point is the point of zero charge, since no vacuum field needs to be generated to satisfy the boundary conditions.

Let's analyze the potential profiles obtained from TranSIESTA calculations to see if they match our expectations. Figure 1.18 shows the potential profiles at V = 0 eV and V = 1 eV. The first thing to note is that the profiles are not as smooth as the ones for the symmetric capacitor, specially on the left electrode. As discussed previously, the details at the interface in these profiles are not very meaningful as they depend on the smoothing. In this case, it just so happens that the profile is harder to smooth

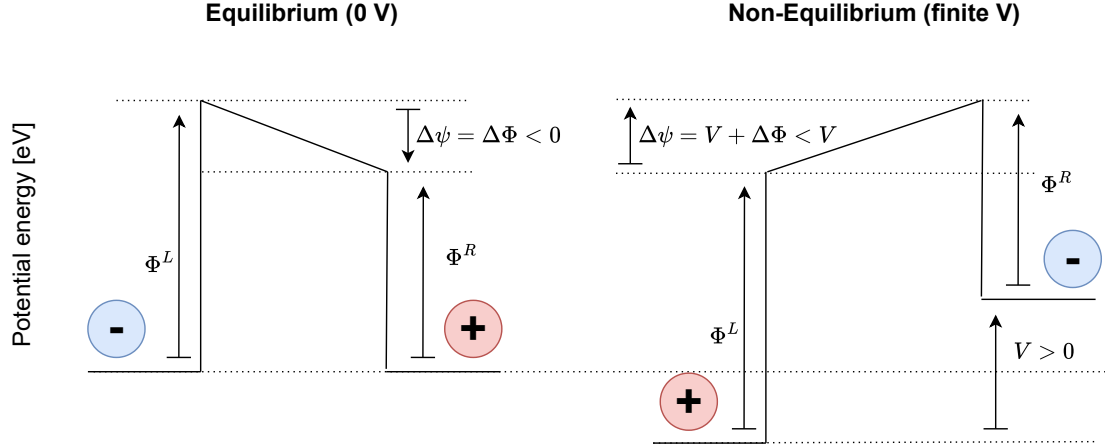Figure 1.17: Expected potential profiles of the Graphene — vacuum — Au capacitor at 0V and with an applied voltage. Even at 0V, there is a field in the vacuum region due to the work function difference between the electrodes. To generate the required field in the vacuum region, the surface of the electrodes must acquire a charge excess, which is represented by the + and - circles. A positive charge is deficit of electrons, and a negative charge is an excess of electrons.

out. I will therefore focus just on the differences between electrode and vacuum profiles, as well as the field in the vacuum region. The behavior is certainly as expected, although with a small discrepancy in the actual values. At 0V, the potential drop in the vacuum region is $\Delta \Phi = -0.39$ eV, which is not exactly the work function difference obtained from the calculation with the slab dipole correction. This discrepancy is probably due to the fact that electronic structure changes slightly once you add the dipole correction. In any case, the value is close enough and most importantly of the same sign. The profile at V = 1 eV shows how the offset in the vacuum potential is honoured in the non-equilibrium situation. Even though the applied voltage is 1 eV, the potential drop in the vacuum region is reduced to $\Delta \psi = V + \Delta \Phi = 0.61$ eV. Regarding the net charge of each layer, they are very similar to the symmetric case and the total charge density of each electrode is consistent with the vacuum field they generate.

Let's discuss the differences between the V = 1 eV and V = 0 eV cases by looking at the differences shown in Figure 1.19. As discussed with the symmetric case, taking the difference on the potential takes away some features of the original profiles. Although the potential profiles are very different, the potential difference looks remarkably close to the one I showed in Figure 1.13 for the symmetric capacitor. At the risk of repeating myself, it is very important to always keep in mind that the potential difference is not the potential that a particle will feel in the system, it is just a delta that is useful for discussion. Having said that, the change in potential is constant up to the graphene layer for the left electrode. This means that the electrode with graphene is also a very good screener of external fields, like the raw gold electrode. It is interesting that in this case it is not gold that accomodates the extra charge. Since graphene is also a conductor, it can also accomodate the extra charge, which is what both the charge difference profile and the layer charges show. If I had chosen to add a layer of insulator like boron nitride instead of graphene, that would be a different story. In any case, the exact response of a given system is not something that one can trivially predict, and it shows the importance of having an abinitio non-equilibrium method like TranSIESTA.

The tendency of the charge response to the applied voltage is shown in Figure 1.20. In this case, the excess charge density should follow:

$$\sigma = (V + \Delta \Phi)\frac{\epsilon}{d} = V\frac{\epsilon}{d} + \Delta \Phi \frac{\epsilon}{d} \tag{1.29}$$

therefore the slope of a linear fit should be $\frac{\epsilon}{d}$ as in the symmetric case, but now we should be able to also extract the work function difference $\Delta \Phi$ from the intercept. In other words, $\Delta \Phi$ is minus the voltage at which the charge excess is zero, i.e. the point of zero charge. After fitting the data, I get $d = 25.00$ Å and $\Delta \Phi = -0.41$. Again, TranSIESTA determines that the vacuum region is not as wide

(a) Potential profile at V = 0 eV.

(b) Potential profile at V = 1 eV.
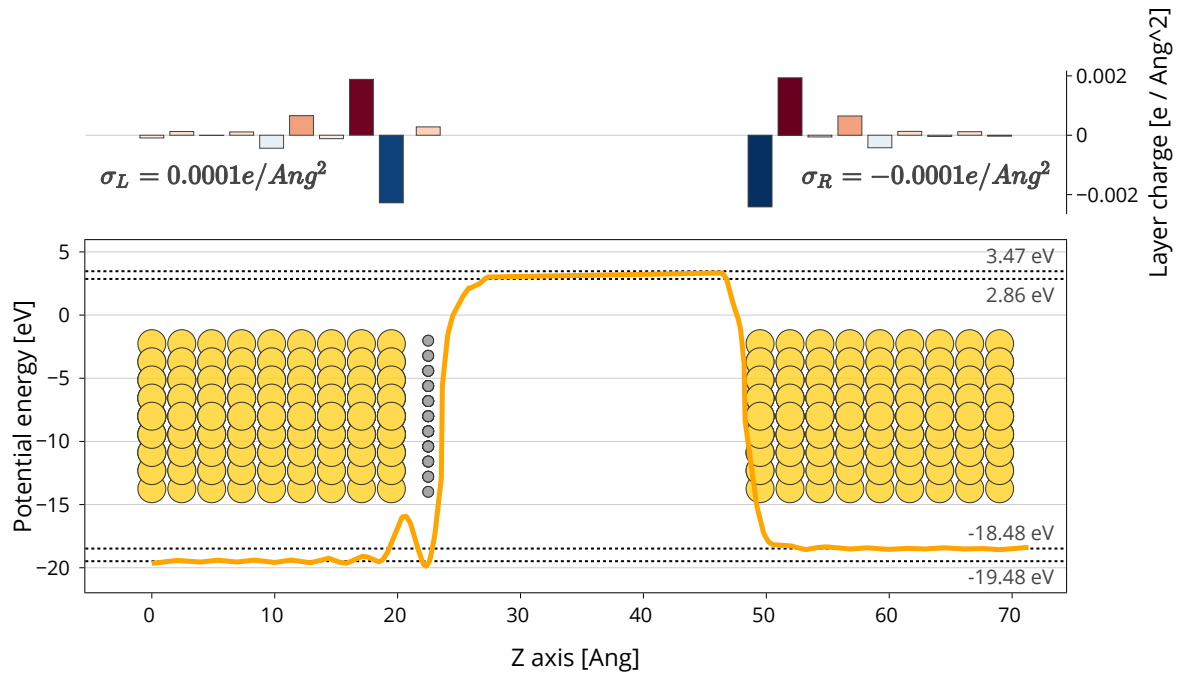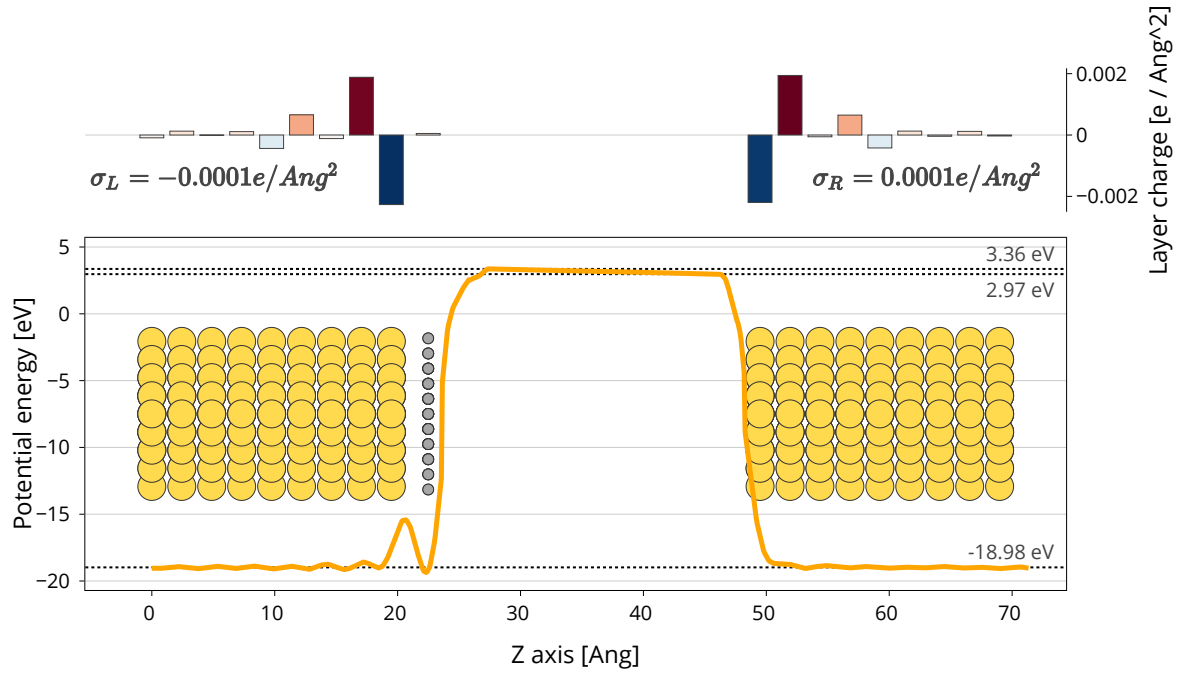
Figure 1.18: Potential profiles obtained from TranSIESTA calculations of the asymmetric Graphene — vacuum — Au capacitor. On top of the profile, the hirshfeld net charge of each layer is shown with a bar graph (positive charge is deficit of electrons). The total net charge density σ of each side is also displayed.
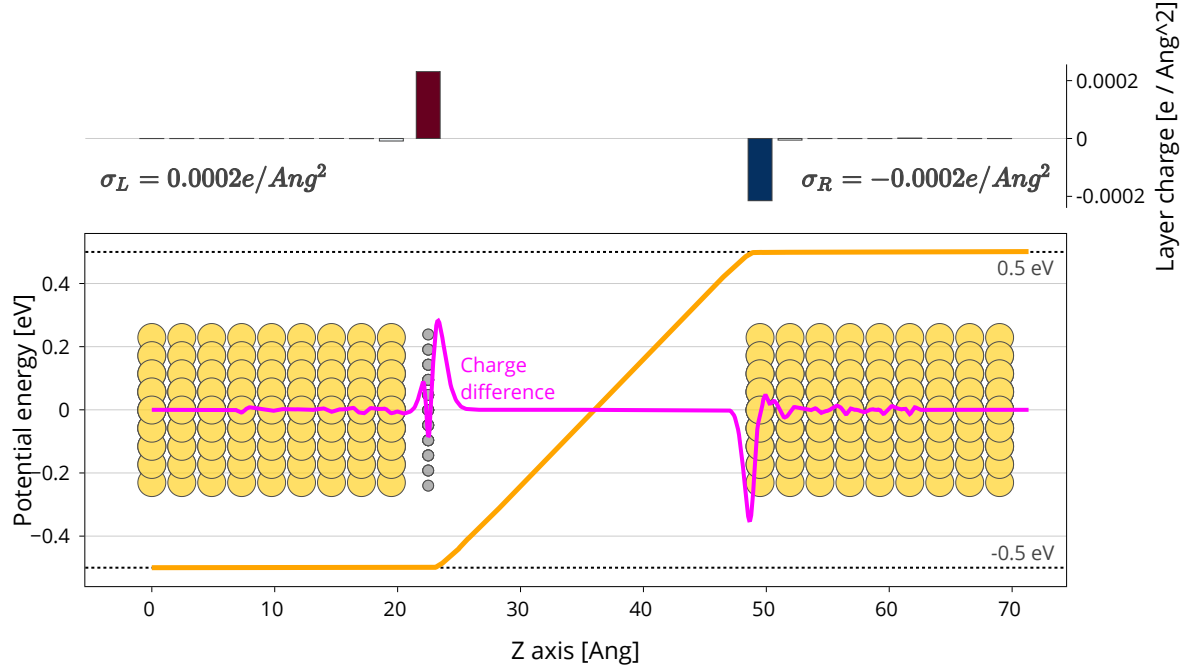
Figure 1.19: Profile of the difference between TranSIESTA calculations at V = 1 eV and at V = 0 eV (1 eV quantities minus 0 eV quantities). Along with the potential difference profile (in orange), the charge difference profile is also plotted in magenta with no corresponding axis, just to visualize the localization of the charge differences. On top of the profile, the difference in hirshfeld charges of each layer is shown with a bar graph. The total charge density difference σ of each side is also displayed. Positive charge is deficit of electrons.

as one could naively think, since the distance between last atom centers is 27 Å. Regarding the work function difference, it is not exactly the same as the one obtained from the slab dipole correction or the potential profile at 0V, but it is close enough.
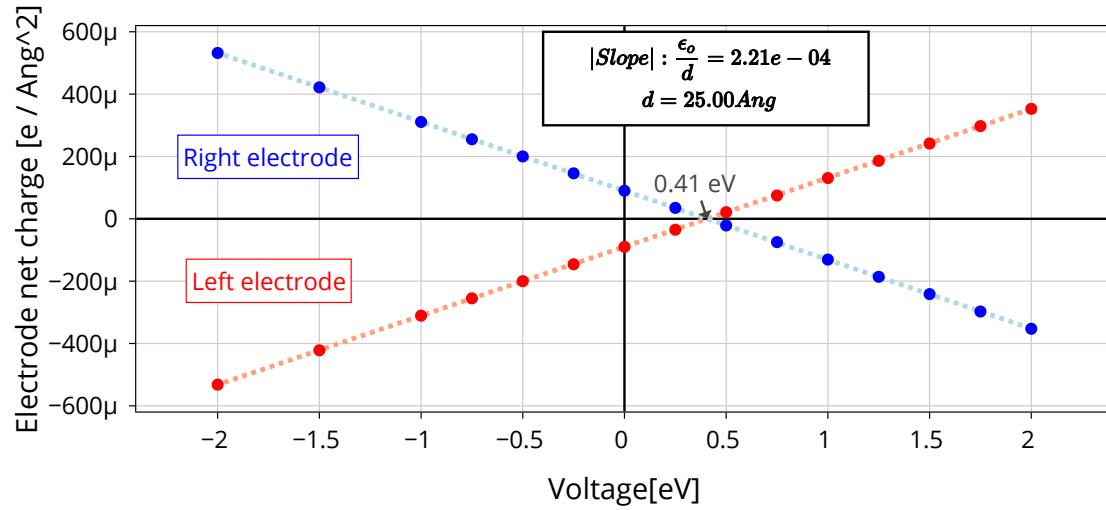
Figure 1.20: Plot of electrode excess charge vs voltage. Opposite electrodes acumulate opposite charges and cross at V = 0.41 eV, where they are both charge neutral. With a linear fit, we can get the slope of the line and estimate the capacitor width d.

1.2.3 A word on non-vacuum media

The choice of vacuum for the first examples was of course not arbitrary. The less variables there are to consider, the simplest the explanation. In my happy ignorance of the deepest physics that govern nature, I consider vacuum to be nothingness. And by definition there can be anything with less variables than nothingness. The next step in the path to understanding is to add as little "something" as possible. In that regard, we can choose to add a medium that electronically responds to the charging of the capacitor plates by screening the field. This is, of course, a dielectric medium with permittivity $\epsilon > \epsilon_o$. Such a medium would have the effect of forcing the capacitor to accumulate more charge in order to keep the same potential drop between plates, as predicted by equation Equation 1.25. To derive that equation I summed the potential contributions along the profile assuming that the medium was vacuum, so why does it still hold? The energetic price you pay to leave the new medium is the same energy that you received when you entered it, therefore both terms cancel out ($\Delta\bar{\mu}_{e^-} - \Delta\bar{\mu}_{e^-} = 0$). That means that the equations that I derived are still valid. The only effect on the potential profile is a rigid shift of the central region, but the potential drop between plates is the same. Well, unless... the price to get out is not exactly the amount you received to get in, of course. I will explore that possibility in the next section.
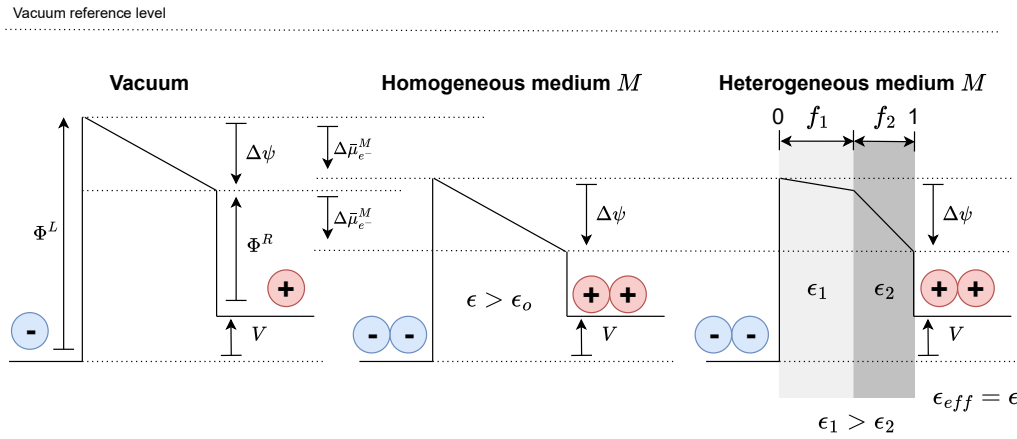
Figure 1.21: Potential profiles for capacitors that have a medium other than vacuum between the electrodes.

Before that, let me mention the possibility that the medium is not a homogeneous dielectric, but instead is composed of multiple regions with different permittivities. An example of such a system is shown in Figure 1.21, where there is a medium with two regions of different permittivities. For the equations that we have seen until now, this is not an issue. The potential drop in vacuum, as well as the excess charge in the electrodes, can be calculated in exactly the same manner, except that ϵ is now an effective permittivity of the medium (ϵ_{eff}). This effective permittivity is calculated by combining the permittivities of the components and the exact formula depends on how the components are distributed. For example, for the case illustrated in Figure 1.21, it is given by:

$$\epsilon_{eff} = \frac{\epsilon_1 \epsilon_2}{(1 - f_1)\epsilon_1 + f_1\epsilon_2} \tag{1.30}$$

The main difference when having an inhomogeneous medium is in how the potential drop between the plates is distributed. The field will not be constant anymore and instead will be smaller in the regions with higher permittivity.

In summary, the quantities that we have analyzed until now do not depend on the medium between the electrodes, as long as there is no net energy cost in traveling across the medium. The choice of using vacuum was simply to make sketches and explanations simpler. However, other properties of the system might depend on the medium. For example in the case of an inhomogeneous medium, a particle will not see the same field if it is in the high permittivity region or in the low permittivity region.

1.2.4 Towards liquid water: A rigid water capacitor

At this point, I could already throw the reader into the wilderness that liquid water is. However, I have chosen to take a more gentle approach. In an attempt to blur the line between vacuum and the "real world", I am now introducing a rigid water system. The system is depicted in Figure 1.22 and it consists of a 3x3x4 grid of water molecules all oriented in the same direction, sandwiched between two fcc [111] gold electrodes. This system serves the purpose of a playground where I can explore the effects of a polarized medium without having to deal with the statistics of thermodynamics. After understanding how the polarization of the water molecules affects the physics of the system, analyzing liquid water will hopefully feel like a less daunting challenge.
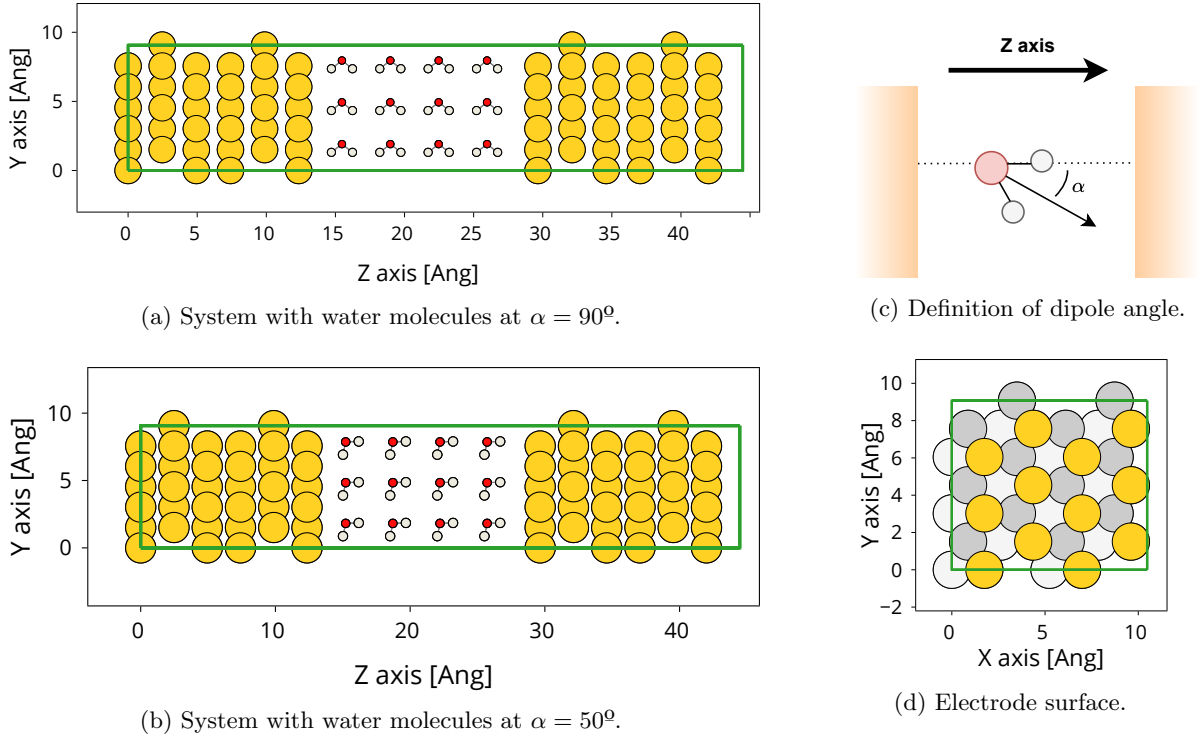

(a) System with water molecules at $\alpha = 90^{\circ}$.

(c) Definition of dipole angle.

(b) System with water molecules at $\alpha = 50^{\circ}$.

(d) Electrode surface.

Figure 1.22: Description of the rigid water system used for this section. It is a capacitor made of two symmetric gold electrodes and a 3x3x4 grid of water molecules all oriented in the same direction. The system mimics the liquid water capacitor that will be used in the next section, but liquid water is replaced by the rigid grid of water molecules. (a) and (b) show the view across the capacitor for different dipole angles α. There is a distance of 17.3 Å between the centers of the last atom of each electrode. (c) shows the definition of the dipole angle α, which is the angle between the molecule's bisector and the electrode's surface normal vector. (d) shows the surface of the electrodes, which is a 2x3 supercell of the minimal orthogonal cell for a gold fcc [111] surface.

In this section, we will see how the orientation of water has two main effects that can be treated separately. The simplest one, which I will treat first, is that for each orientation the medium is different. This is something that even a primary school student would be able to point out by comparing the pictures in Figure 1.22a and Figure 1.22b. It doesn't involve any knowledge on the electronic structure of water to have an intuition that the media might react differently to external fields just because the atoms are positioned differently in each of them. For the second effect that I will discuss, we just have to consider the fact that water is a polar molecule. By orienting all the water molecules on the same direction, we will create a net polarization within the water grid, that is, a potential drop between its ends. This extra potential drop will certainly affect our equations.

1.2.4.1 Different orientation, different permittivity

Let's for now ignore all the knowledge that we have about the polarity of water and treat the different orientations simply as different atomic lattices. What happens when these atomic lattices need to respond to an external field? Well, as any sane lattice would do, they will try to move their charges in a way

that screen the external field and therefore minimize energy as much as possible. Little do the lattices know, though, that in this case they are at the mercy of the electrodes. The electrodes are on a mission to impose a potential drop on the lattice and will acquire as much charge as needed to counteract the resistance of the lattice to the field. The more the lattice resists, the more charge the electrodes will have to acquire to impose a certain potential drop. And this "resistance" of the lattice is of course known as the permittivity ϵ.

It is not a stretch to imagine that different orientations of water will be able to fight the external field differently, just like you and I find it harder to fight gravitational fields (in other words, keep balance) in certain positions and easier in others. In a precarious situation, our human instinct would tell us to change the position of our body to make balancing easier. As far as we know, water molecules have no instinct. Yet, they use a similar mechanism. They reorient their dipoles in order to screen the external field better. In this section, however, we will apply an external field and not allow molecules to reorient. Water molecules will have to respond to the field only by redistributing their electrons. This might seem like a tough restriction to impose, but it is just what TranSIESTA (or any DFT code, for that matter) does for each self consistent calculation of the electronic structure and the atomic forces, in what is known as the Born-Oppenheimer approximation.

(a) Charge response of water at 90°.

(b) Charge difference in the electrodes.

(c) Charge response of water at 0°.

(d) Relative permittivity.
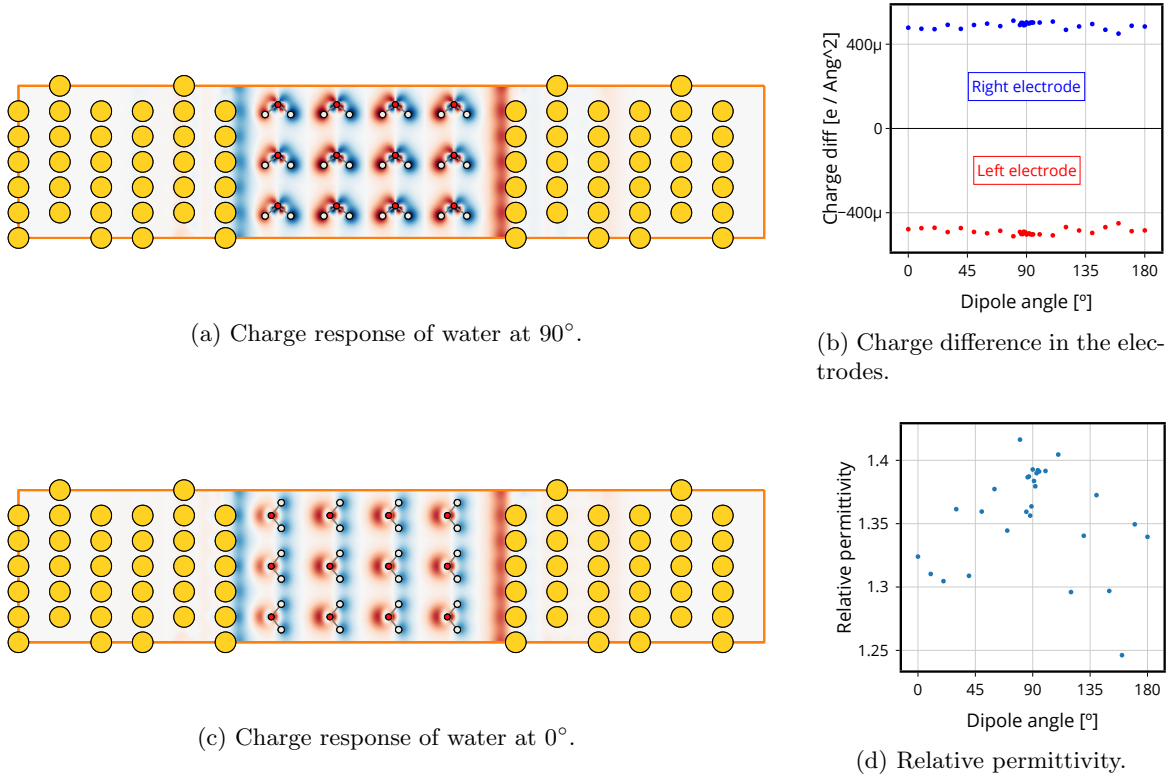
Figure 1.23: Electronic response of water to an external field at different orientations. All quantities are the result of taking the difference between a TranSIESTA calculation at V = -1 eV and a calculation at V = 0 eV. That is, the figures show the effect of shifting the potential of the left electrode up by 0.5 eV and the potential of the right electrode down by 0.5 eV. (a) and (b) show the charge response of the system, with red areas indicating a positive net charge (lack of electrons) and blue areas indicating a negative net charge (excess of electrons). (c) shows the total charge accumulated in the electrodes. (d) displays the relative permittivity computed from the electrode charges using a capacitor width of $d = 15.3$ Å.

In Figure 1.23 I show the results of running TranSIESTA calculations at V = -1 eV and V = 0 eV and taking the difference, with the aim to analyze the electronic response. In Figure 1.23a and Figure 1.23c I show the charge response of the system to the applied bias with dipole angles of 90°and 0°. As usual, electrodes charge at the surface to keep the potential flat inside the metal and all the potential drop takes place in the region between plates. In particular, the left electrode charges negatively and the right electrode charges positively. For both dipole angles, we see how the water molecules react by creating fields with opposite sign. That is, they charge positively on the left side and negatively on the right

side. It is obvious that the mechanism of charge distribution is different for each orientation. In the 90°case, the molecules break the symmetry between the two hydrogen atoms to accumulate opposite charge on each of them. On the other hand, the 0°case takes some fraction of electrons out of the oxygen atom and gives it to the hydrogen atoms. The question then is, are these two mechanisms equally effective at screening the external field? In Figure 1.23d I try to answer this question by computing the relative permittivity of each orientation. Since the left and right electrodes are the same, we can use Equation 1.27 (the typical capacitor equation) to get an expression for the relative permittivity ϵ_r:

$$\epsilon_r = \frac{\sigma d}{\epsilon_o V} \tag{1.31}$$

Using a capacitor width of $d = 15.3$ Å (the distance between last atom centers minus 2 Å), I get relative permittivities between 1.25 and 1.42. It makes sense that the permittivities are a bit higher than for vacuum ($\epsilon_r = 1$), since vacuum doesn't have any electrons to redistribute. Although the relative permittivities plot looks quite noisy, it seems that angles around 90°result in a slightly higher permittivity that those at the edges. I could fill pages talking about the details of the charge response, but that would deviate us from the main point. The message that I want the reader to take home is, again, that there is no way to know the exact response of a given snapshot beforehand. We can only get it by solving the electronic structure self-consistently.

It is fair, however, to ask if these differences in permittivity are relevant in comparison to the ability of water to generate a much bigger response by reorienting its dipoles. I will try to address this question in the next section.

1.2.4.2 Considering water polarity

Up until now, I have disregarded the effects of the net dipole caused by the orientation of the water grid. Taking the difference between calculations has allowed me to momentarily sweep that issue under the rug. But sooner or later, if we want to understand the complexity of liquid water, we will have to face the fact that water is a polar molecule. The orientation of its dipoles can (and most likely will) create a net dipole in the medium. This net dipole has a property that makes it different from the dipoles that we have seen in Figure 1.23. It is not caused by an external field. It is instead intrinsic to the atomic positions. From now on, I call this dipole χ_o, with the idea that when running a DFT calculation this is the "initial dipole" of the medium. The final calculated dipole will be $\chi_o + \delta\chi$, where $\delta\chi$ is the electronic response to the external field (considered by the permittivity and therefore not explicit in the equations). In a sense, χ_o can be understood as an offset on the response of the medium to an external field.

