Physics Area - Ph.D. course in
Theory and Numerical Simulation of Condensed Matter

HIGH-PRESSURE HYDROGEN PHASE DIAGRAM FROM QUANTUM MONTE CARLO AND MACHINE LEARNING

Candidate:
Giacomo Tenti

Supervisors:
Michele Casula
Sandro Sorella
Co-supervisors:
Kosuke Nakano
Stefano Baroni

Academic year 2023/2024

To Sandro

ACKNOWLEDGMENTS

In un percorso fatto di alti e bassi, desidero ringraziare chi ha reso i momenti difficili più lievi e quelli belli ancora più speciali. Grazie a Martina, per essere la persona che è e per ricordarmi ogni giorno quanto sono fortunato. Grazie alla mia famiglia, per avermi sempre dato libertà e supporto in ogni scelta. Grazie ad Alfredo, Cesare, Enrico e Florian, amici prima che colleghi, per gli innumerevoli pranzi e partite di ping pong insieme. Grazie a tutte le persone che ho incontrato durante il mio dottorato, sia in SISSA che fuori: Kosuke, Stefano, Andrea, Otto, Abhishek, Miguel e Shiwei. Grazie a Michele, per avermi "adottato" scientificamente in un momento difficile e per tutto ciò che mi ha insegnato. Infine, grazie a Sandro, da cui ho imparato cosa significa amare questo lavoro.

ABSTRACT

High-pressure hydrogen is of paramount importance in several fields, including planetary science, condensed matter physics, and energy production applications. Despite its significance, many properties of this system are still not fully understood, due to the difficulty of realizing the required extreme conditions in the laboratory and probing the compressed samples. Numerical results are thus extremely valuable. Quantum Monte Carlo (QMC) algorithms have been proven to be among the most effective methods for describing the physics and properties of high-pressure hydrogen, although their large computational cost limits their applicability to small systems. In this thesis, we discuss techniques that aim at combining the accuracy of QMC methods with the efficiency of machine learning potentials (MLPs). In particular, we employ the Δ-learning framework together with kernel ridge regression, and train models on the difference between QMC reference calculations and a computationally cheaper "baseline potential", which in our case was obtained with the density functional theory (DFT) method. This approach allows us to reach a higher accuracy with relatively small datasets, a crucial feature for resource-heavy algorithms like QMC. We also analyze the bias affecting both forces and pressures within the variational Monte Carlo (VMC) method when the wave function employed is not fully optimized, and propose a suitable correction.

We present two applications of our framework to high-pressure hydrogen. In the first one, we determined the deuterium Hugoniot with MLPs trained on both variational and diffusion Monte Carlo. We find a good agreement with experiments, even though our results suggest a slightly more compressible system for large pressures. In the second application, we study the hydrogen liquid-liquid phase transition (LLPT). We discuss results obtained with two MLPs trained on VMC and DFT data, respectively. For the latter, we employed MACE, a message passing neural network, to study the order of the transition in the thermodynamic limit. Our results predict a first-order transition between a defective molecular solid and an atomic liquid close to the melting line, and a liquid-liquid crossover at higher temperatures.

PUBLICATIONS

The contents of this thesis are based on the following publications:

1. Andrea Tirelli, Giacomo Tenti, Kousuke Nakano, and Sandro Sorella. "High-pressure hydrogen by machine learning and quantum Monte Carlo." In: *Phys. Rev. B* 106 (4 2022), p. L041105. DOI: 10.1103/PhysRevB.106.L041105.

2. Kousuke Nakano, Michele Casula, and Giacomo Tenti. "Efficient calculation of unbiased atomic forces in ab initio variational Monte Carlo." In: *Phys. Rev. B* 109 (20 2024), p. 205151. DOI: 10.1103/PhysRevB.109.205151.

3. Giacomo Tenti, Kousuke Nakano, Andrea Tirelli, Sandro Sorella, and Michele Casula. "Principal deuterium Hugoniot via quantum Monte Carlo and Δ-learning." In: *Phys. Rev. B* 110 (4 2024), p. L041107. DOI: 10.1103/PhysRevB.110.L041107.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AAD adjoint algorithmic differentiantion

ACE atomic cluster expansion

ACSF atom-centered symmetry functions

AGP antisymmetrized geminal power

AIMD *ab initio* molecular dynamics

AIRSS *ab initio* random structure search

AO atomic orbital

BO Born-Oppenheimer

BOMD Born-Oppenheimer MD

ccECP correlation consistent effective core potential

CEIMC coupled electron-ion Monte Carlo

CPMD Car-Parrinello molecular dynamics

DFT density functional theory

DAC diamon anvil cells

DMC diffusion Monte Carlo

EOS equation of state

FDM finite-differences method

FN fixed-node

FPS furthest point selection

FSC finite-size corrections

FT-DFT finite-temperature DFT

GEO geminal embedded orbital

GFMC Green's function Monte Carlo

GGA generalized gradient approximation

GTO Gaussian type orbital

HSE Heyd-Scuseria-Ernzerhof

JSD Jastrow correlated Slater determinant

JAGP Jastrow correlated AGP

KRR kernel ridge regression
KZK Kwee-Zhang-Krakauer
LDA local density approximation
LM linear method
LLPT liquid-liquid phase transition
LRDMC lattice regularized diffusion Monte Carlo
MAE mean absolute error
MC Monte Carlo
MD molecular dynamics
ML machine learning
MLP machine learning potential
MO molecular orbital
MPNN message passing neural network
MSD mean squared displacement
MSE mean squared error
NQE nuclear quantum effect
NN neural network
opt-JSD optimized JSD
PAW projector augmented-wave
PBC periodic boundary conditions
PBE Perdew-Burke-Ernzerhof
PES potential energy surface
PILE path integral Langevin equation
PIMD path integral molecular dynamics
PIMC path integral Monte Carlo
PIOUD path integral Ornstein-Uhlenbeck dynamics
RH Rankine-Hugoniot
RMC reptation Monte Carlo
RMSE root mean squared error
RPIMC restricted path integral Monte Carlo
SCAN-L strongly constrained and appropriately normed
SD Slater determinant
SR stochastic reconfiguration
SSCHA stochastic self-consistent harmonic approximation
SOAP smooth overlap of atomic positions
QMC quantum Monte Carlo

VMC variational Monte Carlo

WF wave function

1 INTRODUCTION

Hydrogen is the most abundant element in the universe, constituting about 75% of the mass of ordinary matter, and the first one to form after the Big Bang. It also has the simplest structure and its properties have been studied since the very beginning of quantum mechanics [1]. Already in 1935, Wigner and Huntington postulated that hydrogen in a metallic solid state may exist at high pressure [2]. Since then, the interest in hydrogen phase diagram has never stopped, stimulating countless experimental and theoretical studies.

One of the main drivers for hydrogen research is represented by astrophysical applications such as modeling giant gas planets or brown dwarfs. The internal structure of these celestial objects is inferred indirectly by gathering data from space probes, like the one of the Juno mission studying Jupiter [3–6] or the Cassini mission studying Saturn [7–9], and then using models and other assumptions to determine the core mass, extension, and composition. These quantities are extremely sensitive to the accuracy of the hydrogen equation of state, and even a discrepancy as small as a few percent can produce very different predictions [10, 11]. Accurate results of hydrogen properties are also crucial for technological applications, such as inertial confinement fusion [12–14] or green energy production [15]. Last but not least, the high-pressure physics of this element is relevant for the understanding of a large class of hydrogen-rich high-temperature superconductors, which display a record-breaking critical temperature but are only stable at very high pressures [16–18]. Solid metallic hydrogen itself was indicated as a possible high-temperature superconductor by Ashcroft already in 1968 [19].

It may then sound surprising that, despite their relevance, many of the questions regarding the hydrogen phase diagram are not fully understood, even at the qualitative level. The reason for this lies in the difficulty of realizing the necessary physical conditions in a laboratory, even considering modern-day equipment. Two main experimental techniques have been used to study high-pressure hydrogen: static compression methods and shock experiments. In the first approach, a sample of material is squeezed inside diamon anvil cells (DAC) [20, 21], to achieve extreme pressures. The temperature is controlled using either a resistor or lasers, and the properties of the system are studied using optical measurements such as Raman, IR spectroscopy, or x-ray scattering. This technique allows for good control of the sample conditions while it can achieve temperatures up to $\sim$ 1000 K and pressures up to 500 GPa, with the possibility of reaching even larger

compression as new improvements of the apparatus are made [22]. The second approach is represented by dynamic compression techniques [23]. These experiments use shock waves to achieve higher pressures and temperatures than those obtained with static compression methods, namely $p \sim 1$ TPa and $T \sim 50000$ K. Both static and dynamic experiments are challenging for a variety of reasons. In DAC methods, chemical reactions between the container and the sample are enhanced with pressure and can thus contaminate the latter. Moreover, hydrogen can easily escape by penetrating into cracks and favor the growth of fractures in the anvils. On the other hand, dynamic compression measurements have to be performed very rapidly and are in general less accurate than DAC ones. Quantities such as temperature are not directly measured and have to be inferred using models, giving rise to further uncertainties.

Because of all of this, available experimental data on high-pressure hydrogen are often accompanied by large error bars or even provide contrasting results, making their interpretation hard. One of the topics that are still intensely debated is the aforementioned Wigner-Huntington insulator to metal transition in solid hydrogen; even though different groups have claimed its experimental realization [24–26], a general consensus is still missing [27] and lots of details are unclear to this day, such as the exact symmetries of the different solid phases or if the transition coincides with the stabilization of the atomic crystal. Similar uncertainties are present in the liquid part of the phase diagram where yet another metal-insulator transition is predicted between an atomic and a molecular fluid. Also in this case an experimental agreement on the location of this transition or its specific character (i.e., if it is first-order or a crossover) cannot be definitively established.

Because of the difficulty of obtaining reliable experimental data, computer simulations have been extremely valuable in this context. Indeed, numerical approaches can provide the physical properties of the system without being restricted by any experimental constraint, allowing for an extensive study of the phase diagram. Thanks to its simple structure, hydrogen systems have been used as a testing ground for a large variety of algorithms, assessing their relative accuracy and stimulating new developments in the field [28]. Models based on effective interactions were among the first to be applied. These "chemical models" [29–33], which can be constructed using general physical considerations or fitted in order to reproduce available data, are fast but usually have poor transferability and/or accuracy. Much better results can be achieved with "first-principles" approaches, which directly solve the quantum mechanical problem of the system. One of the most popular of such methods is *ab initio* molecular dynamics (AIMD), where at each time step the physical properties of the system are calculated by using electronic structure

methods such as density functional theory (DFT) [34]. Performing a DFT calculation is several orders of magnitude more computationally expensive than the evaluation of simpler atomistic potentials, limiting the applicability of AIMD to small/medium-sized systems and short simulations. Moreover, DFT calculations depend on the choice of the exchange-correlation functional of the effective one-body problem. Many of the properties of high-pressure hydrogen, such as the location of the liquid-liquid transition or the solid part of the phase diagram, strongly depend on the functional, which, together with the large uncertainty of experimental measures, makes it very hard to determine which choice is the most accurate [35]. A better description of electronic many-body correlations is given by quantum Monte Carlo (QMC) techniques [36]. These algorithms, which include methods such as variational Monte Carlo (VMC), diffusion Monte Carlo (DMC), coupled electron-ion Monte Carlo (CEIMC), and path integral Monte Carlo (PIMC), have been proven to be extremely effective in describing hydrogen [28, 35]. Although more accurate, QMC methods have an even higher cost than DFT, thus further reducing the possible system sizes that can be realistically simulated.

In the past decade, a big revolution took place in the field of molecular dynamics (MD) simulations with the introduction of machine learning (ML) techniques. In particular, ML models aimed at reproducing the potential energy surface of a given system, also called machine learning potentials (MLPs), have made it possible to bridge the accuracy of AIMD with the efficiency of atomistic potentials [37, 38]. If properly constructed (said, trained), MLPs provide results almost indistinguishable from the electronic structure method used to build them at a drastically cheaper computational cost, while also showing good transferability across multiple thermodynamic conditions. Most MLPs in the literature are trained on DFT reference calculations, since within this method one can generate relatively large datasets, necessary to obtain accurate potentials. These models are very effective in describing the properties of several systems, but, as previously mentioned, DFT is often unsatisfactory for hydrogen. The study and development of accurate MLPs trained on QMC data and their application to high-pressure hydrogen will be the main topic of this work. As we will see, several reasons make this a complex task, and this type of approach was explored only recently [39–44]. This thesis is structured as follows. In Chap. 2, the main electronic structure methods we used are discussed, with a particular focus on the QMC approach. A procedure to correct the self-consistency error in VMC forces and pressures is also presented, following Ref.[45]. Chap. 3 includes a description of sampling methods such as classical MD and path integral molecular dynamics (PIMD). We will also introduce MLPs, such as kernel ridge regression and MACE models. Chap. 4 contains an application of an

MLP trained on QMC data for studying the principal Hugoniot curve of deuterium, as published in Ref.[41]. The correction introduced in Chap. 2 was also applied to the dataset of this Hugoniot model, as shown in Chap. 5. Finally, Chap. 6 contains another application of QMC-based MLPs on the liquid-liquid transition, following Ref. [39]. Here, we also discuss a recently obtained MLP trained with DFT data which we used to clarify the recent debate on the first-order character of the transition.

Before going further, we will briefly summarize the main features of the high-pressure hydrogen phase diagram.

HYDROGEN PHASE DIAGRAM: A QUICK OVERVIEW

Our current understanding of the hydrogen phase diagram derives from the combination of a large number of experimental measures and numerical predictions. Here, we will focus on temperatures below 10^5 K and pressures below 1 TPa, roughly corresponding to the thermodynamic range studied in this thesis. The phase diagram is shown in Fig.1. At ambient conditions hydrogen is a molecular gas; the H_2 molecule is among the most stable, with a bonding energy of 4.52 eV and an interatomic distance of approximately 0.74 Å [46]. On the contrary, the intermolecular interaction is very weak, and the low-pressure phase diagram is characterized by very cold boiling and melting temperatures, with the gas-liquid critical point being at about 33K and 13 bar, and a gas-liquid-solid triple point at 13.8K and 0.07 bar [47].

Solid hydrogen exists in several different phases. At temperatures above 100 K and pressures up to 160 GPa, hydrogen is in phase I, a solid with spherically disordered molecules in an hcp structure [48, 49]. At lower temperatures and upon compression, phase I transforms into phase II, or broken symmetry phase. This transition has a large isotope dependence, happening at $\sim$ 60 GPa and $\sim$ 25 GPa for hydrogen and deuterium respectively, signaling the crucial importance of nuclear quantum effects. Candidate symmetries for this phase, obtained with the *ab initio* random structure search (AIRSS) [50], are $Pca2_1$-8, $P6_3/m$-16, and $P2_1/c$-24. At higher pressures and for temperatures lower than 300 K, solid hydrogen transitions into phase III, which is believed to have a monoclinic $C2/c$-24 symmetry [50–53], also confirmed by a recent study including nuclear quantum effects within the stochastic self-consistent harmonic approximation (SSCHA), and QMC energies [54]. The same study also supported the transition at $\sim$ 410 GPa to the conductive phase VI found in DAC experiments [26], with probable Cmca-12 symmetry, and the transition to the long-sought atomic hydrogen at $\sim$ 570 GPa, in a pressure range not yet explored by experiments. At temperatures above 300 K, other phases have been proposed, namely phase I′ [55, 56], isostructural

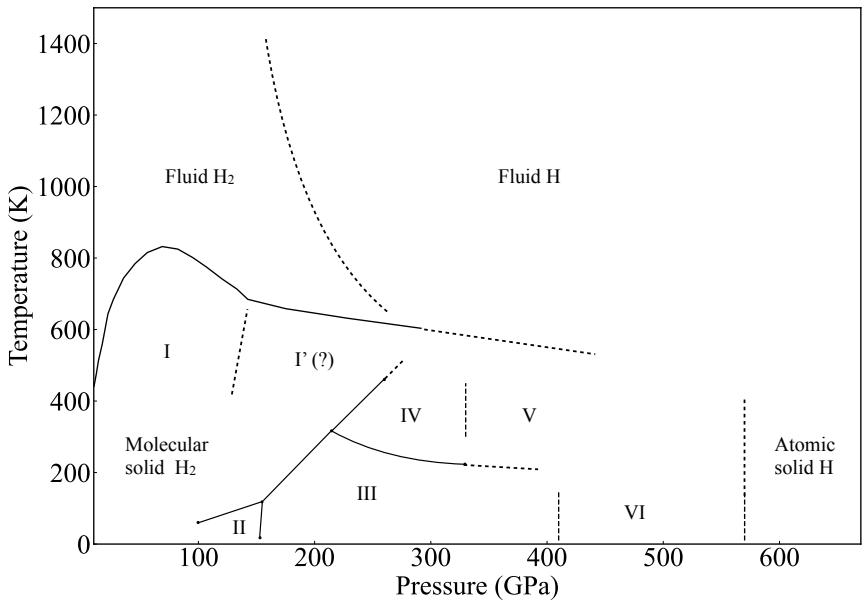

Figure 1: Hydrogen phase diagram. Experimental phase boundaries are taken from Refs.[59, 67] and indicated with solid black lines, while the melting line follows the results of Ref.[63]. Vertical dashed lines between phase III and phase VI and the transition to metallic atomic hydrogen are taken from Ref.[54]. The LLPT location is in qualitative agreement with QMC results [39, 68–70].

to phase III and whose existence is debated [57], phase IV [58–60], a layered molecular structure stable approximately between 200 and 325 GPa, and phase V [61, 62], which is thought to be another precursor of atomic hydrogen.

The hydrogen melting line was recently measured up to 300 GPa using Raman spectroscopy [63]. Similar to other alkali metals, the curve presents a maximum in the $p - T$ diagram at around $p = 70$ GPa and $T = 820$ K, even though recent numerical results using an MLP trained on QMC data have questioned its location [40]. Above the melting temperature, fluid hydrogen undergoes a liquid-liquid phase transition (LLPT) between a molecular insulator and an atomic metal. The location and character of the transition have been studied both experimentally and numerically, but large uncertainties still remain. Most numerical results indicate this transition to be weakly first order with a critical endpoint located at ~ 2000 K, even though the nature of the LLPT was recently debated [64–66].

2

AB INITIO ELECTRONIC STRUCTURE METHODS

Electronic structure methods are aimed at solving, either exactly or approximately, the many-body Schrödinger equation

$$i\hbar \frac{d}{dt} |\Psi\rangle = \hat{\mathcal{H}} |\Psi\rangle ,$$

where $\hat{\mathcal{H}}$ is the system Hamiltonian and $|\Psi\rangle$ is the many-body wavefunction. From now on we will restrict ourselves to collections of N nuclei, with positions $\mathbf{R}_I$, mass M_I, and atomic numbers Z_I ($I = 1, \ldots, N$), and N_e electrons, with positions $\mathbf{r}_i$, mass m, and spins σ_i ($i = 1, \ldots, N_e$). We can write the (non-relativistic) Hamiltonian in no external fields as

$$\begin{aligned}
\hat{\mathcal{H}} =& \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ee} + \hat{V}_{en}, \\
\hat{T}_n =& -\sum_{I=1}^{N} \frac{\hbar^2}{2M_I} \nabla_I^2, \\
\hat{T}_e =& -\sum_{i=1}^{N_e} \frac{\hbar^2}{2m} \nabla_i^2, \\
\hat{V}_{nn} =& \sum_{I<J} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}, \\
\hat{V}_{ee} =& \sum_{i<j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}, \\
\hat{V}_{en} =& -\sum_{i=1}^{N_e} \sum_{I=1}^{N} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|},
\end{aligned}$$

where $\hat{T}_n$, $\hat{T}_e$ are the (total) kinetic energy of ions and electrons, and $\hat{V}_{nn}$,$\hat{V}_{ee}$ and $\hat{V}_{en}$ are the ion-ion, electron-electron, and electron-ion electrostatic potentials, respectively. For pure hydrogen systems, $Z_I = 1$, $N_e = N$, and M_I is equal to the hydrogen or deuterium mass.