The question now is: how does this dipole affect the potential profile and the equations that we have derived until now? The answer to this is not straightforward. However, by drawing the profiles as I do in Figure 1.24, one realizes that the effect of χ_o is simpler than expected. Since χ_o does not depend on voltage or the electrode charge, it just becomes another offset on the voltage drop between plates, as we had seen with $\Delta\Phi$, the work function difference. By going from the bulk of the left electrode to the bulk of the right electrode summing the different contributions, we get:

$$V = -\Delta\Phi + \chi_o + \Delta\psi \tag{1.32}$$

where I feel the need to remind the reader that $\Delta\psi$ is the potential drop between plates caused by the excess charges on the surface of the electrodes. We can now impose different conditions to see how the system behaves. To get cleaner expressions, I will use the fact that our capacitor of interest is symmetric and therefore $\Delta\Phi = 0$, but one could do the same exercise for an asymmetric capacitor. There are mainly two interesting conditions to impose:

No excess charge in the electrodes. This is probably the first natural question to ask. In what conditions do electrodes remain uncharged? With the asymmetric vacuum capacitor we have already seen that this is called the point of zero charge and it does not necessarily correspond to V = 0. To have an uncharged capacitor, the condition that we need to impose is $\Delta\psi = 0$. With that condition, we get the voltage of zero charge as $V = \chi_o$. In other words, if the bias that we want to apply is equal to the potential drop caused by the dipole, the electrodes do not need to contribute with an extra potential drop $\Delta\psi$ as the job is already done by the dipole. This point of zero charge might seem like a point with special stability, but in fact the dipole is quite unhappy in this situation as it sees its own depolarizing field. If left in this condition, the dipole will just become smaller and the capacitor will eventually get charged.
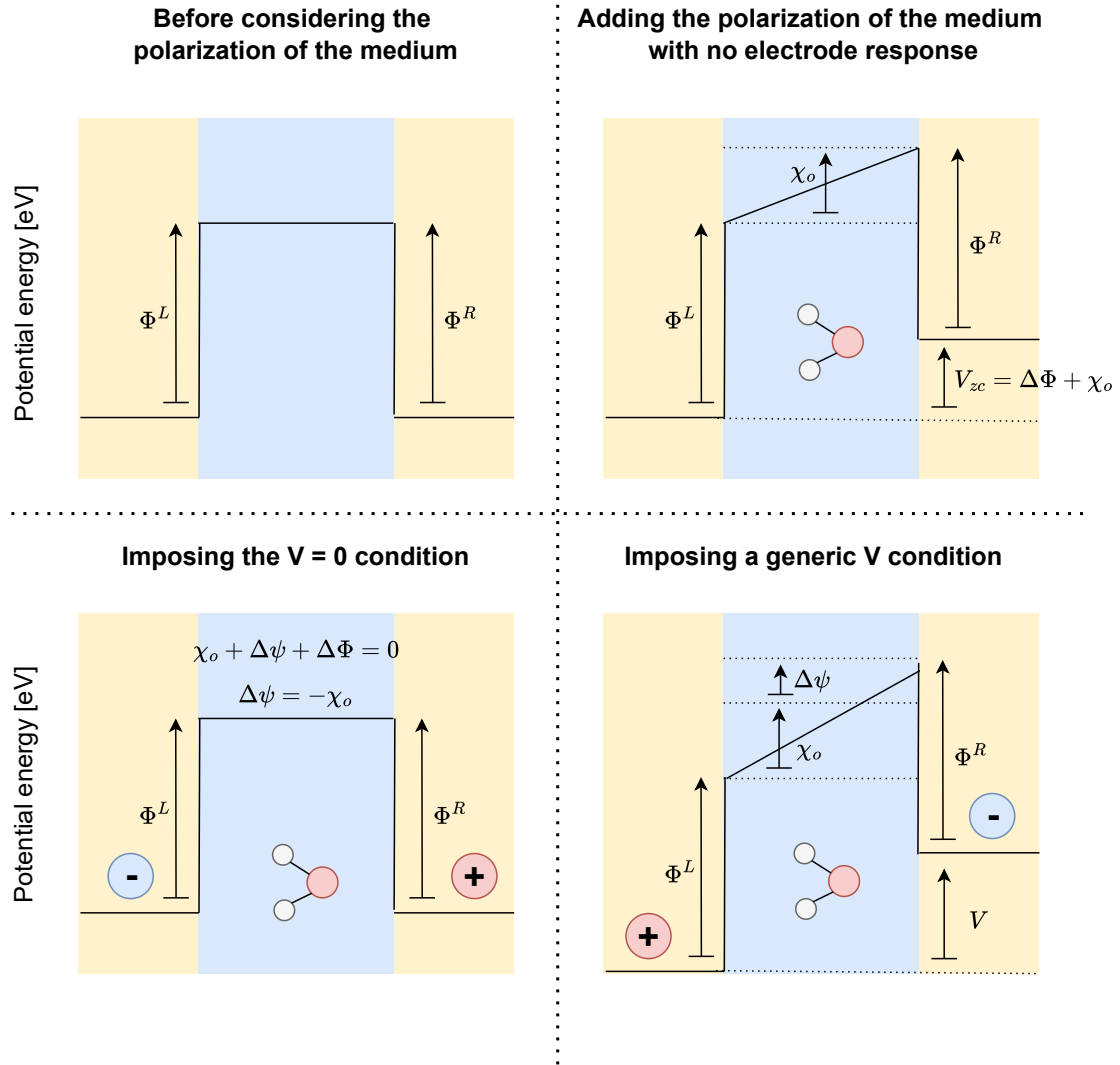
Figure 1.24: Expected potential profiles with a medium that has a net polarization χ_o due to the atomic positions. We work our way up from the unpolarized medium, to adding a net polarization in the medium, to adding the response of the electrodes given a certain V condition. Excess charge on the electrodes for the particular shape of the drawn profiles is represented by the + and - circles. Positive charge is lack of electrons and negative is excess of electrons.

No field inside medium. The other natural question to ask is: what conditions do I need to impose to completely screen the field inside the medium? This no-field scenario is an important scenario, because it means that the dipole of the medium is stabilized. In other words, the water molecules don't feel a force that drives them to change their orientation. For this condition to be fulfilled, the potential drop in the medium must be zero. The condition that we need to impose therefore is $\chi_o + \Delta\psi = 0$, which leads to $\Delta\psi = -\chi_o$. As expected, this means that to counteract the dipole of the medium, the electrodes must acquire a charge that generates the opposite potential drop. What is perhaps less intuitive is the applied voltage that allows us to do that. By introducing our condition in Equation 1.32, we get $V = 0$. That is, the bias that we need to apply to perfectly screen the dipole of the medium is zero. This result tells us that the system in equilibrium will perfectly screen any dipole, regardless of how large it is, provided that the electrodes can accommodate enough charge. If this is not shocking enough, it implies that any voltage other than 0 will overshoot the screening of all dipoles and therefore induce the medium to polarize even more. Eventually, the system will be induced to reach its saturation dipole (all molecular dipoles completely aligned) when the voltage is not exactly 0, no matter how small the voltage is! This is certainly not true in real life, and soon we will see the consideration that we have not taken into account yet.

(a) Excess charge at V = 0 eV.

(b) Excess charge at V = -1 eV.
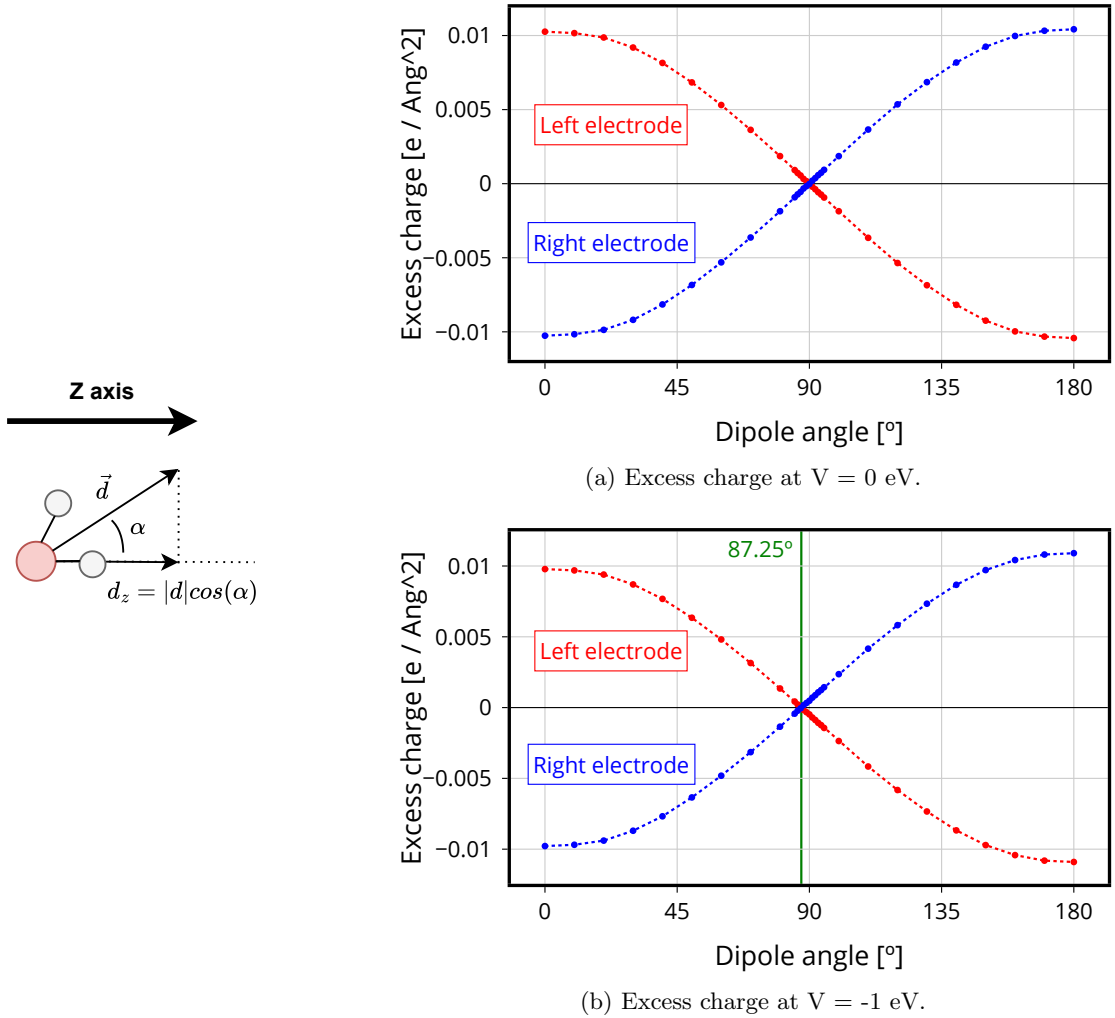
Figure 1.25: Excess charge in the electrodes when running TranSIESTA simulations at V = 0 eV and V = -1 eV. for the rigid water system at different angles. For the V = -1 eV, the green line indicates the point of zero charge at $\alpha = 82.5^\circ$. A sketch of the dipole component in Z is shown to highlight that the dipole in Z and therefore χ_o varies as the cosine of the dipole angle. Therefore, the accumulated charge in the electrodes should also vary as the cosine of the dipole angle.

To check the validity of my analytical results, I will now show the results of running TranSIESTA calculations with the rigid water system for different dipole angles at two different voltages: V = 0 eV and V = -1 eV. First, let's check the excess charge at the electrodes. Figure 1.25 shows how the excess charge

in the electrodes evolves with the dipole angle. At both voltages, the relation between dipole angle and excess charge is a cosine. This is the expected result, since the Z component of the dipole and therefore χ_o varies as the cosine of the dipole angle. We also see the expected result that the dipole angle 90°is not the point of zero charge for V = -1 eV. Instead, V = -1 eV is the voltage of zero charge for a dipole angle of 87.25°. This is, of course, because the dipole potential of the 87.25°configuration is $\chi_o^{87.25} = -1$ eV. By making these two observations, we have checked that the system behaves as expected regarding the accumulation of charge. I should also make a comment on the comparison between the total accumulated charge shown in Figure 1.25 and the charge differences with voltage shown in Figure 1.23. If 87.25°is the point of zero charge, we can also say that the excess charge at 87.25°and V = 0 eV was simply the effect of applying $\Delta V = 1$ eV from the point of zero charge. Or what is the same, the charge that we see at V = 0 eV for any dipole angle with a dipole potential χ_o is equivalent to what we would see in a capacitor with an unpolarized medium with a relative permittivity ϵ_r at $V = -\chi_o$, with ϵ_r being the relative permittivity of the medium that we computed in Figure 1.23d. Remember that the ϵ_r that I computed and that I am referring to here is just due to the response of the electrons, it does not consider water reorientation. With this comparison, ϵ_r certainly feels important since the total accumulated charge depends linearly on it. What about the applied voltage? Is it also important? By comparing the charges in Figure 1.25a and Figure 1.25b, we can see that the vertical shift in the plot induced by V = -1 eV is quite insignificant for dipole angles far from 90°. In fact, the applied voltage will only make a significant difference in the electronic distribution when $|V|$ bigger than or close to $|\chi_o|$. The question is now if it ever happens in real life that $|\chi_o| >> V$, which would be the case in which the applied voltage is not important for the electronic structure. I will answer this in the next section.

Let's now move to understanding the internal field in the medium. To do that, in Figure 1.26 I plot the LDOS(E) across the capacitor. The figure might be a bit overwhelming, so I will try to guide the reader through it. What we can see in that figure is, for dipole angles going from 90°to 180°, the LDOS across the capacitor at a voltage of V = 0 eV and V = -1 eV. At the edges of the system, we can see the states of the electrodes, which cross the Fermi level, indicated by the magenta lines. Between the electrodes, we can see the states of the four layers of water. In fact, we see only the HOMO of water along with some other similar energy states, since they are all below the Fermi level. The main point of these figures is that the energy ramp of the water states can be interpreted as the internal field, assuming that the states of water are just rigidly shifted in energy by the local potential. For example, with the 90°angle we can see that at V = 0 eV the states are all in a flat line (same energy), while for V = -1 eV there is a potential drop in water that is roughly equivalent to the shift of the electrodes, i.e. -1 eV. This is consistent with the equations, so it should make us very satisfied! However, the satisfaction does not last long if one looks at the LDOS for different dipole angles at V = 0 eV. As the dipole potential χ_o increases, the 0V condition is not able to screen the field inside the medium. My equations predicted that a 0V condition should be able to screen any dipole, no matter how big. Furthermore, the equations also predicted that a voltage different from 0 would never be able to stabilize any dipole, and the voltage drop in the medium should be equal to the applied bias. Yet, the dipole angle 120°, with a dipole potential of $\chi_o = 11$ eV, shows an almost flat potential at V = -1 eV. In other words, the dipole angle 120°is stable under the applied bias of -1 eV. Assuming that TranSIESTA calculations are right, the discrepancy between the equations and the results is surely a sign that I am missing something. But what could I be missing? Can the potential profile be explained completely by electrostatics, or is this a sign that I should consider other chemical effects? It was not trivial for me to find the answer, and I hope that future TranSIESTA electrochemists can profit from the conclusions that I will derive. The answer is that I am not considering something as seemingly insignificant as the gap between the electrodes and the medium, what I will call the "interface gap".
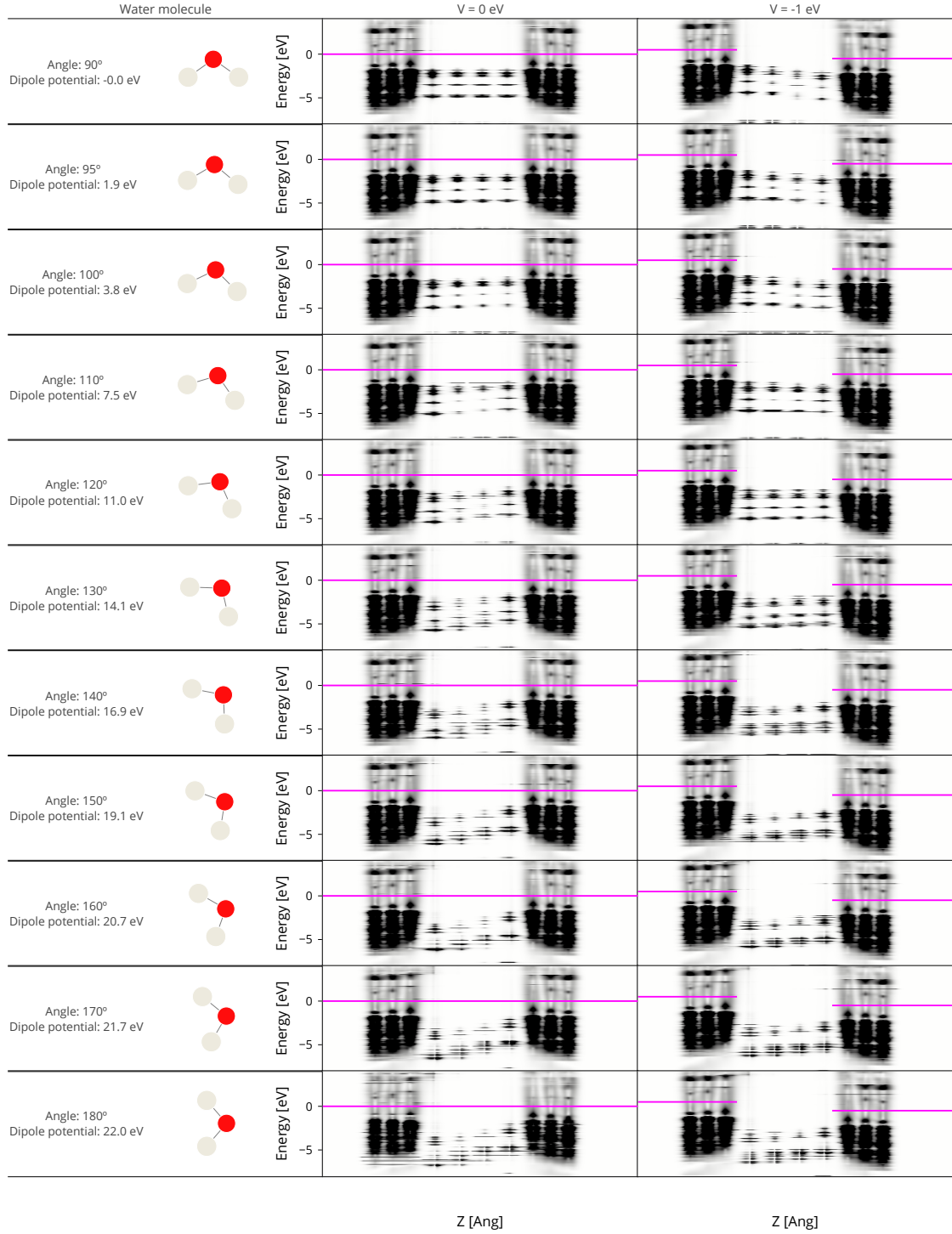
Figure 1.26: Local density of states computed with TranSIESTA for the rigid water capacitor at different dipole angles and voltages. The X axis is the position across the capacitor, and the Y axis is the energy of the states. Black areas indicate regions of space with high density of states at the energy of the Y axis, while white areas are regions with no density of states and gray is something in between. At the left and right side of the LDOS plots there are three clear black "columns". Those are the states of the last three layers of gold. The region between the gold electrodes corresponds to the states of water. Magenta lines over the LDOS indicate the Fermi level, which is constant for the V = 0 case and only well defined at the electrodes for the V = -1 eV case. The dipole potential for each dipole angle has been calculated with a SIESTA calculation of the isolated water grid (no electrodes) using the slab dipole correction.

1.2.4.3 The interface gap: an unexpected key player.

Let's take the profiles in Figure 1.24 and add a vacuum gap between the electrodes and the medium, i.e. the interface gap. The new profiles are shown in Figure 1.27. Since the interface gap has no intrinsic dipole, adding the interface does not affect equations like Equation 1.32 that summed the different potential drops in the system. However, it shows that the zero field condition that I imposed ($\Delta\psi = -\chi_o$) was too simplistic, since once the electrodes are charged the potential drop between the electrodes does not fall entirely at the medium. There is also some potential drop in the vacuum regions, which is indicated in the profiles as PD_v. The total potential drop between electrodes can be expressed in terms of the potential drop in the vacuum region and the potential drop in the medium (PD_m):

$$V = PD_v + PD_m \tag{1.33}$$

here we have used that the capacitor is symmetric and therefore the total potential drop is the applied voltage V. It is straightforward to realize now that the condition of zero field inside the medium, i.e. the one that stabilizes the dipole, is:

$$PD_m = 0 \tag{1.34}$$

which leads us to the simple result:

$$PD_v = V \tag{1.35}$$

that is, in the situation of zero field inside the medium all the voltage drop happens at the interface gap. That sounds really obvious, and all we had to do is allow for an interface gap to exist. In other words, by introducing the interface gap the system now has a way to "leak" some potential drop so that V is not restricted to $V = PD_m = 0$ in the zero field situation. Let's now see what this leakage of potential drop represents in terms of the dipole that a given voltage can screen. We can write expressions for the potential drop in the different regions as a function of the electrode charge σ, the distances d_v (vacuum gap) and d_m (polarized medium width), and the dipole potential χ_o:

$$\begin{cases} PD_v = \frac{\sigma}{\epsilon_o} 2d_v \\ PD_m = \frac{\sigma}{\epsilon_m \epsilon_o} d_m + \chi_o \\ PD_v + PD_m = V \end{cases} \tag{1.36}$$

where ϵ_m is the relative permittivity (due to electronic response only) of the medium, which we have computed previously for the rigid water system. In vacuum, where I have aggregated the potential drop at both interfaces, the only field is the one created by the electrode charges, while in the medium we also have the potential drop caused by χ_o. In the condition of zero field, $PD_v = V$ and therefore the electrode charge is (using the expression for PD_v):

$$\sigma = \frac{V}{2d_v} \epsilon_o \tag{1.37}$$

If we plug this into the expression for PD_m and simplify we get:

$$V = -\chi_o \frac{2d_v}{d_m} \epsilon_m \tag{1.38}$$

which indicates that the voltage of zero field is no longer 0! Instead, we need to apply some bias in the opposite direction of the dipole to compensate for the leakage of potential drop in the interface gap. How much bias we need to apply depends of course on how big the vacuum region is in comparison to the medium region, but it also depends on the relative permittivity of the medium. Again, the electronic response to external fields, which seemed insignificant compared to the orientational response of water, plays a big role. Let me now plug some numbers into the equations. In Figure 1.26 we saw that V = -1 eV was able to stabilize a dipole of $\chi_o = 11$ eV. Using a rough average of the relative permittivity that we computed for rigid water, $\epsilon_m = 1.35$, I get that $\frac{2d_v}{d_m} = 0.067$. If we consider a capacitor width of $d = d_m + 2d_v = 15.3$ Å (2 Å less than the distance between last atoms centers, consistent with our previous observations) I get that the total vacuum is $2d_v = 1.02$ Å, in our symmetric sketch that would be $d_v = 0.51$ Å on each interface. That sure sounds reasonable!

One could also describe the dipole potential as the product of a polarization density P_m and the width of the medium d_m, that is $\chi_o = P_m d_m$. By using this in equation Equation 1.38 we can get the polarization that is stabilized by a certain voltage:

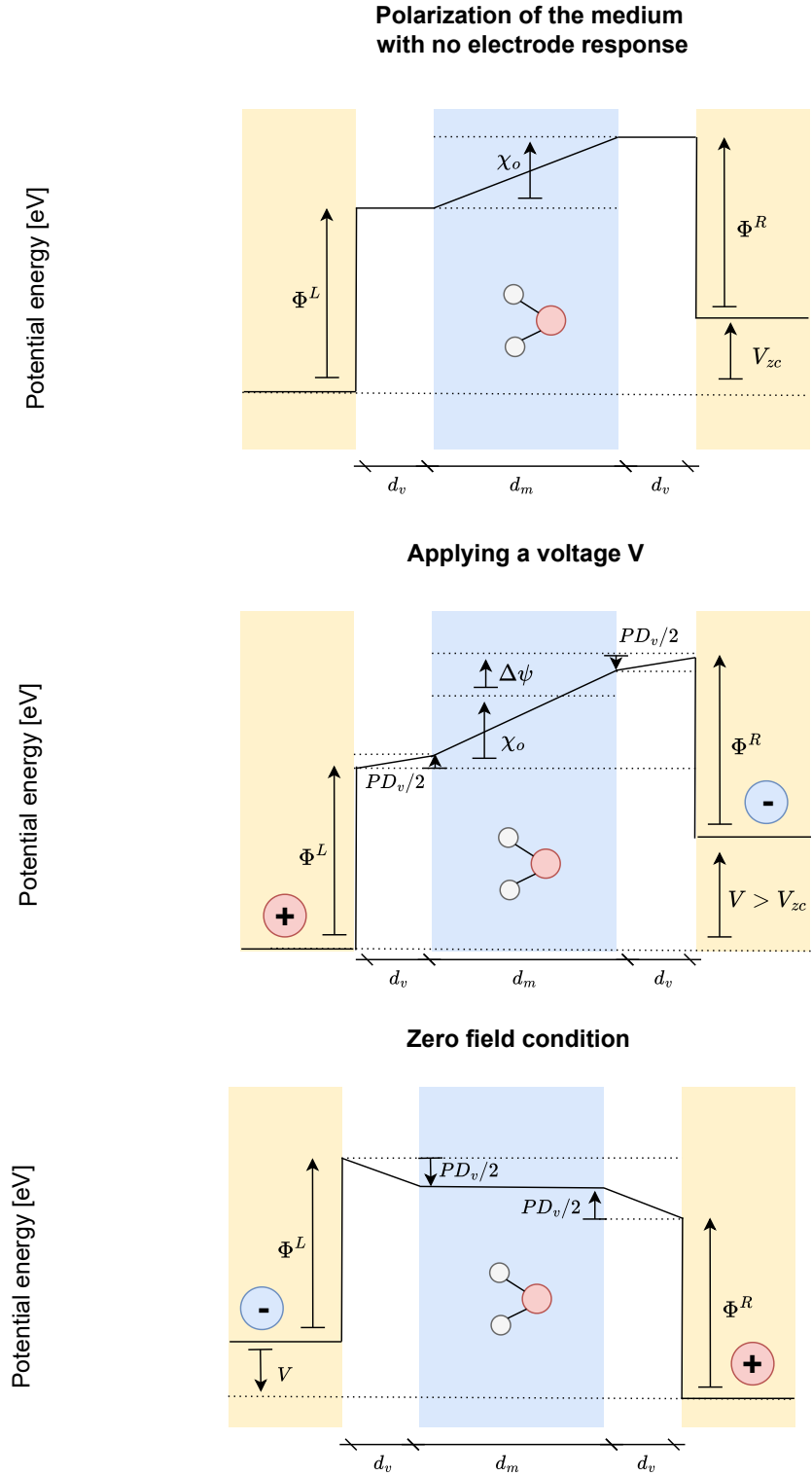$$P_m = -\frac{V}{2d_v \epsilon_m} \tag{1.39}$$

Figure 1.27: Potential profiles for the rigid water system with an interface vacuum gap between the electrodes and the medium.

which does not depend on how big the capacitor is, just on the width of the interface gap and the optical relative permittivity of the medium.

To close this section, I would like to discuss the LDOS once more, but this time focusing only on the zero field cases and relating it with the analytic results just discussed. In Figure 1.28 I show the LDOS for the zero field conditions with V = 0 eV and V = -1 eV. We can take the V = 0 eV case as a reference, and determine that the HOMO of water is -1.82 eV below the Fermi level. If we applied V = -1 eV and allowed the system to relax by rotating its dipoles only, we would get the zero field case for V = -1 eV which is the right panel in Figure 1.28. In this situation, the HOMO of water molecules in the left side is -2.2 eV lower than the Fermi level of the left electrode and at the right electrode the difference is -1.2 eV. There is therefore a relative shift of -0.38 eV and 0.62 eV respectively. The positive sign of the right side is because the voltage drop is from the molecule to the electrode (opposite as in the left side). The asymmetry in the magnitude of the shifts is expected: although in the previous discussion we have assumed both interface gaps to be of equal size, the view of the system in Figure 1.22 shows that the interface at the left side is smaller. Since the field created by the charged capacitor plates is constant, smaller vacuum regions will have a smaller potential drop, and therefore molecules on that side will experience a smaller shift. In any case, the interest of the shifts lies in the fact that when the energy levels of the molecules change their alignment with the energy levels of the electrodes different chemistry might happen. With the shift, different orbitals might interact. Or even better, molecular levels might cross the Fermi level of the electrodes. And it is at that point that the magic of electrochemistry will happen. All because of a minuscule interface gap.
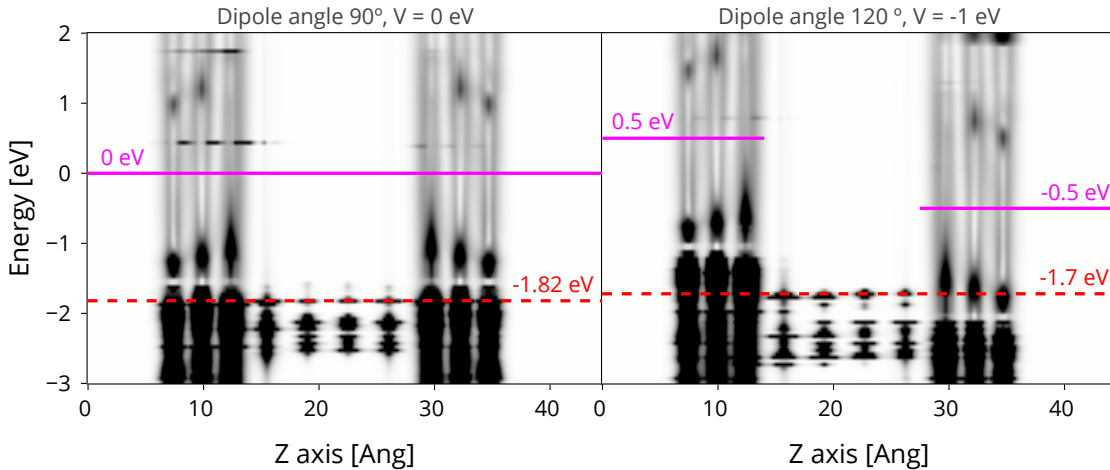

Figure 1.28: Close up comparison of the LDOS of a reference system ($\alpha = 90°, V = 0eV$) and the V = -1 eV system that satisfies the 0 field condition. Magenta solid lines indicate the Fermi levels, while dashed red lines indicate the approximate position of the HOMO of the water molecules.

1.2.5 Getting things moving.

Until this point, I have only discussed static systems, i.e. systems where atoms don't move. With rotating rigid water, I tried to get a grasp of how a dynamic system might behave, but it was still all static situations. The next calculations that I discuss will already be molecular dynamics of liquid water. This section serves as an interlude, and tries to emphasize on the main problematics that moving atoms will bring without discussing any system in particular.

First, let me state that completely uniform dipoles as discussed with rigid water do not exist in the liquid phase. However, the excess charge response of the electrodes is not dependent on the distribution of the dipole. When determining the extra charge on the surface of the electrode, the only thing that matter is the total net dipole between the electrodes. In Figure 1.29 I illustrate this concept by showing a uniformly polarized system like rigid water and a system that has a net polarization because of asymmetric surfaces, without any bulk polarization. In both cases, the potential profile between the electrodes is not important to determine the excess charge, and the only thing that matters is χ_o. Of

course, the potential profile will be an important factor to consider for other things. A charged particle traveling through the medium, for example, will certainly care about what is the potential profile.
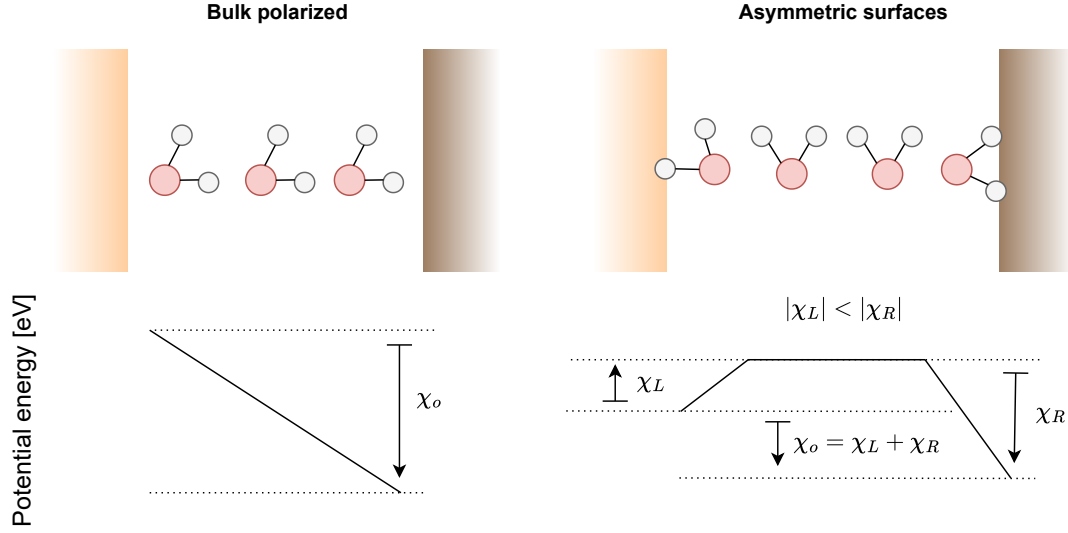

Figure 1.29: Two different examples of a polarized medium. To the left, a bulk polarization with all the molecules oriented in the same direction. To the right, a bulk with zero net polarization but with asymmetric interfaces that result in a net polarization.