Owing to their large mass difference (i.e., $M_I \gg m$), one can often adiabatically decouple electronic and ionic degrees of freedom by performing the Born-Oppenheimer (BO) approximation [71]. In particular, one considers the ions to move inside an effective potential given by the solution of the electronic Schrodinger equation

$$\begin{aligned}
&\hat{H}_{el} (\mathbf{R}) |\Psi_{el,n} (\mathbf{R})\rangle = E_n (\mathbf{R}) |\Psi_{el,n} (\mathbf{R})\rangle, \\
&\hat{H}_{el} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{nn} + \hat{V}_{en},
\end{aligned} \tag{1}$$

where the ionic coordinates $\mathbf{R} \equiv (\mathbf{R}_1, \ldots, \mathbf{R}_N)$ are now fixed parameters, and $\Psi_{el,n}$ is the electronic wavefunction corresponding to the

n-th energy level $E_n(\mathbf{R})$. Moreover, for temperatures smaller than the Fermi temperature T_F, one can effectively discard the excited states in the electronic part and consider only the ground state. In particular Eq.(1), becomes

$$\hat{H}|\Phi_0(\mathbf{R})\rangle = E_0(\mathbf{R})\,|\Phi_0(\mathbf{R})\rangle, \tag{2}$$

where $\Phi_0 \equiv \Psi_{el,0}$ and $E_0(\mathbf{R})$ are the ground state wave function and energy, and we indicated $\hat{H}_{el}$ as simply $\hat{H}$. As we will see in Chap. 3, once the electronic problem is solved, the motion of the ions can be treated using either classical or quantum mechanics. In this Chapter, we will present the electronic structure methods we used for the calculations presented in the thesis. This will also allow us to introduce most of the notation in the next Chapters. A more complete discussion on the topics covered here can be found in Ref. [34] and Ref. [36].

The density functional theory method will be briefly described in Sec. 2.1. We will then discuss the quantum Monte Carlo (QMC) approach (Sec. 2.2). The contents of this section include the variation Monte Carlo method (Sec. 2.2.1), wave function optimization techniques (Sec. 2.2.2), a description of different types of variational ansatz (Sec. 2.2.3), the lattice regularized diffusion Monte Carlo method (Sec. 2.2.4), the fixed-node approximation (Sec. 2.2.5), and a summary of the typical workflow for a QMC calculation (Sec. 2.2.6).

Finally, in Sec.2.3, the problem of the self-consistency error in QMC forces and pressures will be discussed, following Ref. [45]. In particular, the bias correction will be introduced in Sec. 2.3.1, followed by some simple applications on dimers (Sec. 2.3.2) and cubic boron nitride (Sec. 2.3.3). The contents of this Chapter are summarized in Sec. 2.4.

2.1 DENSITY FUNCTIONAL THEORY

A widely used approach to solve the electronic problem in Eq.(1) is density functional theory (DFT). This method, originally introduced by Hohenberg, Kohn and Sham in two seminal papers [72, 73], maps the original interacting many-body problem, into a simpler, mean-field effective system that can be solved self-consistently. In particular, this "auxiliary" system is constructed to have the same ground-state electronic density, $n(\mathbf{r})$, as the many-body one. This mapping results in the Kohn-Sham energy functional

$$E[n] = -\frac{\hbar^2}{2m}\sum_{i=1}^{N_e}|\nabla\psi_i(\mathbf{r})|^2 + \frac{1}{2}\int d^3\mathbf{r}d^3\mathbf{r}'\frac{e^2 n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + V_{nn} + E_{xc}[n], \tag{3}$$

where $\psi_i(\mathbf{r})$ are the eigenstates of the mean-field problem, V_{nn} is the nuclear interaction, and $E_{xc}[n]$ is the exchange-correlation func-

tional. The explicit form of $E_{xc}[n]$ is not known and several approximations have been developed over the years. These range from simpler forms, like the local density approximation (LDA) [74] and the generalized gradient approximation (GGA) [75] to more refined ones, such as metaGGA, hybrid functionals, or those incorporating long-range van der Waals interactions. As a consequence, both the accuracy and the efficiency of DFT calculations strongly depend on the specific choice of the functional.

To represent the one particle orbitals $\psi_i(\mathbf{r})$, the plane-wave basis set is often employed due to its effectiveness in describing delocalized states in bulk systems. In practice, the number of basis functions is determined by fixing an energy cutoff E_{cut}, and considering only wave vectors $\mathbf{G}$ with $\frac{\hbar^2}{2m}|\mathbf{G}|^2 < E_{cut}$. Alternatively, a localized basis of Gaussian type orbitals (GTOs) can be used, which can also be generalized for periodic systems (see, for example, the CRYSTAL basis [76]).

In order to get a smooth density around the nuclei, the ion-electron interaction is renormalized by means of pseudopotentials. This approach is standard even for hydrogen atoms (where only one electron is present), to deal with the Coulomb interaction.

Finally, within periodic boundary conditions (PBC), Bloch's theorem for non-interacting particles can be applied, and the thermodynamic limit is simulated by considering $\mathbf{k}$-points integration with a sufficiently dense grid.

For high temperatures, i.e., $T \sim T_F$, the approximation of Eq.(2), where we considered electrons to be in their ground state, does not hold anymore. DFT can consider the effect of finite temperature electrons thanks to Mermin's extension of the Hohenberg and Kohn theorems to non-zero temperature [77]. In particular, this is done by occupying the bands of the system according to the Fermi-Dirac distribution and minimizing the Helmholtz free energy $A[n] = E - TS[n]$, where $S[n]$ is the entropy functional, given by

$$S[n] = -\sum_i f_i \log f_i + \sum_i (1 - f_i) \log (1 - f_i),$$

where f_i is the occupation of the state $\psi_i(\mathbf{r})$. Explicitly temperature-dependent DFT functionals are often used in this case to reach a higher accuracy [78–80].

2.2 QUANTUM MONTE CARLO METHODS

An alternative to DFT is represented by quantum Monte Carlo (QMC) methods [81], a vast class of algorithms that better describe highly correlated materials and depend on more controllable approximations. Here we will describe two QMC approaches: variational Monte Carlo (VMC) and lattice regularized diffusion Monte Carlo (LRDMC). Both algorithms are implemented in the TurboRVB package [82].

2.2.1 *Variational Monte Carlo*

As the name suggests, variational Monte Carlo is based on the quantum variational principle, which states that, for any wave function (WF) ψ, the expectation value of the Hamiltonian on ψ cannot be lower than the ground state E_0:

$$\frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} \geqslant E_0, \quad \forall\psi, \tag{4}$$

where the equal sign holds only if $\psi = \Phi_0$. An approximation of the ground state can thus be obtained by considering an ansatz $\psi(\{\alpha\})$, which depends on N_p variational parameters $\{\alpha\} = \alpha_1, \ldots, \alpha_{N_p}$, and by minimizing the left-hand side of Eq.(4):

$$E \equiv \min_{\alpha_1,\ldots\alpha_{N_p}} E(\{\alpha\}) = \min_{\alpha_1,\ldots\alpha_{N_p}} \frac{\langle\psi(\{\alpha\})|\hat{H}|\psi(\{\alpha\})\rangle}{\langle\psi(\{\alpha\})|\psi(\{\alpha\})\rangle}. \tag{5}$$

For given parameters, the variational energy $E(\{\alpha\})$ in Eq. (5) can be written as:

$$E(\{\alpha\}) = \int d\mathbf{x}\pi(\mathbf{x},\{\alpha\})E_L(\mathbf{x},\{\alpha\}) \tag{6}$$

where $\mathbf{x} = (x_1, ..., x_{N_e})$, with $x_i = (\mathbf{r}_i, \sigma_i)$, is the collection of all spacial and spin coordinates of the electrons, and we defined the probability distribution π and the local energy E_L as

$$\pi(\mathbf{x},\{\alpha\}) = \frac{|\psi(\mathbf{x},\{\alpha\})|^2}{\langle\psi(\{\alpha\})|\psi(\{\alpha\})\rangle},$$
$$E_L(\mathbf{x},\{\alpha\}) = \frac{\hat{H}\psi(\mathbf{x},\{\alpha\})}{\psi(\mathbf{x},\{\alpha\})},$$

respectively. In VMC, the expectation value in Eq.(6) is computed stochastically by generating a series of configurations $\mathbf{x}_1, \ldots, \mathbf{x}_M$, e.g., through the Metropolis-Hastings algorithm [83, 84], which are distributed according to π:

$$E(\{\alpha\}) = \langle E_L(\{\alpha\})\rangle_{\pi(\{\alpha\})} \approx \frac{1}{M}\sum_{j=1}^{M} E_L(\mathbf{x}_j,\{\alpha\}). \tag{7}$$

Notice how, contrary to deterministic methods, this statistical average will have an associated noise $\sigma_E \sim M^{-1/2}$, which can be reduced by increasing the number of independent samples M. Remarkably, if $\psi(\{\alpha\}) = \Phi_0$, the local energy is identically equal to the ground state energy, $E_L \equiv E_0$, and $\sigma_E = 0$ (zero variance property). Notice that an expression equivalent to Eq.(7) can also be obtained for the expectation value of any local operator $\hat{O}$ on the wave function ψ:

$$\langle\hat{O}\rangle \approx \frac{1}{M}\sum_{j=1} O_L(\mathbf{x}_j), \tag{8}$$

with $O_L(\mathbf{x}) \equiv \frac{\langle\mathbf{x}|\hat{O}|\psi\rangle}{\psi(\mathbf{x})}$.

2.2.2 *Wave function optimization*

Finding the optimal wave function that minimizes $E(\{\alpha\})$ is not a trivial task. The number of parameters $\{\alpha\}$ is often large ($\sim 10^4$-10^5), thus creating an energy landscape with several local minima. Moreover, within VMC, the generalized forces $f_{\alpha_k} = -\frac{\partial E(\{\alpha\})}{\partial \alpha_k}$ also have an associated noise, which further complicates the optimization. In particular,

$$\mathbf{f}_{\alpha_k} \approx -2\Re\left[\frac{1}{M}\sum_{j=1}^{M} E_L^*(\mathbf{x}_j)\left(\mathcal{O}_k(\mathbf{x}_j) - \overline{\mathcal{O}}_k\right)\right],$$

where we defined the operator $\hat{\mathcal{O}}_k$ such that $\langle \mathbf{x}'|\hat{\mathcal{O}}_k|\mathbf{x}\rangle = \mathcal{O}_k(\mathbf{x})\delta(\mathbf{x} - \mathbf{x}')$, with $\mathcal{O}_k(\mathbf{x}) = \frac{\partial \ln \psi(\mathbf{x},\{\alpha\})}{\partial \alpha_k}$ and $\overline{\mathcal{O}}_k = \langle \hat{\mathcal{O}}_k \rangle$. The logarithmic derivative in $\mathcal{O}_k(\mathbf{x})$ can be efficiently computed using the adjoint algorithmic differentiantion (AAD) method [85].

Two of the most effective schemes for WF optimization are the stochastic reconfiguration (SR) and the linear method (LM). In the SR approach [86], the generalized forces $\mathbf{f}_{\{\alpha\}} = (f_{\alpha_1}, \ldots, f_{\alpha_{N_p}})$ are modified using a stochastic positive-defined preconditioning matrix S, and then used to update the parameters

$$\begin{aligned} \alpha_k &\to \alpha_k + \Delta\left(S^{-1}\mathbf{f}_{\{\alpha\}}\right)_k \quad , \\ S_{k,k'} &\approx \Re\left[\frac{1}{M}\sum_{j=1}^{M}\left(\mathcal{O}_k(\mathbf{x}_j) - \overline{\mathcal{O}}_k\right)^*\left(\mathcal{O}_{k'}(\mathbf{x}_j) - \overline{\mathcal{O}}_{k'}\right)\right]. \end{aligned} \tag{9}$$

In practice, the S matrix is often ill-conditioned and needs to be regularized by shifting its diagonal elements [87], e.g., $S_{kk} \to S_{kk}(1+\varepsilon) + \varepsilon'$, with $\varepsilon, \varepsilon' > 0$. The SR approach can be intuitively understood by interpreting S as the metric on the space of variational parameters which measures the distance of the underlying WFs [88]. In particular, parameters that weakly affect the variational energy will be moved "faster" than those that have a stronger impact on the WF, improving the optimization. Notice how, in this way, the SR method can be interpreted as a noisy natural gradient descent.

An alternative approach to the SR is the linear method [89–91]. Within the LM, the WF is first expanded up to linear order in the variation of the parameters $\{\alpha\} \to \{\alpha + \delta\alpha\}$, which can be done using the semi-orthogonal basis:

$$|\psi(\{\alpha + \delta\alpha\})\rangle \approx z_0|\widetilde{\psi}(\{\alpha\})\rangle + \sum_{k=1}^{N_p} z_k\left(\hat{\mathcal{O}}_k - \overline{\mathcal{O}}_k\right)|\widetilde{\psi}(\{\alpha\})\rangle, \tag{10}$$

with $|\widetilde{\psi}(\{\alpha\})\rangle \equiv \frac{|\psi(\{\alpha\})\rangle}{||\psi(\{\alpha\})||}$. The expectation value of the energy E on the WF in Eq.(10) is then minimized with respect to the $N_p + 1$ dimen-

sional vector $\mathbf{z} = (z_0, \ldots, z_{N_p})$, which gives rise to the generalized eigenvalue equation:

$$\overline{\mathbf{H}}\mathbf{z} = E\overline{\mathbf{S}}\mathbf{z}, \tag{11}$$

where the two matrices $\overline{\mathbf{S}}$ and $\overline{\mathbf{H}}$ can be estimated stochastically as

$$\overline{\mathbf{S}}_{k,k'} \approx \frac{1}{M}\sum_{j=1}^{M}\left(\mathcal{O}_k(\mathbf{x}_j) - \overline{\mathcal{O}}_k\right)^*\left(\mathcal{O}_{k'}(\mathbf{x}_j) - \overline{\mathcal{O}}_{k'}\right)$$

$$\overline{\mathbf{H}}_{k,k'} \approx \frac{1}{M}\sum_{j=1}^{M}\left(\mathcal{O}_k(\mathbf{x}_j) - \overline{\mathcal{O}}_k\right)^* \cdot \frac{\partial E_L(\mathbf{x}_j, \{\alpha\})}{\partial \alpha_{k'}}$$

When solving Eq.(11) the eigenvector that maximizes $|z_0|$ is chosen to have a more stable optimization. Finally, the parameters are updated according to

$$\alpha_k \to \alpha_k + \Delta\frac{z_k}{z_0},$$

with a value of $\Delta \sim 1$. The linear method is generally slower (for a single optimization step) than the SR one and might be unstable for large numbers of parameters, but the full optimization is often faster because of the reduced number of steps necessary to converge to the variational minimum. To improve its efficiency, the LM can also be applied to a restricted space of parameters, by considering only α_k with a large signal-to-noise ratio, and/or a fixed number of natural gradient directions [87] (see Eq.(9)).

A key quantity for both approaches is the dimension of the sample M, also called "batch size", used to compute all the quantities used in the minimization. In principle, to have a meaningful accuracy, M should be much larger than the number of parameters, e.g., $M \gtrsim 10 \times N_p$. However, for large systems, this is often impractical, and a smaller number has to be used. This is made possible by the regularization of the preconditioning matrix $\mathbf{S}$ previously introduced, and the resulting optimization is reminiscent of the stochastic gradient technique used in the machine learning community [92].

2.2.3 *Variational wave functions*

The explicit form of the variational ansatz ψ is a key ingredient, which determines the accuracy and performance of the VMC method. Here we will describe the types of WF implemented in the TurboRVB package [82] that we used to generate all the QMC data in this work. We start by expressing ψ as the product of a (bosonic) function $\exp(J)$, symmetric under particle exchange, also called Jastrow part, and an antisymmetric (fermionic) part Φ_{AS}, necessary to have the correct total symmetry:

$$\psi = \exp(J) \times \Phi_{AS}. \tag{12}$$

The Jastrow factor J effectively describes the correlation effects between the electrons. In particular, J is decomposed into one-, two-, and three-/four- body terms

$$J = J_1 + J_2 + J_{3/4}.$$

The one-body Jastrow is further written as a sum of a homogeneous J_1^h and an inhomogeneous J_1^{inh} term. The former is used to satisfy the electron-ion Kato cusp conditions [93], and can be expressed as

$$J_1^h(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \sum_{i=1}^{N_e} \sum_{a=1}^{N} \left(-(2Z_a)^{3/4} u_a \left((2Z_a)^{1/4} |\mathbf{r}_i - \mathbf{R}_a| \right) \right),$$

where $u_a(r)$ is a short-range function containing a parameter b_a, i.e., $u_a(r) = \frac{b_a}{2}\left(1 - e^{-r/b_a}\right)$, which can be independently optimized for each different atomic species. The latter is

$$J_1^{inh}(x_1, \ldots, x_N) = \sum_{i=1}^{N_e} \sum_{a=1}^{N} \left(\sum_l M_{a,l}^{\sigma_i} \chi_{a,l}(\mathbf{r}_i) \right),$$

where l runs over the atomic orbitals (AOs) $\chi_{a,l}$ centered around the atom a and $\{M\}$ are variational parameters. Within TurboRVB, the AOs can be either of Gaussian or Slater type [82].

The two body Jastrow describes correlation effects between pairs of electrons and it is defined as

$$J_2(x_1, \ldots, x_N) = \sum_{i<j} v^{\sigma_i, \sigma_j}\left(\left|\mathbf{r}_i - \mathbf{r}_j\right|\right).$$

Several choices of the function v are possible; in our applications, the following spin-independent form has been used:

$$v(r) = \frac{1}{2} r \cdot (1 - F \cdot r)^{-1}$$

with F being another variational parameter. Notice how J_2 also automatically satisfies the electron-electron cusp conditions. Finally, further correlations (e.g. ion-electron-electron or ion-ion-electron-electron) are included within $J_{3/4}$:

$$J_{3/4}(x_1, \ldots, x_N) = \sum_{i<j} \left(\sum_{a,l} \sum_{b,m} \widetilde{M}_{\{a,l\},\{b,m\}}^{\sigma_i, \sigma_j} \chi_{a,l}(\mathbf{r}_i) \chi_{b,m}(\mathbf{r}_j) \right),$$

where we again introduced the variational parameters $\left\{\widetilde{M}\right\}$. In the above equation, a, b run over the atomic indices and l, m run over the basis set elements. In practice, we can often disregard the coefficients corresponding to different atoms (i.e. those for which $a \neq b$), since this rarely affects the variational energy while significantly reducing the number of parameters to optimize.