The surfaces depicted in Figure 1.29 are also useful to discuss another point. It is easy to imagine that water molecules interacting with a surface might have preferential interaction sites and orientations. Their stability will therefore not be dictated simply by the vanishing of the mean field. The minimum energy of these molecules will be at a compromise between aligning with the field and interacting in favourable ways with the surface. In bulk liquid water this is also the case, there is some contribution to the energy that aims to keep the structure of the liquid, i.e. have as many hydrogen bonds as possible. The steady state scenario (I would say "equilibrium scenario", but remember that we are dealing with non-equilibrium calculations), is probably not at 0 mean field. In the previous discussion, introducing in the system any part that doesn't need 0 mean field to be stable would have deviated us from the horrific scenario that $V \neq 0$ drives the system into complete saturation. I just chose to use vacuum because, again, it is easy to deal with. However, the reader must keep in mind that the vacuum interface gap is the only reason that two entities that are apparently in close contact can shift their energy levels when voltage is applied. Otherwise, the energy levels of both sides of the interface would be pinned and would not experience any relative shift.

Finally, I'd like to take into account that, whatever the stable situation is for a given system, if there is temperature there will be random motion. This motion will temporarily get the system out of the most stable situation just to get back some time later in time. In practice, there will be oscillations around the stable situation, with the frequency and the amplitude of these oscillations depending on many factors. The reader must be confused as to why I am now explaining general thermodynamics. Well, the point I want to make is that even the interface energy shifts that I have discussed are not rigid. Using the definitions in Equation 1.36, one can get an expression for the potential in the vacuum, this time with no zero field condition:

$$PD_v = (V - P_m d_m) \frac{2 d_v \epsilon_m}{d_m + 2 d_v \epsilon_m} \tag{1.40}$$

which tells us that, at a fixed voltage, the potential difference in vacuum depends linearly on the polarization density of the medium. The derivative of the potential drop in vacuum with respect to the polarization density of the medium is:

$$\frac{\partial (PD_v)}{\partial P_m} = -\frac{d_m 2 d_v \epsilon_m}{d_m + 2 d_v \epsilon_m} \tag{1.41}$$

which for large systems ($d_m >> 2d_v\epsilon_r$) is approximately:

$$\frac{\partial(PD_v)}{\partial P_m} \approx -2d_v\epsilon_m \tag{1.42}$$

Therefore, if the polarization density oscillates, the energy shifts between the molecule and the surface will also oscillate proportionally, with the proportionallity factor depending on the width of the interface gap and the optical relative permittivity of the medium. If the system was at finite temperature, it would be important to take these oscillations into account to understand possible reactions at the interface, not just the stable value.

These equations are simple enough because the rigid water system is a simple system. In a complex environment where there there are innumerable factors that can affect the fields and potential drops, analytical solutions might be too cumbersome and it is necessary to resort to numerical methods. It is very important to trust the numerical methods that we use, because once we go into the regime where analytically solving everything is not feasible, we will have no other option. As this section ends, I hope to have provided enough proof that the numerical method in TranSIESTA reproduces the analytical results well enough so that the community can advance into progressively more complex systems with confidence. Not only that, but I hope I have provided enough insight into what is important to look at in TranSIESTA simulations to make relevant findings in the field of electrochemistry.

1.3 Water molecular dynamics

As we depart from the playground I designed to understand NEGF/DFT and move into "real life" simulations, the reader must be prepared to give up on understanding everything that is going on in the simulations in an exact way. The atomic jiggling caused by temperature will blur everything and sometimes it will be hard to access the small details of the system. This is not to say that we should not aim at perfectly understanding the main features of the system. Otherwise, the simulations would become useless for electrochemists. But the focus should now be shifted towards statistical averages (and deviations) instead of specific snapshots of the medium. This will allow us to paint the bigger picture and avoid getting lost in small blurry details. By doing the effort of zooming out, we will get very interesting physics that bring us one step closer to pushing the boundaries of electrochemistry.

For the first time, not only in this thesis but in electrochemistry literature, I will show and analyze the results of running ab initio molecular dynamics with NEGF/DFT with a focus on the consequences of non-equilibrium for electrochemistry. In particular, I show TranSIESTA molecular dynamics of liquid water between gold electrodes under applied voltage. The system is shown in Figure 1.30. This systems contains a total of 264 atoms and 3080 orbitals (using a double zeta polarized basis, the standard in SIESTA). Running the MD in the now gone Marenostrum 4 supercomputing cluster at the Barcelona Supercomputing Center (BSC) using 96 CPUs, TranSIESTA simulates this system at a rate of approx 1000 MD steps per day. With a timestep of 1 fs, that is 1 ps per day. After a previous equilibration of the system with SIESTA, annealing MD at 300K were run for 3/4 ps at different voltages. By doing the math, the reader can realize that these are dynamics that took "only" 3 days to run. Certainly, I could have run longer dynamics. But since this is an initial exploration of TranSIESTA for electrochemistry, the aim is to understand whether the simulations can be interpreted in a way that is useful for electrochemistry. In that regard 3 ps is enough time, as with such small system the liquid is equilibrated and we can even see thermal oscillations. Longer dynamics are left for future electrochemists, and in other chapters I work towards the goal of accelerating NEGF/DFT simulations.
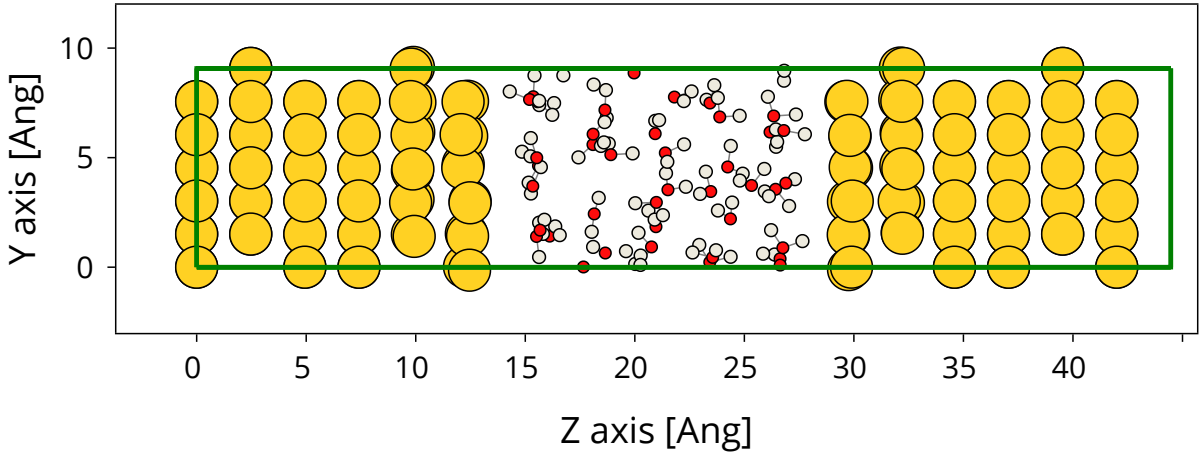

Figure 1.30: View of the liquid water system. It is the same system as for rigid water, but now filled with liquid water. The electrodes are fcc gold with a [111] surface. The space between gold electrodes is 17.3 Å and it is filled with 40 water molecules. The system is periodic in the X and Y directions, while the Z direction is connected to the electrodes. The dimensions of the unit cell (green box) are 10.48 x 9.07 x 44.47 Å.

One of the first things that one can look at in the molecular dynamics is the net electrode charge, since it is a relatively simple analysis. Figure 1.31 shows the net charge at the electrodes for MD at V = 0 eV and V = 2 eV. The electrodes get much more charged at V = 2 eV. If I use the capacitor equation (Equation 1.31) to estimate the relative permittivity of the medium using the mean net charge and $d = 15.3$ Å (2 Å less than the distance between atom centers at the surface) I get $\epsilon_r = 10.45$. We saw in Figure 1.23d that when the medium only responds by redistributing electrons $\epsilon_r \approx 1.35$. The response in this case is 10 times bigger, therefore the liquid must possess a much better mechanism to screen external fields. This mechanism is of course dipole orientation, as we will see later. Another important point to highlight is the fact that the response is not as strong as that of bulk water, which is known to have a relative permittivity of around 80. This decrease in response is a well known effect. The confinement of water reduces its dielectric response due to both the interface gap and the limited

freedom of the interfacial water molecules [11], [12]. Regarding the V = 0 eV dynamics, Figure 1.31 shows that there are charge oscillations. As I discussed by the end of the previous chapter, these charge oscillations will somehow translate into oscillations in the energy level shifts at the interface. Since the charge oscillations are equal or bigger than the charge that the capacitor would have at V = 2 eV if the medium was vacuum, it is easy to think that the magnitude of the energy level shifts at certain moments will be at least comparable. This certainly means that a highly polar medium like water at finite temperature will drive some significant activity at the interface even at 0V. Whether or not there will be a net result from this activity is a different question and will depend on the details of the system.
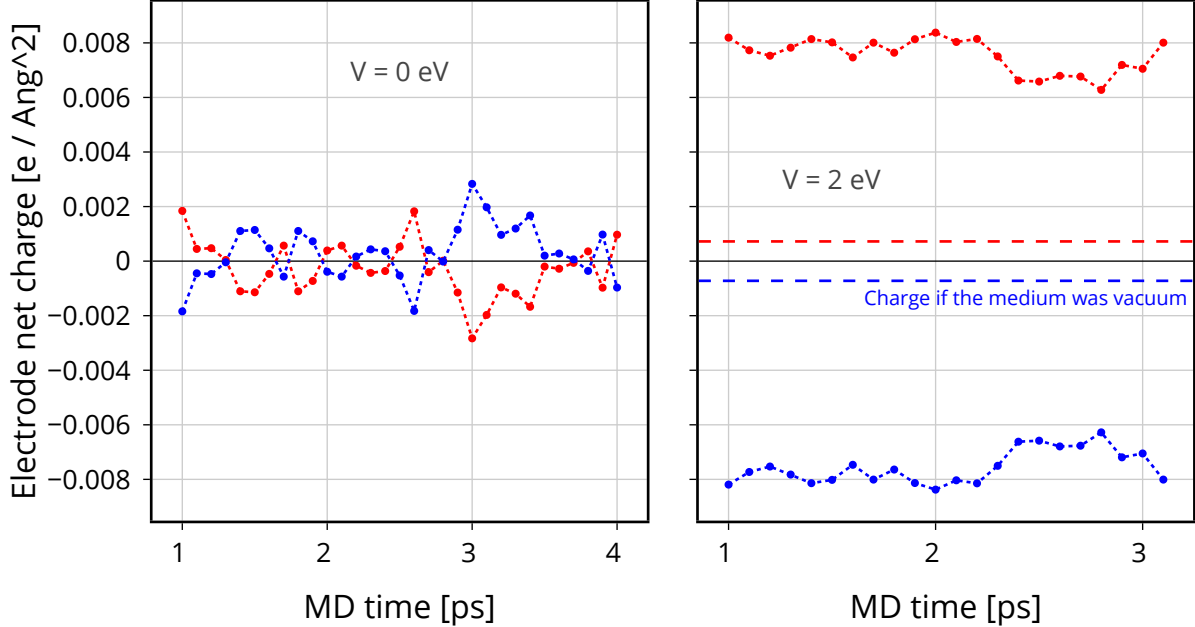

Figure 1.31: Excess net hirshfeld charge at the electrodes during TranSIESTA MD at V = 0 eV and V = 2 eV. In red, the charge at the left electrode, in blue the charge at the right electrode. Positive values mean that the electrode has lost electrons. In the V = 2 eV case, I display as a reference the charge that the capacitor would accumulate if water was removed and the medium was instead vacuum.

1.3.1   Structural properties.

Since the net charges show no clear tendency, I will assume that the MD are at a steady state which just oscillates around some stable point. It seems clear that longer dynamics would give a better representation of the statistical quantities by considering more oscillations. However, this sampling is good enough to start exploring what TranSIESTA has to offer to describe electrochemical systems. Therefore, I will now proceed to analyze the structure of water during the molecular dynamics.

The first natural question to ask regarding structure is: where is each atom located in the system? Do some atomic species prefer certain locations, or do they just distribute in a uniform way? Figure 1.32 shows the density of each species as a function of the z coordinate averaged over the V = 0 eV molecular dynamics. At the electrodes, we see how the atoms at the outermost gold layer oscillate more than those in the second layer, resulting in wider peaks. After the 3rd layer, the atoms are constrained to stay fixed during the MD so that the electrode (4th to 6th layers) stays as bulk-like as possible. Having said that, the interest of the density profiles lies mainly on the water region. In bulk water, the density of both Hydrogen and Oxygen atoms is uniform. The profiles show that there is a very small unstructured region (similar to the bulk) in the middle of the capacitor, between $z = 19$ Å and $z = 23$ Å. Closer to the surfaces, we see some layering, which is symmetric at both sides. In contact with the electrodes, 3.1 Å away, there is a very sharp Oxygen peak followed by a complete depletion. This indicates the presence of a very well defined first layer of water molecules. The peak of Hydrogen at the same position is spread almost symmetrically to both sides, showing that there is no significant polarization of the water layer. After the first layer, there is a second layer which is not as pronounced and fades into the bulk.

Having established the equilibrium atomic density, it is now time to discuss how a voltage of V = 2 eV affects the structure of water. The density profiles for V = 2 eV are shown in Figure 1.33. A positive
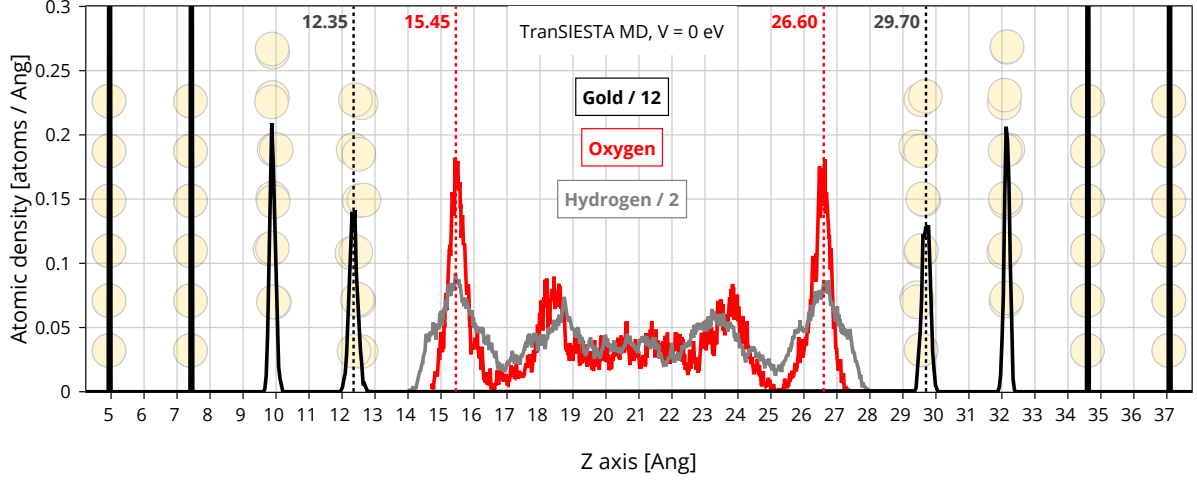
Figure 1.32: Atomic density for different species during a TranSIESTA molecular dynamics at V = 0 eV. Computed by averaging over all steps from $t = 1ps$ to $t = 4ps$.

V, as defined by our previously agreed conventions, results in a positive charge at the left electrode and a negative charge at the right electrode. It is therefore natural to expect that hydrogens, which have a partial positive charge, will be attracted to the right electrode and oxygens, which have a partial negative charge, will be attracted to the left electrode. I will show later how this unfolds in terms of angles that define molecular orientation, but in the density profiles one can already see the trend. At the left surface, the Oxygen peak of the first layer has moved 0.15 Å closer to the surface, while the symmetry of the Hydrogen peak has been broken because the H atoms that were closer to the surface have moved away from it. At the right surface, the Oxygen peak has stayed at the same position while the Hydrogen atoms have shifted towards the surface feeling attracted by the negative charge. At the second water layers, if we read the profiles from left to right, the Hydrogen density clearly shows a shift to the right with respect to the Oxygen density. Pending the analysis of the angles, all these signs point to a orientation of the water molecules to align their dipoles in order to counteract the external field. It is also notable that the bulk region seems to have more structure now. By forcing the water molecules to align their dipoles, the medium has to give up on some of its randomness, resulting in an extra density peak at the center of the water region. This extra layering is probably a consequence of the confinement and bulk water should still show a flat density profile in spite of the preferential orientation of the dipoles.

In order to further support the conclusions drawn from the density profiles, let's now discuss the molecular orientation as illustrated in Figure 1.34. I chose to describe the orientation of water molecules using two representative angles. First, the dipole angle, which is a measure of the dipole alignment with the field created by the electrodes in the z direction. As defined, the dipole angle is 0° when the Hydrogen atoms are pointing completely to the right and 180° when they are pointing to the left. Then, there is the plane angle, which measures the alignment of the water plane with the electrode surface. A plane angle of 0° means that the water molecule is completely flat on the surface, while 90° means that the water molecule is lying perpendicular to the surface. The reader must be aware that these two angles are not completely independent from each other. A molecule that is completely flat (plane angle = 0º), for example, has no way to align its dipole with the field. In general, in a geometry with a plane angle φ the dipole angle will be restricted to the range $[90 - \varphi, 90 + \varphi]$. Armed with this knowledge, we can start to analyze the angle distributions that come out of V = 0 eV and V = 2 eV MD. For a meaningful analysis, I divide the system into three regions based on the density profile plots: bulk, left interface, and right interface (Figure 1.34a).

At 0V (Figure 1.34c), the bulk region shows a uniform distribution of water orientations for both angles. A uniform distribution of configurations results in sine-shaped distributions of angles because of statistical reasons. The greater the alignment of the dipole to Z (closer to 0 or 180º), the less configurations that have that alignment. For example, there is only one configuration with the dipole perfectly parallel to Z (0º) A similar argument can be made for the plane angle. The bigger the plane
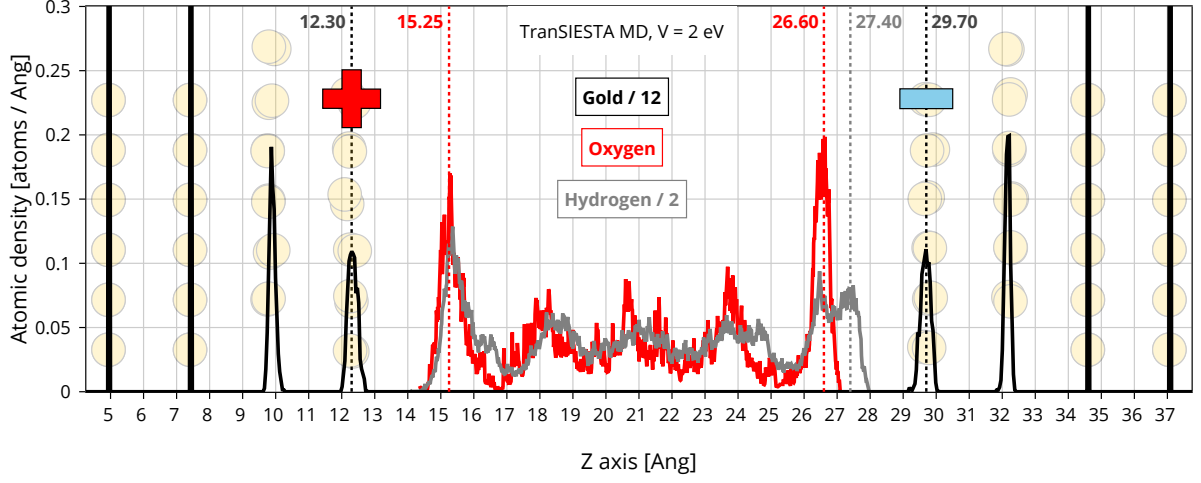
Figure 1.33: Atomic density for different species during a TranSIESTA molecular dynamics at V = 2 eV. Computed by averaging over all steps from $t = 1ps$ to $t = 3ps$. Plus and minus signs show the sign of the accumulated charge at each electrode. Positive means lack of electrons.

(a) Definition of different regions: left (red), bulk (black), right (blue).

(b) Definition of angles

(c) Distribution at V = 0 eV

(d) Distribution at V = 2 eV
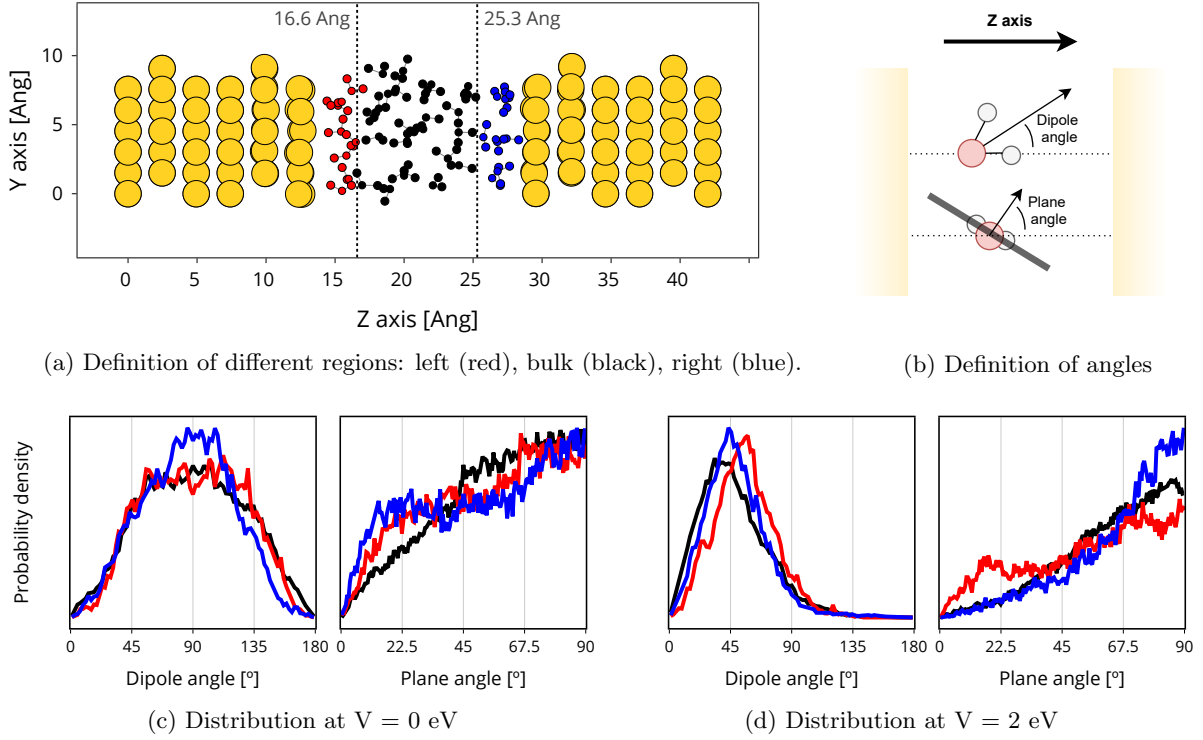
Figure 1.34: Orientation of water molecules for dynamics at V = 0 eV and V = 2 eV. The distribution of angles is shown for three different regions: bulk (black), left interface (red), right interface (blue). The definition of the regions is shown in (a). Water molecules are determined to be interfacial if their oxygen atom has a smaller z coordinate than 16.6 Å (left) or a z coordinate larger than 25.3 Å. These thresholds are the positions of minimum oxygen density after the first layer, as shown in Figure 1.32.

angle, the more configurations that have that angle. In summary, there is no preferential orientation of water molecules in the bulk at 0V, which leads to net zero dipole. The interfaces, however, clearly deviate from the uniform distribution. At the interfaces, the plane angle shows a strong preference for angles around 20º. This means that the surface likes that water molecules sit on a rather flat configuration. As

discussed before, this restricts the dipole angle to be between 70⁰ and 110⁰, which alters a little bit the dipole angle distributions.

When applying a voltage of V = 2 eV (Figure 1.34d), things change dramatically. The bulk water has no longer uniform angle distributions. Dipoles are oriented with the H atoms pointing to the right to counteract the field, and this is reflected with a distribution that now peaks at 45⁰. The plane angle distribution is also affected indirectly by this tendency. The alignment of the dipoles with the field in bulk water is fully expected, there is nothing surprising about it. The most interesting result of V = 2 eV is the fact that we can not discuss both interfaces in the same way. There is a clear asymmetry between the left and right interfaces. At the right interface, where H atoms feel attracted to the negative surface, things look very similar to the bulk. Molecules orient to align with the field. At the left interface, however, the surface doesn't seem to particularly like that the water molecule comes "Oxygen first", and still insists on the 20⁰ plane angle. This results on a limitation of the dipole alignment capabilities of the left interface and the dipole distribution is shifted towards the center with respect to the bulk and the right interface. To give the reader a quantitative measure of the differences between regions, the mean dipole angles are 47.5⁰ for the bulk, 55.7⁰ for the left interface, and 49.0⁰ for the right interface. The reader, like I did, might have the impression that this is an unequivocal sign that the regions that have less alignment with the field (angles closer to 90⁰) have a lower permittivity. This, however, is not directly true. The polarization density of a medium does not depend only on the orientation of its molecules but also on the density of molecules. The molecular density of water close to the interface is not uniform and only a more detailed analysis could determine the differences in permittivity between regions.

As a summary, regarding the structure of water in this system, I would like the reader to take home three messages:

- There is a clear difference in the behavior of the water molecules at the bulk of the liquid and at the interface.

- Interfaces are not "dead layers" that don't react to external fields, although their response might be a bit weaker than that of the bulk.

- The interfaces with a positively and a negative charged surface have different response to external fields.

1.3.2 Energy levels.

For an electrochemist, reactions are the most important thing. If I just look at the structure of water, I can't say I am participating in the field of electrochemistry. Although the structure of water at the interface might catalyze reactions, reactions are triggered by applied voltage mainly because the applied voltage shifts the energy levels at the interface. In the simple examples of the vacuum capacitor and the rigid water capacitor that I presented in the previous section, it was possible to estimate what the energy level shifts would be based on pure electrostatics. With liquid water, however things are not as simple. I could fill pages and pages trying to determine the exact structure of water and relating it to the energy levels. However, as I warned at the beginning of this section, we should not get lost in details. The point of using numerical methods like TranSIESTA is that we know that they implement the laws of physics correctly, so we don't have to worry about double checking every single detail. I already did enough checks on the previous section with the simpler systems, and now we can simply enjoy the benefits of computation. Instead of trying to predict the energy level shifts, we can probe them directly in the TranSIESTA simulations. Then, of course, we should try to understand whether they make sense.

To understand energy level shifts, I will look at the density of states of the system. I already showed plots of the local density of states (LDOS), which allow us to visualize how the energies of the states change across the capacitor. Now I will also introduce the projected density of states (PDOS) analysis, which allows us to project the states on atomic orbitals so that we have a clearer understanding of which atoms are contributing to the states in regions of space where there are multiple contributions. By combining both, I get a plot like the one in Figure 1.35, which shows the density of states averaged over the MD at 0V. There, the LDOS looks quite flat on the water regions indicating that there is no average electric field in the bulk of water. It is imperative for me to remind the reader that this is an average, we have already seen in Figure 1.31 that the fluctuations are quite big. At the interfaces, there seems to be a slight slope that decreases the potential when approaching the surface. This is likely to be related to the fact that there is a small preference for H atoms to be a bit closer to the surface on average than

O atoms, as we have seen on the angles discussion. But this might as well be just a statistical deviation from the average. Apart from that, there is nothing much interesting in the density of states at 0V. However, I will ask the reader to keep in mind the position of the first peak of the interfacial water in the PDOS at -3.4 eV. That is, 3.4 eV lower than the Fermi level of the electrodes.
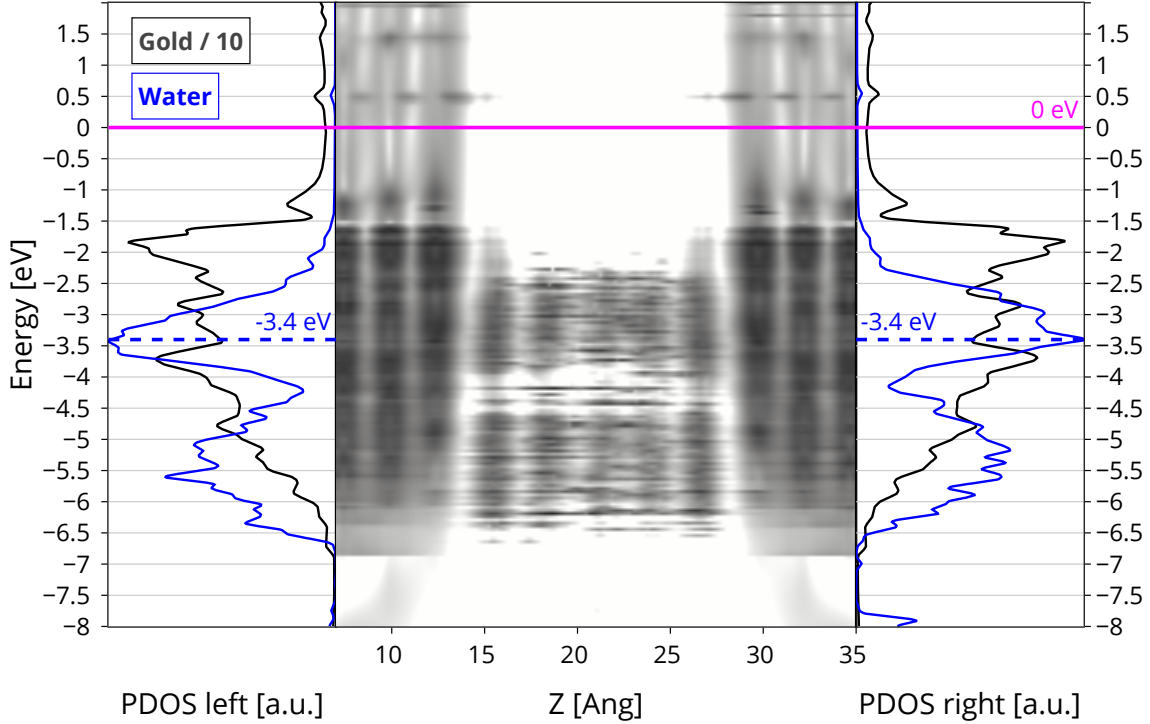

Figure 1.35: Density of states at V = 0 eV averaged over 30 MD steps uniformly distributed over 3 of dynamics (from $t = 1ps$ to $t = 4ps$). At the center, the LDOS across the capacitor is shown. The PDOS of the gold electrode (black) and the interfacial water (blue) is shown at each side of the LDOS.

Before I move on to see the results of the density of states in the V = 2 eV run, allow me to discuss how do we expect the HOMO of water to be shifted at the interface. Figure 1.36 depicts the energy levels at 0V, and two extreme cases in which the energy levels could behave when applying voltage. When applying V = 2 eV, the levels of the left electrode will shift rigidly by -1 eV and those at the right electrode will shift by +1 eV. If the potential falls only inside water, there is no potential drop at the interface. This results in the interfacial states of water being completely pinned to those of gold, and therefore interfacial water would suffer the same rigid shift as the electrodes. This situation would be catastrofic for electrochemists, since nothing interesting is happening with the application of voltage. On a completely opposite situation, the potential drop inside water could be zero if water is able to perfectly screen the field. Then, the potential would drop only at the interfaces, resulting in a complete decoupling between the states of the electrode and the states of interfacial water. This is a much more interesting situation for electrochemists, since water at the left interface has now its HOMO 1 eV closer to the Fermi level of its electrode, while at the right side water levels have moved 1 eV away from the Fermi level. The relative shifts can induce new interesting chemistry. Let me also point out that in the case of an asymmetric capacitor, these shifts could be triggered by a difference in work functions, which can be seen as an "effective voltage in a symmetric capacitor". That is, the effect of applying a voltage V on an asymmetric capacitor is $V_{eff} = \Delta\psi = V + \Delta\Phi$.