Similarly to the Jastrow factor, also the antisymmetric part of WF Φ_{AS} can be chosen in several ways. The most common and simple choice is to use a Slater determinant obtained with mean-field methods such as Hartree-Fock or DFT, i.e., $\Phi_{AS} = \Phi_{SD}$, and to optimize only the variational parameters in the Jastrow. In this case the WF ψ is called the Jastrow correlated Slater determinant (JSD) wave function. For many applications, the JSD ansatz might not be accurate enough, and an improved WF is thus necessary. One way to do this is by using an antisymmetrized geminal power (AGP) [94]

$$\Phi_{AGP}(\mathbf{x}) = \hat{A}\left[g(x_1, x_2)\, g(x_3, x_4) \cdots g(x_{N_e-1}, x_{N_e})\right], \tag{13}$$

where we supposed for simplicity an even number of electrons N_e and indicated with $\hat{A}$ the antisymmetrization operator. The function g is called the pairing function, and for spin unpolarized systems it can be taken as

$$g(x_i, x_j) = f(\mathbf{r}_i, \mathbf{r}_j) \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \delta_{\sigma_i, -\sigma_j}. \tag{14}$$

The resulting WF, $\psi = \exp(J) \times \Phi_{AGP}$, is called the Jastrow correlated AGP (JAGP). Notice that, with the use of different pairing functions, including, for example, other spin terms, the expression in Eq.(13) can also describe other types of WFs, such as the Pfaffian [95]. The spatial part of g is expanded in terms of AOs as

$$f(\mathbf{r}_i, \mathbf{r}_j) = \sum_{a,l} \sum_{b,m} \lambda_{\{a,l\},\{b,m\}} \phi_{a,l}(\mathbf{r}_i)\, \phi_{b,m}(\mathbf{r}_j), \tag{15}$$

where the indices l, m run over the atomic orbitals $\phi_{a,l}, \phi_{b,m}$ belonging to the AGP basis set (in general, different from the Jastrow one) centered around the a-th and b-th atoms, respectively. For periodic systems, $\phi_{a,l}$ is generalized to fulfill periodic or twisted boundary conditions [82]. We can recast the previous equation in a more compact form

$$f(\mathbf{r}_i, \mathbf{r}_j) = \sum_{\mu=1}^{L} \sum_{\nu=1}^{L} \lambda_{\mu,\nu} \phi_\mu(\mathbf{r}_i)\, \phi_\nu(\mathbf{r}_j), \tag{16}$$

where the indices $\mu, \nu = 1, \ldots, L$ run over the entire set of AOs of dimension L. The full (or partial) optimization of the AGP variational parameters $\{\lambda\}$ is often a formidable task and requires a good initialization. Notice how Eq.(16) can be formally written in terms of L molecular orbitals (MOs) as

$$f(\mathbf{r}_i, \mathbf{r}_j) = \sum_{k=1}^{L} \lambda_k \left(\Phi_k^{MO}(\mathbf{r}_i)\right)^* \Phi_k^{MO}(\mathbf{r}_j), \tag{17}$$

where

$$\Phi_k^{MO}(\mathbf{r}) = \sum_{\mu=1}^{L} c_{\mu,k} \phi_\mu(\mathbf{r}), \tag{18}$$

with coefficients $c_{\mu,k}$. When expressed in this way, it is apparent how the simpler single Slater determinant case can be recovered from Eq.(17) when we only consider $N_e/2$ MOs in the expansion, i. e., $\lambda_k = 1$ for $1 \leqslant k \leqslant N_e/2$ and $\lambda_k = 0$ elsewhere. This formal equivalence can be exploited to initialize the variational parameters in Eq.(16) from the MOs obtained with a DFT calculation. Moreover, the AGP can also be projected at each optimization step into a given number n of MOs, with $N_e/2 \leqslant n \ll L$ [96]. This restricted optimization is more stable than the original "unrestricted" one and can be effectively used to improve on the JSD solution. We call this WF JAGPn, which in the case of $n = N_e/2$ reduces to an optimized JSD (opt-JSD), i.e., where both the antisymmetric part and the Jastrow factor are optimized. Thus, in our notation, the opt-JSD differs from the JSD WF for the VMC optimization of the Slater determinant.

Other types of partial optimization can also be employed. For example, an efficient way to improve the variational energy while keeping the number of variational parameters tractable consists of optimizing only a subset of the $\lambda_{\{a,l\},\{b,m\}}$ in Eq.(15), namely those corresponding to atoms a and b which are closer than a given cutoff radius ("locality approach").

A complementary approach for making the optimization of Eq.(16) more feasible is to reduce the dimension of the basis set L, using geminal embedded orbitals (GEOs) [97]. Notice that the pairing function in Eq.(16) can be decomposed into atom-projected contributions U^I_{proj}:

$$f(\mathbf{r}_i, \mathbf{r}_j) = \sum_{I=1}^{N} U^I_{\text{proj}}(\mathbf{r}_i, \mathbf{r}_j) = \sum_{I=1}^{N} \sum_{\mu,\nu} A^I_{\mu,\nu} \phi_\mu(\mathbf{r}_i)\, \phi_\nu(\mathbf{r}_j),$$

where $A^I_{\mu,\nu} = 0$ when μ does not refer to the I-th atom. In practice, the $A^I_{\mu,\nu}$ can be obtained with DFT, i. e., through Eq.(17). The projected pairing function for atom I can be then truncated and expressed in terms of p orthonormal GEOs $\widetilde{\phi}_{I,\beta}(\mathbf{r}_i)$

$$\widetilde{U}^I_{\text{proj}} = \sum_{\beta=1}^{p} \sigma^I_\beta \widetilde{\phi}_{I,\beta}(\mathbf{r}_i) \psi^{\text{env}}_{I,\beta}(\mathbf{r}_j),$$

where we also introduced the orbitals $\psi^{\text{env}}_\beta(\mathbf{r}_j)$ describing the environment surrounding the I-th atom. Each GEO is defined in terms of the AOs

$$\widetilde{\phi}_{I,\beta}(\mathbf{r}_i) = \sum_{l=1}^{L} \mu^{\text{GEO}}_{I,l} \phi_{I,l}(\mathbf{r}_i). \tag{19}$$

The coefficients $\mu^{\text{GEO}}_{a,l}$ are finally obtained by minimizing the Euclidean distance between the truncated $\widetilde{U}^I_{\text{proj}}$ and the original U^I_{proj} [82, 97]. By appropriately choosing p, this new basis set significantly reduces the number of variational parameters in Eq. (16), with

the $\lambda_{\mu,\nu}$ matrix size lowered from $L \times L$ to $pN \times pN$, while retaining an accuracy comparable to the original AOs basis.

The locality approach can be combined with the GEO contraction for an efficient optimization of the WF determinant.

2.2.4 *Lattice regularized diffusion Monte Carlo*

The optimal variational WF, corresponding to the minimum of Eq.(5), might already be an excellent approximation of the true system ground state $|\Phi_0\rangle$, but a further improvement can be achieved using a projection technique. Lattice regularized diffusion Monte Carlo (LRDMC) [98] is one of these methods, based on Green's function Monte Carlo (GFMC) [99]. In the GFMC approach, the ground state is filtered out from a given state ψ by repeatedly applying the operator

$$\hat{\mathcal{G}} = \left(\hat{\Lambda} - \hat{H}\right), \tag{20}$$

where $\hat{\Lambda} = \lambda \mathbb{1}$ for some large number λ. In particular, if we write $|\psi\rangle$ in the basis of eigenstates $|\Phi_n\rangle$ of the Hamiltonian $\hat{H}$ with eigenvalue E_n, i.e., $|\psi\rangle = \sum_n a_n |\Phi_n\rangle$, we have that

$$\begin{aligned}
&\lim_{m\to+\infty} \hat{\mathcal{G}}^m |\psi\rangle = \\
&= \lim_{m\to+\infty} (\lambda - E_0)^m \left[a_0 |\Phi_0\rangle + \sum_{n\neq 0} a_n \left(\frac{\lambda - E_n}{\lambda - E_0}\right)^m |\Phi_n\rangle \right] \propto |\Phi_0\rangle,
\end{aligned} \tag{21}$$

with the requirement that $\frac{\lambda - E_n}{\lambda - E_0} < 1$ for $n \neq 0$ and $a_0 \neq 0$. Notice that, in the $\lambda \to +\infty$ limit, a continuous time formulation can be derived from Eq.(21) [100]. To apply the GFMC method to the ab initio Hamiltonian in Eq.(1), within the LRDMC we define a discretized $\hat{H}^a$, defined on a lattice with spacing a. This can be done by replacing the kinetic and potential terms with

$$\begin{aligned}
\nabla^2_{i,a} f(x_i, y_i, z_i) &\equiv \frac{1}{a^2} \{[f(x_i + a) + f(x_i - a) - 2f(x_i)]\} + \dots \\
&\quad + \{\leftrightarrow y_i\} + \{\leftrightarrow z_i\}, \\
V^a(\mathbf{x}) &\equiv V(\mathbf{x}) + \frac{1}{2} \left[\frac{\sum_{i=1}^N \left(\nabla^2_{i,a} - \nabla^2_i\right) \psi_G(\mathbf{x})}{\psi_G(\mathbf{x})} \right],
\end{aligned}$$

where $\psi_G(\mathbf{x})$ is a guiding function, usually equal to the optimal VMC wave function. For further details on how to construct the discretized Hamiltonian $\hat{H}^a$, we refer the reader to Refs. [94] and [36]. Finally, the WF evolution

$$\psi_{n+1}(\mathbf{x}') = \sum_{\mathbf{x}} \mathcal{G}_{\mathbf{x}',\mathbf{x}} \psi_n(\mathbf{x})$$

is simulated through a Markow chain. In practice, the importance sampling technique is used in order to reduce statistical fluctuations and improve the efficiency of the algorithm:

$$\psi_G(\mathbf{x}')\psi_{n+1}(\mathbf{x}') = \sum_{\mathbf{x}} \widetilde{\mathcal{G}}_{\mathbf{x}',\mathbf{x}}\psi_G(\mathbf{x})\psi_n(\mathbf{x}), \tag{22}$$

where we defined the operator $\widetilde{\mathcal{G}}_{\mathbf{x}',\mathbf{x}} = \mathcal{G}_{\mathbf{x}',\mathbf{x}}\frac{\psi_G(\mathbf{x}')}{\psi_G(\mathbf{x})}$, using the guiding function ψ_G. The LRDMC algorithm implements Eq.(22) by statistically evolving a "walker", with initial position $\mathbf{x}_0$ and weight $w_0 = 1$, according to a transition probability $p_{\mathbf{x}',\mathbf{x}} = \widetilde{\mathcal{G}}_{\mathbf{x}',\mathbf{x}}/b_{\mathbf{x}}$, with $b_{\mathbf{x}} = \sum_{\mathbf{x}'}\widetilde{\mathcal{G}}_{\mathbf{x}',\mathbf{x}}$. By using the discretized Hamiltonian $\hat{H}^a$, there are $(6N_e + 1)$ configurations that can be reached from a given $\mathbf{x}$, including $\mathbf{x}$ itself[1], thus allowing the calculation of $b_{\mathbf{x}}$ even in a continuous model. At each step, the weight is also updated according to $w_{n+1} = b_{\mathbf{x}_n}w_n$. In practice, we can perform a simulation of length N_{sim}, much larger than the equilibration time needed to project the ground state, and compute an accumulated weight G_n^p corresponding to a projection of length p:

$$G_n^p = \prod_{j=1}^{p} b_{\mathbf{x}_{n-j}}.$$

The ground state energy is then computed using

$$E_0 \approx \frac{\sum_n G_n^p \widetilde{E}_L(\mathbf{x}_n)}{\sum_n G_n^p}, \tag{23}$$

where $\widetilde{E}_L(\mathbf{x}) = \sum_{\mathbf{x}'} H^a_{\mathbf{x}',\mathbf{x}}\frac{\psi_G(\mathbf{x}')}{\psi_G(\mathbf{x})}$ is the local energy within importance sampling. In particular, Eq.(23) satisfy the zero variance property when $\psi_G = \Phi_0$. Ground state expectation values of other operators $\hat{O}$ can also be estimated with LRDMC [36]. An approximate expression is given by the mixed average

$$O_{MA} = \frac{\langle\psi_G|\hat{O}|\Phi_0\rangle}{\langle\psi_G|\Phi_0\rangle} \approx \frac{\sum_n G_n^p \widetilde{O}_L(\mathbf{x}_n)}{\sum_n G_n^p},$$

where $\widetilde{O}(\mathbf{x}) = \sum_{\mathbf{x}'} O_{\mathbf{x}',\mathbf{x}}\frac{\psi_G(\mathbf{x}')}{\psi_G(\mathbf{x})}$. This approximation is exact when Φ_0 is an eigenstate of $\hat{O}$. In the other cases, a good estimate of the ground state expectation value is obtained with $2O_{MA} - O_{VMC}$. A many-walker approach, together with the branching (reconfiguration) scheme [101] can also be implemented to further increase the efficiency of the computation of Eq.(23), as done in the TurboRVB package.

Finally, notice how the results in the LRDMC approach will depend on the lattice constant a used for the discretized Hamiltonian, and

1 Notice that, in the $\lambda \to +\infty$ continuous time limit, $\mathbf{x}$ is excluded from the configurations that can be reached after one step because a "residency" time on $\mathbf{x}$ can be estimated *a priori*.

an extrapolation to the $a \to 0$ limit must be performed [98]. This is similar to the $\tau \to 0$ limit done in the standard DMC approach, with τ being the imaginary time step of the Suzuki-Trotter expansion. Within the LRDMC method, the extrapolation in a is usually very smooth and reliable, and often a single small value of a is sufficient.

2.2.5 *Fixed-node approximation*

For fermions, the $\mathcal{G}$ operator in Eq.(20) cannot be made positive everywhere and thus cannot be used to define the transition probability $p_{x',x}$. In the QMC community, this is referred to as the sign problem. An effective solution is given by the fixed-node (FN) approximation. On the lattice, this is equivalent to considering a modified Hamiltonian $\hat{H}^{FN}$ such that the new operator $\widehat{\mathcal{G}}^{FN}$ is strictly positive. In particular, the new transition probability keeps the walkers inside the nodal pockets of the WF ψ_G used for the approximation. As a consequence, the FN projected WF Φ_0^{FN} will have the same nodal surface as ψ_G. Moreover, it can be shown [102] that the FN energy is bounded by the true ground state energy, i.e.

$$E_0 \leqslant E_0^{FN} \leqslant E_G$$

with E_G being the variational energy of ψ_G. The FN approach strongly benefits from a good guiding function ψ_G and yields exactly E_0 when ψ_G has the correct nodal structure.

2.2.6 *Typical workflow*

We end this section by summarizing the typical workflow for a VMC and LRDMC calculation using the TurboRVB package.

1. Once the basis sets for both the Jastrow factor and antisymmetric part are chosen, the WF is initialized by running a DFT calculation.

2. The variational parameters of the WF are optimized using methods such as the SR or the LM, or a combination of the two. Those may include only the Jastrow factor parameters, as in the JSD WF, or also those relative to the antisymmetric part (JAGP). In practice, the final "optimal" parameters are obtained by averaging each α_i after convergence.

3. The optimized WF ψ_{VMC} is used to compute energy and other observables using Eq.(7) or Eq.(8).

4. Finally, the LRDMC method, using ψ_{VMC} as the guiding function, can be used to improve on the VMC results. By running the calculations at different lattice sizes a, the $a \to 0$ result can be extrapolated.

Besides the energy, the workflow just described can output ionic forces and virial pressure, which are needed to perform QMC-driven geometry relaxation or molecular dynamics. These quantities are also important to generate accurate machine learning potentials, as we will discuss in Chap. 3.

2.3 UNBIASED FORCES AND PRESSURE IN VMC

The ionic forces and virial stress are obtained by differentiating the total energy E of the system with respect to the atomic coordinates $\mathbf{R}$ or the cell parameters, respectively. Since we know how to obtain E in QMC, an unbiased value of both quantities can be calculated using the finite-differences method (FDM), i.e., by running simulations with displaced variables and computing the derivative numerically. This is impractical for several reasons. First, the FDM requires multiple $\sim N$ independent QMC calculations to compute forces and virials, which makes this approach computationally expensive. Second, estimating the derivative may not be trivial due to the statistical noise associated with each point. A better fit may be obtained by computing more displacements, with a further increase in computing time.

Several alternatives to the use of the FDM have been proposed for both diffusion and variational Monte Carlo. Here we will mainly focus on forces obtained with the latter method; the calculation of forces within fixed node DMC is a much more complex and debated topic [103–110]. Within VMC, by directly differentiating Eq.(5) we obtain

$$\mathbf{F}_a = -\frac{dE}{d\mathbf{R}_a} = -\left\langle \frac{\partial E_L}{\partial \mathbf{R}_a} \right\rangle \tag{24a}$$

$$-2\left\langle (E_L - E)\frac{\partial \log \psi}{\partial \mathbf{R}_a} \right\rangle \tag{24b}$$

$$-\sum_{i=1}^{N_p} \frac{\partial E}{\partial \alpha_i}\frac{d\alpha_i}{d\mathbf{R}_a}. \tag{24c}$$

The three terms in Eq.(24) are called Hellmann-Feynman, Pulay, and variational terms, respectively. An analogous equation applies for the (isotropic) virial pressure, where the ionic position is replaced by the cell volume V. In practice, the term in Eq.(24c) is often discarded, yielding the following expression for the VMC force

$$\mathbf{F}_a^{VMC} = -\left\langle \frac{\partial E_L}{\partial \mathbf{R}_a} \right\rangle - 2\left\langle (E_L - E)\frac{\partial \log \psi}{\partial \mathbf{R}_a} \right\rangle. \tag{25}$$

As it turns out, the application of Eq.(25) is not trivial, and obtaining efficiently a statistically meaningful $\mathbf{F}_a^{VMC}$ with a finite variance requires the application of techniques such as the zero-variance zero-bias principle [111], the space warp transformation [112] and

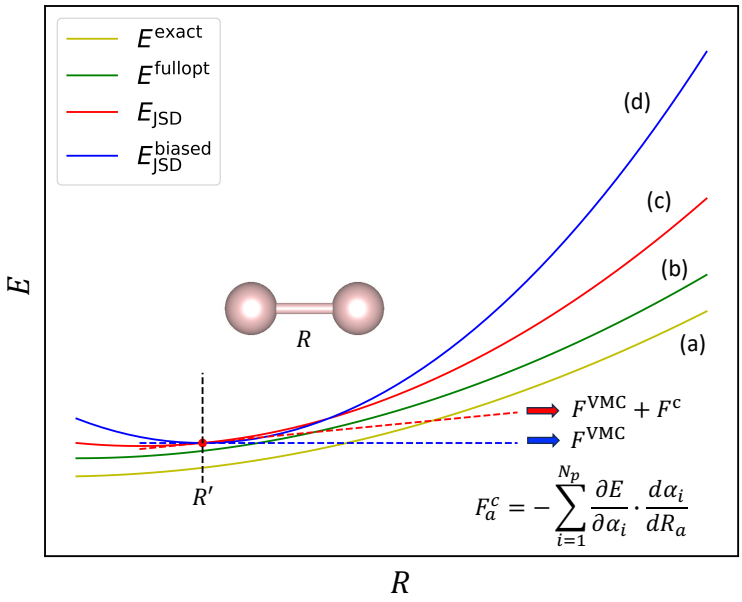

Figure 2: Schematic illustration of PESs as a function of the dimer bond length R. (a) The exact PES of the system, which is in practice inaccessible. (b) VMC PES obtained by a full optimization of all the variational parameters. (c) VMC PES obtained with the JSD WF, with the determinant part obtained at each R with a DFT calculation. (d) VMC PES obtained with the JSD, with a fixed determinant part obtained at $R = R'$ with DFT.

reweighting [85, 104, 111, 113–115]. We also note here that a good estimate for DMC forces can be obtained using the Reynolds approximation [103], by evaluating Eq.(25) with the mixed average between ψ and the projected WF Φ_0.

Neglecting the variational term is only justified when the WF is at the variational minimum for all the parameters $\{\alpha\}$. In this case $\frac{\partial E}{\partial \alpha_i} \approx 0$ for all i, and Eq. (24c) gives no contribution to the force. However, for many applications, only the Jastrow part of the WF is optimized while taking the antisymmetric part equal to a "frozen" Slater determinant. The advantage of this JSD WF comes from the fact that the Jastrow term is usually much easier and faster to optimize reliably than the determinant part. When using Eq. (25) for a JSD WF, the unoptimized parameters in the Slater determinant cause a bias in the value of the force, which we indicate as the "self-consistency error" [115, 116].

In order to illustrate the problem, let us consider as an example the potential energy surface (PES) of a dimer. In this case, the energy will only depend on the relative distance between the atoms R. According to the level of theory, we can identify different potential energy sur-

faces as illustrated in Fig.2. Every electronic structure method aims, in principle, to find the exact PES of the system E^{exact}. This is in practice impossible, except for very simple nodeless ground states. Within VMC the best PES we can get is E^{fullopt}, obtained by choosing a variational ansatz and optimizing all parameters. We have mentioned how a full optimization can rapidly become infeasible, especially for large systems. With the JSD WF, a good compromise is achieved by only optimizing the Jastrow factor and taking the determinant part equal to the DFT solution at each bond distance R, resulting in the PES E_{JSD}. In general we will have $E_{\text{JSD}}(R) \geqslant E^{\text{fullopt}}(R) \geqslant E^{\text{exact}}(R)$ for all R values, because of the variational principle. For the JSD WF, the slope given by F^{VMC} in a given point $R = R'$ is, in general, not consistent with the PES E_{JSD}, but corresponds instead to $E^{\text{biased}}_{\text{JSD}}$, obtained by using the DFT orbitals at R' for all the other R values. Notice that because of the derivative mismatch, in this case, there exist values R'', close to R', for which $E^{\text{biased}}_{\text{JSD}}(R'') < E_{\text{JSD}}(R'')$.

2.3.1 *Bias correction*

Our aim here is to correct F^{VMC} such that the resulting force is consistent with $E_{\text{JSD}}(R)$. In the JSD WF, the only parameters that contribute to the term in Eq.(24c) are those contained in the Slater determinant. By using the expansion in terms of the AOs in Eq.(16) the force correction takes the form

$$\mathbf{F}^{c}_{a} = -\sum_{\mu=1}^{L}\sum_{\nu=1}^{L}\frac{\partial E}{\partial \lambda_{\mu,\nu}}\frac{d\lambda_{\mu,\nu}}{d\mathbf{R}_a}. \tag{26}$$

In this case, since the AGP reduces to a Slater determinant, the $\lambda_{\mu,\nu}$ are related to the coefficients of the MOs in Eq.(18) , i.e, $\lambda_{\mu,\nu} = \sum_k c^*_{\mu,k} c_{\nu,k}$. The first term in Eq.(26) is the generalized force used in the WF optimization and can be efficiently computed via the AAD. On the other hand, the second term may be estimated with the FDM through finite differences, i.e.,

$$\frac{d\lambda_{\mu,\nu}}{dR^{u}_{a}} \approx \frac{1}{2\Delta R^{u}_{a}}\left(\lambda^{R^{u}_{a}+\Delta R^{u}_{a}}_{\mu,\nu} - \lambda^{R^{u}_{a}-\Delta R^{u}_{a}}_{\mu,\nu}\right) \tag{27}$$

for all three Cartesian coordinates $u = x, y, z$, with $\mathbf{R}_a = (R^x_a, R^y_a, R^z_a)$. The coefficients $\lambda^{\mathbf{R}}_{\mu,\nu}$ are deterministically obtained with DFT so that the evaluation of Eq.(26) for all ionic forces of the system requires $6N$ of such calculations. This is still cheaper than directly using the FDM to obtain the force, since the cost of DFT is mainly fast Fourier transform bound, with a favorable $O(N^2 \log N)$ scaling, and thus better than the VMC one. Notice that the application of the FDM for computing $\frac{d\lambda_{\mu,\nu}}{d\mathbf{R}}$ is only possible thanks to to the gauge invariance of the $\lambda_{\mu,\nu}$, a property inherited from the close relation of the AGP with the one-body density matrix [117]. In fact, it is trivial to demonstrate that

those coefficients are invariant under both a phase rotation of the MOs ($\Phi_k^{MO} \to e^{i\theta}\Phi_k^{MO}$) and a possible MOs degeneracy (that can produce a discontinuous change in $\mathbf{R}$ of the $c_{\mu,k}^{\mathbf{R}}$).

The derivatives $\frac{d\lambda_{\mu,\nu}}{d\mathbf{R}}$ can also be obtained without the FDM, by using linear response theory [118] or solving the coupled perturbed Kohn-Sham equations [119]. These methods will be possibly faster and avoid numerical instabilities related to the ionic displacement $\Delta\mathbf{R}$, and are currently under implementation in our workflow. Finally, Eq.(26) can be cast in a suitable form to be computed in VMC:

$$\mathbf{F}_a^c = -2\Re\left\{\left\langle E_L \sum_{\mu=1}^{L}\sum_{\nu=1}^{L}\left[\left(\mathcal{O}_{\mu,\nu} - \overline{\mathcal{O}}_{\mu,\nu}\right)\frac{d\lambda_{\mu,\nu}}{d\mathbf{R}_a}\right]\right\rangle\right\}, \tag{28}$$

where $\mathcal{O}_{\mu,\nu} = \frac{\partial \log|\psi|}{\partial\lambda_{\mu,\nu}}$ and $\overline{\mathcal{O}}_{\mu,\nu} = \langle\mathcal{O}_{\mu,\nu}\rangle$. From Eq.(28), the variance scaling of the correction with respect to N can also be obtained. Using the fact that the variance of the local energy E_L and the logarithmic derivative $\mathcal{O}_{\mu,\nu}$ scale like $O(N_e)$ and $O(1)$ respectively [36, 85], we can see that $\mathrm{Var}(\mathbf{F}_a^c)$ is bound by $O(L^2N_e)$, with the L^2 term coming from the summation over the AGP variational coefficients. One can, however, exploit the locality of the $\lambda_{\mu,\nu}$ to reduce the variance by a factor of L. In fact, $\lambda_{\mu,\nu}$ corresponding to pairs of atoms that are far away, will usually give negligible values of $\frac{d\lambda_{\mu,\nu}}{d\mathbf{R}_a}$, and may be effectively discarded from the sum. Given that both L and N_e scale proportionally with N, we can conclude that the variance of the bias correction has a $O(N^2)$ scaling in the $N \to +\infty$ limit, which is N times larger than the standard VMC force of Eq. (25)[85]. In practice, for systems of moderate size, we noticed that the variance of the corrected force $\mathbf{F}_a$ is often similar to the one of $\mathbf{F}_a^{VMC}$, indicating that $\mathrm{Var}(\mathbf{F}_a) = \mathrm{Var}(\mathbf{F}_a^{VMC}) + \mathrm{Var}(\mathbf{F}_a^c) \approx O(N) + \varepsilon O(N^2)$ with a small ϵ. Before showing some application of Eq. (28), notice that the correction can also be applied to more general WFs beyond the JSD one, for which only a subset of parameters are optimized, e.g. those satisfying the locality condition within a given cutoff radius. In this case, only the unoptimized coefficients will enter the sum of Eq.(26).

2.3.2 *Applications to the H_2 and Cl_2 molecules*

Here we show two simple applications of the force correction to determine the equilibrium structure of the H_2 and Cl_2 molecules. For the calculations, we used the correlation consistent effective core potentials (ccECPs) [120–123] accompanied with uncontracted cc-pVDZ basis sets. The DFT MOs were prepared using the PySCF package [124, 125] with an LDA exchange-correlation functional [74] and then converted into a format suitable to TurboRVB using the TurboGenius [126] package via TREX-IO files [127]. The Jastrow factor (including one-, two-, and three-body terms) was then opti-

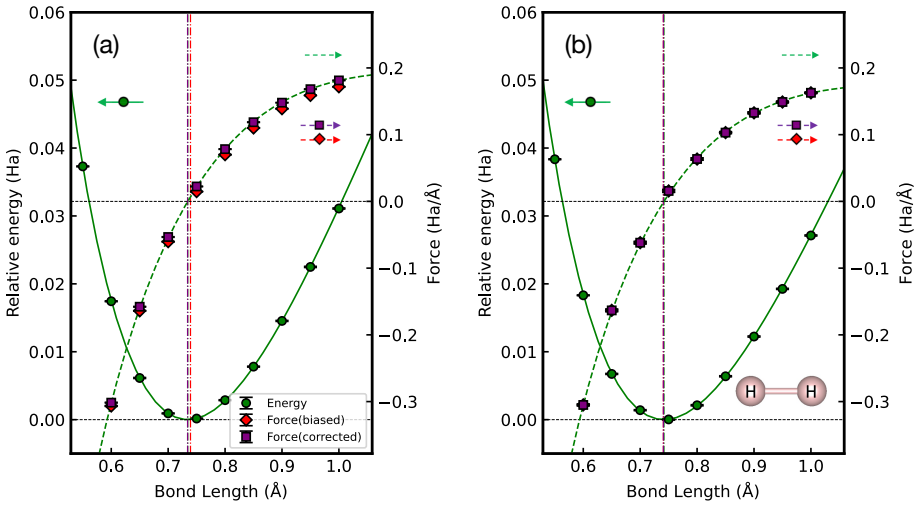

Figure 3: H_2 fitted PESs (solid green lines) with their associated numerical derivatives (dashed green lines). The regular VMC forces and corrected ones are indicated with red and purple symbols, respectively. The calculations were performed with (a) small [1s] and (b) large [4s2p1d] Jastrow basis sets. The PESs and forces are computed from R = 0.30 Å to R = 2.00 Å with 18 equally spaced datapoints plus 5 additional datapoints (0.55 Å, 0.65 Å, 0.75 Å, 0.85 Å, and 0.95 Å). The vertical dashed lines represent equilibrium bond lengths obtained by fitting the PES (forces) with a 11th (10th) order polynomial. The plotted forces are F_x acting on the left atom of the dimer, where the x axis is aligned with the direction of the molecular bond.

mized using the linear method. The coefficients derivatives $\frac{d\lambda_{\mu,\nu}}{d\mathbf{R}}$ were computed using a displacement $\Delta R = 0.001$.

In the first application, we tested the force correction on the simple H_2 dimer. This is a very special case, for which the JSD solution is in principle exact if the Jastrow basis is converged. In particular, this allows us to study different situations, from a poor Jastrow to a refined one. Here we considered two Jastrow basis sets, one consisting only of a [1s] orbital and the other including [4s2p1d] orbitals. The results are shown in Fig.3. The equilibrium distance as obtained from the PES (with the FDM), from the standard VMC force of Eq.(25) and the corrected force, are reported in Tab.1. Notice how, in the case of the under-converged Jastrow factor, the correction mitigates the self-consistency error, while, for the converged one, the Jastrow factor completely compensates for the DFT determinant and the correction is negligible. This example demonstrates that the variational term in Eq.(26) not only removes the bias due to frozen MOs, but also the one caused by an unconverged Jastrow.

We also applied the correction to the Cl_2 dimer. Fig.4 again shows the PES of the system obtained from VMC, using a JSD WF and a [3s1p]

Table 1: The equilibrium bond distances r_{eq} (Å) of the H_2 and Cl_2 molecules obtained from the PESs, the regular VMC force, and the corrected force. The corresponding PESs are shown in Fig. 3 and Fig. 4. Experimental results are taken from Ref. [128].

Dimers	Source	r_{eq} (Å)
H_2 (Jas. [1s])	PES	0.7344(2)
	VMC force	0.7392(1)
	Corrected force	0.7341(1)
	Experiment	0.741
H_2 (Jas.[4s2p1d])	PES	0.7418(3)
	VMC force	0.7408(6)
	Corrected force	0.7408(6)
	Experiment	0.741
Cl_2 (Jas.[3s1p])	PES	1.987(1)
	VMC force	1.9979(1)
	Corrected force	1.9864(1)
	Experiment	1.987

Jastrow basis set. The equilibrium geometries are also reported in Tab.1. Here we can see that the self-consistency error is larger, which is consistent with Ref. [116], reporting that the bias increases with the effective nuclear charge. The results in Tab.1 illustrate that the proposed force correction also works for molecules with higher atomic numbers.

2.3.3 *Pressure correction for cubic BN*

Like the forces, also the virial pressure $P = -\frac{dE}{dV}$ suffers from the self-consistency error and can thus be corrected. The pressure correction has exactly the same form of Eq.(26) with the derivatives $\frac{d\lambda_{\mu,\nu}}{d\mathbf{R}_\alpha}$ replaced with $\frac{d\lambda_{\mu,\nu}}{dV}$. To demonstrate this, we considered cubic boron nitride (cBN). For the calculations, we used the ccECP with an uncontracted cc-pVDZ basis set. The basis set elements with small exponents (less than 0.20 a.u.) were cut in order to solve the linear dependency of the basis set and suppress the statistical error on forces and pressures [129]. A $2 \times 2 \times 2$ supercell (comprising 256 electrons in the simulation box) at the gamma point $\mathbf{k} = \Gamma$ was used. The molecular orbitals in the WF were prepared via the built-in DFT code of TurboRVB, using a LDA [74] exchange-correlation functional. For the B and N atoms, we used a [3s1p] basis for the Jastrow, which we then optimized with the LM at each volume. The variational coefficients

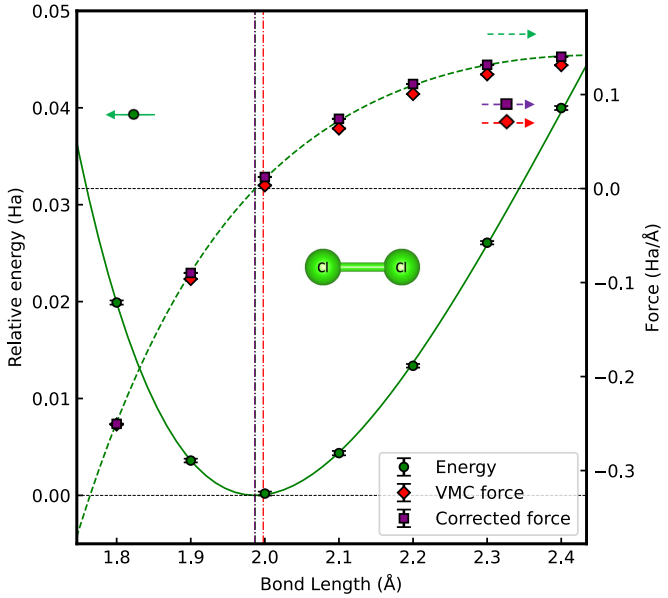

Figure 4: Cl_2 fitted PES (solid green line) with its numerical derivative (dashed green line). The regular VMC forces and corrected ones are indicated with red and purple symbols, respectively. The calculations were performed with a [3s1p] Jastrow basis set. The PES and forces are computed from $R = 0.30$ Å to $R = 2.00$ Å with 18 equally spaced datapoints plus 5 additional datapoints (0.55 Å, 0.65 Å, 0.75 Å, 0.85Å, and 0.95 Å). The vertical dashed lines represent equilibrium bond lengths obtained by fitting the PES (forces) with a 11th (10th) order polynomial. The plotted forces are F_x acting on the left atom of the dimer, where the x axis is aligned with the direction of the molecular bond.

Table 2: Equilibrium lattice parameters and volumes per atom obtained by fitting the EOS, and from the regular VMC pressure and the corrected one. Experimental values are taken from Ref. [128]. Zero-point energy and temperature effects are not included.

Source	Lattice (Å)	Volume (Bohr^3)
EOS	3.5962(3)	5.814(1)
VMC pressure	3.5800(1)	5.7353(7)
Corrected pressure	3.5943(2)	5.8042(7)
Experiment	3.594	

derivatives $\frac{d\lambda_{\mu,\nu}}{dV}$ were computed with the FDM using relative volume variations $\Delta V = \pm 0.3\%$. In Fig. 5 the BN EOS and biased/corrected pressures are shown. In Tab. 2 we also report the equilibrium lattice parameters and volumes obtained by the different approaches. Notice how the self-consistency error introduces a pressure shift ~ 5 GPa on the whole volume range. Our method gives corrections that bring the estimated pressure very close to the unbiased values derived from the EOS.

2.4 WRAP-UP

This Chapter introduced two first-principles frameworks: DFT and QMC. Within the latter approach, we described the WF forms used in this thesis to achieve accurate results in the variational Monte Carlo (VMC) method. We also showed how these WFs can be further improved by means of projection techniques such as LRDMC. These algorithms are implemented in the TurboRVB code [82], which we used for all the QMC calculations in this thesis. We then addressed the problem of obtaining unbiased energy derivatives (i. e., forces and pressures) within the VMC framework, in case of wave functions that are not fully optimized. This is particularly important, since one of the most used WFs is the JSD, where the parameters of the Jastrow part are at the variational minimum, while the Slater determinant part is taken directly from DFT calculations. We demonstrated the effectiveness of our correction in both molecular and bulk systems, by removing the bias in the force and pressures, respectively. As we will explore in Ch. 5, the availability of unbiased quantities, computed at a relatively cheap cost given by the combination of the VMC method and the JSD WF, opens the possibility to generate affordable datasets for training QMC-based machine learning potentials.

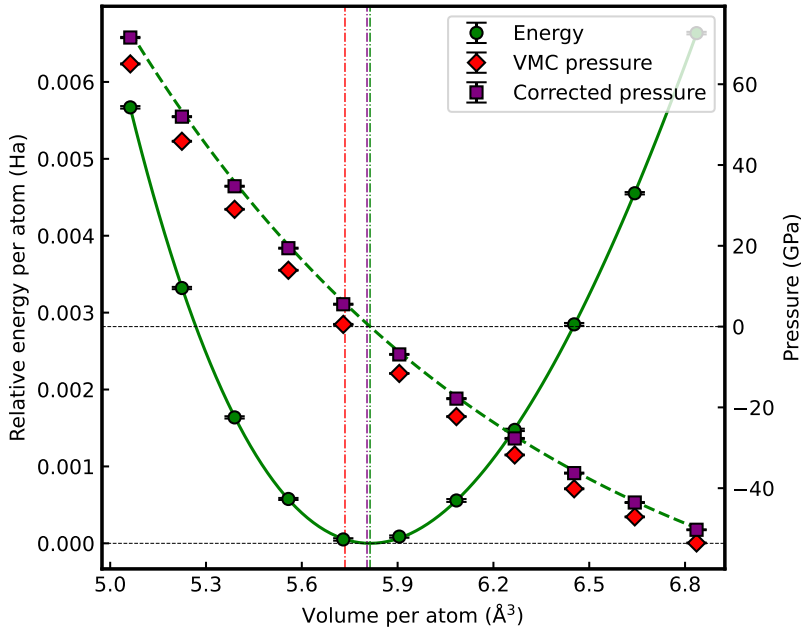

Figure 5: Cubic BN equation of state (solid green line) and negative derivative (dashed green line). The regular and corrected VMC pressures are indicated with red diamonds and purple squares, respectively. The vertical dashed lines represent equilibrium volumes obtained by fitting the EOS and pressures with the Vinet forms [130].

3

ACCELERATING MOLECULAR DYNAMICS SIMULATIONS VIA MACHINE LEARNING POTENTIALS

Electronic structure methods, such as DFT or QMC, can be used to compute the PES of the system, i. e., $E_0(\mathbf{R}) = E_0(\mathbf{R}_1, \ldots, \mathbf{R}_N)$ in Eq.(2), for any given $\mathbf{R}$. Once this is done, thermodynamic average quantities can be obtained by sampling the ionic configurations according to the desired statistical ensemble. The two most popular sampling methods are molecular dynamics (MD) and Monte Carlo (MC) simulations. In this Chapter only the former will be discussed; the latter method is frequently used in combination with QMC, e. g., in the CEIMC approach [131].

In Sec. 3.1 and Sec. 3.2, we will introduce the classical MD and path integral MD approach, respectively. We will then switch our focus to machine learning potentials (MLPs) and how they can be used to perform MD simulation with the accuracy of first-principles methods at a fraction of the computational cost (Sec. 3.3). The specific features of MLPs trained on QMC reference data will be discussed in Sec.3.3.1. Finally, we will describe the two main machine learning models used in this thesis. The first approach, for which we developed our own implementation, is based on the kernel ridge regression method (Sec. 3.4) in combination with the Δ-learning technique (Sec. 3.4.1), a similarity kernel with the smooth overlap of atomic positions (SOAP) descriptor (Sec. 3.4.2), and a sparsification method to reduce the number of model parameters (Sec. 3.4.3). The second approach, a recently introduced [132] equivariant neural network called MACE, will be discussed in Sec. 3.5. A summary of this Chapter is reported in Sec. 3.6.