Which one of these situations will we encounter when I analyze the density of states for the V = 2 eV case? In principle, one could think that water molecules would choose to reorient in a way that completely neutralizes the field inside the liquid. However, Figure 1.37 shows that this is not the case at all. While it is true that molecules lower their energy by aligning with the field, there is also a structural contribution to the energy. If water molecules align to the field, the network of hydrogen bonds weakens, and this is not energetically favourable . This conflict of interests is the reason why the permittivity of water is not infinite. In the end, liquid water must find a compromise between the two contributions and this is what we see in the LDOS. The potential drop between electrodes is 2 eV, and the potential drop inside the liquid is approximately 1.15 eV, as indicated by the red labels at the bottom of the LDOS

Complete pinning

Potential drops only inside water

V = 0 eV

0 eV   0 eV

-3.4 eV   -3.4 eV

V = 2 eV

1 eV

-1 eV   -2.4 eV

-4.4 eV

Complete decoupling

Potential drops only at the interface gaps

V = 0 eV

0 eV   0 eV

-3.4 eV   -3.4 eV

V = 2 eV

1 eV
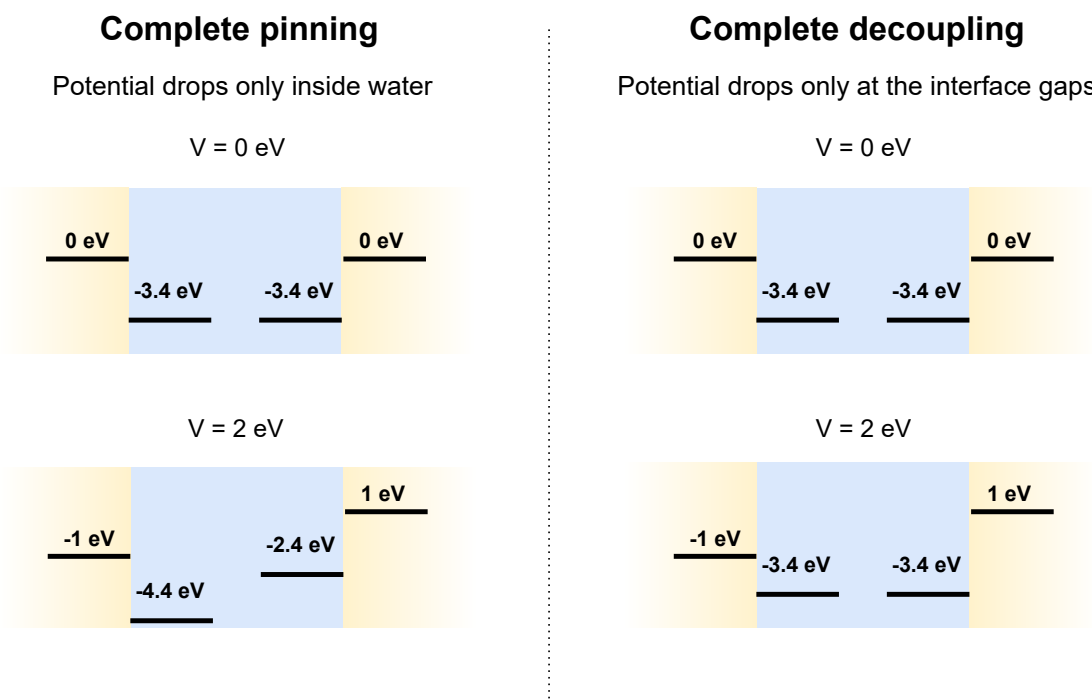
-1 eV   -3.4 eV

-3.4 eV

Figure 1.36: Sketch of the two extreme cases of energy level shifting at the interface. If voltage doesn't induce a potential drop at the interface, the energy levels at both sides just rigidly shift. This can happen for example if there is no interface gap or the medium is happy with the potential drop that the voltage induces (i.e. the applied voltage is a point of zero charge). If there is no net potential drop inside the medium, the energy levels of the medium just by the interface will stay unmoved while electrode levels shift. This can happen if there is zero field inside the medium (the medium can perfectly screen the external field caused by the electrodes) or somehow all the potential drops inside the medium compensate to have a net zero drop.

image, where the drop is easier to quantify because there are no gold states. This difference means that the liquid has screened some field but not all of it. The rest of the potential drop (0.85 eV) should be located at the interface gaps. At the left interface, the first water peak is now -2.7 eV lower than the Fermi level, which represents a relative shift of +0.7 eV of the interfacial water with respect to gold compared to the 0V case. At the right interface, the first water peak is now -3.8 eV lower than the Fermi level, representing a shift of -0.4 eV of water with respect to the electrode. In total, the potential drops at the interfaces sum up to 1.1 eV. If I add this to the potential drop that I determined inside the liquid, it results in a total of 2.25 eV, which is not the expected 2 eV drop between electrodes. By now, the reader will have already realised that it is not easy to quantify things in liquid water. The PDOS peaks are not sharp and their position is not well defined. The LDOS is noisy. But noise is far from the only problem in quantifying things. What if the work functions of the gold surfaces have changed in their interaction with water? What if interfacial water energy levels are moving because there is different chemistry with the surface? What if there is charge transfer between gold and water molecules? There are many factors that play their part in the energy levels shift. The computational electrochemist must, as I am now, be satisfied with being able to roughly explain what is happening in terms of simple arguments. For the exact details of why levels shift as they do, the scientist must carry a deeper exploration of TranSIESTA's results.

In any case, I hope that this discussion has set the ground for future practitioners to test situations in which interesting reactions are expected. A path that I find particularly interesting is to treat the analysis that I did here as a characterization of the medium where reactions could happen. That is, with this system, we now know that the expected shifts at the left and right interfaces are around +0.7 eV and -0.4 eV. Is there something that I can add to the system that should produce interesting chemistry taking this shifts into account? Of course, to make this approach valid, the addition of the extra entity

should have a negligible effect on the potential profiles. This can only happen if the new added entity is at low concentrations, which is only possible if the system is big. This is certainly not the case for the system that I presented here. Although the scaling of TranSIESTA is linear with the number of atoms, it can become quite expensive for big systems. Therefore, the focus is now put into ways of accelerating the simulations while keeping them as reliable as possible, so that we can reach bigger systems. The following chapters are the results of my attempt to contribute to this goal.
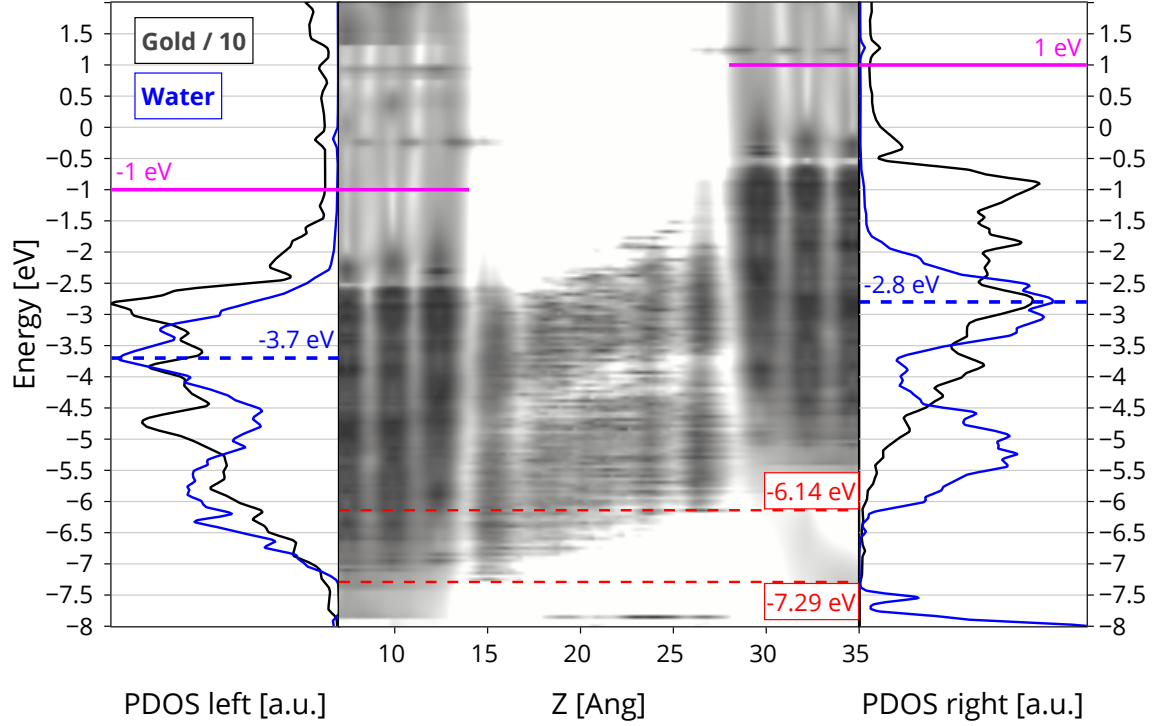

Figure 1.37: Density of states at V = 2 eV averaged over 20 MD steps uniformly distributed over 2 ps of dynamics (from $t = 1ps$ to $t = 3ps$). At the center, the LDOS across the capacitor is shown. The PDOS of the gold electrode (black) and the interfacial water (blue) is shown at each side of the LDOS.

1.4   Conclusions

Finding how electrons distribute in an atomic system is not a trivial task. If they are important for the physics of a problem, you have to know where they are and what they do. Of course, we can always create models that approximately simulate reality and give results as fast as possible. I hope that with this chapter I have conveyed the idea that describing electrons at the level of DFT is paramount to simulating electrified solid-liquid interfaces. Creating simple models for the effects of applied voltage or work function differences is very hard and the models are bound to work only for a particular interface. For this reason, having numerical methods that are able to simulate electrons at the DFT level in out-of-equilibrium situations is key to set computational electrochemists free on the kind of systems that they can simulate. Methods that combine NEGF with DFT, and in particular TranSIESTA, have long been used to successfully simulate out-of-equilibrium situations for electronic transport applications. But these methods have still not found a place in the electrochemistry community.

Throughout this chapter I have tried to, step by step, build the confidence of the reader in TranSI-ESTA. First, I have taken vacuum capacitors and succesfully compared the analytical solutions with the results of TranSIESTA. Then, I have introduced water in a controlled way in a rigid water system and still managed to analitically describe the system in a way that matches TranSIESTA results. Once there was enough confidence in TranSIESTA, I have introduced for the first time molecular dynamics of liquid water under applied voltage using NEGF/DFT. In this system, I must admit, I have found moderate success in describing things analitically. And it is because of this incomplete success that I know I have reached the point when things get interesting: the numerical tool, i.e. TranSIESTA, has become useful!

Using the TranSIESTA liquid water molecular dynamics, I have provided a guide for future practitioners to analyze and interpret this kind of simulations. In particular, NEGF/DFT describes electronic structure in a way that is consistent with the fact that the electrodes don't share a Fermi Level when they are out of equilibrium. This allows us to analyze energy level shifts between the states of the electrode and the molecular states at the interface, which is crucial to understand the electrochemistry of the system. With this analysis, TranSIESTA is positioned as a great method to simulate full electrochemical cells. Although TranSIESTA allows us to run MD of 264 atoms with 3080 orbitals (DZP basis set) at a rate of 1000 steps/day, which is pretty good for DFT standards as of year 2024, it is still not enough to simulate the kind of system size that would represent a full electrochemical cell. The main problem is that unrealistically high concentration of dissolved molecules/ions might drastically affect the global response of the system to applied voltage leading to conclusions that are not transferable to experiments. Therefore, we need to find ways to accelerate the simulations while keeping the essence (i.e. the reliability) of TranSIESTA as much as possible. This need is what lead me into exploring the machine learning methods that are described in the following two chapters.

Summary of conclusions:

- TranSIESTA is a method within the SIESTA open source code that combines Non-equilibrium Green's Functions (NEGF) with DFT to compute electronic structure in out of equilibrium systems in a self-consistent way.

- TranSIESTA can be used to run molecular dynamics under applied voltage.

- TranSIESTA is a great tool for electrochemistry, but it is not yet adopted by the community.

- The speed of TranSIESTA is the speed of DFT codes based on atomic orbitals. Further acceleration is needed to simulate bigger systems, which leads us to the machine learning methods in the next chapters.

Summary of original contributions from the author:

- Exploration of TranSIESTA in simple systems to build confidence in the method.

- Exploration of molecular dynamics of liquid water under applied voltage using TranSIESTA. The original idea of running them was not from the author.

- Development of a visualization module in the python package sisl [13] to visualize general DFT output, along with its tutorials and documentation. The visualization module was used to generate the plots in this chapter.

- Implementation of non-trivial analysis algorithms within the sisl [13] package. For example, a more efficient calculator of electronic density from the density matrix to make the LDOS(E) plots in this chapter possible (now they take around 3 minutes to compute, with the old algorithm it took one day).

Chapter 2

Accelerating DFT simulations by machine learning the density matrix

I closed the previous chapter by highlighting the need for accelerating TranSIESTA calculations. In general, there are several approaches that a computational physicist/chemist can take to speed up atomistic calculations:

1. Wait for the hardware to improve.

2. Optimize the algorithms that the code uses so that they run faster in the existing hardware.

3. Design new algorithms that do the same calculations faster.

4. Find a parametrization that approximates the results of the ab initio calculations but is much faster to compute.

Traditionally, the first approach has proven to be the best in the long term. However, we scientists feel the need to try to tackle the problematics that society faces as soon as we possibly can. For that reason, the developers of DFT codes are constantly working on the second and third approaches. DFT users daily benefit from it by running calculations with that extra 50 atoms that the optimizations allow. What about approach number four, you might ask? Can we use it to get further? Traditionally, the answer to that question was based on whether you cared about explicitly describing electrons or not. If you were not interested in electrons, you could try to handcraft a set of equations based on physics (usually classical dynamics) that, after finding the optimal parameters, would reproduce the atomic forces of the system without having to deal with electrons. These equations are known as "force fields". Although not trivial, force fields were kind of intuitive for humans to design as the atoms in general follow dynamics that are fairly simple to understand. Therefore, force fields have been a widely used tool in computational chemistry for decades. Parametrizing the description of electrons is however a much harder task, so scientists interested in electrons have mainly sticked to approaches one, two, and three.

With the advent of deep learning in recent years, the paradigm shift for force fields was not a conceptual one but just a matter of scalability. The main question in deep learning is: can we make our function depend on so many parameters that we can't even begin to grasp its complexity, but still be able to optimize it to produce the right results? And the answer is yes, as the fact that this sentence has been autocompleted by a deep learning model proves. With this approach, not only can we tackle the parametrization of problems that can be modelled with "simple" analytic expressions , but also those that are too hard for us to manually describe. Knowing this, it was just a matter of time that force fields would evolve into the beasts that they are today. Nowadays, deep learning force fields like nEquip [14] and MACE [15] can very easily reduce the behavior of electrons in a given system to a generic force field function by optimizing hundreds of thousands or millions of parameters with very little human effort. They just need data examples. Recently, a single force field that can be used for a range of systems that it was never optimized for has been created [16] purely by giving enough data to the generic function.

In this scenario, a computational electrochemist might ask the seemingly bizarre question: Am I really interested in electrons? Or do I just want to see how molecules react and produce new molecules? That is certainly a tough question to answer. If there was a model that perfectly predicted any relevant reaction without explicilty using the equations that we know electrons follow, then probably we would not care about electrons. But for now, we are not there yet. Even when we get there, we will probably

need methods that provide the data for force fields to parametrize. Therefore, there is still some interest in making ab initio approaches like TranSIESTA as fast as possible.

Seeing the success of deep learning in creating force fields, it is natural to wonder if deep learning would also help in accelerating ab initio calculations. As a member of the group that develops the SIESTA DFT code, it became an objective of mine to explore this possibility. By accelerating everyone's calculations I would, of course, also accelerate the calculations that I am interested in. But in the meantime, I would be contributing to a much bigger cause. My idea was to find a way of accelerating SIESTA without adding any new uncertainty to it, so that users wouldn't have to trust the deep learning method. In fact, I wanted that users didn't even have to know about it, just as any other software optimization. It was then that I started the journey of predicting electron densities with deep learning, so that I could feed them into the DFT engine to accelerate its convergence. In this chapter I will describe the methodology that I, together with Peter B. Jørgensen, then at Denmark Technical University (DTU), developed to achieve this goal. We took the hard route by predicting the density matrix instead of the electron density in real space, but the results that I will display show that it was worth it.

2.1 Graph2Mat: A universal graph to matrix converter

In this section, I focus on explaining the methodology that Peter B. Jørgensen and I developed to accelerate DFT calculations by predicting the density matrix with deep learning. The piece of code that we developed has been named Graph2Mat and it is open sourced at https://github.com/BIG-MAP/graph2mat. Its main functionality is to convert graphs to matrices, and throughout this section I will explain from scratch how this contributes to making DFT calculations faster. The section does not show any results, as this is reserved for the next section.

2.1.1 Why learning matrices?

I will start by answering the most important question: why did we become interested in predicting matrices in the first place? It all starts with the self-consistency cycle of DFT, which is depicted in Figure 2.1. In DFT, you have to solve the Kohn-Sham equations to get the electronic structure. The problem is tipically tackled as the construction of a Hamiltonian matrix (H) which is then diagonalized to get the Kohn-Sham wavefunctions (ψ) and their energies (E). The Hamiltonian, however, has some terms that depend on the electronic density (ρ) like electrostatic interactions of electrons and the exchange-correlation functionals. And the electronic density can only be computed once the wavefunctions and their energies are known. This fact creates an apparently infinite loop. In practice, this is solved by starting from an initial guess of the electron density and iterating through the loop until self-consistency is reached, i.e. both the Hamiltonian and electron density do not change from the previous iteration (according to some threshold). This loop is known as the self-consistency cycle of DFT and it represents the majority of the computational cost of DFT calculations. If nothing is known about the system that is being simulated, the safest bet is to start from atomic densities. That is, the electron density that each atom would have if it was isolated. Clearly, this is almost always a very bad guess, but it is one that leads into stable convergence most of the time. When running molecular dynamics, the electron density is usually taken from the previous step, i.e. the density converged for the previous atomic positions is used. The previous step density gives a much better guess and leads to a great reduction in the number of iterations needed to reach self-consistency. Extrapolation techniques that use the density of previous N steps ranging from simple linear extrapolation to much more involved methods have shown to reduce even further the number of iterations needed [17]–[19]. When starting a simulation from scratch, however, there is no general methodology that can be used to produce a good initial guess.
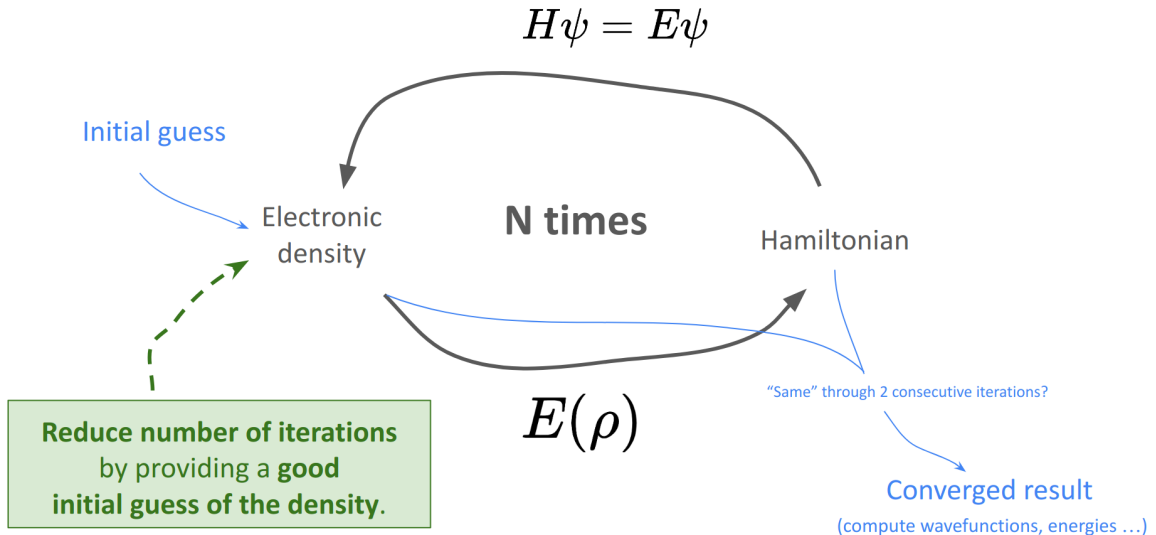

Figure 2.1: The self-consistency cycle of DFT calculations. It starts from an initial guess of the electron density and iterates until the Hamiltonian and electron density converge. The prediction of a good initial guess can reduce the number of iterations needed to reach self-consistency.

As machine learning grew in popularity in the field of materials science, it first stroke the force fields community. Soon enough, though, a zoo of methods to predict electron densities emerged [20]–[26]. By the time we started working on the project, however, no one had shown the usage of predicted densities to accelerate DFT self-consistency. We therefore set out to explore the idea. Initially, the plan was to use the methods that Peter already had developed together with Arghya Bhowmik to predict electron densities in real space, called DeepDFT [22]. However, when discussing the project, I realized that there was a hidden burden in using real space densities: to train the models you need lots of data. Not particularly in terms of the number of examples needed to train the models, but the electron density for each example is a huge 3D grid that occupies a lot of memory. It was then that we started discussing about the density matrix $\rho_{\mu\nu}$. The density matrix is formally an operator representing the projection on the occupied subspace, but in practice can be seen as a decomposition of the electron density into contributions from pairs of basis functions (ϕ_μ, ϕ_ν):

$$\rho(\vec{r}) = \sum_\mu \sum_\nu \rho_{\mu\nu} \phi_\mu(\vec{r}) \phi_\nu(\vec{r}) \tag{2.1}$$

In other words, for machine learning purposes, we can interpret $\rho_{\mu\nu}$ as a matrix of coefficients that will generate the electron density when multiplied by the basis functions. The main point of this realization is that the density matrix is a much more compact encoding of the electron density. As depicted in Figure 2.2 for an example of a water molecule with a minimal basis, there can be several orders of magnitude of difference in the amount of data required to represent the density. And even though the electron density in real space requires a lot of memory, it still represents a discretized approximation that depends on the real space grid fineness. In contrast, the density matrix contains the exact density hat is computed with the support given by the basis function orbitals used in a code like SIESTA. When we started, there was no published method that successfully predicts density matrices for a wide chemical space. At the time of writing, after 2 years (which in the ML field is a long time), such a method has not been published yet. There are, however, methods that work good for predicting density matrices in very reduced chemical spaces, e.g. a single molecule in different configurations [27], [28]. In any case, armed with the technical knowledge that I had regarding the inner workings of SIESTA, we decided to embark on the journey of predicting density matrices for large chemical spaces with deep learning. The next natural question was: how do we do it?
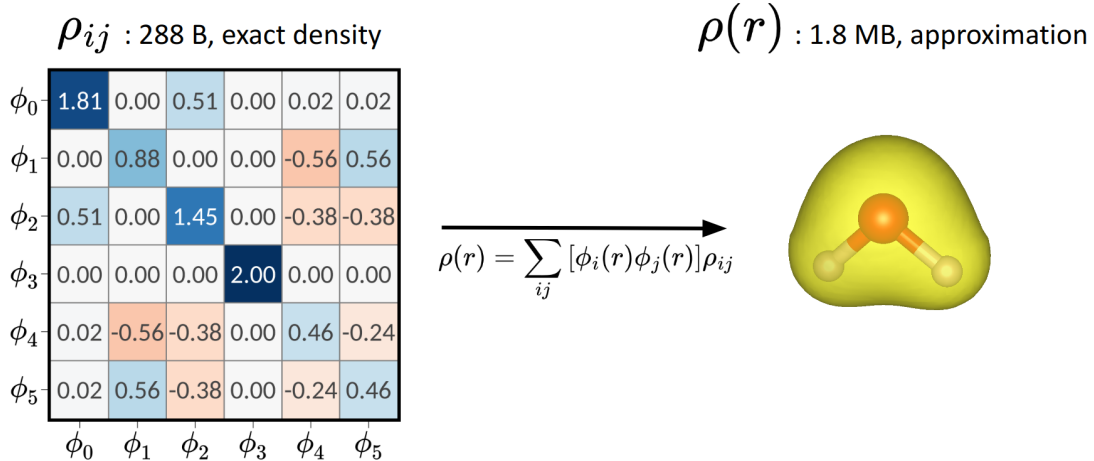

Figure 2.2: Comparison of the density matrix ($\rho_{\mu\nu}$) with the real space density $\rho(\vec{r})$. The depicted density matrix is computed with the minimal basis set (4 orbitals for Oxygen and 1 for each Hydrogen), while the real space density is computed in a very small 7 x 7 x 9 Å unit cell and grid voxels of 0.1 x 0.1 x 0.1 Å.

2.1.2 Atomic descriptors

Machine learning models use known features of the system to predict unknown properties. The most simple features that can be used as input for the models are atomic coordinates and atomic numbers, just what you would input to a DFT code. However, machine learning in general has historically shown to be very sensible to the way that the input data is represented, and atomistic ML is no exception. Instead of the raw system features, it has proved very useful to precompute some measure of the chemical environment for each atom based on its neighbors [29]–[32]. The idea is that the more similar environments two atoms have, the more similar their contribution will be to the target property. The field of atomistic ML has been, for a while, evolving these representations to ensure they are smooth and satisfy symmetries like permutation invariance or rotational equivariance. In parallel, the community has developed the idea that target properties, even if they are global, can be predicted as the sum of atomic contributions. This is a useful idea because it doesn't make ML architectures dependent on the size of the system. A typical workflow then is to compute a high dimensional representation for each atom and then convert this representation to the target property, as depicted in Figure 2.3. As the initial representation is just a descriptor, it can be reused for different target properties. Sometimes environment representations are hard coded and sometimes they have learnable parameters that can be optimized for the particular target property during training.
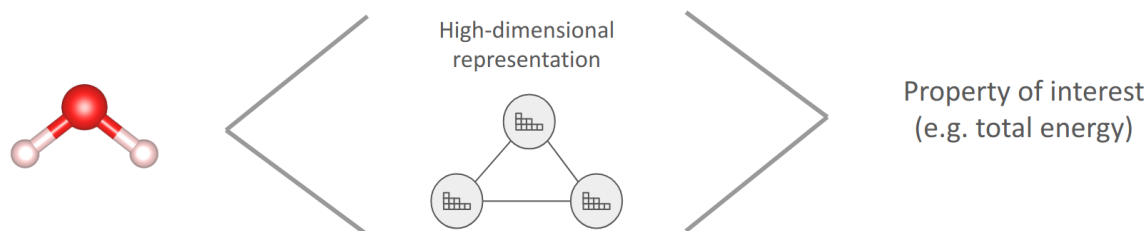

Figure 2.3: Overview of a typical ML workflow to learn properties on atomistic systems. The raw coordinates and atomic numbers are transformed in some way into more complex measures of the atomic environment. This leads into having a high dimensional representation of each atom. After that, this representation is converted (often by reducing dimensionality) to the target property. The target property might be an atom-wise property or a global property, in which case the atomic contributions are tipically summed.

The idea of our methodology is to start from the point of already having a good atomic representation. At this point, all we need to do is to convert the representation to a matrix. For that, we build a graph and we interpret the atomic representation as the node features of the graph. We then use typical Graph Neural Networks (GNNs) techniques to process the graph and convert the node features to matrix blocks.

2.1.3 Anatomy of a density matrix

To understand how a graph can be converted to a matrix, it is important to understand the properties of the kind of matrix we are interested in producing. We are interested in learning the density matrix that is computed by DFT codes based on atomic orbitals, such as SIESTA. This matrix has two key properties that we need to take into account when designing the architecture of the model:

- Although its elements correspond to pairs of atomic orbitals, the matrix can be decomposed into atomic blocks.

- The density matrix is equivariant with respect to the rotation of the atomic coordinates.

These properties are shared with other matrices that are relevant in materials science, the most notorious of them being the Hamiltonian matrix. Our architecture is not specific to the density matrix, and it therefore will be able to produce other matrices that share these properties.

2.1.3.1 Decomposition in atomic blocks

Figure 2.4 shows the shape of a density matrix for a water molecule with a minimal basis set. Black lines delimit the blocks that contain the interactions between all orbitals of a given pair of atoms. This is what will be called atomic blocks. We will refer to them as M_{ij}, being i and j the indices of the atoms that interact. By defining this decomposition, we adhere to the idea of building the target as a combination of atomic contributions, and therefore our model architecture will be independent of the size of the system. Another key feature of the decomposition is that to construct the density in real space as indicated in Equation 2.1 we only need the elements of the density matrix for basis functions that overlap in some point of space. Therefore, we only need edge blocks M_{ij} for atoms that are close enough to interact. In other words, we want to predict a sparse matrix. Graph neural networks are a natural choice for this task. The node n_i of the graph should be trained to produce M_{ii} and the output of an edge $n_i \rightarrow n_j$ (e_{ij}) should be trained to be M_{ij}. In this last sentence, I have divided atomic blocks based on whether they come from interactions within the atom (M_{ii}, node blocks) or between two neighboring atoms (M_{ij}, edge blocks). This is not only a distinction based on the graph, but they also have different properties. If the matrix is symmetric, node blocks are symmetric while edge blocks don't have that restriction, but have to satisfy $M_{ij} = M_{ji}^T$. There is another distinction to be made in practical terms for the architecture of the network. Different species might have a different number of orbitals, and therefore the size of the blocks will vary. We will see how we deal with this fact when discussing the architecture of the model in section 2.1.4.
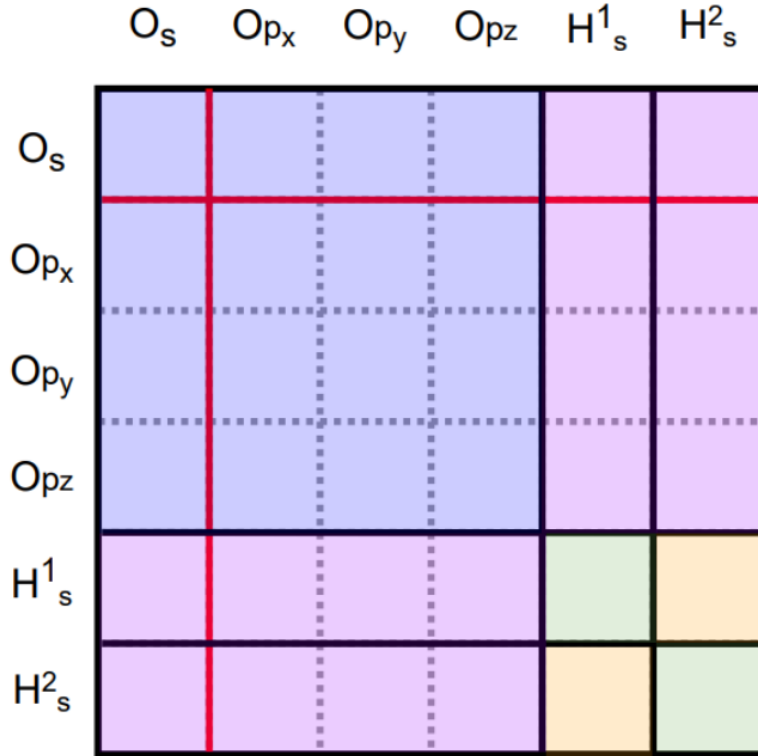

Figure 2.4: Shape of a density matrix for a water molecule with a minimal basis set (one s orbital and one set of p orbitals for Oxygen, and one s orbital for each Hydrogen). Red lines delimit blocks that correspond to interactions between a pair of sets of orbitals (e.g. a set of p orbitals which includes 3 basis functions with a set of s orbitals that includes only 1 basis function). Black lines delimit blocks that correspond to interactions between a pair of atoms, which contain the interactions between all of their orbitals. Colors indicate the different types of atomic blocks, which come from interactions between different atoms.