3.1 CLASSICAL MD SIMULATIONS

For sufficiently high temperatures and low densities, the motion of the nuclei is well approximated by classical mechanics, and the equations of motion can be integrated, e. g., using the velocity Verlet algorithm. This approach can be used to effectively simulate a system with constant total energy and represents the simplest example of classical MD. It is often more convenient to work in alternative ensembles, such as the isothermal-isocore (NVT) or isothermal-isobaric (NPT) ones. For NVT simulations, several thermostats are available, including the Nosé-Hoover one [133–135] and the stochastic velocity rescaling method [136]. In NPT molecular dynamics, a barostat is also required to regulate pressure, with the Parrinello-Rahman baro-

stat [137] being a commonly used choice. An alternative approach for simulating the canonical ensemble is given by Langevin dynamics [138, 139], i. e.,

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{F} - \gamma\mathbf{P}(t) + \boldsymbol{\eta}(t), \tag{29}$$

where $\mathbf{P} \equiv (\mathbf{P}_1, \ldots, \mathbf{P}_N)$ with $\mathbf{P}_\alpha = M_\alpha \frac{d\mathbf{R}_\alpha}{dt}$, γ is a friction coefficient and $\boldsymbol{\eta}_\alpha(t)$ is a random force term. Thanks to the fluctuation-dissipation theorem [140], the correct canonical distribution at temperature T can be achieved by taking $\langle\boldsymbol{\eta}(t)\rangle = 0$ and $\langle\eta_i(t)\eta_j(t')\rangle = 2\gamma k_B T\delta(t-t')\delta_{ij}$.

Running an MD simulation requires the calculation of the ionic forces at each time step, which can be done using a variety of approaches. Fast-to-evaluate atomistic potentials allow for simulations of a large number of particles and very long trajectories (of the order of several ns or even µs), but their accuracy is usually limited. On the contrary, electronic structure methods require a much larger computational time, but they can better capture the physics of the system. For example, ionic forces can be computed within DFT by using the Hellmann-Feynman theorem [141, 142]. This approach is known as *ab initio* molecular dynamics (AIMD) or Born-Oppenheimer MD (BOMD) [143]. A more computationally efficient method, still within the DFT framework, is represented by Car-Parrinello molecular dynamics (CPMD) [144]. In CPMD, a modified Lagrangian is introduced, containing both ionic and electronic degrees of freedom, and assigning to the latter a fictitious mass parameter µ. For an appropriate choice of µ, the electronic subsystem "oscillates" around the BO PES during the dynamics, thus removing the need for a full DFT calculation at each step.

QMC methods can also be used as a driver for MD, further improving the accuracy of the results. Notice how, in this case, forces have an associated noise. This feature can be exploited to include noisy QMC forces in the Langevin dynamic framework, by appropriately modifying the stochastic term $\boldsymbol{\eta}(t)$ in Eq.(29) to also include non-diagonal elements describing the correlation between different components [113, 145, 146].

3.2 PATH INTEGRAL MD

As the temperature of the system is decreased, the classical description of the ions starts to break down and nuclear quantum effects (NQEs) have to be considered. For hydrogen, this is particularly important since, due to its light mass, NQEs remain significant up to relatively high temperatures ($T \leqslant 2000K$). One of the most common methods to include quantum effects is the PIMD approach [147]. Within PIMD the original quantum problem is mapped into a classical system made of M replicas ("beads") of the original system,

interacting among themselves with a harmonic potential and at a temperature $M \times T$. In particular, the quantum partition function $Z = \mathrm{Tr}\left[\exp(-\beta\hat{H})\right]$ is factorized using a Trotter decomposition

$$Z \approx \frac{1}{(2\pi\hbar)^{MN}} \int \left(\prod_{j=1}^{M} d^{3N}\mathbf{R}^{(j)} d^{3N}\mathbf{P}^{(j)} \right) e^{-\tau H_M(\{\mathbf{R}\},\{\mathbf{P}\})}, \tag{30}$$

where we set $\tau = \beta/M$ and indicated with $\{\mathbf{R}\} = \left(\mathbf{R}^{(1)}, \ldots, \mathbf{R}^{(M)}\right)$ and $\{\mathbf{P}\} = \left(\mathbf{P}^{(1)}, \ldots, \mathbf{P}^{(M)}\right)$ the collection of all coordinates and momenta of the beads. The Hamiltonian of the classical system is given by

$$\begin{aligned} H_M(\{\mathbf{R}\},\{\mathbf{P}\}) = &\sum_{i=1}^{N}\sum_{j=1}^{M} \left[\frac{1}{2}\left(\mathbf{P}_i^{(j)}\right)^2 + \frac{1}{2} M_i \omega_M^2 \left(\mathbf{R}_i^{(j)} - \mathbf{R}_i^{(j-1)}\right)^2 \right] \\ &+ \sum_{j=1}^{M} V\left(\mathbf{R}^{(j)}\right), \end{aligned} \tag{31}$$

where V is the potential energy of the system, $\omega_M = M/\beta\hbar$ and $\mathbf{R}^{(0)} \equiv \mathbf{R}^{(M)}$. Once the quantum-to-classical isomorphism is established, a thermostat can be applied to control the temperature [147–149]. An efficient choice is given by the path integral Ornstein-Uhlenbeck dynamics (PIOUD) method [150]. This approach applies a Trotter breakup on the Liouville operator, separating the physical modes of the system from the harmonic ones, and explicitly integrating the latter together with the thermostat, which is done analytically. This is at variance with the path integral Langevin equation (PILE) algorithm [147], whose Liouvillian contains an additional Trotter breakup between the integration of the harmonic modes and their thermostatting step. Compared to PILE, the PIOUD method shows better stability in both the simulation time step and the number of replicas M, while also being able to incorporate noisy forces directly in the thermostat.

3.3 MACHINE LEARNING POTENTIALS

The large computational burden of first-principles methods strongly limits both the size of the system and the length of the trajectories that one can realistically simulate with MD. For example, AIMD simulations are only possible for a few thousand atoms and ~ 100 ps. The problem is even more severe if one wants to use QMC as an energy and force driver; in this case, the simulation of hundreds of atoms for a few pico-seconds is already challenging. For this reason, the range of applicability and popularity of AIMD have been limited compared to those of simulations using atomistic potentials.

This situation dramatically changed during the last two decades with the introduction of machine learning potentials (MLPs) [38]. The basic goal of an MLP is to approximate the PES $E(\mathbf{R})$ of a given N

atom configuration, starting from a set of reference electronic structure data (e.g. energies, forces), called the training set, usually computed with *ab initio* methods such as DFT or QMC. In practice, an MLP is a very flexible function containing a large number of parameters to be optimized, such as a neural network (NN), and, contrary to physical atomistic potentials, does not contain any a priori assumption on the system. When the model is properly trained, i.e., its parameters are optimized, an MLP can predict energies and energy derivatives on unseen configurations, with similar accuracy to the original *ab initio* method, but at a computational cost several orders of magnitude smaller. MLPs were first applied in the context of simulating small molecules, reactive molecular systems, and interactions of molecules with frozen (i. e., with fixed positions) metallic surfaces [37, 151]. These "first-generation" MLPs, following the classification of Ref. [152], were limited to small systems, since they could only be applied to configurations with the same number of atoms as the training ones. A big improvement, in particular regarding the application to extended systems and MD, was made in the so-called "second-generation" MLPs [153]. In these models, the basic assumption is that E can be expressed as a sum of atomic contribution, each depending on the relative coordinates of all the other atoms with respect to the central one:

$$E = \sum_{i=1}^{N} e(\mathcal{R}_i) \quad \text{where} \quad \mathcal{R}_i = (\mathbf{R}_{i1}, \ldots, \mathbf{R}_{iN}) \quad \text{with} \quad \mathbf{R}_{ij} \equiv \mathbf{R}_j - \mathbf{R}_i. \tag{32}$$

Here, we will refer to $\mathcal{R}_i$ as the local environment around the i-th atom and to $e(\mathcal{R}_i)$ as the local atomic energy, and suppose the minimum image convention holds when periodic boundary conditions are applied. The function $e(\mathcal{R}_i)$ in Eq.(32) can often be considered local, i. e., to only depend on the atoms closer than a certain cutoff radius r_c. Different MLPs can be distinguished by their specific representation of the local atomic energy. A wide range of techniques has been successfully applied during the years, including neural networks, gaussian approximation potentials [154, 155] , kernel ridge regression [39, 156] , message-passing networks [132, 157] and more. Irrespective of the specific architecture, $e(\mathcal{R}_i)$ will usually depend on a set of parameters $\{\beta\}$ to be optimized:

$$e(\mathcal{R}_i) \equiv e(\mathcal{R}_i; \{\beta\}). \tag{33}$$

The vast majority of MLPs implement a functional form that inherently satisfies some of the symmetries required by physical constraints, such as permutational and rotational invariance. This is often done by considering a mapping from the real coordinates $\mathbf{R}_1, \ldots, \mathbf{R}_N$ into a set of symmetry-preserving input variables, called descriptors. Like the ML architectures themselves, nowadays there exists a wide choice of descriptors such as atom-centered symmetry functions (ACSF) [153],

smooth overlap of atomic positions (SOAP) [158], atomic cluster expansion (ACE) [159] and more.

The parameters (or weights) in Eq. (33) are optimized by minimizing the loss function $\mathcal{L}(\{\beta\})$, which quantifies the error of the prediction with respect to the training data. In particular, we indicate as $\mathcal{R}_i^\mu \equiv (\mathbf{R}_1^\mu - \mathbf{R}_i^\mu, \ldots, \mathbf{R}_{N_\mu}^\mu - \mathbf{R}_i^\mu)$ the i-th atomic environment belonging to the μ-th configuration of the training set (of total dimension N_t), with $\mu = 1, \ldots, N_t$, having N_μ atoms and volume V_μ. Moreover let E_{ref}^μ, $\mathbf{F}_{ref}^{\mu,j}$ and P_{ref}^μ be the reference values of the total energy, the force acting on the j-th atom and the (isotropic) virial pressure of the configuration. A possible choice of the loss function is then given by

$$\begin{aligned}\mathcal{L}(\{\beta\}) = w_E \frac{1}{N_t} \sum_{\mu=1}^{N_t} \left[\frac{1}{N_\mu}\left(E_{pred}^\mu(\{\beta\}) - E_{ref}^\mu\right)\right]^2 \\ + w_F \frac{1}{N_t} \sum_{\mu=1}^{N_t} \frac{1}{3N_\mu} \sum_{j=1}^{N_\mu} \left[\mathbf{F}_{pred}^{\mu,j}(\{\beta\}) - \mathbf{F}_{ref}^{\mu,j}\right]^2 \\ + w_P \frac{1}{N_t} \sum_{\mu=1}^{N_t} \left[P_{pred}^\mu(\{\beta\}) - P_{ref}^\mu\right]^2,\end{aligned} \tag{34}$$

where the predicted energies, forces, and pressures are computed as

$$E_{pred}^\mu(\{\beta\}) = \sum_{i=1}^{N_\mu} e(\mathcal{R}_i^\mu; \{\beta\}),$$

$$\mathbf{F}_{pred}^{\mu,j}(\{\beta\}) = -\nabla_j \left[\sum_{i=1}^{N_\mu} e\left(\mathcal{R}_i^\mu; \{\beta\}\right)\right], \tag{35a}$$

$$P_{pred}^\mu(\{\beta\}) = -\frac{\partial}{\partial V_\mu}\left[\sum_{i=1}^{N_\mu} e\left(\mathcal{R}_i^\mu; \{\beta\}\right)\right]. \tag{35b}$$

The three terms in Eq.(34) are the mean squared error (MSE) of the energy per atom, ionic forces, and virial isotropic pressure, respectively, and w_E, w_F, and w_P are tunable weights multiplying the different MSEs. Other functional forms of the loss function are possible, which can also include additional physical quantities (e.g., charges, all 6 independent components of the stress, etc...). The loss function minimization can be performed with several algorithms, depending on the specific ML method implemented. For example, within NNs, the gradient of $\mathcal{L}(\{\beta\})$ is usually computed using only a small batch of training configurations, and the parameters are optimized using stochastic gradient descent. The derivatives of the loss function with respect to the parameters can be computed with adjoint algorithmic differentiation, which is often called backpropagation [160] in this context.

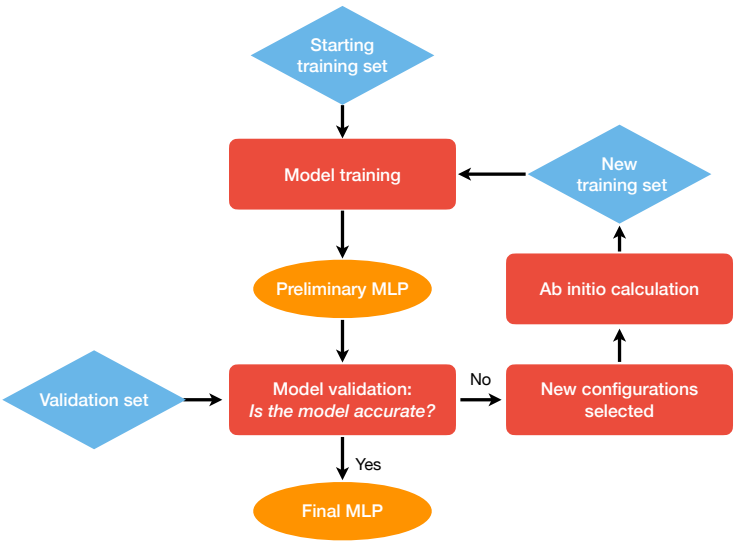

Figure 6: Schematic flowchart of the active learning scheme used to construct machine learning potentials. The model is initially trained on a starting set of *ab initio* data points and then validated using a validation set. New configurations are iteratively added, filling the holes in the training set, until a good accuracy of the model is attained.

The selection of the configurations belonging to the training set is one of the most crucial tasks in the construction of an MLP. In particular, an incomplete sampling of the relevant phase space or "holes" in the dataset can result in incorrect predictions. For this reason, it is essential to validate the potential by checking its accuracy on another set of configurations, different from the training one, called validation or test set. This is often done by computing the root mean squared error (RMSE), which for a given quantity X reads

$$\mathrm{RMSE}_X = \sqrt{\frac{1}{N_{\mathrm{test}}} \sum_{\mu=1}^{N_{\mathrm{test}}} \left(X^{\mu}_{\mathrm{pred}} - X^{\mu}_{\mathrm{ref}} \right)^2},$$

with N_{test} being the test set dimension. At this point, if the performance of the MLP is not satisfactory, other configurations are added to the training set, and the whole process is repeated. This iterative procedure, sketched in Fig.6, is called active learning in the ML community. The new configurations can be obtained in several ways, e.g. by using different MLPs trained on the same training set but with a different random initialization and studying the variance of their predictions ("query by committee") [161]. In the end, the process is stopped when the model is considered to be accurate enough, i. e., when no more configurations need to be added.

3.3.1 *QMC-based MLPs*

Now that we have described some generalities about MLPs and their construction, let us switch our focus to models trained using QMC data points. These MLPs represent a very small fraction of those present in the literature, which are usually trained on DFT datasets. We can identify three main aspects that differentiate QMC-based MLPs.

- Contrary to DFT, any dataset generated using QMC is fundamentally noisy. This is not new in the ML community, where noisy data are commonly encountered. Recent studies and applications have shown that this "unbiased noise" is not detrimental and even QMC data with large error bars can be effectively used to build good quality MLPs [42].

- QMC calculations usually require about $10-100$ the computational time of a DFT calculation. This makes the generation of large datasets (10^4-10^5 configurations) extremely challenging. Possible solutions to this problem include selecting only the most relevant structures in the training set, using particularly data-efficient ML architectures, or employing techniques such as Δ-learning (see Sec.3.4.1).

- As discussed in the previous chapter, the calculation of QMC energy derivatives is not straightforward. Indeed, within VMC, the usual expression for ionic forces and pressure is affected by the self-consistency error when the WF is not at the full variational minimum. Even if an MLP produces energy derivatives that are by definition consistent, the inclusion of biased data in the training set translates into a strong model dependence on the different weights w_E, w_F and w_P in Eq.(34), and can produce inaccurate models (see Ch. 5). This is particularly important when the magnitude of the bias approaches that of the target quantity.

In the rest of this Chapter, we will describe the specific architectures and techniques we used to construct the MLPs presented in this thesis, namely kernel ridge regression and MACE.

3.4 KERNEL RIDGE REGRESSION

The first approach we will explore is a combination of kernel ridge regression (KRR) and the Δ-learning technique. This method was originally employed to study the liquid-liquid transition in hydrogen in Ref. [39] (see Ch. 6). We will also discuss the details of our implementation of this framework, named TurboML, developed using a combination of Fortran and Python.

3.4.1 *Δ-learning approach*

The Δ-learning method [162] is a strategy that expresses the target *ab initio* PES as the sum of a computationally cheaper baseline potential and a correction given by an MLP. In our applications, the target method is QMC (variational or diffusion) while the baseline is given by DFT:

$$E^{QMC}(\mathbf{R}) = E^{DFT}(\mathbf{R}) + \Delta^{MLP}(\mathbf{R}).$$

The advantage of this approach is based on the observation that constructing an MLP that reproduces the correction Δ^{MLP} usually requires a much smaller training set than learning E^{QMC} directly. This is particularly useful for expensive electronic structure methods. Indeed, a sufficiently high accuracy can often be obtained with less than 10^3 data points.

When running MD simulations within the Δ-learning approach, the total energy, forces, and pressure are computed at each time step with both the baseline potential and the MLP giving the correction. Most of the time this implies that the computational cost of the simulation will be given by that of the baseline, i. e., DFT. Although expensive, the resulting dynamics are much faster than an *ab initio* QMC MD, e. g., about three/four orders of magnitude for systems of a few hundred atoms. This type of approach will be particularly effective when high accuracy is needed and relatively short simulations are sufficient. Alternatively, a faster framework can be obtained by replacing the DFT baseline with cheaper methods or even with an MLP trained on DFT data.

3.4.2 *Kernel regression with SOAP*

A crucial aspect to consider if one wants to construct an accurate MLP with limited training points is the choice of the particular architecture of the model and the description of the local environment around each atom. In Refs [39, 41], we used KRR, an approach that expresses the local atomic energy in Eq.(33) as

$$e(\mathcal{R};\{\beta\}) = \sum_{\mu=1}^{N_{env}} \beta_\mu \mathcal{K}(\mathcal{R},\mathcal{R}_\mu), \tag{36}$$

where we indicated with $\mathcal{K}(\mathcal{R},\mathcal{R}_\mu)$ the "normalized kernel" between the local environment to be predicted $\mathcal{R}$ and one of the N_{env} local environments belonging to the training set $\mathcal{R}_\mu$. We can interpret $\mathcal{K}(\mathcal{R},\mathcal{R}_\mu)$ as a measure of the similarity between the two local environments. In particular, for any pair $\mathcal{R}_i, \mathcal{R}_j$, we have that $0 \leqslant \mathcal{K}(\mathcal{R},\mathcal{R}_\mu) \leqslant 1$, with $\mathcal{K}(\mathcal{R}_i,\mathcal{R}_j) = 0$ representing "ortogonality", and $\mathcal{K}(\mathcal{R}_i,\mathcal{R}_j) = 1$ indicating that $\mathcal{R}_i$ and $\mathcal{R}_j$ are equivalent. The kernel we used is based

on the smooth overlap of atomic positions (SOAP) [158, 163]. Within SOAP, each environment $\mathcal{R}_i$ is described by a smooth density

$$\rho(\mathbf{r}, \mathcal{R}_i) \propto \sum_{|\mathbf{R}_{ij}| \leqslant r_c} f_c\left(|\mathbf{R}_{ij}|\right) \exp\left(-\frac{|\mathbf{r} - \mathbf{R}_{ij}|^2}{2\sigma^2}\right), \tag{37}$$

where $f_c(r)$ is a function that goes smoothly to zero for r larger than the cutoff radius r_c and σ is a hyperparameter of the model. Notice how in Eq.(37) we only considered one atomic species for simplicity. The similarity kernel between two local environments $\mathcal{R}_i$, $\mathcal{R}_j$ is then written as

$$K\left(\mathcal{R}_i, \mathcal{R}_j\right) = \int_{O(3)} d\hat{U} \left[\int d^3\mathbf{r} \rho(\mathbf{r}; \mathcal{R}_i) \rho(\mathbf{r}; \hat{U}\mathcal{R}_j)\right]^n, \tag{38}$$

where we integrated over all possible rotations $\hat{U}$ (including reflections) the n-th power (n being another hyperparameter) of the overlap of the two densities. Evaluating Eq.(38) can be done using the bispectrum [158]. The TurboML code implements an alternative form of the kernel using only a discrete subgroup of N_{sym} symmetry operators $\hat{U}_k$

$$\widetilde{K}\left(\mathcal{R}_i, \mathcal{R}_j\right) = \frac{1}{N_{sym}} \sum_{k=1}^{N_{sym}} \left[\int d^3\mathbf{r} \rho(\mathbf{r}; \mathcal{R}_i) \rho(\mathbf{r}; \hat{U}_k\mathcal{R}_j)\right]^n. \tag{39}$$

The set of symmetries can, for example, be taken equal to the cubic or icosahedral symmetry groups, and then Eq.(39) can be evaluated trivially. This approach is reminiscent of the implementation of non-local pseudopotentials in QMC [164] where a similar angular integration to the one in Eq.(38) appears. Finally, the normalized kernel $\mathcal{K}\left(\mathcal{R}_i, \mathcal{R}_j\right)$ is given by

$$\mathcal{K}\left(\mathcal{R}_i, \mathcal{R}_j\right) = \left[\frac{\widetilde{K}\left(\mathcal{R}_i, \mathcal{R}_j\right)}{\widetilde{K}\left(\mathcal{R}_i, \mathcal{R}_i\right) \widetilde{K}\left(\mathcal{R}_j, \mathcal{R}_j\right)}\right]^\eta, \tag{40}$$

with a power η that can be used to tune the selectivity of the kernel.

The training of the model weights $\{\beta\}$ in Eq. (36) is performed by minimizing a loss function like the one in Eq.(34). This is equivalent to solving the linear system:

$$\mathbf{A}\boldsymbol{\beta} = \mathbf{b}, \tag{41}$$

where $\mathbf{A}$ is a $N_{env} \times N_{env}$ matrix. When inverting $\mathbf{A}$ a regularization is necessary to obtain a stable solution for the coefficients $\{\beta\}$ and avoid ill-conditioning.

Our implementation TurboML is parallelized using a hybrid OpenMP/MPI paradigm to compute the total energy in Eq. (32) and the kernel in Eq. (39), distributing over the local environments and symmetries, respectively. The training uses a parallel implementation of the conjugate gradients method to invert the matrix in Eq. (41) and computes local atomic energy derivatives with AAD.

3.4.3 *Local furthest point sampling sparsification*

When constructing an MLP using KRR, a careful selection of the reference points to be included in the training set is crucial. In fact, the computational cost of both model training and evaluation scales polynomially with the training set dimension N_{env}, and discarding "useless" local environments can drastically improve the model efficiency. To do this, a common approach is the so-called sparsification, i.e., the selection of a representative subset of points S from a larger dataset X. In our implementation, the furthest point selection (FPS) algorithm was used, which can be described as follows:

1. In the first iteration, select a random point $x \in X$ and add it to the point selection S.
2. At each iteration, compute the quantity
$$d(y, S) = \max_{x \in S} \mathcal{K}(x, y) \tag{42}$$
for each point $y \in X \backslash S$, which measures the maximum similarity of y to the points already selected.
3. Chose a new point $\bar{y}$ from $X \backslash S$ such that
$$\bar{y} = \arg\min_{y \in X \backslash S} d(y, S).$$
4. Repeat from point 2 until S has the desired dimension N_{env} or $d(y, S) \geqslant \gamma$ for all $y \in X \backslash S$, where $0 < \gamma < 1$ is a given threshold.

In our code, the evaluation of Eq. (42) is performed in a distributed way and the allocation of the kernel matrix $K(x, y)$ can be split over several processes. Notice how, by applying the FPS to a dataset of N_t configurations, only a subset of $n_\mu \leqslant N_\mu$ local environments will be selected from the μ-th configuration with N_μ atoms. In Eq.(34) we will only keep the conditions corresponding to configurations that have $n_\mu \neq 0$. With this local version of the FPS, we can efficiently train MLPs on a small subset of local environments, achieving accuracy comparable to models trained on the entire dataset. In particular, a model trained on a sparsified set with $N_{env} \simeq 5 - 10 \times 10^3$ environments often has an RMSE on energy and forces only $\sim 5\%$ larger than a model trained on the full training set, while being ~ 10 times faster to evaluate (see Fig.7).

3.5 MACE MODELS

MACE [132] is a recently introduced framework combining message passing neural network (MPNN) [157, 165] and high body-order messages to construct very accurate MLPs. In the context of atomistic simulations, an MPNN describes the system as a labeled graph, where

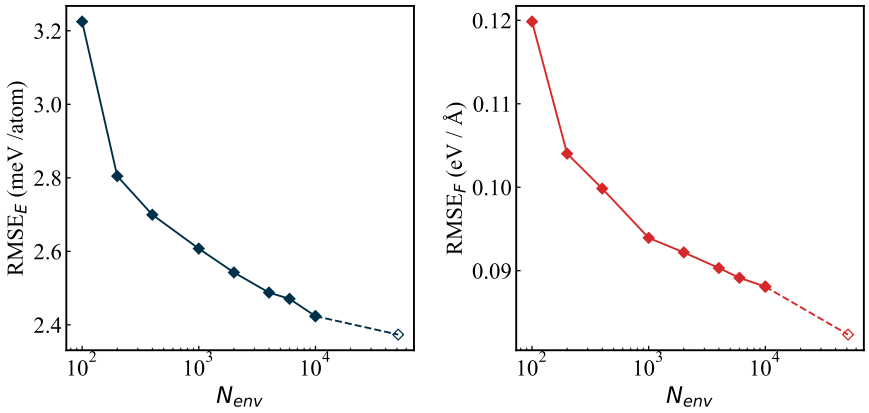

Figure 7: Comparison of the accuracy of MLPs trained on sparsified datasets with different numbers of local environments N_{env}, measured using the RMSE relative to energy and forces on the test set. The training and test sets comprised 400 configurations and 50 configurations, respectively, with 128 hydrogen atoms each. The models were trained on the difference between reference data obtained within DFT but with two different functionals, namely the Perdew-Burke-Ernzerhof (PBE) and LDA functionals, respectively. The last point in each panel (empty symbol) corresponds to the model trained on the full set of local environments.

each atom i corresponds to a node. The state of each node σ_i in a given layer t of the NN is represented by a tuple

$$\sigma_i^{(t)} = \left(\mathbf{R}_i, Z_i, \mathbf{h}_i^{(t)}\right),$$

where $\mathbf{R}_i$ and Z_i denote the position and atomic number of atom i, $\mathbf{h}_i^{(t)}$ is a learnable feature vector and $\mathbf{h}_i^{(1)} = Z_i$. In equivariant MPNNs, the feature vectors $\mathbf{h}_i^{(t)}$ transform as

$$\mathbf{h}_i^{(t)}\left(\hat{U}\left(\mathbf{R}_1, \ldots, \mathbf{R}_N\right)\right) = D\left(\hat{U}\right)\mathbf{h}_i^{(t)}\left(\mathbf{R}_1, \ldots, \mathbf{R}_N\right)$$

for a general rotation $\hat{U}$, where $D(\hat{U})$ is a Wigner D-matrix belonging to an irreducible representation of the $O(3)$ group. Notice how the usual "invariance" is obtained in the special case where $D(\hat{U}) \equiv \mathbb{1}$. The graph's edges connect nodes corresponding to atoms within the model cutoff radius r_c. A forward pass of the network is performed by first constructing the messages

$$\mathbf{m}_i^{(t)} = \bigoplus_{|\mathbf{R}_{ij}|<r_c} M_t\left(\sigma_i^{(t)}, \sigma_j^{(t)}\right),$$

where $\bigoplus_{j|\mathbf{R}_{ij}<r_c}$ is a permutational invariant operation over the neighbours of the i-th atom and M_t is a message function. Both $\bigoplus_{j|\mathbf{R}_{ij}<r_c}$ and M_t are learnable. The features of the next layer are then updated as

$$\mathbf{h}_i^{(t+1)} = F_t\left(\sigma_i^{(t)}, \mathbf{m}_i^{(t)}\right),$$

where F_t also depends on trainable weights. Finally, after T forward steps, the local atomic energy of the system is obtained using a set of learnable "readout" functions G_t

$$e(\mathcal{R}_i) = \sum_{t=1}^{T} G_t\left(\sigma_i^{(t)}\right).$$

In this last step, only invariant features are selected, such that $e(\mathcal{R}_i)$ has the correct transformation property. The training procedure is performed by minimizing a loss function as the one in Eq. (34).

Within MACE, the messages include many-body interactions between the nodes up to a given order $(\nu+1)$, i. e.,

$$\mathbf{m}_i^{(t)} = \sum_j \mathbf{u}_1\left(\sigma_i^{(t)}, \sigma_j^{(t)}\right) + \sum_{j_1,j_2} \mathbf{u}_2\left(\sigma_i^{(t)}, \sigma_{j_1}^{(t)}, \sigma_{j_2}^{(t)}\right) + \ldots + \sum_{j_1,\ldots j_\nu} \mathbf{u}_\nu\left(\sigma_i^{(t)}, \sigma_{j_1}^{(t)}, \ldots \sigma_{j_\nu}^{(t)}\right). \quad (43)$$

The supposedly exponential scaling of Eq. (43) with respect to the body-order ν is solved by making use of the atomic cluster expansion (ACE) approach [159], which allows for the computation of a complete set of basis functions of any order at a constant cost, independent of ν. This means that Eq. (43) can be effectively evaluated as a linear combination of features (see Ref. [132]). The ACE approach has been demonstrated to be complete [166], i. e., to be capable of describing any $(\nu+1)$ order equivariant function in the limit of infinite features, and includes as special cases many atomic environment descriptors, such as ACSF and SOAP. By incorporating higher order messages, MACE models achieve faster learning rates compared to other MPNNs, while also needing fewer message-passing steps T to converge, namely $T = 2$. As a result, MACE MLPs have been proven to be extremely accurate, while also showing good scalability and generalization capabilities.

3.6 WRAP-UP

In this Chapter, we discussed how machine learning potentials (MLPs) allow for molecular dynamics (MD) simulations with the accuracy of first-principles methods but at a fraction of the computational cost. We introduced the Δ-learning approach, a framework that can be utilized to construct accurate models with a relatively small number of data points. This is particularly important for QMC-based MLPs, for which the generation of large datasets is challenging. We also presented the two architectures employed in our work. The first one is kernel ridge regression (KRR), for which we developed our own parallel implementation, TurboML, based on a modified SOAP kernel and a local sparsification approach to reduce the number of model param-

eters. Applications of this method for studying the deuterium Hugoniot curve and the hydrogen liquid-liquid phase transition (LLPT) will be presented in Chs. 4 and 6, respectively. The second approach, MACE [132], is an equivariant message passing neural work, which, thanks to higher-order messages, is able to reach a very high accuracy compared to other ML architectures. In Ch. 6, we will apply a MACE model trained on DFT data to study the first-order character of the LLPT.

4

QMC CALCULATION OF THE DEUTERIUM HUGONIOT CURVE

Dynamic compression techniques [23] have been extensively used during the last decades to study high-pressure hydrogen and have provided invaluable insights into the physics of this system. These experiments measure the properties of a sample after it is compressed by one or multiple shock waves, which can be generated using several methods such as gas guns [167–169], lasers [170–178], magnetically driven flyer plates [179–181] and explosives [182–186]. If the resulting wave is sufficiently planar and the process adiabatic, the end state of the system after one shock will satisfy a set of conservations laws, called Rankine-Hugoniot (RH) equations [187], that define the principal Hugoniot curve. Deuterium is often used, because its higher density and lower shock impedance with respect to hydrogen allow for the realization of higher pressures.

The experimental data for the deuterium Hugoniot have been used to benchmark the different equations of state (EOSs) obtained using theoretical or numerical methods. These include EOSs based on chemical models [29–31, 33] or first-principles methods, such as DFT [188–197], restricted path integral Monte Carlo (RPIMC) [12, 198] or CEIMC [199, 200]. However, this comparison is difficult due to the large uncertainties of the experimental data. In fact, measures on the sample have to be done extremely quickly and the value of quantities like density, pressure, and temperature is often inferred using other observations. This often resulted in discrepancies between different experiments [201]. Recently, more accurate measures of the Hugoniot have been made available [170, 178, 180, 181], with a relative error as small as 2% for the compression peak in the region of $\sim$ 50 GPa, which allow for better benchmarks of numerical EOSs.

In this chapter, we will show an application of the KRR and Δ-learning approach described in Sec. 3.4 to the construction of QMC-based MLPs for determining the principal Hugoniot curve of deuterium [41]. In Sec. 4.1 we will describe the simulations we used for calculating the Hugoniot. We will then explain how finite temperature effects can be approximately taken into account within the Δ-learning approach (Sec. 4.2). The results obtained with our MLPs will be presented in Sec. 4.3. Next, we will provide further details on our calculations, by showing the method we used for computing the energy of the unshocked state (Sec. 4.4), a comparison with previous QMC results (Sec. 4.5), a description of the dataset we used to train our MLPs (Sec. 4.6), and the details of the QMC reference calcu-

lations (Sec. 4.7). Finally, the accuracy of our models will be assessed in Sec. 4.8. The results of this Chapter are summarized in Sec. 4.9.

4.1 HUGONIOT DETERMINATION: GENERAL ASPECTS

To estimate the principal Hugoniot from MD simulations, we made use of the RH jump equation [187]:

$$H(v, T) = e(v, T) - e_0 + \frac{1}{2}(v - v_0)[P(v, T) + P_0] = 0, \tag{44}$$

where v, T, $e(v, T)$, $p(v, T)$ and v_0, T_0, e_0, p_0 are the volume per atom, temperature, energy per atom, and pressure of the final and initial states, respectively.

In particular, we ran molecular dynamic simulations in the NVT ensemble for a system of $N = 128$ atoms at several temperatures and densities, corresponding to the range where the zero of $H(v, T)$ was expected, accumulating thermodynamic data over trajectories of a few picoseconds. Following the Δ-learning scheme, these simulations were performed using MLPs trained on the difference between QMC (both variational and diffusion) and DFT reference calculations with both the local density approximation (LDA) [74] and the Perdew-Burke-Ernzerhof (PBE) [75] functionals. At each step, the energy, forces, and pressure were calculated at the DFT level using the Quantum Espresso package in its GPU accelerated version [202–204] with the chosen functional (PBE in most cases), and then summed with those predicted by the MLP. The resulting dynamics have the same cost as standard DFT AIMD simulations, which is roughly 10^3 times smaller than the original QMC one. For the DFT simulations, a 60 Ry plane-wave cutoff with a projector augmented-wave (PAW) pseudopotential [205] was used together with a $4 \times 4 \times 4$ Monkhorst-Pack $\mathbf{k}$-point grid, while for the dynamics we used a time step of 0.25 fs and a Langevin thermostat [113, 206] with damping $\gamma = 0.13$ fs^{-1}. For each temperature T, Eq.(44) was solved to determine the Hugoniot coordinates (v^*, p^*). This is done by fitting the Hugoniot function $H(v, T)$ and the pressure $p(v, T)$ with a spline function, and by numerically finding v^* and the corresponding p^*. The reference state energy, volume, and pressure appearing in Eq.(44) are fixed by the experimental setup and have to be estimated separately (see Sec. 4.4). Here we used conditions corresponding to a reference mass density $\rho_0 = M_D/v_0 = 0.167$ g/cm^3 (M_D being the deuterium atomic mass), i. e., the value used in Refs. [180, 181, 199, 200]. We decided to perform simulations in the temperature range between 4000 K and 35000 K, which includes the Hugoniot compressibility maximum and roughly corresponds to the experimental measures performed in the Z-facility [207] of Ref. [181]. Across these conditions, deuterium goes from a molecular state at low compression, to an atomic one at higher pressures. From a numerical perspective,

this range of temperatures justifies treating the nuclei classically (thus dropping the distinction between deuterium and hydrogen), as it has been shown [200] that for $T \geqslant 4000$ K the effect of NQEs starts to be negligible. Moreover, we verified that converged results can be obtained with relatively small systems, given that we are far above the critical temperature of both the proposed liquid-liquid and solid transitions.

4.2 FINITE TEMPERATURE ELECTRONIC EFFECTS

The finite temperature effects on the high-pressure part of the deuterium Hugoniot have been extensively studied at the DFT level and led to the development of very accurate T-dependent exchange-correlation functionals [78–80, 195, 208]. For the Hugoniot calculation at temperatures above 10 kK, we again took advantage of the Δ-learning technique to estimate the change in the PES due to thermalized electrons. In particular, for a given configuration, we can write the target QMC energy $E_{QMC}^{T=T_H}$ at temperature T_H as

$$E_{QMC}^{T=T_H} \approx E_{DFT}^{T=T_H} + \left(E_{QMC}^{T=0} - E_{DFT}^{T=0}\right), \tag{45}$$

with the assumption that the difference between the values at $T = T_H$ and $T = 0$ is the same for both methods. Notice that the term in parenthesis in Eq.(45) is exactly what has been learned by the model, which in this way does not need to explicitly depend on temperature. The resulting dynamics are thus performed by simply adding the correction to the finite-temperature DFT (FT-DFT) result at each step. Here we used the zero temperature PBE functional for the simulations; even if not rigorous, recent FT-DFT calculations of the Hugoniot have shown that for $T \lesssim 40$ kK this introduces a negligible error with respect to a an explicitly temperature-dependent GGA functional [195]. Therefore, in our application, temperature is taken into account through the thermal occupation of the one-body DFT levels. At $T = 10$ kK and $T = 15$ kK, we considered 120 bands in the calculation, while 150 bands were used for $T = 20$ kK and $T = 35$ kK.

We remark that this approach can also be applied when an MLP trained on DFT is used as the baseline, in place of an *ab initio* calculation. Indeed, in this case, finite temperature effects can be directly estimated from the DFT density of states [209].