2.1.3.2 Rotational equivariance.

Rotational equivariance is currently a hot topic in atomistic ML. Physical properties are usually equivariant with respect to atomic coordinates, unless something like an external electric field or spin polarization breaks the symmetry. This means that when the system is rotated the property rotates in the same way. It is the case for the density matrix, as I will show briefly. Rotational equivariance is therefore an undeniable fact that our target data satisfies. The debate nowadays is on whether the models should by design guarantee the equivariance of the output [33] or not [34]. Using equivariant architectures seems to lead to more robust models that are easier to train since they require no data augmentation. State-of-the-art machine learning potentials, for example, are currently based on equivariant models. But, although equivariant models satisfy the "physical law" that equivariance is, they don't do it like nature does. They do it by restricting internal features to behave like a certain order of spherical harmonics ℓ. For practical reasons, models can only deal with a finite number of ℓ, tipically $\ell = 2$ or $\ell = 3$, before it becomes too memory demanding. However, nature has at its disposal an infinite number of ℓ. The restriction that equivariant models impose results in a decreased expressivity in comparison to non-equivariant models. One argument against equivariant models, for example, is that the atomistic ML community can not benefit directly from improvements in other fields where they don't take equivariance into account. In any case, I will now describe the equivariance of the density matrix, as it is a property of the data. When predicting the density matrix, we can choose to use this property or not. In particular, we have indeed used equivariant models to take advantage of the existing state-of-the-art machine learning methodologies.

In DFT codes based on atomic orbitals, wavefunctions (ψ_i) are represented as linear combinations of atomic orbitals (ϕ_μ).

$$\psi_i = \sum_\mu c_{i\mu} \phi_\mu \tag{2.2}$$

where $c_{i\mu}$ are the coefficients of the linear combination. The density matrix is simply built by taking the outer product of the coefficients array for each wavefunction, and weighting it by the occupation of the wavefunction (n_i).

$$\rho = \sum_i c_i \otimes c_i^* n_i \tag{2.3}$$

To understand the equivariance properties of the density matrix, then, it is useful to first understand the equivariance properties of the wavefunction coefficients. Then, the density matrix will satisfy the properties of a product of such equivariance. Figure 2.5 illustrates how the wavefunction coefficients transform when the system is rotated. The transformation of each coefficient depends on the basis function that it applies to. For example, coefficients of s orbitals ($\ell = 0$) don't change at all with rotations, because the $\ell = 0$ spherical harmonics are spherically symmetric. On the other hand, coefficients of p orbitals ($\ell = 1$) behave like vectors. They transform with the same rotation matrix R that induced the rotation of the system. These ($\ell = 0$ and $\ell = 1$) are simple examples that can be easily illustrated, but the general fact is that coefficients from orbitals with any ℓ will transform with the representation of the rotation R in the order of spherical harmonics ℓ. This representation corresponds to the Wigner D^ℓ matrices, i.e. $R(\ell = x) = D^x$. For $\ell = 0$, $D^0 = 1$, which results in the coefficient invariance that I just discussed. For $\ell = 1$, $D^1 = R$. For further spherical harmonics, the transformation is more complex but it is still well defined.

We can therefore assume the fact that the wavefunction coefficients will transform like spherical harmonics according to the (ℓ, m) of the basis function. With this, I will now discuss the equivariance of the density matrix, which is composed of outer products of the coefficients array. Figure 2.6 illustrates the equivariance of an outer product of two vectors, which is tipically explained in literature due to its simplicity. This would correspond to the outer product between two sets of p ($\ell = 1$) orbitals (p_x, p_y, p_z). The outer product will create a 3×3 matrix containing all possible products. If we want to predict this matrix, we could treat each of the elements of the matrix individually, loosing all notion of what the equivariance of the data was. But the matrix can be decomposed into a set of arrays with known equivariance. This is known as the irreducible representation (from now on, *irrep*). In particular, the irrep of the $p \times p$ matrix consists of three features: a $\ell = 0$ feature (the trace, which is the scalar product), a $\ell = 1$ feature (related to the symmetry of the matrix, the cross product) and a $\ell = 2$ feature (rest of the matrix).

Although the discussion of a $p \times p$ block is nice, a given atom might contain more than one set of orbitals. The irrep of the outer product between two atoms with any basis is simply the sum of the
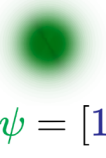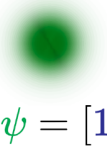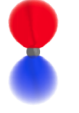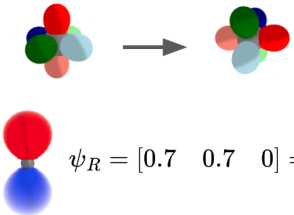
	Basis	Wavefunction	Wavefunction with rotated basis
s orbitals $\ell = 0$	ϕ_s	$\psi = [1]$	$\psi = [1]$
p orbitals $\ell = 1$	(p_x, p_y, p_z)	$\psi = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$	$\psi_R = \begin{bmatrix} 0.7 & 0.7 & 0 \end{bmatrix} = R\psi$

Figure 2.5: Equivariance of wavefunction coefficients for s orbitals ($\ell = 0$) and p orbitals ($\ell = 1$). The first column illustrates the basis functions for each case. The second shows an example wavefunction and how it is described in terms of coefficients for the given basis. In the third column, I rotate the basis, which is equivalent to rotating the wavefunction by rotating the system. Then I show how the coefficients transform in each case. For s orbitals, the coefficient remains the same, while the coefficients of the p orbitals transform with the same rotation R that was used to rotate the basis.

irreps of all possible products between sets of orbitals. E.g. given two atoms with s and p orbitals:

$$irreps(M_{ij}) = irreps([s+p] \times [s+p]) = irreps(s \times s) + irreps(s \times p) + irreps(p \times s) + irreps(p \times p) \quad (2.4)$$

Of course, the irreducible representation is different depending on the atom species that interact to produce a given matrix block M_{ij}.

As discussed at the beginning of this section, understanding the irreducible representation of the density matrix is nice to understand how the matrix transforms with rotations of the system. One can then choose to design models that formally guarantee this equivariance, or not.
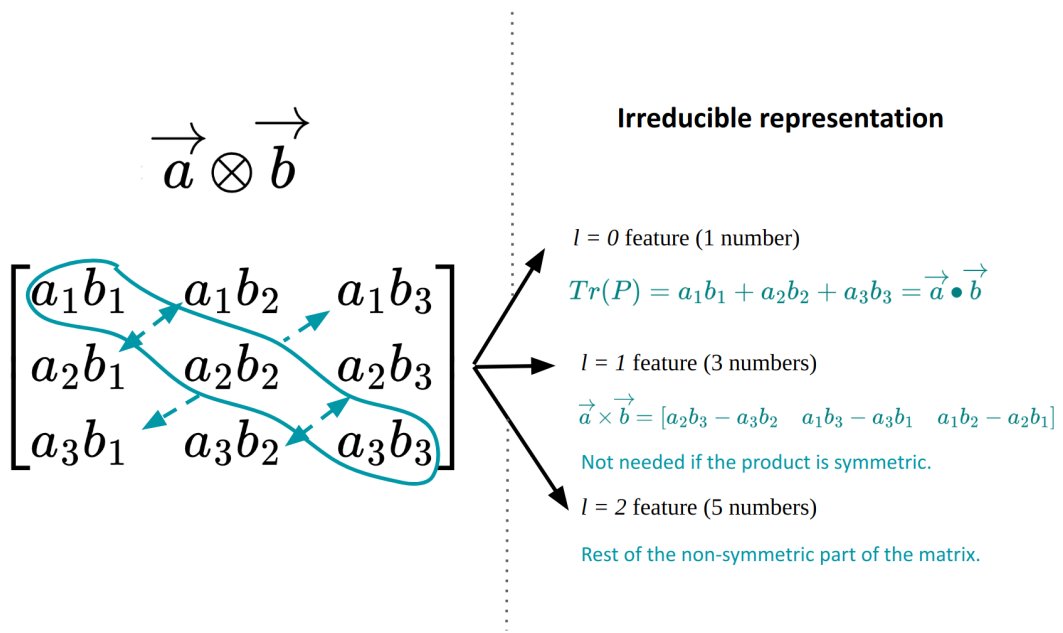
Irreducible representation

$l = 0$ feature (1 number)

$$Tr(P) = a_1b_1 + a_2b_2 + a_3b_3 = \vec{a} \bullet \vec{b}$$

$l = 1$ feature (3 numbers)

$$\vec{a} \times \vec{b} = [a_2b_3 - a_3b_2 \quad a_1b_3 - a_3b_1 \quad a_1b_2 - a_2b_1]$$

Not needed if the product is symmetric.

$l = 2$ feature (5 numbers)

Rest of the non-symmetric part of the matrix.

Figure 2.6: Equivariance of an outer product between two $\ell = 1$ tensors, $\vec{a}$ and $\vec{b}$. The trace of the resulting matrix corresponds to the scalar product, which is invariant with a rotation of both vectors. The cross product $\vec{a} \times \vec{b}$ defines how asymmetric the matrix is. If $\vec{a}$ and $\vec{b}$ are parallel the cross product is zero, and the matrix is symmetric. The cross product rotates with rotations of the $\vec{a}$ and $\vec{b}$ vectors, but remains invariant when they are inverted. This is known as a pseudovector. The rest of the degrees of freedom in the matrix correspond to a $\ell = 2$ feature.
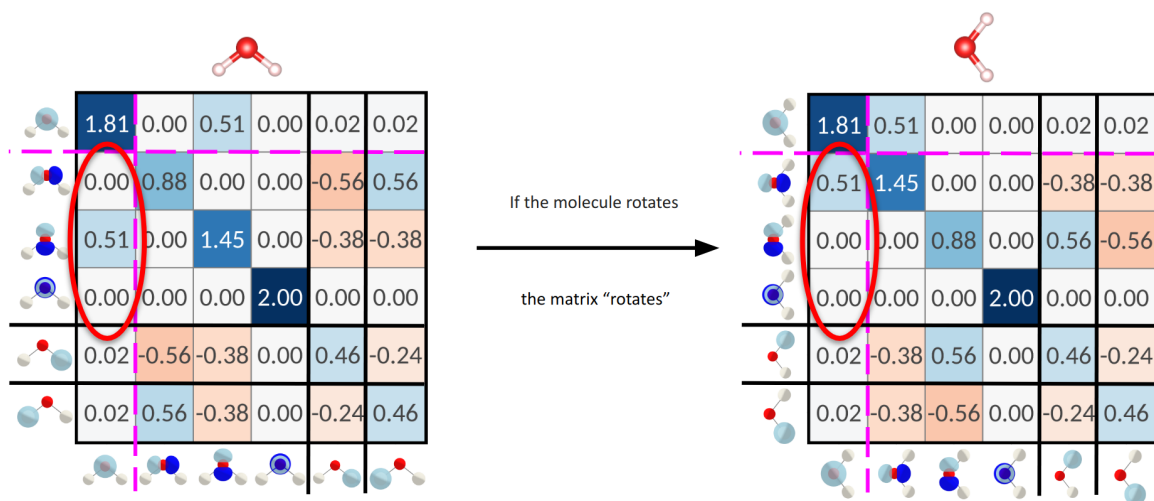
Figure 2.7: Example of the equivariance of the density matrix by comparing two real density matrices computed with SIESTA for a water molecule. The colorscale is just a visual reinforcement of the matrix values. The matrix on the right corresponds to the molecule rotated 90º with respect to the one on the left, which essentially corresponds to a swap between the horizontal and the vertical axis. Circled in red, the interaction between s and p orbitals of the Oxygen atom, which transforms as a vector and therefore has swapped the first and second components.

2.1.4 The architecture of Graph2Mat

At this point, the reader hopefully understands why do we want to predict matrices and why do we use graphs to achieve it. As I said, the first step of our workflow is to compute atomic environments and build a graph with the atomic environments as node features. In this section I discuss how we then convert this graph into our actual target matrix. We have developed an open source python package `graph2mat` that, apart from providing very useful tools for manipulating sparse equivariant matrices in a machine learning context, it implements a generic function Graph2Mat that converts graphs to matrices.

As depicted in Figure 2.8, Graph2Mat is a skeleton that can host arbitrary functions to customize the processing of the graph. It manages all the annoying nuances regarding the shape and the symmetry of the matrix, so that a user can focus only on the computation part. Graph2Mat takes as an input simply the graph description and the node features, which are likely to come from an atomic environment representation. It then divides the computation in all the different kinds of block types that the matrix has. For example, Oxygen blocks are computed with a different function than Hydrogen blocks. And an Oxygen-Oxygen edge block is computed with a different function than a Oxygen node block. By defining functions f_n and f_e, the user can tweak the computation of node and edge blocks. By default, functions f are simple tensor products as directly defined in the `e3nn` [33] package. f_n is a `TensorSquare` and f_e is a `FullyConnectedTensorProduct`. Graph2Mat determines what is the irreducible representation of each block type and then lets the function know that it should produce an output with that shape. When all computation is done, Graph2Mat merges all blocks to produce the full matrix.

A small detail that I would like to highlight is that Graph2Mat is not actually itself equivariant and it therefore does not take into account irreducible representations. The sketch that I show in Figure 2.8 is a version (subclass) of Graph2Mat that is equivariant and that is provided within the `graph2mat` package. I have decided to illustrate the equivariant version because it is the one that we have used to produce the results that I will show in the next section.
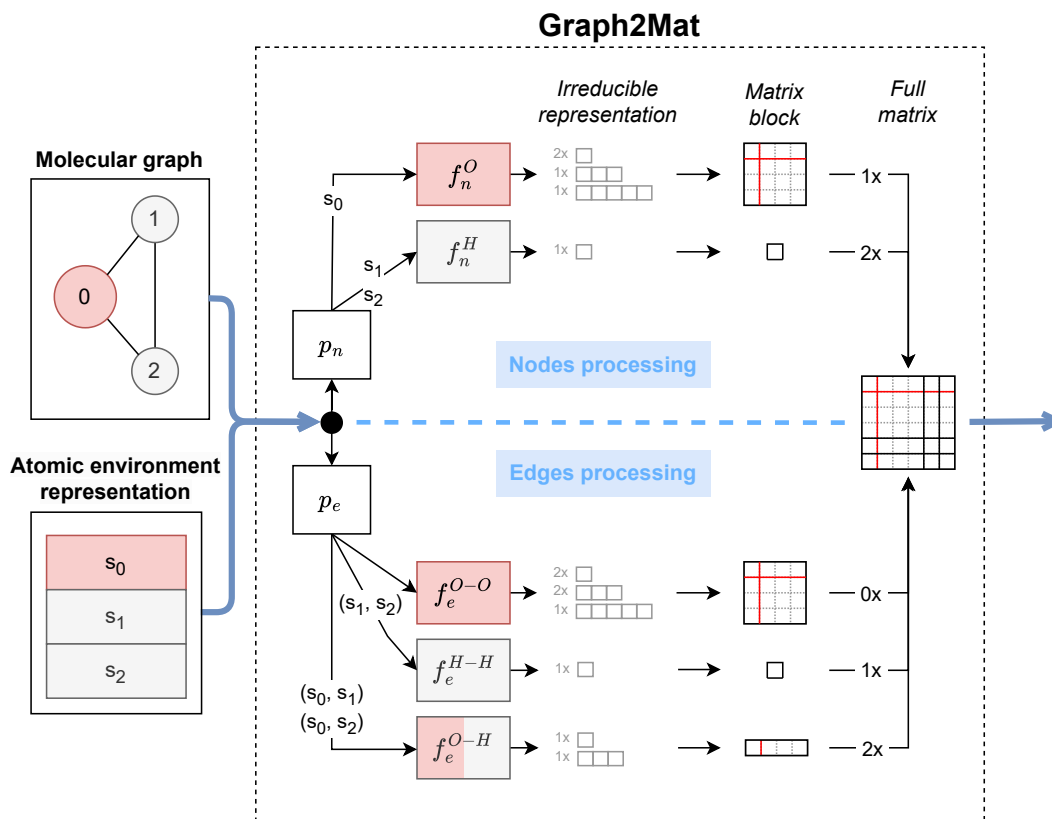
Figure 2.8: Architecture of the Graph2Mat function for a dataset consisting of structures with O and H atoms using a SZ basis. The figure also illustrates the passing of a water molecule through it. The two inputs are the molecular graph and the atomic environment representation for each node in the graph. Then, the processing splits in two: the computation of node blocks (M_{ii}) and the computation of edge blocks ($M_{ij}, i \neq j$). Preprocessing functions p compute the final states that will be passed to the block generating functions f. Each function f produces an output with the shape of an irreducible representation, and that representation is then converted to a matrix block M_{ij}. Finally, all blocks are merged to generate the full matrix. The number of blocks that are gathered from each type to generate the full matrix is indicated as "0x", "1x" or "2x".

2.2   MACE + Graph2Mat: End-to-end learning of matrices

Having given Graph2Mat the task to generate matrices from graphs, we only needed to decide which input to give to Graph2Mat. We decided to use MACE [15] as the atomic environment descriptor, because of its apparent compatibility with Graph2Mat: MACE is also an equivariant graph neural network based on e3nn, so we could share the preprocessed inputs (e.g. the description of the graph) between both parts of the workflow. Furthermore, MACE has shown to obtain astonishing results in the fitting of force fields. We were curious to see if it could perform equally well in the prediction of matrices. The complete model, combining MACE and Graph2Mat, is referred to as MACE+Graph2Mat.

Although the MACE+Graph2Mat model can predict any matrix, here I will show only the results of learning the density matrix (DM). The DM is the most interesting matrix for us, as the goal of this chapter is to produce good estimates of the electronic density. As a reminder to the reader of what we are trying to achieve, in Figure 2.9 I display the interplay between machine learning and DFT using the matrix predictions. From the atomic environment representation, a machine learning potential would compute the numbers that are trained to be properties like energies, dipole moments, etc... Instead, in our approach we convert the representation to a matrix which is trained to be the DM. Then, we take the DM prediction for the system and feed it into the DFT engine, which will compute the ground state properties of the system. Depending on whether you ask the DFT engine to converge the density or to take the predicted density as the truth, you will either get the "true" properties of the system or get predictions of the properties based on the predicted density.
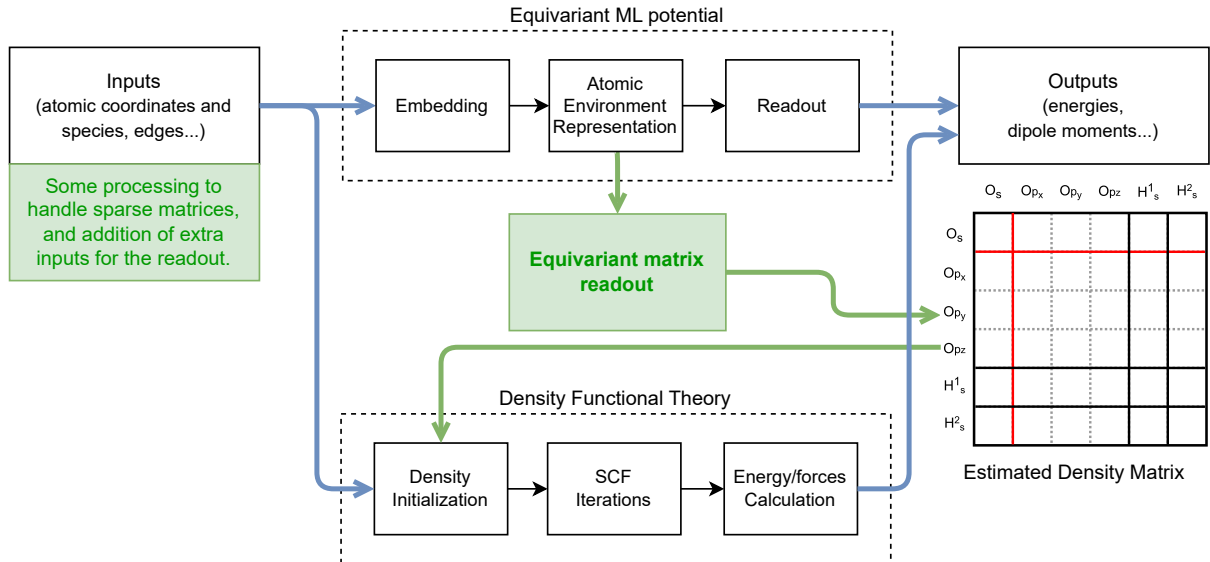

Figure 2.9: Integration between ML and DFT

Let me also state that one of the powers of our modular approach that separates the atomic environment description from the matrix generation is that we can easily swap the environment descriptor for another one. Although MACE has shown to give very good results, as you will see in this section, we are also playing with the idea of incorporating long range descriptors through LODE [35]. There are no results to show in that regard yet, but the ease with which we have been able to carry first experiments with LODE has clearly shown that keeping things modular pays off in the long term.

2.2.1   Training

2.2.1.1   The datasets

The datasets used for this work were computed with SIESTA [36], [37]. We used pseudodojo's [38] psml pseudopotentials and all of SIESTA's default settings, including the default split-valence double zeta polarized (DZP) basis set. In the dataset, along with the coordinates and atomic numbers of each structure, we stored the target matrix (or matrices). In particular, the datasets we use contain the density matrix, overlap, hamiltonian, and energy density matrix. However, I only show the results on predicting the DM, because as I said it is the focus of this chapter. The results I present in the following sections involve two different datasets:

QM9 QM9 [39], [40] is a dataset of 134k small organic molecules with up to nine "heavy" atoms (C, N, O, or F) and is widely used for benchmarking machine learning models for molecular property predictions. 10k randomly sampled structures were saved for testing.

Ethylene Carbonate (EC) The Ethylene Carbonate dataset is generated from a molecular dynamics simulation of a liquid ethylene carbonate electrolyte consisting of 8 Ethylene Carbonate molecules in a unit cell. The dataset was developed for benchmarking the equivariant DeepDFT model [22], [41]. It consists of 12k snapshots from the MD run, 1k of which have been saved for testing.

The two datasets are available at [42]. We selected these two datasets because:

- They are representative of two different situations. QM9 is a dataset of molecules while EC requires the treatment of periodic boundary conditions.

- They were both used as benchmarks for the DeepDFT model [22] developed previously by my collaborators Peter Jørgensen and Arghya Bhowmik, that predicts the density on real space grids. Therefore, they provide a good comparison point to assess the performance of DM prediction.

2.2.1.2   Training details

During our learning workflows, we train the MACE+Graph2Mat models end-to-end from coordinates and atomic numbers to the DM. To avoid the overhead of computing the electron density in real space, we compute the loss function on the values of the density matrix directly. Element-wise mean average error (MAE) is used as the loss function, ignoring errors that are lower than the convergence threshold in SIESTA. Our tests have shown that using the mean squared error (MSE) instead of MAE leads to decreased performance. An Adam optimizer with a learning rate of 0.005 is used to optimize the parameters of the model. The training presented here has been performed on a single NVIDIA GeForce RTX 3090 card. However, our code is ready for multi-GPU distributed training, and we have tested for it.

2.2.2   Predicting densities

Although my main goal is to make DFT calculations faster, we can not ignore the fact that throughout the way we have developed a model that can predict electron densities. Predicting densities in itself is useful, for example to model the electronic response of metals, estimate bonding strength or to use it as a starting point to predict other properties like energies and forces. To evaluate the performance of MACE+Graph2Mat at predicting densities, Figure 2.10a compares the learning curves for the QM9 and EC datasets with those obtained by DeepDFT, which as explained before is a grid-based approach. In order to be able to compare both, I use an error metric that is computed on the real space grid, the Normalized Density Error (NDE):

$$NDE = \frac{\int |\rho(\vec{r}) - \rho_{pred}(\vec{r})| d\vec{r}}{\int \rho(\vec{r}) d\vec{r}} \tag{2.5}$$

This might seem unfair for MACE+Graph2Mat, which has never been trained to minimize real space density errors. But ultimately, a scientist that is interested in predicting electron densities does not care about the basis used and/or the entries of the matrix, they will only care about the precision of the density in real space. What the learning curves show is that given a number of training examples, MACE+Graph2Mat and DeepDFT perform similarly, although the density matrix approach seems to give slightly better results.

To the reader, the improvements in prediction precision might appear too modest to justify the use of the density matrix. However, as explained previously in Figure 2.2 the DM is a much more compact representation. Figure 2.10b illustrates this point by comparing the size of the QM9 and EC datasets when using the DM or real space grids. The QM9 density matrix dataset is more than 1 order of magnitude smaller than the grid dataset, while EC shows a 1 order of magnitude decrease. The difference in the benefits for both systems is due to QM9 molecules being surrounded by vacuum, which makes the sparse representation even more efficient. In any case, this means that the matrix-based approach is much more data efficient: it can acheive the same performance using much less data per structure. This might seem like a not very important fact, but anyone that has had to deal with a 1 TB dataset will certainly appreciate the fact that they can work with a 30GB dataset and acheive the same results, as is the case for QM9.
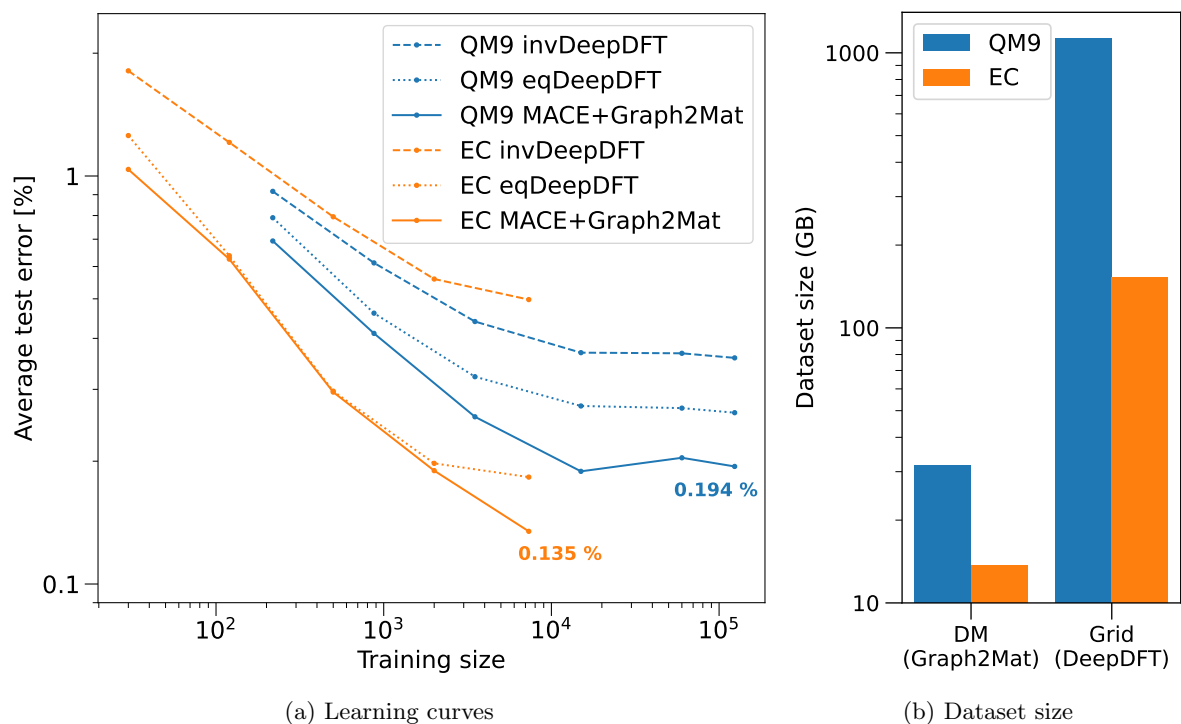
(a) Learning curves

(b) Dataset size

Figure 2.10: Comparison with DeepDFT (a grid based approach) on two large scale benchmarks: QM9 in blue and EC in orange. We compare to both invariant DeepDFT (invDeepDFT) and equivariant DeepDFT (eqDeepDFT). (a) shows the learning curves with respect to the number of structures used for training. The performance is very similar, although MACE+Graph2Mat performs a bit better. (b) shows the comparison of dataset size when using the DM vs when using grids to store the density. The compressed size is displayed for grids, since DeepDFT workflows uncompress them on the fly.

2.2.3 Using predicted densities

After determining that MACE+Graph2Mat can predict electron densities with good precision, we can now move on to understanding how to use these predictions. For this purpose, we used only the QM9 test set. Figure 2.11 shows the performance of the MACE+Graph2Mat model in various downstream tasks that use the density predictions.

2.2.3.1 Reducing the number of SCF iterations

For our initial goal that was derived from the motivation to accelerate NEGF/DFT calculations, this is the most important downstream task. Without a doubt, the predicted densities are a much better estimate for the true electron density than the atomic densities. However, whether this translates into faster SCF convergence is not a trivial question to answer. In Figure 2.11a I compare the number of SCF steps needed to converge when starting from the predicted densities against when starting from atomic densities. Even for the model that has been trained only on 218 molecules, the predicted densities lead to an average of 27% less iterations needed to converge. This shows that even a rough estimate of the density is enough to improve on the atomic densities. As the model is trained on more examples the performance improves, until it reaches saturation around 15k training examples, where the reduction in the number of steps is around 40%. This is also the point where Figure 2.10a shows a saturation on the electron density error not only for the MACE+Graph2Mat model but also for DeepDFT, indicating that there is probably not much more to learn from QM9 by increasing the training set size further.

In any case, the results are very promising. They show that the approach has potential to significantly reduce the computational cost of DFT calculations. In this case, I must say, the molecules are so small that the reduction in SCF steps does not translate into faster simulations. The prediction of the density introduces an overhead of around 1-2s (running on 1 CPU), while each SCF step in SIESTA takes around 0.2s (running on 10 CPUs) for the smallest QM9 molecules. This is not a concern, since the model can be accelerated by simply running on GPU, optimizing the code, or compiling the model to run in production. Also, systems as small as the molecules in QM9 are not the target for DFT acceleration, and the ML model scales linearly with the number of atoms while SIESTA scales at least quadratically with the number of orbitals (unless you are in systems big enough where linear scaling algorithms pay off). In other words, the overhead of the ML model is not expected to be a problem for the systems that need acceleration. Another thing to take into account is that QM9 molecules are very easy to converge. There are other systems that are much harder to converge, or even don't converge at all starting from atomic densities. In the future, it would be very interesting to check if the electron density predictions can mitigate these convergence problems.

2.2.3.2 Predicting system properties from the density

Apart from using the predictions as an accelerating tool for DFT, one can compute many properties from the electron density. In this section, I explore the performance that we get when we compute ground state properties of the system directly from the predicted densities, without further refining from DFT.

Single SCF step energy   We can use SIESTA to compute the total energy from the predicted unconverged density. Figure 2.11b shows the mean energy error going from 2.2 eV for the 218 molecules training set to 0.019 eV for the largest training sets. The error of the largest training sets is not in the same ballpark as state-of-the-art machine learning interatomic potentials, which have reported mean absolute errors of around 5 meV [43]–[45]. But it is still a significantly low error taking into account that the model was never trained to optimize energies. An interesting difference from ML potentials is that the energy errors reported in this work have all the same sign. The variational principle ensures that an unconverged (but idempotent and correctly scaled) DM will always give higher energy than the converged solution.