4.3 RESULTS ANALYSIS

Fig. 8 shows the Hugoniot curves obtained with two MLPs, trained on the VMC-PBE and LRDMC-PBE difference, respectively, together with several experimental values for pressures below 150 GPa [168, 177, 178, 180, 181]. We also report the FT-DFT results we obtained with the PBE functional [74] and the DFT points reported in Ref.[181] using the

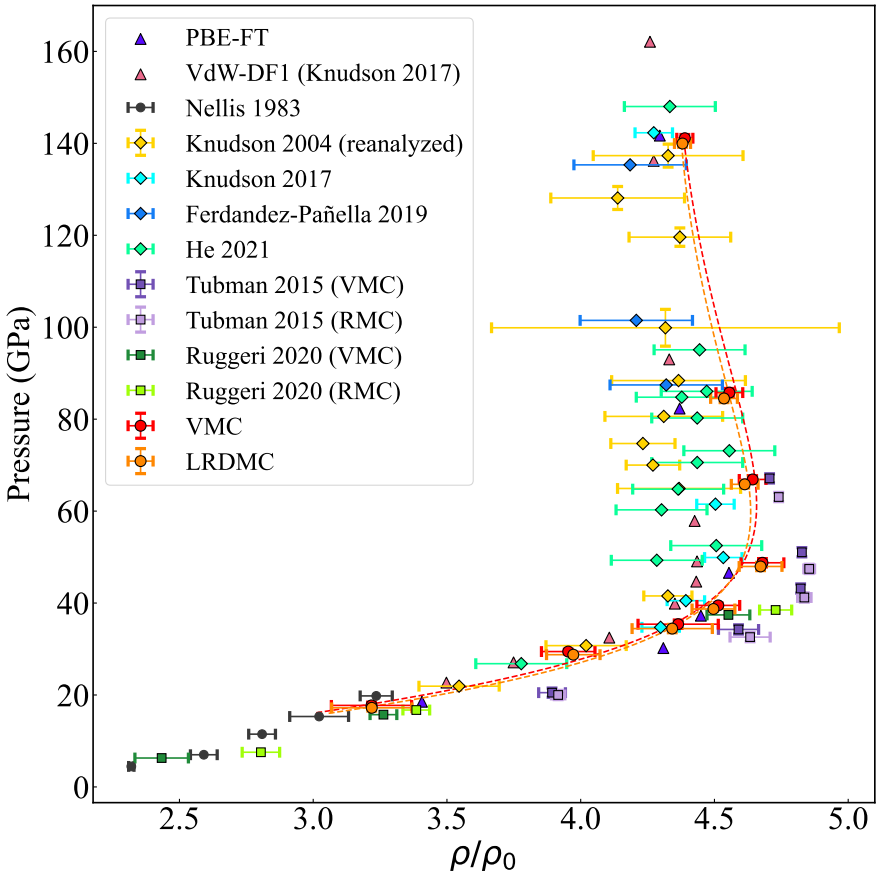

Figure 8: Principal Hugoniot in the density-pressure space. Blue and pink triangles are DFT results obtained with the PBE functional by us, and the VdW-DF1 functional computed in Ref. [181], respectively. Black dots are the gas gun results from Ref. [168]. The Hugoniot points measured in recent experiments are indicated with diamonds of different colors, in particular yellow [180], cyan [181], blue [177] ($\rho_0 = 0.173(1)$ g/cm^3), water green [178] ($\rho_0 = 0.170(1)$ g/cm^3). CEIMC results of Ref. [199] and [200] based on VMC and RMC are reported in purple and green squares. Red and orange circles are the results obtained with our MLPs trained on VMC and LRDMC datapoints, respectively, and a PBE baseline. The dashed lines are guides for the eye.

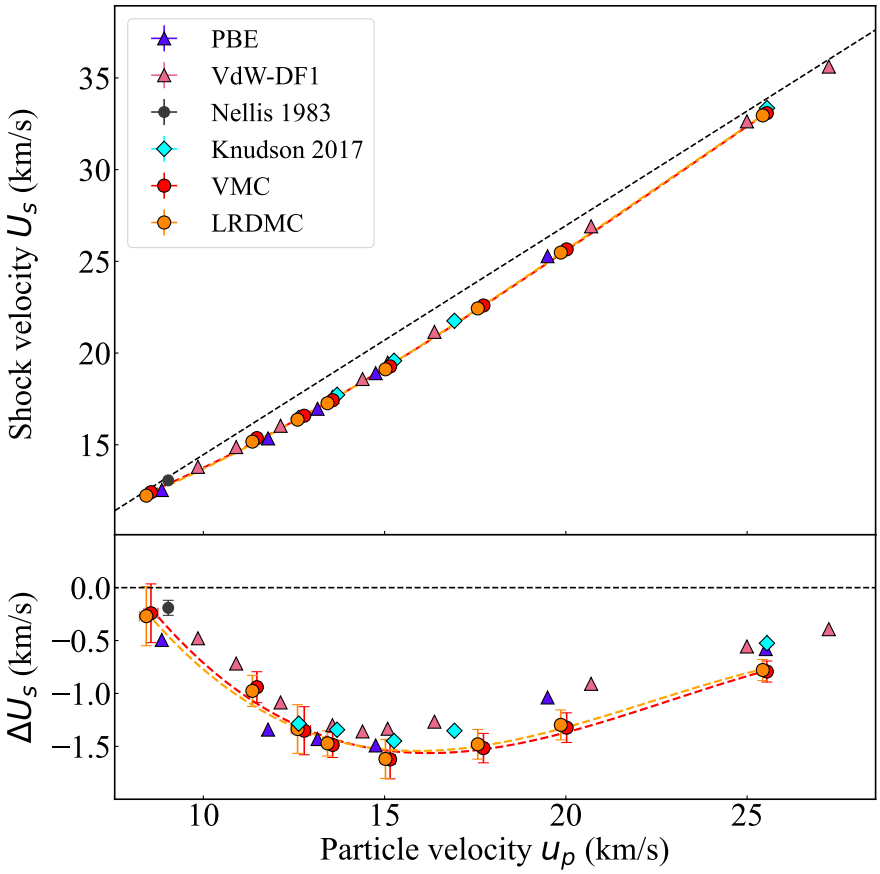

Figure 9: [top panel] Hugoniot in the u_p–U_s space. The black-dashed line is the re-analyzed gas-gun fit reported in Ref. [181], i. e., the shock velocity extrapolated from measures of molecular deuterium at lower pressures [168]. [bottom panel] Relative shock velocity with respect to the gas-gun fit. Only the experimental points of Refs. [181] and [168] are reported. Symbols are the same as in Fig.8.

VdW-DF1 [210, 211] functional. The CEIMC results of Ref. [199, 200], at both the VMC and reptation Monte Carlo (RMC) level of theory, are also shown for comparison. For T larger than 10 kK, our MD simulations employed FT-DFT as previously described. Both the VMC and LRDMC models give a very similar Hugoniot line, well reproducing the experiments in the low density-low pressure region. With respect to the most accurate experimental data of Ref. [181], our estimate of the relative density ρ/ρ_0 at the compressibility peak is $\sim 3-4\%$ larger, still within the error bars. For larger pressures, we predict a Hugoniot in agreement with the experiments but systematically more compressible. However, in this regime, the correspondingly larger uncertainties in the measures prevent a clear-cut assessment of our outcome. Our results are compatible with the low-temperature CEIMC ones reported in Ref. [200] within the statistical accuracy, while being significantly less compressible than older results [199]. As discussed later in Sec. 4.5, the biggest discrepancy between our points and Ref.[200] at $T = 8000$ K can be explained by the differences in the equation of state between the two methods.

Fig. 9 displays the same points in the $u_p - U_s$ space, where u_p is the particle velocity and U_s is the shock velocity, the two being calculated using the RH relations

$$u_p = \sqrt{(p + p_0)(\rho_0^{-1} - \rho^{-1})},$$
$$U_s = \rho_0^{-1}\sqrt{\frac{p + p_0}{\rho_0^{-1} - \rho^{-1}}},$$

where ρ and ρ_0 are the mass densities of the final and reference state. The difference ΔU_s between these points and the linear fit on the gas-gun data re-analyzed in Ref. [181] is also shown (bottom panel of Fig. 9). Notice that the drop in the slope of U_s relative to u_p coincides with the onset of the molecular-atomic transition, while the magnitude of the ΔU_s minimum relates to the position of the relative compression peak. In particular, the PBE Hugoniot curve manifests a premature start of the dissociation, while it correctly predicts the magnitude of the compressibility maximum. Our QMC results correctly predict the position of the peak and starting slope, while showing some discrepancies for $u_p \gtrsim 10$ km/s with respect to the data of Ref. [181]. In this regime, DFT, and in particular the result obtained with the VdW-DF1 functional, seems to be in better agreement with experiments, thanks to a favorable error cancellation in the Hugoniot [212].

Noticed again how, in our results, the discrepancy with the experiments is much milder than the value reported by previous QMC calculations at densities and pressures close to the compressibility peak [199] (see Fig. 8). This could be due to the direct optimization of the WF nodal surface used for the QMC calculations on the dataset

(discussed in Sec.4.7), which reduces the fixed node error mentioned in Ref. [212], the only approximation left in projective Monte Carlo methods, such as LRDMC and RMC. We later argue (see Sec. 4.5) that this discrepancy can be also due to a biased pressure estimate.

The molecular-atomic transition is further investigated in Fig. 10, where we report the radial distribution function, $g(r)$, calculated on trajectories obtained with the LRDMC model for several temperatures at densities close to the Hugoniot curve. The inset of Fig. 10 displays the value of the molecular fraction m, defined as the percentage of atoms that stay within a distance of 2 Bohr (roughly corresponding to the first $g(r)$ minimum after the molecular peak) from another particle for longer than a characteristic time, here set to 6 fs. The results indicate a distinct atomic character for $T \geqslant 10$ kK and a clear molecular peak at lower temperatures. The LRDMC model shows slower decay of the molecular fraction with temperature than the PBE and VdW-DF1 ones, being compatible with the latter for temperatures above 10 kK.

4.4 REFERENCE STATE CALCULATIONS

A crucial part of the numerical determination of the Hugoniot is to estimate the reference state energy per atom, e_0, and pressure, p_0. In particular, having a precise value of e_0 within the target method is important to take advantage of possible error cancellation effects and remove biases related to finite basis sets. To estimate the reference state energy and pressure, we followed a procedure similar to Ref. [200]. We performed a PIOUD simulation [150] (see Sec. 3.2) on a system of $N = 64$ deuterium atoms at $T_0 = 22$ K and $\rho_0 = 0.167$ g/cm^{-3}, in order to account for quantum effects, which are required because of the light deuterium mass and low temperature. Forces and energy were calculated with DFT through the Quantum Espresso package [202–204]. We checked the dependence of thermodynamic quantities on the number of replicas (or beads) M and on the choice of the DFT functional, by studying the value of the (average) quantum kinetic energy T for several values of M using the BLYP [213] and PBE functionals. In particular, we considered two estimators for the instantaneous value of T, namely the virial and primitive (or Barker) estimators, given respectively by

$$T_{M,\mathrm{vir}} = \frac{N}{2\beta} + \frac{1}{2M}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\mathbf{R}_i^{(j)} - \overline{\mathbf{R}}_i\right)\cdot\frac{\partial V}{\partial \mathbf{R}_i^{(j)}}, \tag{46}$$

$$T_{M,\mathrm{pri}} = \frac{3NM}{2\beta} - \frac{mM}{2\beta^2\hbar^2}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\mathbf{R}_i^{(j)} - \mathbf{R}_i^{(j-1)}\right)^2, \tag{47}$$

where M is the number of replicas used in the PIOUD simulation, $\mathbf{R}^{(j)} = \left(\mathbf{R}_1^{(j)}, \ldots, \mathbf{R}_N^{(j)}\right)$ are the coordinates of the system belonging

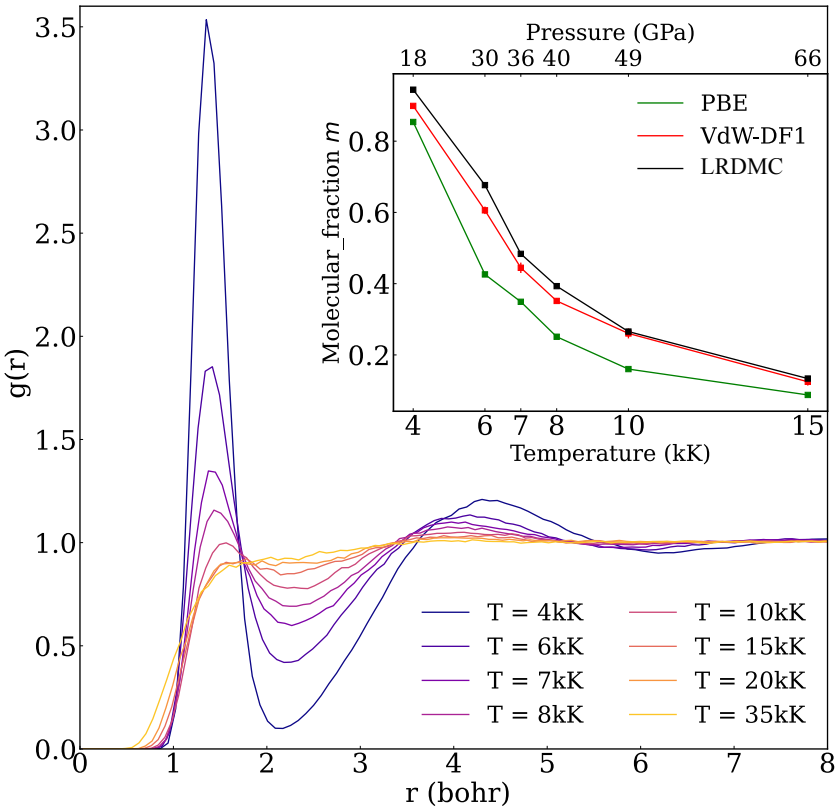

Figure 10: $g(r)$ for several temperatures and densities close to the principal Hugoniot, obtained using the LRDMC model. The molecular fraction value, m, is reported in the inset for each value of temperature up to 15kK. On the top axis, the corresponding pressure at the Hugoniot is also shown. The values obtained with AIMD using the PBE and VdW-DF1 functionals are reported for comparison at the same temperatures (notice that a pressure and density mismatch between methods can be present in this case due to different equations of state). Figure from [41].

to the j-th bead, $\overline{\mathbf{R}}_i = \frac{1}{M}\sum_{j=1}^{M}\mathbf{R}_i^{(j)}$ is the centroid position and $\beta = k_B T_0$. The results are shown in Fig. 11. We notice that a very large number of replicas is necessary for having a sufficiently converged result, while the value obtained with the PBE and BLYP functionals is extremely similar for all values of P. In the end, we chose to use the PBE functional and $P = 128$ replicas to have a reasonable trade-off between convergence and computational cost. For the DFT calculations, we used a 60 Ry plane waves cutoff and a $2 \times 2 \times 2$ Monkhorst-Pack k-point mesh. For the dynamics, we used a time step of 0.3 fs and let the system thermalize for 0.3 ps. We then extracted one configuration from a randomly chosen bead every 10 MD steps, for a total of 170 snapshots. Finally, the potential energy of these configurations was calculated using the appropriate method (PBE, VMC, or LRDMC). We then estimated e_0 for each method as

$$e_0 = \frac{1}{N}\left(\langle E_{pot}\rangle_{sample} + T^{PBE}_{256,pri}\right), \tag{48}$$

using the value of the primitive estimator at $M = 256$ beads as the best guess for the converged value of the kinetic energy. The approximation for the potential energy has been checked by running PBE *ab initio* simulations on this set and confirming that the "true" mean value (as calculated by averaging over the beads and the trajectory) is consistent with our model estimate obtained by averaging over the sample. Results for e_0 for the various methods are reported in Tab. 3. The reference state pressure p_0 is not reported, since it is two orders of magnitude smaller than the shocked pressure, and thus irrelevant for the Hugoniot determination.

Finally, we studied the effect of varying e_0 within its confidence interval on the Hugoniot density and pressure. In doing so, we also took into account the possible uncertainty on the energy difference $e(v, T) - e_0$ originating from the finite batch size we used for estimating energy gradients in the WF optimization. We estimated this uncertainty by running optimizations of increasing batch size on three different 128-atom configurations. The results indicate an error < 0.4 mHa/atom on $e(v, T) - e_0$. Taking everything into account, varying the energy within a standard deviation leads to shifts in the final principal Hugoniot which are smaller than the uncertainty produced by the MLPs, later estimated in Sec. 4.8.

4.5 COMPARISON WITH PREVIOUS QMC CALCULATIONS

The EOSs at $T = 8$ kK obtained with our MLPs trained on VMC and LRDMC data, together with the ones reported in Ref. [200] for both VMC and RMC, and those computed within DFT using the PBE and vdW-DF1 functionals are shown in Fig. 12. The differences between the methods are here more apparent than in the Hugoniot curve,

	e_{pot} (Ha/atom)	e_0 (Ha/atom)
PBE	−0.58217(2)	−0.58055(2)
VMC	−0.58622(2)	−0.58460(2)
LRDMC	−0.58660(2)	−0.58498(2)
VMC + FSC	−0.58503(2)	−0.58342(2)
LRDMC + FSC	−0.58542(2)	−0.58380(2)
RMC (Ref. [200])	−0.58570(6)	−0.58385(6)

Table 3: Estimated potential (e_{pot}) and total (e_0) energies per atom of the reference state at $\rho_0 = 0.167$ g/cm^3 and $T_0 = 22$ K for different methods, with and without finite-size corrections (FSC). RMC results (with FSC) from Ref. [200] are also reported.

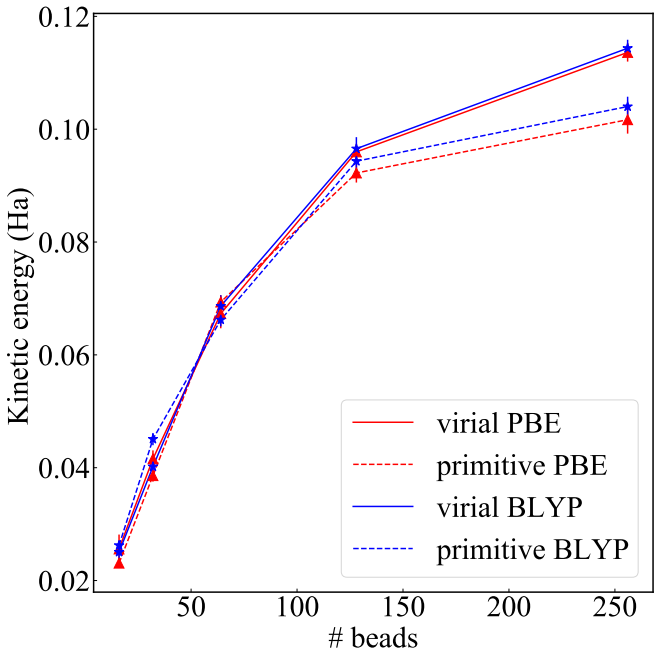

Figure 11: Convergence of virial and primitive estimators for the quantum kinetic energy, as computed with Eqs. (46) and (47), respectively, as a function of the number of replicas (beads) used in the PIOUD simulations.

where error cancellation occurs [212], and it is the known cause of the robustness of the DFT Hugoniot against the choice of the functional. Indeed, from Fig. 12 we can observe a $\sim 10\%$ difference in the pressure between the VdW-DF1 and PBE functional, even though the Hugoniot coordinates $(\rho^*/\rho_0, p^*)$ predicted by the two methods only differs by a few percent.

By comparing our results with the QMC data reported in Ref. [200] we found an excellent agreement in the value of the energy per atom. In particular, notice how our LRDMC reference energy e_0, including finite-size corrections (FSC), is perfectly compatible with their RMC value (also including FSC), as shown in Tab. 3. However, both the VMC and the LRDMC MLPs consistently predict a larger pressure than the one in Ref. [200] by ~ 4 GPa, which causes the slightly different compression found at $T = 8$ kK. By shifting the data reported in Ref. [200] by this amount, we can see an almost perfect match of the Hugoniot positions (Fig. 13). The origin of this residual discrepancy might be found in the different methods used to estimate pressure in the two QMC frameworks. In our case, the VMC and LRDMC pressures, which we used for training the two models, were calculated using the adjoint algorithmic differentiation to obtain directly the derivative of the total energy with respect to the cell parameters. In Ref. [200] a virial estimator was used, which can in principle produce discrepancies of the order of magnitude observed here, as shown in Ref. [212]. We also point out that both our models and the vdW-DF1 functional predict pressures slightly larger than those given by the PBE in the proximity of the Hugoniot position, in contrast with the EOS reported in Ref. [200], which predicts lower pressures. Although the QMC pressures are consistently lower than PBE ones on a given configuration, here we found that the MD simulations driven by our MLPs sometimes yield larger average pressures. This owes to the fact that different dynamics sample different distributions in the phase space. To demonstrate this, we extracted two samples of configurations from an *ab initio* PBE and a LRDMC-MLP trajectory, respectively, and then computed both PBE pressures and MLPs corrections (corresponding to the difference between LRDMC and PBE). The results shown in Tab. 4 demonstrate how the increased PBE pressures for the LRDMC-sampled trajectory can compensate, or overtake, the negative contribution given by the MLP correction.