Dipole moment   Dipole moments for molecules can be computed directly from the DM without going through the DFT machinery by evaluating $Tr[DM * X]$, where Tr represents the trace, $*$ means matrix multiplication and X is the expected value of the position operator $X_{ij} = \langle \phi_i | \vec{r} | \phi_j \rangle$. This operation is linear scaling since the full matrix multiplication is not needed, just its diagonal elements. X_{ij} has the same sparsity as the overlap, therefore only the sparse part of the DM (what we predict) is needed. With this approach, Figure 2.11c shows a MAE of 0.13 Debye for the larger training sets, which is one order of
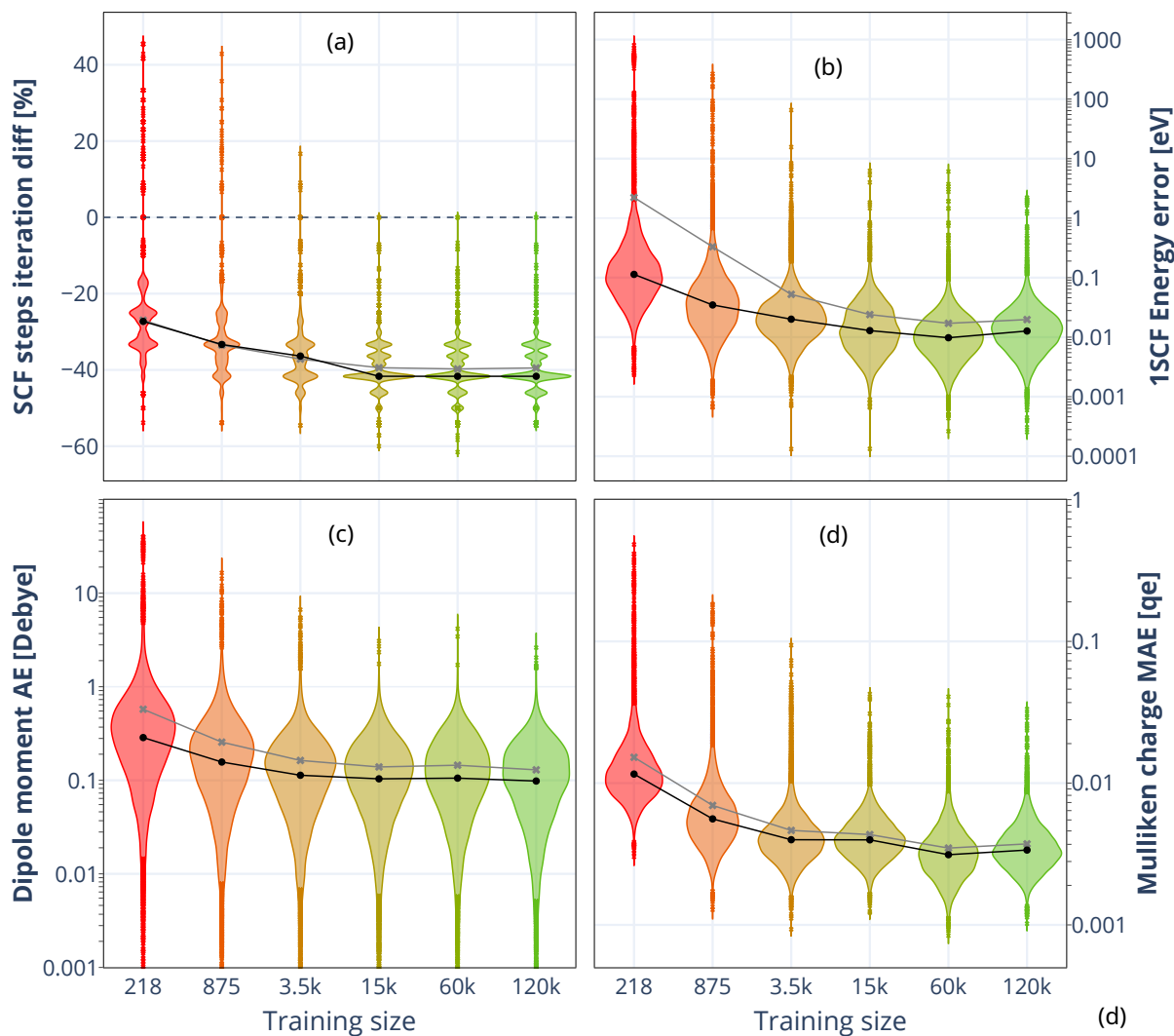
Figure 2.11: Performance of QM9 models in various downstream tasks in the test set as a function of training set size. Violin plots show the distribution while the gray/black lines shows the mean/median. (a) SCF steps taken to converge the SIESTA run starting from the predicted DM (vs starting from atomic densities). (b) Error made on the total energy if we just do 1 SCF step in SIESTA from the predicted DM. (c) Absolute error on the magnitude of the dipole moment. (d) Mean absolute error of the Mulliken charges in a molecule. The top two plots are from integration of the model within DFT, while the bottom ones are directly computed from the DM.

magnitude higher than published end-to-end approaches [44], [45], but still makes for a useful indicator of a molecule's dipole.

Mulliken charges   Mulliken charges [46] can also be computed directly from the DM using the overlap matrix. We compute them with the sisl python package [13]. With this approach, the smallest training set in Figure 2.11d shows a MAE of 0.015 e^-, while the MAE converges at around 0.003 e^- with training size.

For all properties there are some large outliers, especially when using the smallest training set. These outliers are common across properties and correspond to structures that are ill represented in the training set. For example, the training set of 218 molecules does not contain structures with F and the obvious outliers (energy errors above 100 eV) correspond to molecules with 3 F atoms. Such outliers distort the mean error, and it might be better to look at the median of the errors to understand how the models perform. In subsection 2.2.4 I discuss how to identify these outliers. It is also notable that, as mentioned with the number of SCF steps, all properties seem to saturate at 15k or 60k training examples.

2.2.4   Uncertainty estimation

Although showing good performance is always very nice and often a requirement for a ML model to be relevant, I personally really like the results that I am going to present in this section, which have nothing to do with performance. Machine learning models are very good at hiding their insecurities. Regardless of how little they know about the question that you are asking them, they will always give you an answer. With humans, a good strategy to tell how uncertain someone is about what they are saying would be to pay attention to their body language. ML models don't have a body (yet), so that won't work. Another (not so nice) approach with humans is to tease them so that they confess if they are making things up. ML models don't have feelings (yet), so that won't work either. At this point, the reader might be thinking that it is pointless to compare ML models to humans. But surprisingly, there are techniques that work for both humans and ML models alike. The first one is to take a bunch of them that have similar knowledge and ask them the same question separately. It is very unlikely that the answers match if they are just making them up. With ML models, this is called uncertainty estimation by model ensemble. Model ensembles are widely used because they have almost no requirements on the model, they are very easy to implement and they work.

But there is another way of uncertainty estimation that is a bit smarter. Imagine sending a kid to the store to buy a loaf of bread. When they get back, you ask them how much they paid for it so that you can give them the money back. The kid tells you "It was 50€.". You know that no matter how ecological the bread is, it can't be 50€ (in year 2024). You don't know the exact answer, but you determine that the kid is lying simply because their answer doesn't fall into a reasonable range. With ML models you can do the same. In particular, there might be a simple rule that answers should follow in order to be reasonable, but you can choose not to tell the model about it. In that way, the model is likely to give unreasonable answers when it is uncertain. For ML models that predict the density in periodic systems like the ones we simulate with DFT, the simple rule is charge neutrality (or total charge). By not telling the ML model what the total charge of the system is and not designing it in a way that it is constrained to give the correct total charge, the model is likely to be off on the total charge when it is uncertain. Figure 2.12 shows how the error that the model makes on the number of electrons is clearly correlated to the error made on the properties computed from the density, in particular the total energy. Essentially, the plot is telling us that a density with the wrong number of electrons is a bad density and we shouldn't use it to compute other things. This sounds obvious, but it is something that you wouldn't have if the model was constrained to give the correct number of electrons, as the predictions could be extremely wrong while giving the perfect number of electrons. This uncertainty estimation method based on simple charge counting does not suppose practically any computational cost in comparison to training and running an ensemble of models.

Although I don't have a human analogy for them, there are also other useful uncertainty proxies that come up if one SCF step is performed from the predicted density inside the DFT engine. These are particularly useful when the model is deployed for molecular dynamics, where within SIESTA we have to do at least one SCF iteration to compute the forces because of the non-orthogonality of the basis set (appendix A in [36]). The best metric that I have found for this job is the maximum error on the self-consistency of the Hamiltonian matrix (dHmax), which is printed by SIESTA on each SCF step. The bottom row of Figure 2.12 shows its correlation to energy and forces errors. The energy error depends

quadratically on the self-consistency error, while the forces error scales linearly, which is the expected behavior [47].
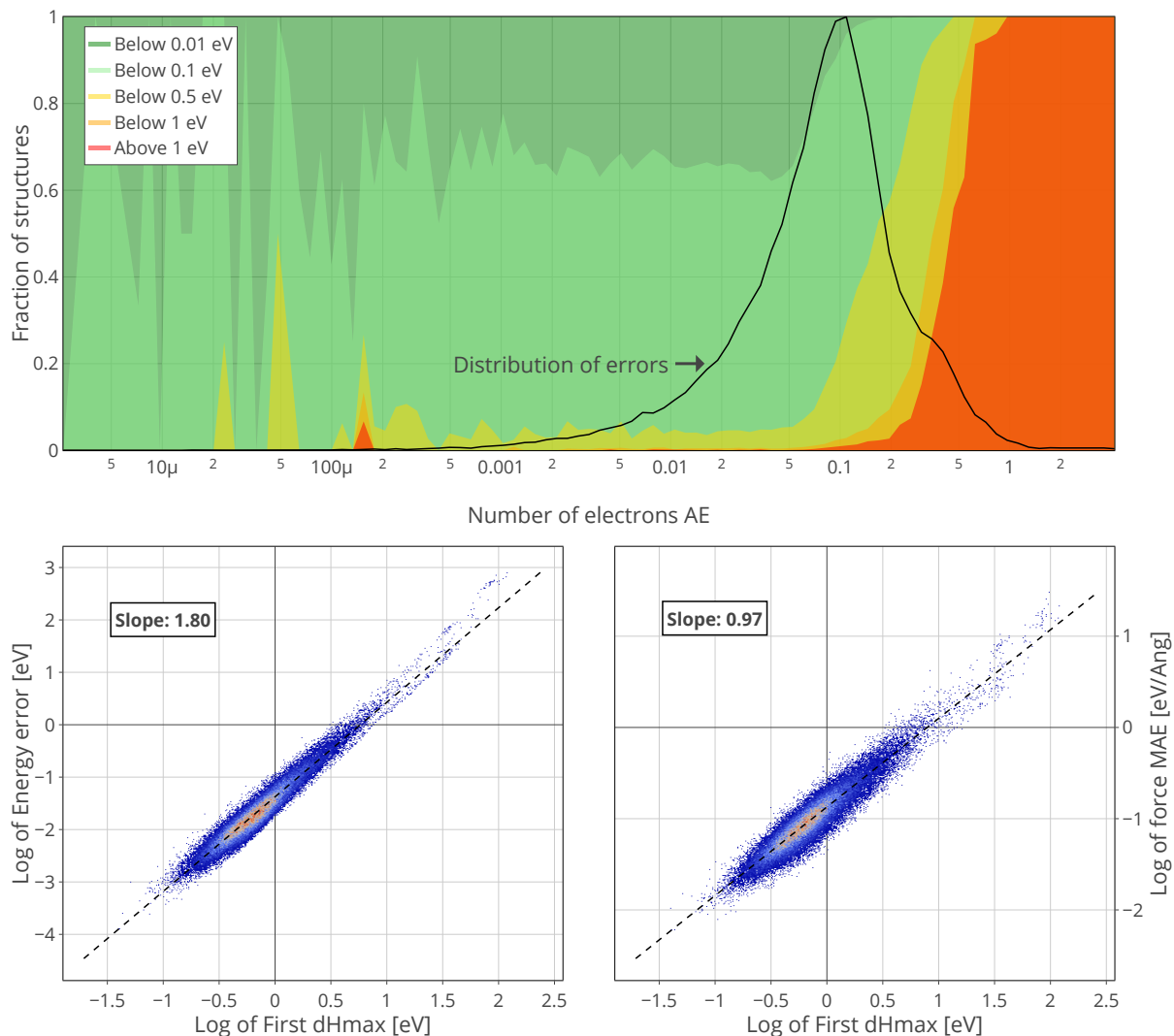

Figure 2.12: Uncertainty proxies from the QM9 test set. Data from all training sizes is used to generate the figures. (Top) For a given error on the number of electrons, fraction of structures that are below a certain energy error threshold. The black line shows the distribution of errors on the number of electrons. The number of electrons of a DM can be computed without any SCF step. (Bottom) Correlation of the maximum self consistency error of the Hamiltonian at the first SCF step with the error of 1-SCF energies and forces. The self consistency error is only available after 1 SCF step. Since the axes are logarithmic, the slope of the linear fit indicates the degree of the polynomial that correlates the X and Y variables.

2.3 Conclusions

When I started the project presented in this chapter, I had one goal: accelerating DFT calculations. By designing and implementing a modular approach to machine learn the density matrix, I have been able to reuse a preexisting environment descriptor, MACE, to achieve this goal. The results have shown that there is potential in using a ML model that predicts the density matrix to speed up DFT simulations. Even with very few training examples, the model can predict rough estimates of the electron density for QM9 molecules which are already much better than the initial guess typically used in DFT simulations. These predictions are good enough to reduce the number of SCF iterations needed to converge the DFT calculations of molecules in the test set.

MACE has been surprisingly successful in supporting foundational models [48] that can predict fairly good atomic forces for virtually any system of interest. Therefore, it doesn't feel like a stretch to think that our MACE+Graph2Mat model can also be used to create foundational models for electron density. Producing rough estimates of the electron density for any general system could result in the acceleration of DFT calculations performed daily by many scientists. This is a path that I would like to explore in the future.

Inevitably, the models will sometimes produce bad estimates. I have shown that by not constraining the model to predict electron densities with the correct total number of electrons, we have tools to estimate uncertainty and discard predictions if they are likely to be bad. These tools are computationally much less demanding than creating model ensembles, and I envision that they could be very useful for active learning strategies in the future.

Summary of conclusions:

- The density matrix is an alternative target for machine learning electron densities which is much more data-efficient than grid-based methods.

- The MACE+Graph2Mat model shows slightly better performance than state-of-the-art grid-based methods on the prediction of electron densities.

- The MACE+Graph2Mat model can be used to accelerate DFT calculations by providing electron density estimates that facilitate the convergence of the SCF cycle.

- Since the model is unconstrained with respect to the total charge, the error on the total charge can be used as an uncertainty metric to detect bad predictions.

Summary of original contributions from the author:

- Design and development of an open source python package (*graph2mat*) that converts molecular graphs into sparse equivariant matrices for ML applications. Also wrote tutorials and documentation to facilitate community adoption.

- Generation and publishing of matrix datasets computed with SIESTA, available at [42].

- Training of density matrix models that have proved the concept that targeting the density matrix is a good approach to machine learn electron densities and accelerate DFT simulations.

- Sending of an article as first author to the Machine Learning Science and Techonology (MLST) journal, which is currently under review.

Chapter 3

Using learned densities to accelerate NEGF-DFT

When we work hard on a project, it is often easy to lose track of why we started it in the first place. While working on Graph2Mat and the proof of concept that general DFT calculations can be accelerated by learning the density matrix, I was tempted to follow many interesting ramifications. However, I always kept in mind that my personal fulfillment would come from the application of these techniques to simulations that are relevant for electrochemistry. By applying Graph2Mat to simulations that involve TranSIESTA, I would feel that the journey was worth it and the circle was (kind of) closed. Since the time of a PhD is limited and the journey to get here was not easy, I only had time to come up with a humble proof of concept. In this short chapter, I present the results that I obtained with two objectives in mind: (1) feeling personally fulfilled and (2) trying to set the ground to trigger future research in this direction.

3.1 QM/MM: An easy target

Once the determination to apply Graph2Mat to TranSIESTA simulations was made, the next step was to identify a specific problem to tackle. At that time, my colleague Ernane de Freitas Martins was running the simulations depicted in Figure 3.1. These are QM/MM simulations of a gold capacitor with liquid water between its plates at applied voltage. The system is very similar to the one I presented in section 1.3, but in this case the QM/MM method allows sampling much bigger sizes. The gold electrodes are treated with a full quantum mechanical description with TranSIESTA (QM part), while the water molecules are treated classically with the TIP3P model (MM -molecular mechanics- part).
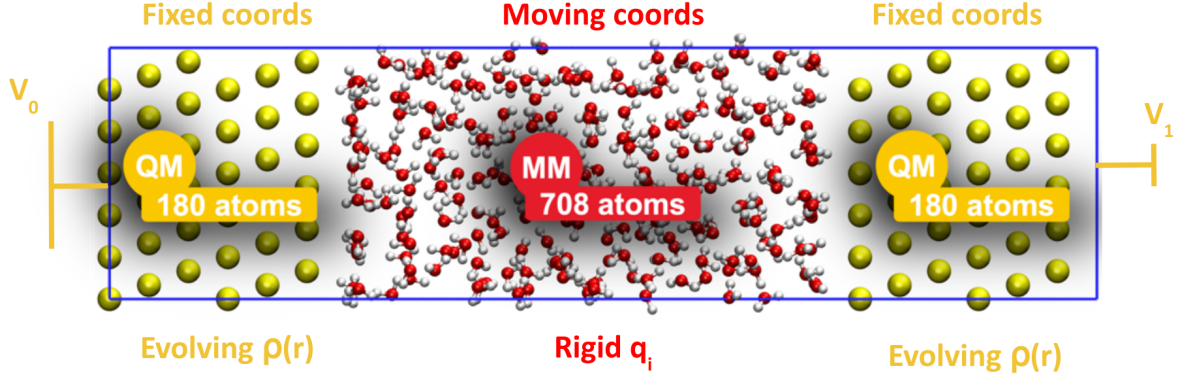
Figure 3.1: The QM/MM system that I explore for ML acceleration in this chapter. The system is coupled to gold electrodes on the left and right sides (Z direction), while the other two directions are periodic. Gold electrodes have a [111] surface. The unit cell, which has dimensions 15.7 x 15.1 x 61.8 Å, contains a 3x5 supercell of the orthogonal primary cell for Au[111]. The distance between plates is 34.6 Å.

In the kind of molecular dynamics that my colleague Ernane was doing, gold atoms are kept at fixed coordinates while the water molecules move. Essentially, in order to accelerate these calculations, the machine learning model needs to learn the electron density of fixed electrodes in the presence of moving water molecules. In practical terms, the fact that the electrodes are fixed makes the ML task much easier because:

- We can take the "unperturbed" density of the electrode (the electrode in vacuum) as a reference and just learn the response to the water molecules, which is much smaller in magnitude and therefore easier to machine learn up to a certain absolute error.

- Atomic forces on the MM region due to the QM electron density are purely electrostatic and therefore can be directly computed from the electron density. This is in contrast to calculating the forces in the QM region, where at least one SCF step in TranSIESTA is required due to the nonorthogonality of the basis set. Since QM region atoms are not moved, the forces in the QM region are not needed and molecular dynamics can be ran without entering the self-consistency loop.

Not surprisingly, this problem has also been identified by others as a target for ML models of the electron density. Andrea Grisafi and Mathieu Salanne have recently published a paper accelerating similar QM/MM simulations with SALTED, their ML model for electron densities [49], [50].

3.2 Exploring the training data

In order to run the proof of concept application, I got myself a small dataset of 200 example configurations that come from a previous QM/MM molecular dynamics run of 20ps for the system at 0V. The configurations are uniformly spaced 0.1 ps apart in the MD. Out of those 200 configurations, I randomly sampled 180 for training, 10 for validation and 10 for the testing that I will show in following sections. In Figure 3.2, I characterize this small dataset so that the reader understands what I am dealing with. The metallic response to liquid water can be split into two well defined contributions. There's a short-range interaction where electrodes mainly create image charges of the water molecules. Bonding could also happen in this short-range interaction. The variability of short-range interactions in the training set is hard to quantify. However, the variability of long-range interactions is much easier to quantify. The long-range contribution is only due to the total net dipole of liquid water and results in charge excess accumulation in the electrodes.

Figure 3.2a shows the distribution of the total dipole in liquid water. As the reader can see, the dipoles are not symmetrically distributed around zero. This is probably because the MD from which we sampled the configurations was not long enough to sample the full distribution of dipoles. For my equivariant ML model (MACE+Graph2Mat), this shouldn't be a problem. Dipoles with opposite signs will create the opposite response on the electrodes, which is the automatic behavior of an equivariant model. Otherwise the problem would need to be solved with data augmentation, i.e. by including rotated versions of the system in the dataset. In any case, the sampled MD was 20 ps long, which requires a significant amount of computation even with the QM/MM method. It becomes clear that ML acceleration can be helpful to reach a more complete sampling of the dipole distribution and therefore sample appropriately the thermodynamics of the system.
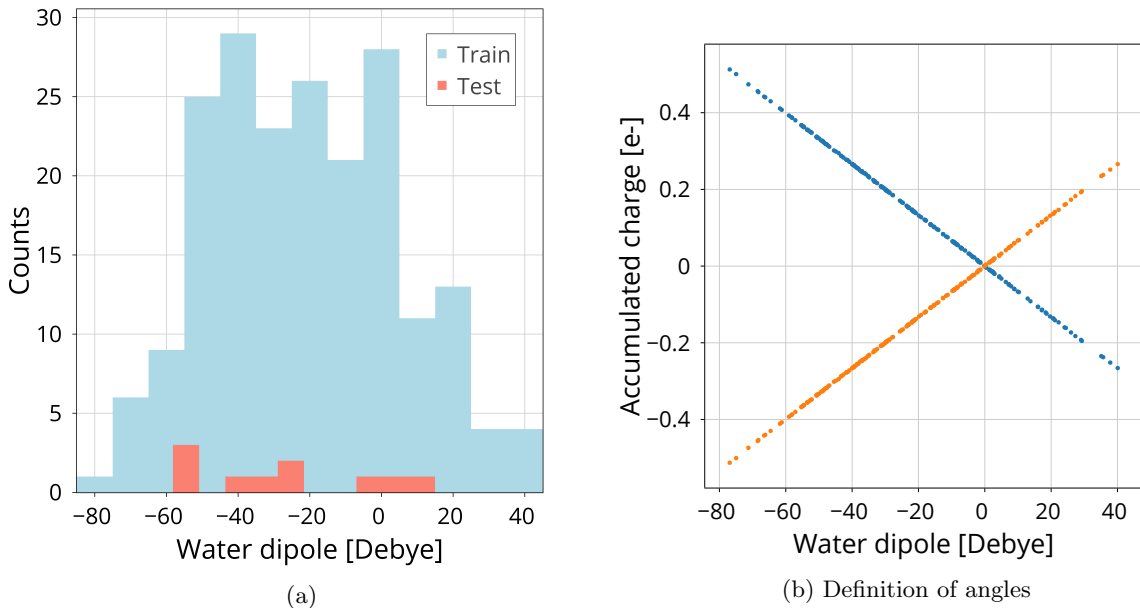

(a)

(b) Definition of angles

Figure 3.2: Characterization of the training data. (a) Distribution of the total dipole in liquid water. (b) Net charge at the electrodes at 0V as a function of the total dipole in liquid water. Blue and orange markers represent data for the left and right electrodes respectively.

Figure 3.2b shows how the accumulated charge on the electrodes is purely a function of the total dipole. In fact, the dependence is linear, as the capacitor equation would predict.

3.3 A global problem

By exploring the training data, an interesting problem comes up. The charge density response of the electrodes depends on the total dipole of liquid water, which is a global property of the system. Nevertheless, my MACE+Graph2Mat workflow uses MACE as the atomic environment descriptor. And MACE is a local descriptor. Therefore MACE will only account for the short-range contributions to the metallic response. Since MACE is a message passing graph neural network, one can make the interaction cutoffs large enough or make the algorithm do enough message passing steps so that the global property is captured for this system size. However, this is not a good general solution, as the model will not extrapolate to larger systems. Another option would be to incorporate established long-range descriptors to the model, but in this case it would be overkill. As Figure 3.2b shows, the only long-range descriptor that we need is the total dipole of liquid water, which explains all the variability in excess charge with a simple linear relation.

As a first approach to solve the problem I have created a modified version of MACE that accepts a global parameter as an extra input that is the same for each node. Then I simply include the total net dipole as an extra atom attribute to all nodes of the graph. The resulting workflow is depicted in Figure 3.3. It is possible that this is also overkill. Maybe the global parameter could be simply an extra descriptor passed to Graph2Mat, resulting in a [MACE and total dipole]+Graph2Mat model.
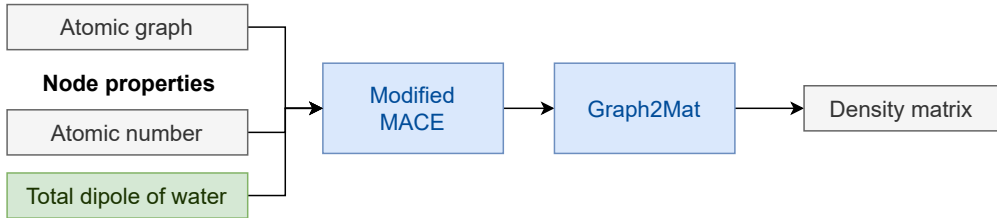

Figure 3.3: Modified MACE that accepts a global parameter as an extra input.

But in fact, one could think of smarter approaches. If the long-range and the short-range contributions to the density could be completely decoupled, then it would be easier to predict them separately. Based on the tendency in Figure 3.2b, the long-range contribution could be predicted separately with a simple linear interpolation as depicted in Figure 3.4. Perhaps it could even be possible to easily extrapolate. However, one would need to check what is the dependence of individual matrix elements on the total dipole. Individual matrix elements are not necessarily linearly dependent on the total dipole, as opposed to the total accumulated charge.
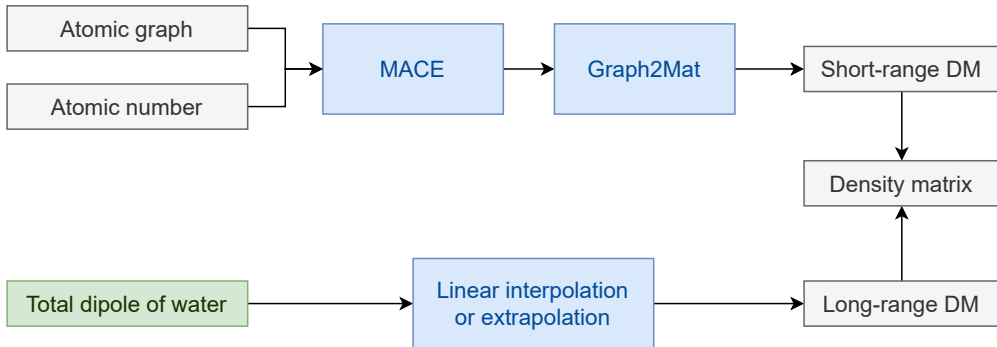

Figure 3.4: Proposed alternative to treat long-range contributions to the density.

The decoupling between long-range and short-range contributions to the density will be discussed in following sections, and I will come back to the issue with an enhanced version of Figure 3.4. To close this section, I would like to mention that to make the proposed model size-independent, one would need to also incorporate the capacitor width d. This is because the accumulated charges do not depend on the potential drop but on the electric field that they must create.

3.4 A workflow to run ML accelerated dynamics

3.4.1 Design

SIESTA, and therefore TranSIESTA, is a very efficient and highly parallelized fortran code. My machine learning models are implemented in pytorch, therefore they run in python. If I want to run molecular dynamics within TranSIESTA using the ML density predictions, I need to find a way of communicating one to the other. In particular, what I need to do is (1) pass the coordinates from fortran to python, (2) pass the predicted density matrix from python to fortran. There are not many examples online on how to communicate fortran with python. Step 2 makes it even more complicated, because the density matrix is a sparse matrix that is distributed over multiple processes in a MPI parallel SIESTA run. Since these are just proof of concept applications, I decided to give up on efficiency and hold on to the simplest solution possible. The result is the workflow shown in Figure 3.5.
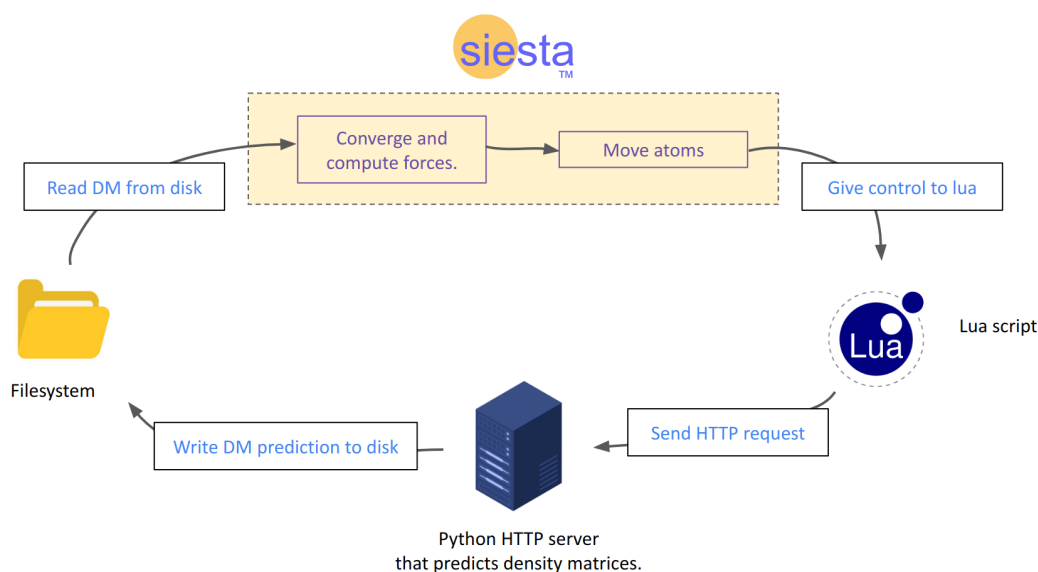

Figure 3.5: Workflow for running ML accelerated molecular dynamics

SIESTA has the possibility of reading the density matrix from a file at each MD step. Therefore, it becomes clear that the simplest approach is to write the predicted DM to a file and let SIESTA read it and distribute it internally as needed. With that, the python to fortran communication, which was the hardest part, is solved. The only thing left to do is to pass the coordinates from SIESTA to python and let SIESTA know when the predicted DM is ready. SIESTA has the possibility of yielding control to an external lua script at certain checkpoints of the calculations. I had to create an extra checkpoint in SIESTA right after the atoms are moved (contributed to the published SIESTA version), and then the rest can be managed externally by a lua script. With that, I have shifted the problem to a communication between lua and python, which is not particularly much easier. However, the main point of this shift is that exploring ideas now doesn't involve recompiling SIESTA and future users can use the mainstream version of SIESTA. The communication between lua and python is simply done by creating a python server and sending HTTP requests from lua whenever there are new coordinates to predict the DM for. The server simply reads the coordinates from a file that SIESTA writes at each MD step and when the prediction is ready it responds with OK. The lua script then receives the OK response and gives back control to SIESTA, which continues with the dynamics. This is certainly not a simple workflow, but the problem was not simple to begin with. Furthermore, this workflow allows for easy experimental modifications as it doesn't require the recompilation of SIESTA.

3.4.2 Performance

As I have discussed, the main strength of the implemented workflow is not efficiency. However, it is instructive to look at the performance metrics to see where future improvements should focus. In Figure 3.6, I show the timings of the ML accelerated dynamics with no SCF steps and compare them to regular QM/MM dynamics at 0V and 1V. The numbers are computed using a machine with 40 CPUs, running TranSIESTA in parallel with 40 processes and the ML model in a single CPU. They are therefore representative of the performance in a local cluster with no GPU available. In the ML accelerated dynamics, there are two sections that completely disappear: the SCF steps of course, but also the computation of the QMMM potential. The latter is the potential that the MM region causes on the QM region, and it is treated as an external potential for SIESTA to solve the electron density. Since ML directly predicts the density, this potential is not needed.
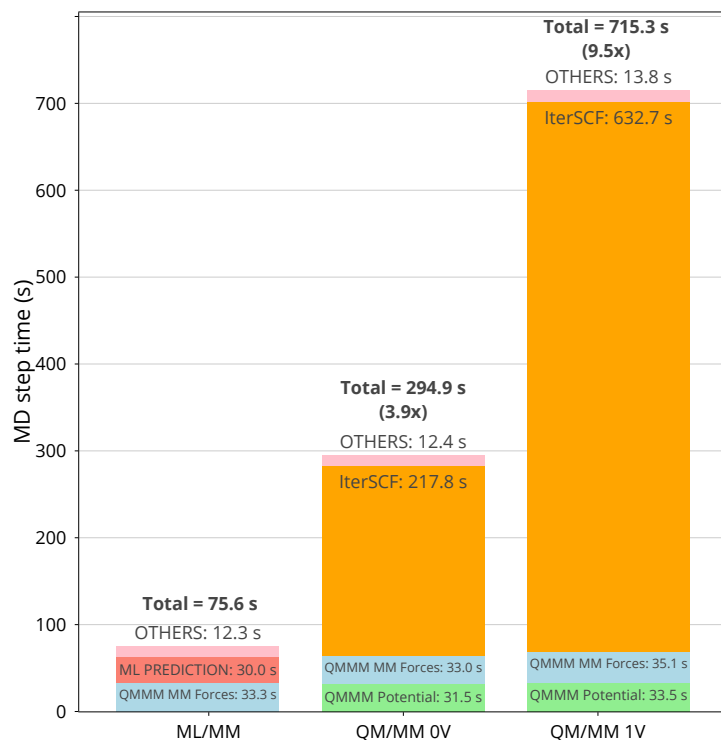

Figure 3.6: Timings for a single MD step in the ML accelerated dynamics (ML/MM) and in the regular QM/MM dynamics at 0V and 1V. TranSIESTA is run in parallel on 40 CPUs while the ML model is run on a single CPU.

There are then three main sections that take time in the ML accelerated dynamics.

- **ML prediction**: The time that this step takes can be reduced by (from simplest to hardest): running the model on GPU, using smaller models, optimizing the Graph2Mat part of the model which takes half of the time or parallelizing the model over CPUs.

- **QMMM MM forces**: This is the computation of the electrostatic MM forces due to the QM electron density. For the particular case of the capacitor, forces for MM atoms that are far from the electrodes are only due to the electric field caused by the electrode excess charge, which is constant. The field could be computed once and applied to all atoms, instead of computing the field atom by atom as it is done now.

- **Others**: This part involves initialization of several things in SIESTA and it would be hard to reduce, although there are some things that are not needed for the ML accelerated dynamics. For production runs, the best bet would probably be to avoid using SIESTA in the ML accelerated dynamics and implement the computation of the electrostatic forces in a simpler code optimized for the task.

3.5 ML/MM dynamics at 0 V

The first challenge that I set myself to solve is accelerating the QM/MM dynamics of the system at 0V. The problem is already challenging enough without voltage, and the experience at 0V would be for sure useful to tackle the problem at a finite voltage. In general, since the QM part of the simulation is replaced by a ML model, I call this approach ML/MM.

3.5.1 Charge density decomposition

As I discussed when introducing the QM/MM system in section 3.1, the fact that the simulations keep gold atoms fixed allows taking a reference density and computing only its response to the water molecules. In this section I analyze in detail the consequences of decomposing the density like that. I will split the total density $\rho(r)$ in two parts: the density of the capacitor without water $\rho(r)_{vacuum}$ and the difference between the total density and the vacuum density, which corresponds to the response induced by water, $\Delta\rho(r)_{water}$. This splitting is illustrated in Figure 3.7.
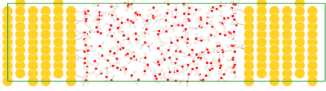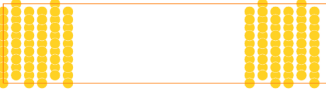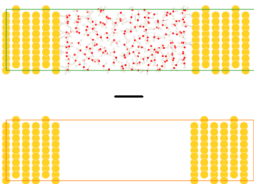

Figure 3.7: Charge density decomposition at 0V

Of course, the vacuum density doesn't change during the dynamics, so it is a constant reference. The only thing that changes is $\Delta\rho(r)_{water}$. Therefore, it is only this delta that the ML model needs to learn. By looking at the magnitude of the means of both density contributions, it is clear that the response to water is a small contribution to the total density. Since the magnitude of the vacuum density is 1000x bigger than the response to water, the first impulse might be to think that the "importance" of each of them is also determined by this factor. If this was the case, the effect of $\Delta\rho(r)_{water}$ would be negligible and we wouldn't even need to learn it. This is not quite true because of the particular distribution of the response, which I show in Figure 3.8. The response is very localized at the surface, the most important region in terms of the influence it has on water, and it does not propagate into the bulk of the electrode. Furthermore, the response involves a charge excess of opposite sign on each surface, which will create a constant electric field between plates and therefore will affect bulk water molecules. Interestingly, the display of the response also illustrates the effect of short-range interactions. If it weren't for short-range interactions, the response would be uniform over the surface, as the excess charge would distribute uniformly. However, not only the response is not uniform, but there are parts of the same surface with different signs of the response. An interesting question that I raised in the previous section and that I will not solve just yet is whether the excess charge is simply an offset that is uniformly distributed over the surface or if it is affected by these local perturbations. In simpler words, whether the short-range and long-range interactions can be decoupled.
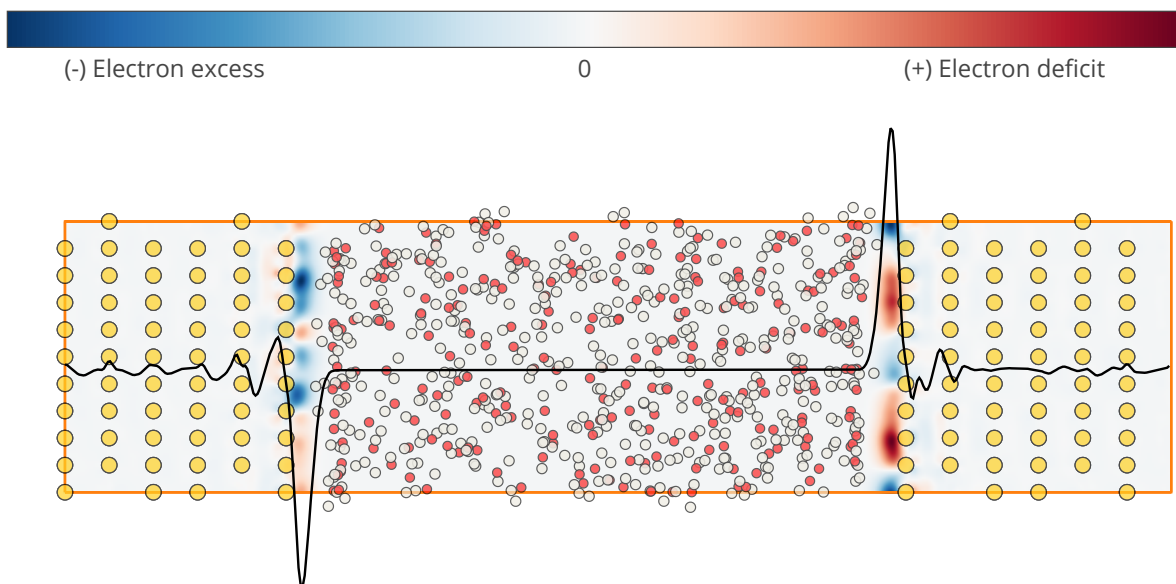
Figure 3.8: Charge density response to water, $\Delta\rho_{water}$, for a particular example. Scale is not shown to focus simply on the shape of the response. The black line shows the response averaged over the plane of the capacitor surface, showing that there is a charge excess of opposite sign on each surface. The response averaged only over one axis is shown in a colorscale, which displays that the response is not uniform over the surface due to short-range interactions.

Analyzing the electron density is a nice exercise that helps understanding the system. However, the final goal is to run molecular dynamics with the forces resulting from the predicted density. It is therefore very important to have a clear picture of how the electron density translates into atomic forces. In Figure 3.9 I show a decomposition of the MM atomic forces based on whether they are already present with the vacuum density or they are directly a result of the response to water. The histograms show that the forces caused by $\Delta\rho(r)_{water}$ are around one order of magnitude smaller than the total forces. This is very important, because it means that the ML model is only responsible for a small correction to the forces, which will minimize the impact of prediction errors. To have an even better picture of what the contribution of $\Delta\rho(r)_{water}$ to the forces is, in Figure 3.10 I illustrate the magnitude of each cartesian component of the forces on the atoms of the system for an example configuration. The X and Y components of the forces (the surface in-plane directions) are only relevant very close to the surface and vanish in the bulk. The Z component is also most relevant near the surface, where there is great variability. However, it doesn't vanish in the bulk, where the constant electric field results in a constant force in O and H atoms (according to their respective charge) as one would expect.
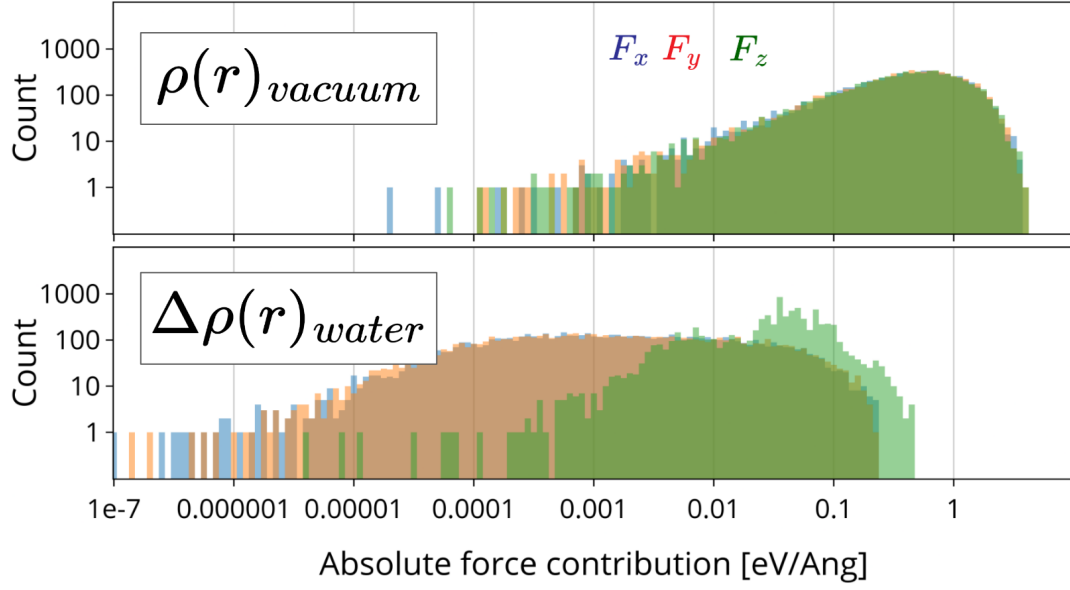
Figure 3.9: Histograms of MM atomic forces decomposition at 0V. At the top, the forces computed with the vacuum density, which contain not only electrostatic forces but also dispersion forces. At the bottom plot, the forces due to the electron density response of the electrode to water. In each plot, the cartesian components of the forces are split in three different colors. Blue (x) and red (y) are the forces in the directions of the plane of the capacitor, while green (z) is the force in the direction across the capacitor.
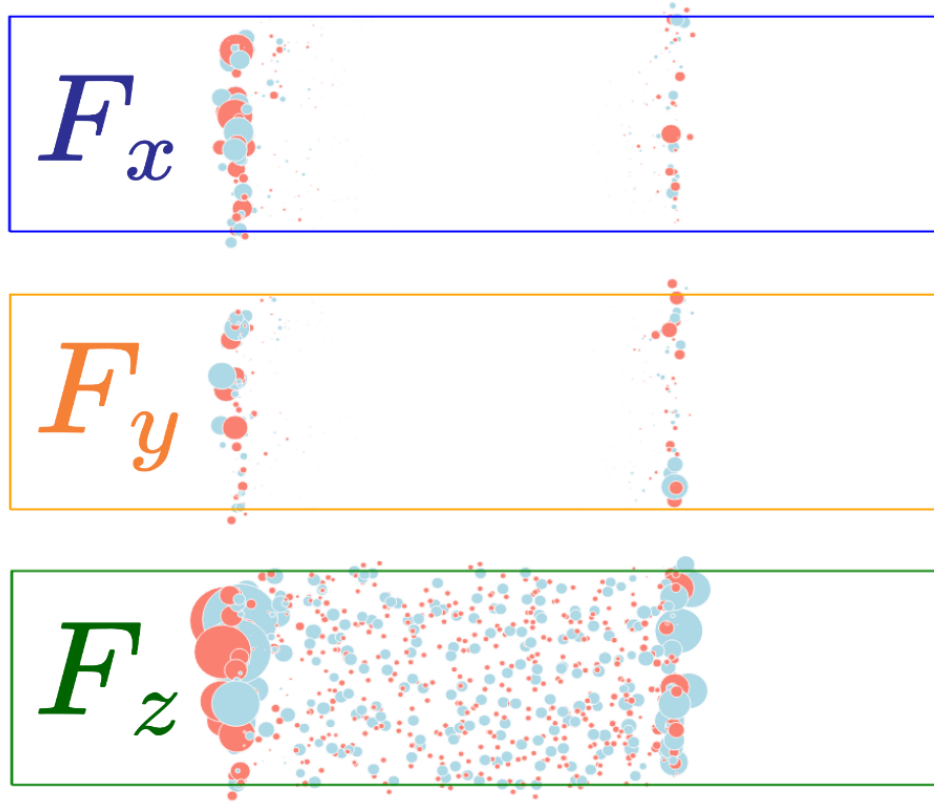

Figure 3.10: Forces caused by $\Delta\rho(r)_{water}$ on an example configuration. Circles represent water atoms inside the unit cell and are styled by the forces that each atom experiences. For each cartesian component of the forces, the magnitude of the force determines the size of the atom, while the color just shows the sign of the force (red for positive and blue for negative).

3.5.2 Results

Having characterized the system, I trained a ML model to learn $\Delta\rho(r)_{water}$ feeling confident that it was the best choice. Before running molecular dynamics with the trained model, I used the test set to evaluate the performance of the model. In Figure 3.11 I compare the ML force errors on the test set with the atomic forces of the system. The histograms show that the magnitude of the errors is around one order of magnitude smaller than the forces that ML is responsible for (those coming from $\Delta\rho(r)_{water}$), and more than 2 orders of magnitude smaller than the total forces. This implies that the errors are less than 1% of the total forces, which sounds very good. However, small errors can drive dynamics off if they are not random. The only way of assessing if the ML predictions will generate the appropriate dynamics is by running the dynamics, which is what I will do next.
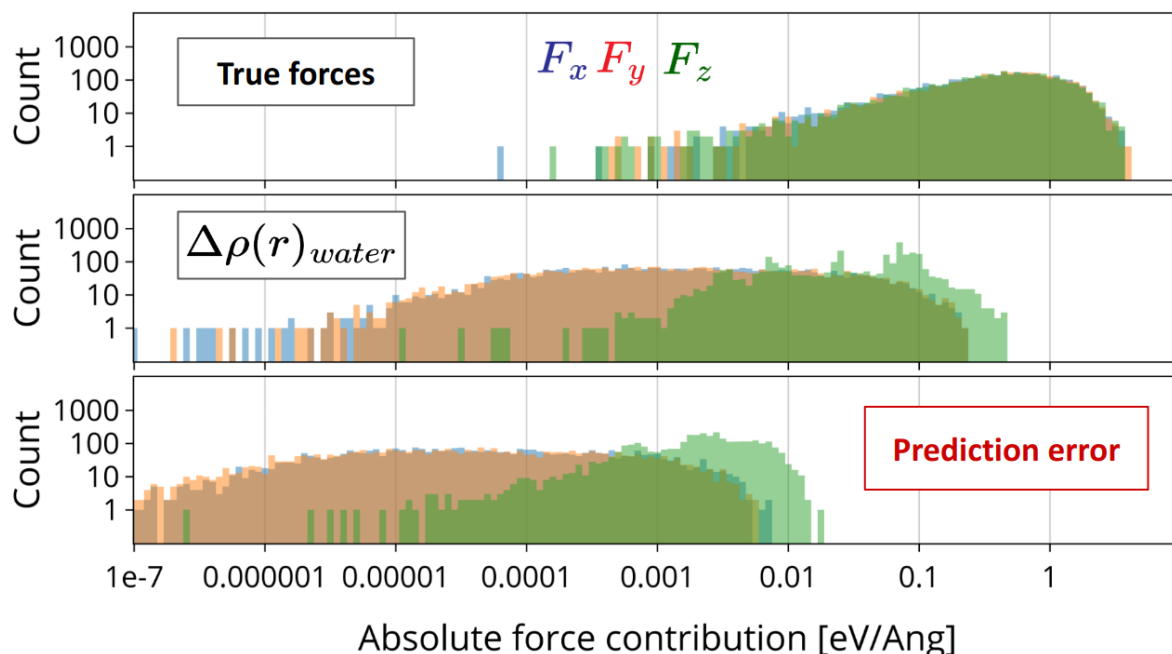

Figure 3.11: Force errors obtained on the test set (bottom plot), compared to the total forces on the system (top) and the contribution of $\Delta\rho(r)_{water}$ (middle). In each plot, the cartesian components of the forces are split in three different colors. Blue (x) and red (y) are the forces in the directions of the plane of the capacitor, while green (z) is the force in the direction across the capacitor.

With the dynamics generated from the ML model, the question now is: "How do I assess if the dynamics are correct?". This is a question that sounds simple but is not. When the dynamics are completely off, it is easy to see. But when there are small drifts in the dynamics, that is not so easy to identify. Ultimately, one must ask: "What do I want to see in the dynamics?". Because the things that you want the ML dynamics to reproduce are probably the things worth checking. In my case, I have to go back to the original problem postulated in chapter 1 to find my answer. For electrochemistry, energy level shifts at the interface are one of the most important things. It is not easy to access the oscillations in energy level shifts, as I would need to run NEGF/DFT calculations for each snapshot. But they have a linear dependence with the total net dipole of water, and that is easy to measure along the dynamics. Therefore I decided to look at how the total dipole of water evolves along the dynamics.

In Figure 3.12 I show the dynamics of the total dipole of water. I compare the dynamics obtained using the ML predictions with (1) the "true" dynamics obtained from QM/MM and (2) the dynamics obtained if the electrode doesn't respond at all to water. The three dynamics start all from the same configuration with a given net dipole. As discussed in subsection 1.2.4 the water dipole creates a field that goes against its own existence (a depolarizing field). With no external help, the dipole feels forces that drives its reduction. This is what we see in the dynamics with no response from the electrode. The dipole quickly decays to zero, and then shows small oscillations around it. If the metallic electrodes respond with the opposite field, however, they can partially stabilize the dipole, making it last for longer. This is what the QM/MM dynamics ("true" response) show. In fact, guided by the characterization of the data in Figure 3.2a it is expected that the dipole does not see full oscillations in the 3ps dynamics that

I show. There is (luckily!) not much else to add to comment the dynamics with the predicted response. The dynamics start exactly like the true QM/MM dynamics and then they diverge a little, but overall they look very similar. However, the dynamics is flirting with the edge of the training dipole distribution and this might be affecting the performance. It seems that the QM/MM and ML/MM dynamics start to diverge a little more once they get to the edge of the training distribution. But the divergence might be just a matter of MD time. Further studies should try to clarify this point.
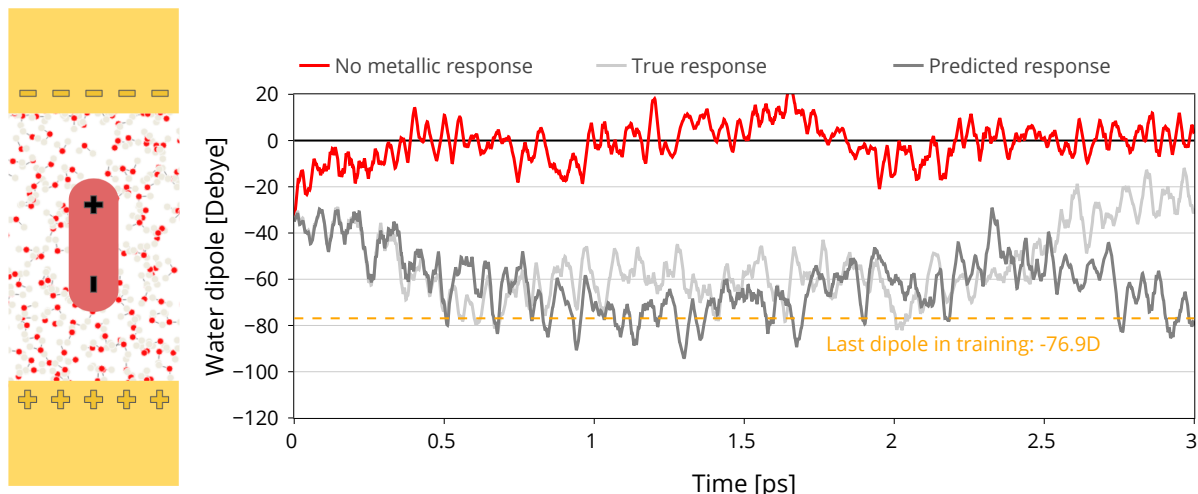

Figure 3.12: ML/MM dynamics at 0 V. The dynamics of the total dipole of water are compared for three different cases: no metallic response to water, QM/MM dynamics (true response) and ML/MM dynamics (predicted response). The dashed orange line indicates the edge of the training dipole distribution.

The evidence shows that the ML/MM dynamics seem to be able to reproduce the QM/MM dynamics at 0V, although they may benefit from a wider distribution of training dipoles. With the proof of concept done at 0V I can now move to the more challenging problem of ML/MM dynamics at 1V.

3.6 ML/MM dynamics at 1 V

In this section, I will follow similar steps as in the ML/MM at 0V section. First I will decompose the charge density and understand the part that needs to be learned, and then I will move on to the results on the test set and the dynamics. Some details that are the same as for 0V will be omitted to avoid repetition.

3.6.1 Charge density decomposition

The approach that I take to decompose the charge density for a given voltage is to take the decomposition at 0V and keep adding more contributions on top of it until I get the total density. The final decomposition is illustrated in Figure 3.13. The first two components, $\rho(r)^{0V}_{vacuum}$ and $\Delta\rho(r)_{water}$ together make the 0V charge density. Now, I need to add the effect of the voltage. The first contribution that I add is the effect of voltage in vacuum $\Delta\rho(r)_{voltage}$. The rationale for adding this contribution is that it is a fixed reference during the dynamics and we don't need to learn it. After adding this contribution, we have the isolated effects of water $\Delta\rho(r)_{water}$ and voltage $\Delta\rho(r)_{voltage}$. The rest of the density are the effects of the interaction between water and voltage, $\Delta\rho(r)_{water+voltage}$. Figure 3.13 shows that this second order contribution is extremely small, two orders of magnitude smaller than $\Delta\rho(r)_{water}$. The first impression is therefore that maybe it can be disregarded. But of course, the magnitude of the forces caused by this contribution will be the deciding factor. Nevertheless, it does not hurt to wonder why is this contribution so small in the QM/MM simulations. There are two main reasons that I can think of:

- The MM water doesn't have polarizable charge, and therefore its optical permittivity is the same as the vacuum permittivity. In other words, the electrode excess charge creates the same potential drop between plates in vacuum as in the MM water.

- There is no charge transfer between water and the electrodes, which could depend on voltage.
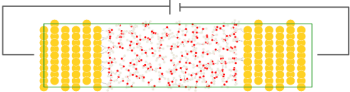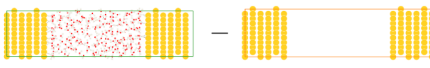

Figure 3.13: Charge density decomposition at finite voltage.

With that said, let's leave speculation aside and look at the facts. In Figure 3.14 I show the effect of each contribution on the atomic forces, just as I did for the case at 0V. I will comment only on the two new contributions, $\Delta\rho(r)_{voltage}$ and $\Delta\rho(r)_{water+voltage}$. The effect of voltage in vacuum is basically a constant field in the z direction, and it is therefore not surprising that it mainly results in forces in the z direction. In fact, the force is the same for all Oxygen atoms and also the same for all Hydrogen atoms, hence the histogram only shows two bins. Although the force that $\Delta\rho(r)_{voltage}$ induces is quite small,

considering this density contribution is "for free" (fixed reference), so there is no reason to disregard it. The second order contribution $\Delta\rho(r)_{water+voltage}$ is a different story. To take it into account, we would need to learn it. Its magnitude is at least three orders of magnitude smaller than the total forces. It is also one order of magnitude smaller than the errors that I reported in the previous section for the predictions of the forces caused by $\Delta\rho(r)_{water}$. Therefore, for these simulations $\Delta\rho(r)_{water+voltage}$ seems safe to disregard. I made the choice to disregard it, and the results that I will show next seem to support this choice.
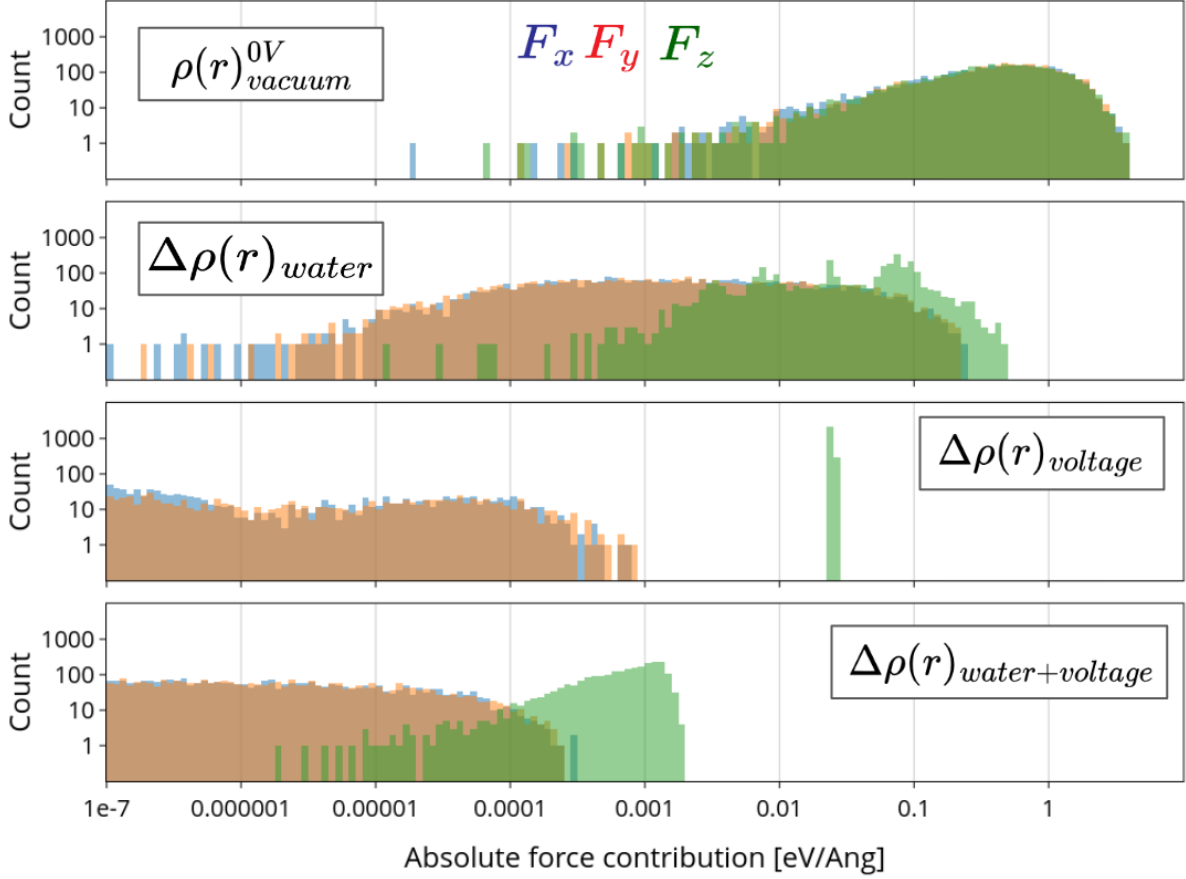

Figure 3.14: Force decomposition at 1V

The reader might have already realised that disregarding the second order contribution $\Delta\rho(r)_{water+voltage}$ has huge practical implications. The only contribution that needs to be learned is $\Delta\rho(r)_{water}$. And we already have a ML model that can predict that! The model that I trained to run molecular dynamics at 0V is the same that I need to run molecular dynamics at 1V. In general, this means that a single model can be trained on 0V data and be used to run dynamics for this system at any voltage. This sounds beautiful, so when I first thought about it I was very excited to see if the results would support this idea. Unlike me, the reader won't have to wait for one day to see if everything holds up, as I will begin presenting the results as soon as I finish this sentence.

3.6.2 Results

As I did in the 0V case, I will begin by showing the results of the ML model on the test set. In Figure 3.15 I show the force errors obtained on the test set. To compute the total forces, I generated the density by adding the predicted $\Delta\rho(r)_{water}$ to the two fixed references $\rho(r)_{vacuum}^{0V}$ and $\Delta\rho(r)_{voltage}$. The force errors (on the total force!) are almost the same at 1V as they were at 0V. This is a very good sign, as the dynamics at 0V were very good.

Let's see if this translates into good dynamics. In Figure 3.16 I show the dynamics of the total dipole of water, as I did for the 0V case. When running the dynamics with no response to water, that is considering only $\rho(r)_{vacuum}^{0V}$ and $\Delta\rho(r)_{voltage}$, the dipole decays to an equilibrium value. This is similar to the dynamics with no response at 0V, except in this case the equilibrium dipole is not 0 because of
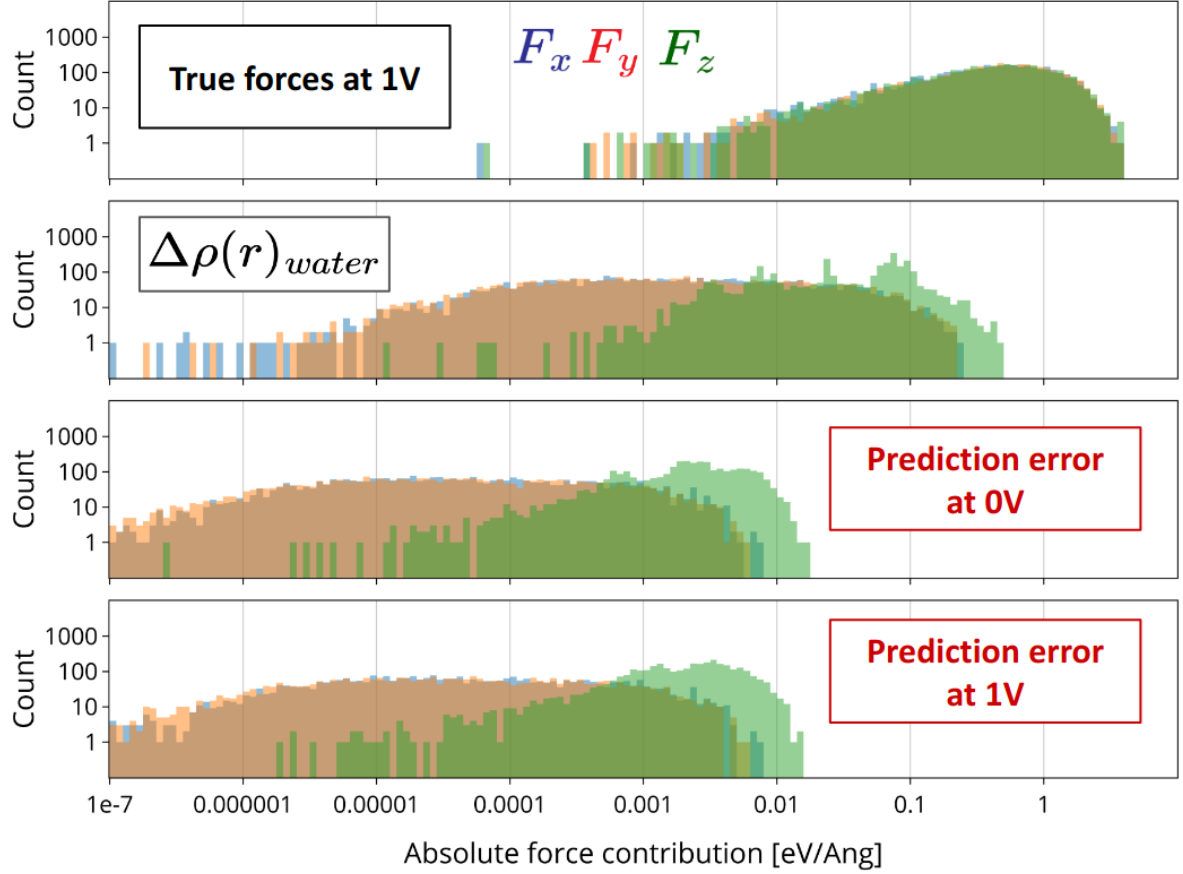
Figure 3.15: Force errors obtained on the test set.

the effect of voltage. The true QM/MM dynamics, as it did for the 0V case, stabilizes bigger dipoles than the no-response dynamics. Except again in the 1V case the dipoles are a bit bigger due to the effect of voltage. This small increase in the dipoles however has a huge impact on the ML/MM dynamics. As they go out of the training distribution, the ML/MM dynamics start to crazily diverge from the QM/MM dynamics and never recover. In fact, they end up saturating the dipole completely after 2.5 ps of dynamics. In this case, no further studies are needed to conclude that getting out of the training distribution broke the dynamics. Since they have clearly gone wrong, I can do a post-mortem analysis to understand if I could have identified the moment where the dynamics started to diverge.

Obviously, it would have been very easy to identify that the dynamics is going out of the distribution by simply looking at the total dipole. But going out of distribution does not necessarily mean that the ML model will completely fail. The model can have good extrapolation capabilities up until a certain point. This is a great case to put the concepts that I discussed in subsection 2.2.4 to the test. In that section, I introduced how the error that the model makes on the total number of electrons can be used to estimate the uncertainty of the model. Figure 3.17 shows the evolution of this proposed uncertainty metric along the dynamics. Clearly, the error on the number of electrons becomes huge (10 electrons) once the dynamics have gone completely wrong. This means that detecting when the model is very uncertain is an easy job for the uncertainty metric. The question is now: how soon could I have detected that the model was showing signs of failure? Realistically, the first slightly anomalous peak of the uncertainty metric is at 1.125 ps, when the error on the number of electrons becomes 0.081. At that point, the dynamics is at a dipole of around -110 D, well outside of the training distribution. It has been 0.4 ps (400 steps) since the dynamics first went out of distribution and started to significantly diverge from the QM/MM dynamics. Therefore, it might seem that the uncertainty metric detected the failure with some delay. In the defense of the uncertainty metric, I must say that although they had gone out of distribution, the ML/MM dynamics had not gone completely wrong until that point. If we continue advancing along the dynamics, there are not significantly bigger peaks until 0.3 ps later, when the error in the number of electrons surpasses the 0.1 mark. However, failure could be identified by looking at
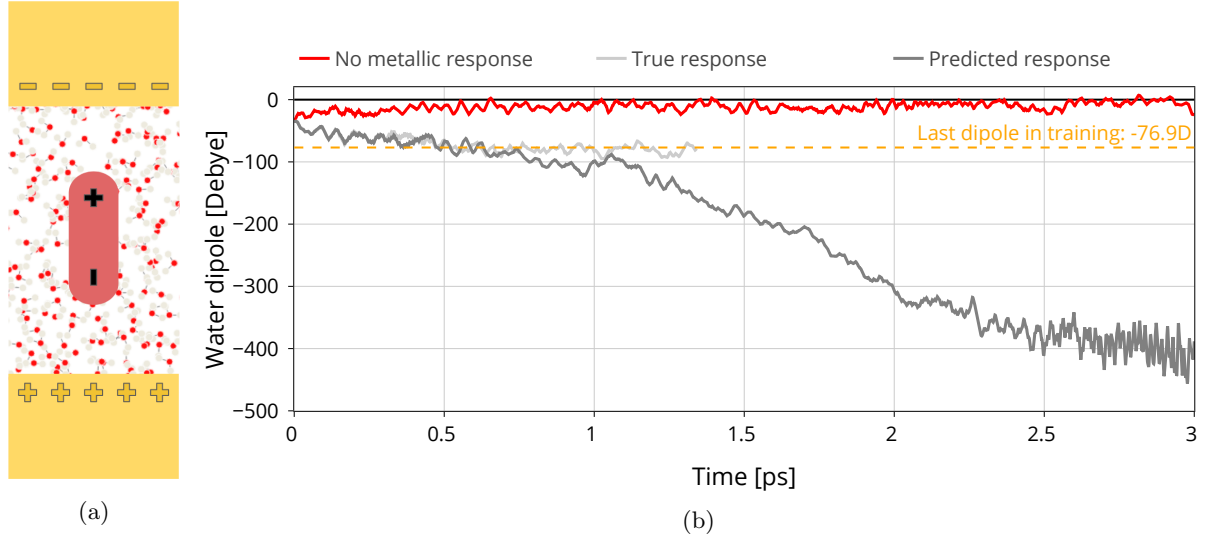
Figure 3.16: ML/MM dynamics at 1 V. The dynamics of the total dipole of water are compared for three different cases: no metallic response to water, QM/MM dynamics (true response) and ML/MM dynamics (predicted response). The dashed orange line indicates the edge of the training dipole distribution.

the local minima of the errors, which have significantly increased. With this, I reach the conclusion that using the error on the number of electrons and some reasonable heuristics, it is possible to detect "soon enough" when the ML/MM dynamics are starting to fail. In an active learning framework, this could be used to add new configurations to the training set and retrain the model before continuing the dynamics. Another heuristic to determine in an active learning framework would be whether it makes sense to rewind the dynamics back some amount of steps and continue from there. For the particular case of this system, the dynamics would probably "recover" from the failure without rewinding, as the dipole would simply be too big and the improved dynamics would reduce it with no problem. However, the failure might affect statistics if the dynamics are not too long. That is to say, there are choices to be made.
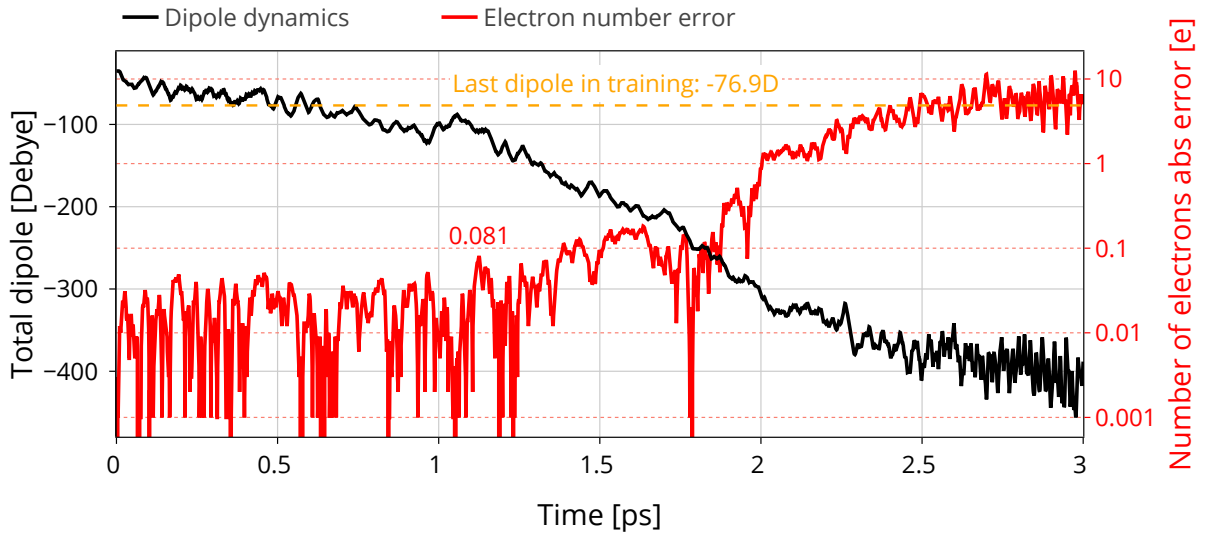

Figure 3.17: Tracking unceraity along the ML/MM dynamics at 1 V. The black line shows the dipole dynamics (refering to the right axis), while the red line shows the error on the number of electrons (refering to the left axis), both as a function of MD time.

The discussion about uncertainty would not be complete if I didn't comment on the small region

between 1.6ps and 1.8ps where the dynamics are already completely wrong but the uncertainty metric goes to 0. If the reader, like me, has felt confused by this, Figure 3.18b explains what is happening. It shows the correlation of the total dipole with the error in the number of electrons. While dipoles are within the training distribution (from -20 to -79.6D) the errors are contained within a certain range. For dipoles to the left of -79.6D, the model is extrapolating. Initially it seems like the model can extrapolate without affecting the error up to approximately -100D. Then, the extrapolation behavior of the model starts to show clear signs of failure. However, the relation between the dipole and the error is not monotonic. The error goes into negative values before exploding up into big positive values. I leave for future work the study of whether the predicted densities are correct when the error crosses 0 or the total charge is correct just by pure chance. I also leave for future work understanding how good are the out-of-distribution predictions which give good uncertainty metric values (up to dipoles of -100D).
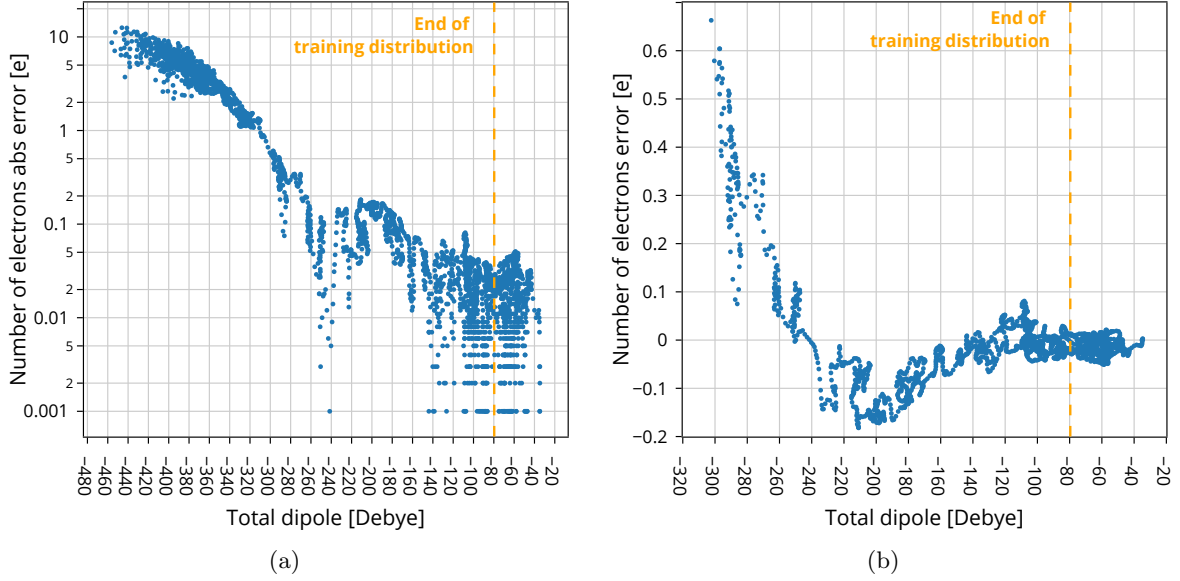

Figure 3.18: Correlation of the error on the number of electrons with the total dipole of water. (a) shows the number of electrons absolute error in a log scale, while (b) shows the error in a linear scale for a reduced range of dipoles so that the crossing at y=0 can be observed.

Leaving uncertainty aside, I would like to make a final comment on the great success that the narrow training set has eclipsed. Before reaching the point where the dynamics go out of distribution, the ML/MM dynamics in Figure 3.16 are very good. This means that our approximation to discard the second order contribution $\Delta\rho(r)_{water+voltage}$ was correct in practical terms. This has two important implications:

- We can use a single ML model to run dynamics at any voltage for this system.

- It hints to the fact that long-range and short-range interactions can be decoupled. The effect of voltage, which we have decoupled from everything else, is the same as the effect of the total water dipole. Both create the need to accumulate charges at the surface of the electrodes to satisfy the boundary conditions. If we use the terms discussed in subsection 1.2.4, both V (the voltage) and χ_o (water's dipole) determine $\Delta\psi$ (the potential drop caused by the charge accumulation at the surface). The expression in fact was $\Delta\psi = V - \chi_o$. The point that I want to make is that the long-range interactions of water can probably be decoupled from the short-range interactions and unified with the effect of voltage.

3.7 Conclusions and future perspectives for ML/MM

In my opinion, this chapter has shown that the future of ML/MM is bright. Although machine learning force fields could also be used to run this kind of dynamics, the fact that the electrostatic forces are computed using the correct physics equations probably makes the ML/MM approach more robust. I make this claim due to the fact that the ML models used in this chapter have been trained on a small uncurated training set of only 180 configurations, and give close to perfect dynamics. Further studies could explore whether smaller datasets are enough to produce the same results. I would not be surprised if the dataset could be reduced, since the bottleneck of the performance on the ML/MM dynamics that I have shown is the range of total dipoles in the dataset. On the other hand, I have shown that the timings of my ML/MM calculations are far from competitive with force fields, which run at rates of tenths of miliseconds per step, as opposed to 70s in my proof of concept ML/MM. However, I have discussed that there is plenty of room for performance optimization. I envision that this ML/MM approach could get to timings that are a good compromise considering the increased robustness. In fact Grisafi et al. have shown ML/MM calculations running at 0.5s per step [50], for similarly sized systems.

Whether ML force fields or ML/MM calculations are chosen as the better tool to simulate electrified interfaces, this chapter has provided a nice path to make the task simpler by decomposing the problem. I have shown that by decoupling the effect of voltage, one can use the same model to run the dynamics at different voltages. Also, as discussed by the end of the last section, it is likely that the effect of water on the electrodes can be decomposed into short-range and long-range effects. Not only that, but the long-range effects of water's net dipole can be unified with the effects of voltage and predicted with a simple linear interpolation. The workflow that I envision for the future is shown in Figure 3.19. By decomposing the problem in this way, the ML model is released of the responsibility of predicting long-range effects that water has on the metal and can focus simply on the short-range effects. This has a huge performance impact, because it means that the ML performance would scale $O(1)$ (constant time) with the distance between capacitor plates as it doesn't need to take bulk water into account. This is not particular to ML/MM calculations, since the same methodology could be used for ML force fields. Although the proposed workflow sounds nice, it might however not work when the area of the electrodes is big compared to the distance between them. In that case, different parts of the electrode surface might see different water dipoles due to the formation of domains that do not average out. This would result in a non-uniform surface charge accumulation. As a result, the effect of the dipole could not be unified with the effect of voltage (which always results in a uniform surface charge accumulation) and should be treated with a separate long range model for water dipoles.
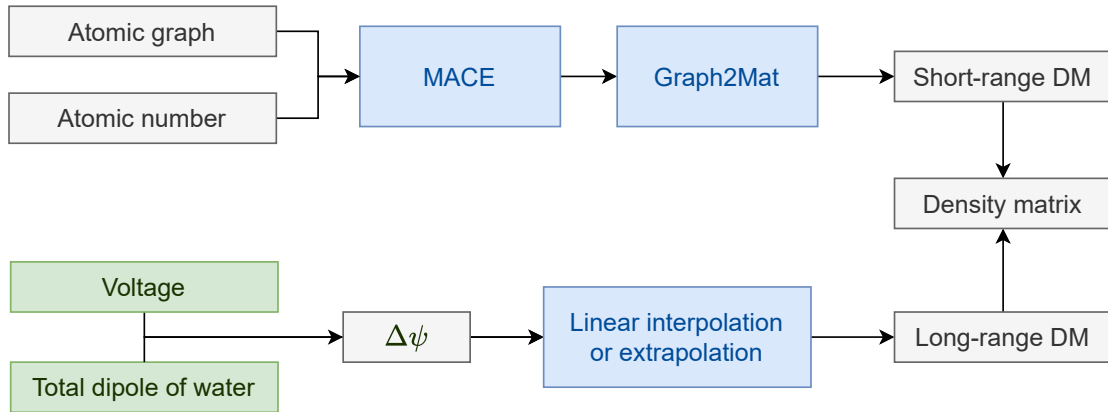
Figure 3.19: Proposed ML/MM workflow for simulating electrified interfaces when the dipole of water averaged over the capacitor's length is uniform (capacitor length much bigger than electrode unit cell area). Voltage and total dipole can then be treated equally and the long range effects of water can be computed with a simple linear interpolation.

Regardless of the specific approach used to run ML/MM dynamics, I see them as a very useful tool to characterize a medium in which reactions should happen. Purely from ML/MM simulations, perhaps taking into account the optical permittivity of water when computing the effect of voltage, one could estimate the energy shifts at the interfaces and their oscillations. Otherwise, ML/MM could be used simply as a sampling method to generate configurations for TranSIESTA calculations to characterize the

system. Using the decomposition of the density that I showed, one could also easily simulate the system under an alternating applied voltage, since it is just a matter of changing the reference density when voltage changes.

Summary of conclusions:

- ML/MM is a promising tool to simulate electrified interfaces.

- When all the medium is MM, a single ML model can be used to simulate the dynamics at different voltages.

- My ML/MM approach using MACE+Graph2Mat gives good dynamics using a very small dataset. However it needs performance optimization in order to be relevant.

- Decoupling long-range contributions from short-range contributions of the influence of water on the metallic electrodes will probably increase the performance of ML/MM calculations.

Original contributions from the author:

- Conceptualization of the decomposition of different effects in QM/MM calculations and illustration in concrete examples.

- Slight modification of MACE to use a global descriptor.

- Implementation of the workflow to run ML accelerated SIESTA molecular dynamics, using electron density predictions from a ML model running in python.

Conclusions

Summary of conclusions:

- Non Equilibrium Green's Functions can be coupled with Density Functional Theory (e.g. the TranSIESTA code) to run ab initio molecular dynamics of solid-liquid interfaces under applied voltage that are relevant for electrochemistry.

- The density matrix can be used as a machine learning target to predict electron densities and accelerate general DFT calculations. I have shown this using a model that involves a well established atomic environment representation (MACE) and a newly developed method to convert the representations to matrices (Graph2Mat).

- The machine learning models designed in this thesis can be used to convert QM/MM molecular dynamics of electrified solid-liquid interfaces into ML/MM molecular dynamics which are faster and can reproduce the results of QM/MM.

Original contributions from the author:

- Thorough exploration of the capabilities of TranSIESTA for electrochemical systems, using some specific analysis tools, including non-trivial electronic structure algorithms, that have been developed by the author and have been made available to the community in a widely used open source python package (*sisl*).

- Developement, in conjunction with collaborators, of the Graph2Mat method to enable easy machine learning of electronic structure matrices at scale. The method, along with documentation and tutorials, has been made available to the community in the form of an open source python package, *graph2mat*, which is published on PyPi for ease of use.

- Design and realization of a workflow to run ML accelerated molecular dynamics within the open source DFT SIESTA code. The workflow can be easily reproduced using the mainstream published SIESTA version, which contains some modifications done by the author.

A single PhD thesis is not going to magically make computational electrochemistry an easy matter. But I hope that with my work here I have put another stone on the path towards the understanding and rational design of electrochemical processes which will be key for the future of society. If anything, I hope that I have presented my work in a way that is clear and interesting enough to inspire others to take the torch and continue the journey to places that I have not been able to reach or even imagine. Much of the success that I have had throughout my PhD has been based on understanding the code that others have written (TranSIESTA, MACE) and reusing it to do new things that the original authors did not have in mind when they wrote it. I could only do that because there was comprehensive documentation and the code was clear (as clear as code can be) and open source. With my novel implementations, I have tried to follow the same principles. The day that someone, inspired by what I have reported here and in the documentation of my implemenations, finds a use for it that I had never imagined, I will deem this work a success.

To release the reader from the responsibility of imagining new paths from scratch, let me use the following lines to draw my own picture of the future, with which the reader might find some overlap. I see a future where the community has adopted TranSIESTA to simulate electrochemical systems. In this future, scientists will be able to simulate full electrochemical cells. But it will not be all thanks to TranSIESTA. Machine learning will play a big role in allowing TranSIESTA to access the realm where a single simulation contains the metallic electrodes, the electrolyte and reactive entities diffusing through the electrolyte. In this future, and this is the hard part, people will not have forgotten about electrons.

Scientists will have found the most optimal way to interface machine learning with TranSIESTA so that TranSIESTA can give electronic insights with a confidence and a level of detail that machine learning can not provide. The community will use this electronic information to propose electrochemical mechanisms and rationally design new materials and processes, accelerating the rate of discovery and innovation in electrochemistry.

It was very easy for me to paint this future, because I am confident that the community will get to some place like that, be it with TranSIESTA or with a similar future code. The question is how long it will take, or what will be the singularity that will suddenly get us very close to it. And of course, I can not answer those questions. But I can propose short term avenues that might eventually trigger that singularity in the mind of some smart researcher. In my opinion, the ML/MM path looks promising in the short term. Something as simple as using the TranSIESTA based ML/MM that I presented in this thesis to simulate the system of water between gold electrodes, but now big enough so that an interesting molecule or ion can be included in realistic concentrations, would already be a big step. It is very important that in that new system we aim to describe how electronic levels behave. As of now, I have no idea to keep the electronic description other than running NEGF for the particular snapshots. But since MACE+Graph2Mat already provides a good estimate of the density, one could try to compute the levels from the predicted densities in a non self-consistent way (i.e. with just 1 SCF step). Regarding the MM part, the classical force fields that I used in this work could of course be replaced by ML force fields in order to increase the flexibility on the liquids that can be simulated. Eventually, it is possible that an explicit prediction of the electron density like the one that I get with MACE+Graph2Mat, is no longer needed because a new ML model that is reliable enough and more efficient comes up. But for every step that we take in the ML direction, we must make sure that the electronic levels can still be described. Only then will scientists be able to discuss the emergent electrochemistry in a rational way, and not depend purely on data-driven approaches.

Without the aim of discouraging future practitioners, I would not end this work with peace of mind if I did not warn the traveller about the obstacles that they might find along the way. Gold is not the only material that I have tried to use as an electrode in TranSIESTA calculations. I have been successful in using other materials like silver and copper, but others have proved much more difficult to converge, like Platinum and graphite. It is possible that MACE+Graph2Mat can also help with these convergence problems by providing good initial guesses that ensure a path to convergence is found. I have to admit also that ML/MM was not my first choice for running ML accelerated dynamics. Initially I wanted to run molecular dynamics that mimic full QM simulations, not QM/MM. However, I encountered the problem that forces can not be computed directly from the density in SIESTA. Due to the non-orthogonality of the basis, one needs to go through 1 SCF step in order to get the forces, with the associated overhead that this represents. I felt like I could never justify this as an alternative to ML force fields, so I went with the ML/MM calculations that I showcased. Nevertheless, I am not sure that there isn't a case in which running 1 SCF step dynamics with MACE+Graph2Mat is worth it. In any case, every cloud has a silver lining, and exploring ML/MM has led me to think that perhaps multiscale methods are the answer. After all, who cares about the electrons at the bulk of water, if what we are interested in is electrochemistry at the interface?

At this point, I could continue describing details of simulations that I have in mind, but I have decided to pass a Gaussian filter over my ideas before writing these lines. The aim is that the image is blurry enough for the reader to not stick to my own details. Hopefully each reader will reconstruct the image with their own algorithm, and the methods presented here will find new flavours and a variety of new applications.

Bibliography

[1] M. Winter, "The solid electrolyte interphase–the most important and the least understood solid electrolyte in rechargeable li batteries," *Zeitschrift für physikalische Chemie*, vol. 223, no. 10-11, pp. 1395–1406, 2009.

[2] A. Wang, S. Kadam, H. Li, S. Shi, and Y. Qi, "Review on modeling of the anode solid electrolyte interphase (sei) for lithium-ion batteries," *NPJ Computational materials*, vol. 4, no. 1, p. 15, 2018.

[3] M. H. Hansen and J. Rossmeisl, "Ph in grand canonical statistics of an electrochemical interface," *The Journal of Physical Chemistry C*, vol. 120, no. 51, pp. 29 135–29 143, 2016.

[4] F. Deißenbeck and S. Wippermann, "Dielectric properties of nanoconfined water from ab initio thermopotentiostat molecular dynamics," *Journal of Chemical Theory and Computation*, vol. 19, no. 3, pp. 1035–1043, 2023.

[5] M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, "Density-functional method for nonequilibrium electron transport," *Physical Review B*, vol. 65, no. 16, p. 165 401, 2002.

[6] N. Papior, N. Lorente, T. Frederiksen, A. García, and M. Brandbyge, "Improvements on non-equilibrium and transport green function techniques: The next-generation transiesta," *Computer Physics Communications*, vol. 212, pp. 8–24, 2017.

[7] K. Stokbro, J. Taylor, M. Brandbyge, J.-L. Mozos, and P. Ordejon, "Theoretical study of the nonlinear conductance of di-thiol benzene coupled to au (1 1 1) surfaces via thiol and thiolate bonds," *Computational Materials Science*, vol. 27, no. 1-2, pp. 151–160, 2003.

[8] G. Calogero, N. R. Papior, B. Kretz, A. Garcia-Lekue, T. Frederiksen, and M. Brandbyge, "Electron transport in nanoporous graphene: Probing the talbot effect," *Nano letters*, vol. 19, no. 1, pp. 576–581, 2018.

[9] G. Calogero, I. Alcón, N. Papior, A.-P. Jauho, and M. Brandbyge, "Quantum interference engineering of nanoporous graphene for carbon nanocircuitry," *Journal of the American Chemical Society*, vol. 141, no. 33, pp. 13 081–13 088, 2019.

[10] J. Böckris, A. Reddy, and M. Gamboa-Aldeco, "Modern electrochemistry, 2a: Fundamentals of electrodics," *Modern Electrochemistry, 2A: Fundamentals of Electrodics*, 2000.

[11] L. Fumagalli *et al.*, "Anomalously low dielectric constant of confined water," *Science*, vol. 360, no. 6395, pp. 1339–1342, 2018.

[12] F. Deißenbeck, C. Freysoldt, M. Todorova, J. Neugebauer, and S. Wippermann, "Dielectric properties of nanoconfined water: A canonical thermopotentiostat approach," *Physical Review Letters*, vol. 126, no. 13, p. 136 803, 2021.

[13] N. Papior, *Sisl: V0.15.1*, 2024. DOI: `10.5281/zenodo.597181`. [Online]. Available: `https://doi.org/10.5281/zenodo.597181`.

[14] S. Batzner *et al.*, "E (3)-equivariant graph neural networks for data-efficient and accurate inter-atomic potentials," *Nature communications*, vol. 13, no. 1, p. 2453, 2022.

[15] I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csányi, "Mace: Higher order equivariant message passing neural networks for fast and accurate force fields," *Advances in Neural Information Processing Systems*, vol. 35, pp. 11 423–11 436, 2022.

[16] D. P. Kovács *et al.*, "Mace-off23: Transferable machine learning force fields for organic molecules," *arXiv preprint arXiv:2312.15211*, 2023.

[17] É. Polack, G. Dusson, B. Stamm, and F. Lipparini, "Grassmann extrapolation of density matrices for born–oppenheimer molecular dynamics," *Journal of Chemical Theory and Computation*, vol. 17, no. 11, pp. 6965–6973, 2021.

[18] A. Niklasson, M. Challacombe, C. Tymczak, and K. Németh, "Trace correcting density matrix extrapolation in self-consistent geometry optimization," *The Journal of chemical physics*, vol. 132, no. 12, 2010.

[19] F. Pes, É. Polack, P. Mazzeo, G. Dusson, B. Stamm, and F. Lipparini, "A quasi time-reversible scheme based on density matrix extrapolation on the grassmann manifold for born–oppenheimer molecular dynamics," *The Journal of Physical Chemistry Letters*, vol. 14, no. 43, pp. 9720–9726, 2023.

[20] J. A. Rackers, L. Tecot, M. Geiger, and T. E. Smidt, "A recipe for cracking the quantum scaling limit with machine learned electron densities," *Machine Learning: Science and Technology*, vol. 4, no. 1, p. 015 027, 2023.

[21] C. Li, O. Sharir, S. Yuan, and G. K. Chan, "Image super-resolution inspired electron density prediction," *arXiv preprint arXiv:2402.12335*, 2024.

[22] P. B. Jørgensen and A. Bhowmik, "Equivariant graph neural networks for fast electron density estimation of molecules, liquids, and solids," *npj Computational Materials*, vol. 8, no. 1, pp. 1–10, Aug. 2022, ISSN: 2057-3960. DOI: 10.1038/s41524-022-00863-y. Accessed: Nov. 8, 2022.

[23] A. Chandrasekaran, D. Kamal, R. Batra, C. Kim, L. Chen, and R. Ramprasad, "Solving the electronic structure problem with machine learning," *npj Computational Materials*, vol. 5, no. 1, p. 22, 2019.

[24] T. Koker, K. Quigley, E. Taw, K. Tibbetts, and L. Li, "Higher-order equivariant neural networks for charge density prediction in materials," *arXiv preprint arXiv:2312.05388*, 2023.

[25] A. Grisafi, A. M. Lewis, M. Rossi, and M. Ceriotti, "Electronic-structure properties from atom-centered predictions of the electron density," *Journal of Chemical Theory and Computation*, vol. 19, no. 14, pp. 4451–4460, 2022.

[26] B. Cuevas-Zuviría and L. F. Pacios, "Machine learning of analytical electron density in large molecules through message-passing," *Journal of chemical information and modeling*, vol. 61, no. 6, pp. 2658–2666, 2021.

[27] X. Shao, L. Paetow, M. E. Tuckerman, and M. Pavanello, "Machine learning electronic structure methods based on the one-electron reduced density matrix," *Nature communications*, vol. 14, no. 1, p. 6281, 2023.

[28] S. Hazra, U. Patil, and S. Sanvito, "Predicting the one-particle density matrix with machine learning," *Journal of Chemical Theory and Computation*, 2024.

[29] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, "Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons," *Physical review letters*, vol. 104, no. 13, p. 136 403, 2010.

[30] A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments," *Physical Review B—Condensed Matter and Materials Physics*, vol. 87, no. 18, p. 184 115, 2013.

[31] R. Drautz, "Atomic cluster expansion for accurate and transferable interatomic potentials," *Physical Review B*, vol. 99, no. 1, p. 014 104, 2019.

[32] M. J. Willatt, F. Musil, and M. Ceriotti, "Atom-density representations for machine learning," *The Journal of chemical physics*, vol. 150, no. 15, 2019.

[33] M. Geiger and T. Smidt, "E3nn: Euclidean neural networks," *arXiv preprint arXiv:2207.09453*, 2022.

[34] S. Pozdnyakov and M. Ceriotti, "Smooth, exact rotational symmetrization for deep learning on point clouds," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[35] A. Grisafi and M. Ceriotti, "Incorporating long-range physics in atomic-scale machine learning," *The Journal of chemical physics*, vol. 151, no. 20, 2019.

[36] J. M. Soler *et al.*, "The siesta method for ab initio order-n materials simulation," *Journal of Physics: Condensed Matter*, vol. 14, no. 11, p. 2745, 2002.

[37] A. García *et al.*, "Siesta: Recent developments and applications," *The Journal of chemical physics*, vol. 152, no. 20, 2020.

[38] M. J. Van Setten *et al.*, "The pseudodojo: Training and grading a 85 element optimized norm-conserving pseudopotential table," *Computer Physics Communications*, vol. 226, pp. 39–54, 2018.

[39] L. Ruddigkeit, R. van Deursen, L. C. Blum, and J.-L. Reymond, "Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17," en, *J. Chem. Inf. Model.*, vol. 52, no. 11, pp. 2864–2875, Nov. 2012.

[40] R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, "Quantum chemistry structures and properties of 134 kilo molecules," en, *Sci Data*, vol. 1, p. 140 022, Aug. 2014.

[41] P. B. Jørgensen and A. Bhowmik, *Ethylene carbonate molecular dynamics dataset*, Jul. 2022. DOI: 10.11583/DTU.16691825. [Online]. Available: https://doi.org/10.11583/DTU.16691825.

[42] A. Bhowmik and P. B. Jørgensen, *Density matrix, hamiltonian matrix, overlap matrix and energy density matrix of qm9 and ethylene carbonate dataset; linked to https://doi.org/10.26434/chemrxiv-2024-j4g21*, Aug. 2024. DOI: 10.11583/DTU.c.7310005. [Online]. Available: https://data.dtu.dk/collections/Density_matrix_Hamiltonian_matrix_overlap_matrix_and_energy_density_matrix_of_QM9_and_ethylene_carbonate_dataset_linked_to_https_doi_org_10_26434_chemrxiv-2024-j4g21/7310005/1.

[43] A. Musaelian *et al.*, "Learning local equivariant representations for large-scale atomistic dynamics," *Nature Communications*, vol. 14, no. 1, p. 579, 2023.

[44] K. Schütt, O. Unke, and M. Gastegger, "Equivariant message passing for the prediction of tensorial properties and molecular spectra," in *International Conference on Machine Learning*, PMLR, 2021, pp. 9377–9388.

[45] S. Zhang, Y. Liu, and L. Xie, "A universal framework for accurate and efficient geometric deep learning of molecular systems," *Scientific Reports*, vol. 13, no. 1, p. 19 171, 2023.

[46] R. S. Mulliken, "Electronic population analysis on lcao–mo molecular wave functions. i," *The Journal of chemical physics*, vol. 23, no. 10, pp. 1833–1840, 1955.

[47] C. T. Chan, K. P. Bohnen, and K. Ho, "Accelerating the convergence of force calculations in electronic-structure computations," *Physical Review B*, vol. 47, no. 8, p. 4771, 1993.

[48] I. Batatia *et al.*, "A foundation model for atomistic materials chemistry," *arXiv preprint arXiv:2401.00096*, 2023.

[49] A. Grisafi, A. Bussy, M. Salanne, and R. Vuilleumier, "Predicting the charge density response in metal electrodes," *Physical Review Materials*, vol. 7, no. 12, p. 125 403, 2023.

[50] A. Grisafi and M. Salanne, "Accelerating qm/mm simulations of electrochemical interfaces through machine learning of electronic charge densities," *arXiv preprint arXiv:2405.07370*, 2024.