4.6 MLPS DATASET CONSTRUCTION

The construction of the dataset used in Ref. [41] followed the active learning scheme described previously. We first performed a first set of AIMD simulations with the PBE functional on a system of $N = 128$ atoms for temperatures in the [4kK: 20kK] range and Wigner-Seitz radii r_s (with $\frac{4\pi}{3} r_s^3 a_0^3 = v$) in the $[1.80, 2.12]$ range, from which we

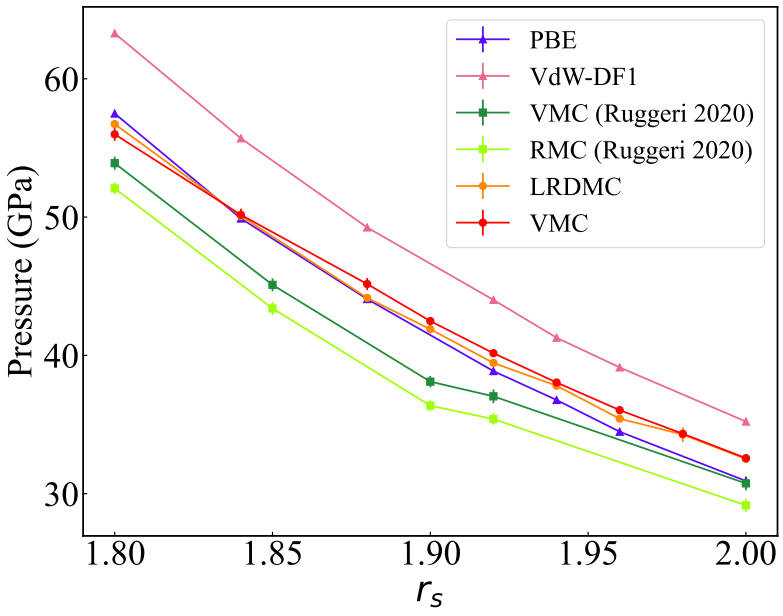

Figure 12: Equation of state close to the Hugoniot position at T = 8000 K, as obtained by different methods: DFT results obtained by our simulations with the PBE and VdW-DF1 functional (blue and pink lines); CEIMC equations of state based on VMC and RMC taken from Ref. [200] (green and light green lines); VMC and LRDMC results obtained with our MLPs (red and orange lines).

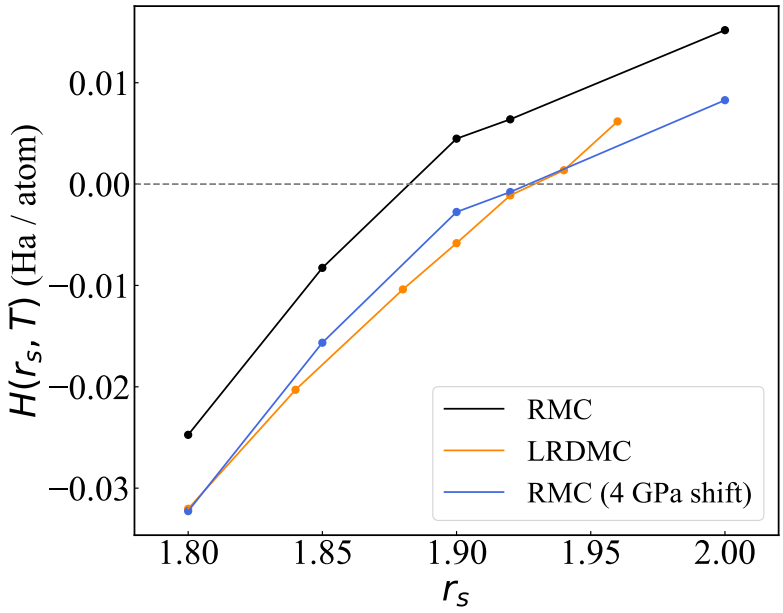

Figure 13: Hugoniot at T = 8000 K obtained with the LRDMC model (orange line), together with the one reported in Ref. [200] (black line) and the same curve shifted by 4 GPa, i.e., the pressure difference at the corresponding density (light blue line).

Virial pressure (GPa)	PBE sample	LRDMC sample
PBE	14.6(4)	17.7(4)
LRDMC	11.1(4)	14.5(5)
MLP Δ-correction	−3.57(4)	−3.17(4)

Table 4: Average value of the virial pressure calculated by DFT-PBE and by the LRDMC-MLP, on sampled trajectories obtained using the two methods, at thermodynamic conditions $r_s = 1.92$ and $T = 8000$ K. The Δ correction is also reported.

extracted a set of uncorrelated snapshots. New configurations were added to this initial dataset by monitoring, for each unseen configuration, the quantity

$$\chi = \frac{1}{N} \sum_{i=1}^{N} \max_{\mu \in \text{train set}} \mathcal{K}(\mathcal{R}_i, \mathcal{R}_\mu).$$

The number χ gives a quantitative measure of "how close" the unknown configuration is to what is already included in the training set. In particular, we stopped adding new data points when χ exceeded a fixed threshold of 0.80 during the dynamics. The final dataset, from which we extracted both the training and the test set, comprised 561 configurations of 128 atoms in total. The final range of temperatures and Wigner-Seitz radii r_s spanned by these configurations was [4 kK, 20 kK] and $[1.80, 2.12]$, respectively. Their distribution in the $r_s - T$ space is shown in Fig. 14.

4.7 QMC CALCULATIONS DETAILS

For the thermodynamic conditions relevant to the deuterium Hugoniot, the self-consistency error in the QMC forces calculation can have a huge impact (see Ch. 2 and Ch. 5) and thus has to be mitigated. In Ref. [41] we adopted an alternative strategy to the one described in Sec.2.3.1, based on an explicit optimization of the antisymmetric part of the wave function. In this way not only the bias of forces and pressure is reduced, but, in principle, also the overall accuracy of these quantities is improved. Here we will describe the details of the QMC calculations performed on both the dataset used for model training and validation, and the sample used to estimate the reference state energy e_0. For the simulations, the JAGP WF was used, with basis sets of [4s2p1d] and [2s2p1d] GTOs for the antisymmetric part and Jastrow part, respectively. The Gaussian orbital exponents were optimized beforehand for our particular system and thermodynamic conditions. The optimal exponents do not vary significantly in the density and temperature range explored and thus we fixed their

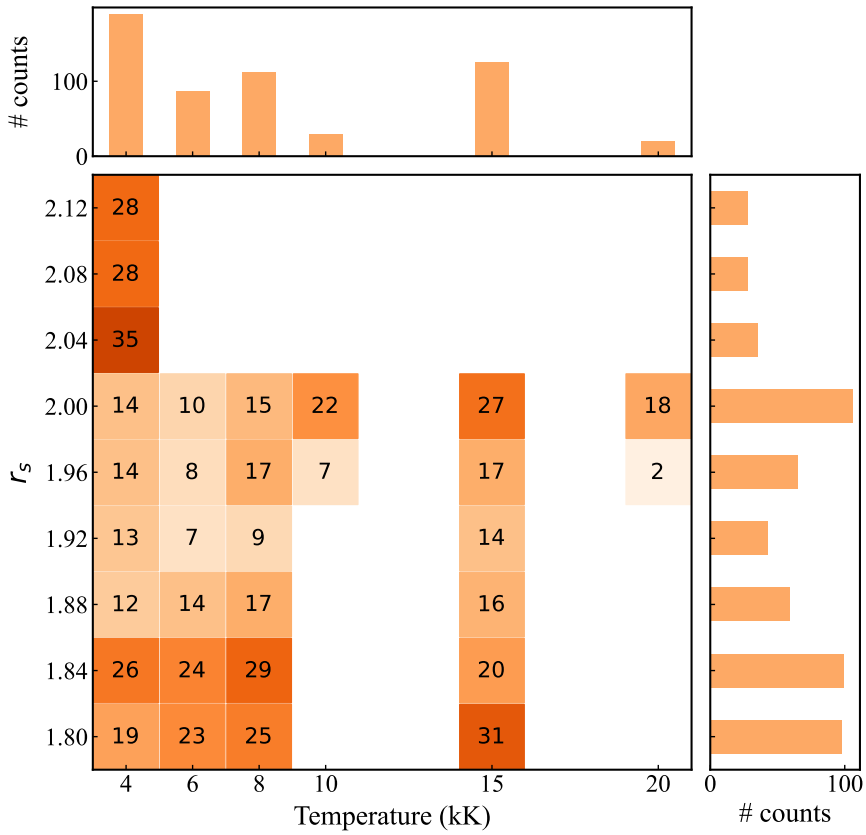

Figure 14: Distribution in the $T - r_s$ space for the dataset used for training and testing of the Hugoniot MLP. For each temperature and r_s the corresponding number of 128 H atoms configurations extracted from MD at those conditions is reported. The histograms at the top and right side represent the distribution in T and r_s, respectively.

value for all the configurations in the dataset. The final basis set employed is reported in Appendix A. The number of variational parameters in the antisymmetric part was further reduced thanks to geminal embedded orbitals (GEOs) [97]. In particular, the WF was first initialized with the built-in DFT code of the TurboRVB package at the Γ point and then projected into a basis of 6 GEOs $\widetilde{\phi}_{a,l}(x)$ for each atom a. Notice that using this basis is essential to perform a reliable WF optimization, which would be unfeasible otherwise for the exceedingly large number of parameters ($\sim 10^6$ without GEOs). The WF in the new basis was initialized by running a DFT calculation with the LDA functional and twisted boundary conditions at the Baldereschi point [214] $\mathbf{k} = (0.25, 0.25, 0.25)$ (in crystal units). Both the Jastrow factor and the antisymmetric part of the WF were optimized using the LM. Among the variational parameters $\lambda_{\{a,l\},\{b,m\}}$ in Eq.(15), we only optimized those corresponding to atom pairs closer than a cutoff $r_c = 4.0$ Bohr. For a stable optimization, we also projected the AGP into N/2 molecular orbitals at each step, such that the resulting optimized WF is an opt-JSD. The details of the constrained optimization technique are written in Ref. [36]. Finally, total energies and forces were calculated at the VMC and the LRDMC levels with the optimized wavefunctions. The LRDMC calculations were performed with a lattice size $a = 0.20$ Bohr, which is sufficient to obtain converged results. The improvement of this optimized WF with respect to a standard JSD (with uncontracted orbitals) can be appreciated by comparing the VMC energy per atom on the dataset configurations, which is shown in Fig.15. Notice how the energy of our opt-JSD WF is always lower or at most equal to the one obtained with the JSD, despite the use of the GEOs. For VMC forces and pressures a recently proposed regularization was used [215] while LRDMC quantities were estimated with the Reynolds [103] mixed average. Finite-size corrections (FSC) were estimated using the Kwee-Zhang-Krakauer (KZK) DFT functional [216]. We verified that these corrections have a very small effect on the Hugoniot position (see Sec. 4.8) and we thus applied them only to the LRDMC quantities. For a given ionic configuration, the final statistical noise on energies, forces and pressures was of the order of 7×10^{-6} Ha/atom, $1 \times 10^{-}3$ Ha/Bohr and 0.05 GPa, for VMC, and 2×10^{-5} Ha/atom, 2×10^{-3} Ha/Bohr and 0.1 GPa, for LRDMC.

4.8 MLP TRAINING AND VALIDATION

We trained four MLPs on the Hugoniot dataset, of which two were used to obtain the Hugoniot curve [41] and two were used for testing. The target method and DFT baseline used for each model, together with their RMSE on the test set are reported in Tab. 5. For these models, Eq.(36) was modified by including both a constant energy shift and

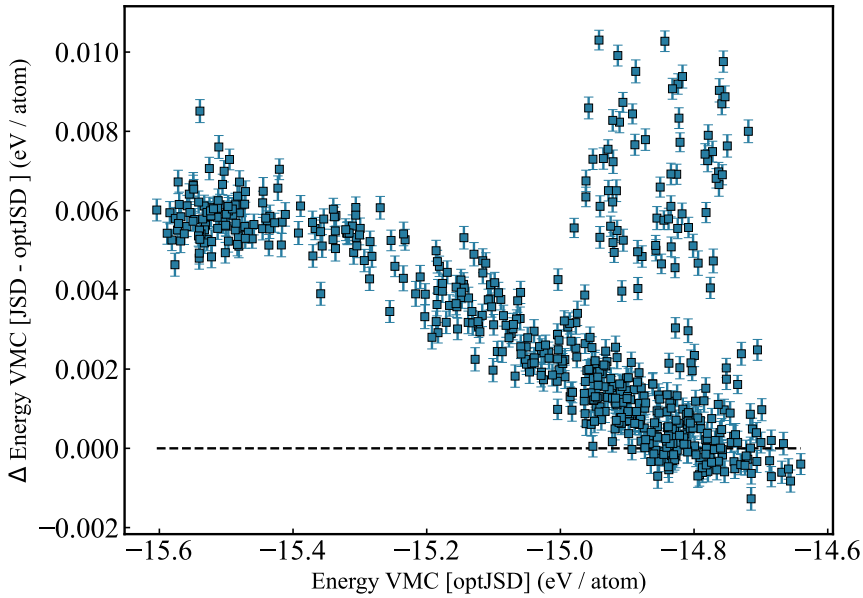

Figure 15: Difference between the VMC energies per atoms on the Hugoniot dataset as obtained using the JSD ansatz (with uncontracted orbitals) and our opt-JSD, respectively, plotted as a function of the opt-JSD energies. The optimization of the nodal surface results in an energy gain most of the time, being at least compatible within three sigmas with the JSD energy.

a pairwise potential inspired by Ref.[217], which is expressed as a linear combination of spline functions $B_\nu(r)$:

$$e\left(\mathcal{R}_i;\{\beta\},\{\alpha\},\delta\right) = \sum_{\mu=1}^{N_{env}} \beta_\mu \mathcal{K}\left(\mathcal{R}_i,\mathcal{R}_\mu\right) + \sum_{j\neq i} \sum_{\nu=1}^{N_{spline}} \alpha_\nu B_\nu\left(\left|\mathbf{R}_{ij}\right|\right) + \delta,$$

where the pairwise term is defined on a grid of points extending from r_{min} to r_{max}. Since the local atomic energy e is still linear in all the variational weights $\{\beta\}$, $\{\alpha\}$, and δ, the training procedure can be performed in the same way. In particular:

- the cutoff radius for the kernel function has been set to $r_c = 4.0$ Bohr, while in Eqs.(37)-(39)-(40) we used $2\sigma^2 = 1.5$ Bohr^2 and $n = \eta = 2$;
- for the pairwise potential, we used $N_{spline} = 11$ spline functions defined on a grid of 10 equally spaced points between $r_{min} = 0.3$ Bohr and $r_{max} = 5.5$ Bohr;
- the loss function weights w_E,w_F and w_P in Eq.(34) have been set to $20 \times 128^2, 10 \times 3 \times 128, 10^7$ and $20 \times 128^2, 10 \times 3 \times 128, 2 \times 10^7$ (in the appropriate inverse atomic units) for the VMC and LRDMC models respectively;
- a regularization of 10^{-5} was used to solve Eq.(41);

	RMSE		
	E (mHa/atom)	F (mHa/Bohr)	P (GPa)
VMC - PBE	0.26	4.9	0.76
LRDMC (+ FSC) - PBE	0.36	5.5	0.72
VMC (+ FSC) - PBE	0.29	5.1	1.1
VMC - LDA	0.35	5.5	0.80

Table 5: Value of the RMSE on different observables as calculated on the test set for the four MLPs trained on the Hugoniot dataset. Model names are in the "target-baseline" format. The first two MLPs were used to obtain the Hugoniot curves in Figs.8 and 9.

- the models were trained on a sparsified dataset containing 6000 local environments taken from ~ 500 configurations of 128 atoms.

To further validate the accuracy of our MLPs, we studied the variance of the results obtained with models trained on the same target, but using different DFT baselines. The Hugoniot function $H(r_s, T)$ and pressure at $T = 8$ kK, calculated using MLPs trained on the difference between VMC and two DFT functionals (i.e., PBE and LDA), are shown in Fig.16. The resulting curve shows a small dependence on the baseline used; in particular, the VMC-LDA model predicts a $H(r_s, T)$ shifted of $\sim 1 - 2$ mHa / atom with respect to the one obtained with the VMC-PBE potential. This seems to be due to a different value for the average energy per atom, as the pressure predicted by the two MLPs is more consistent. Considering both this source of uncertainty and the prediction error of the models (reported in Tab.5), we can assign an absolute error-bar of $\Delta\rho \sim 0.08$ and $\Delta P \sim 1$ GPa to the Hugoniot coordinates predicted by our models, which is reflected by the error-bars in Figs. 8 and 9. These values are obtained by varying the energy per atom and pressure inside their confidence interval and computing the corresponding Hugoniot variation for each temperature, as shown in Fig.17.

We also compared the results of MLPs trained on VMC with and without finite size corrections, with a PBE baseline. The Hugoniot functions at three different temperatures are shown in Fig.18. From this, we can observe that FSC have a small effect on the resulting Hugoniot, at least for the cell sizes considered here. Therefore, we decided to apply them only to the LRDMC dataset.

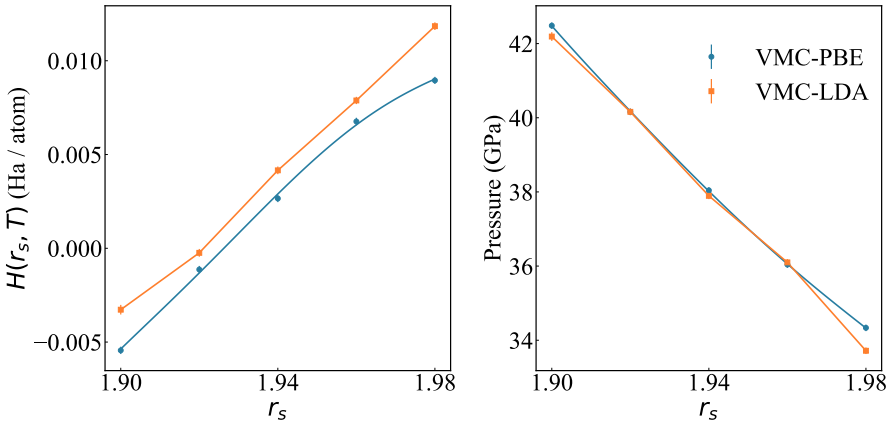

Figure 16: Effect of different baselines in the Δ-learning approach on the deuterium Hugoniot and pressure at T = 8000 K. In this case, we compared two models trained on the difference between VMC and two DFT baselines, using the PBE and LDA exchange-correlation functionals.

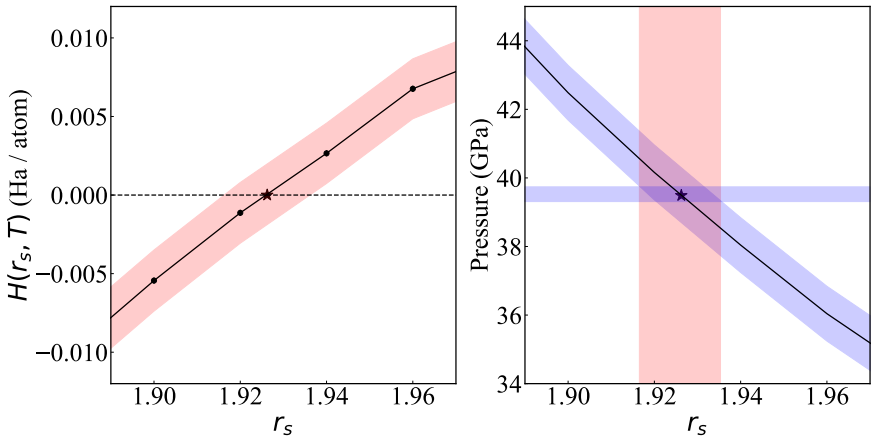

Figure 17: Variation of the Hugoniot function at $T = 8000$ K, as $e(r_s, T)$ and $P(r_s, T)$ are moved within their confidence interval (given by $\Delta_e = \pm 0.5$ mHa/atom and $\Delta_P = \pm 0.8$ GPa). The intersection of the red shaded area with the $y = 0$ axis (left panel) gives the uncertainty on the Hugoniot positions r_s^* (represented by the star symbol), which in turn is reflected on the Hugoniot pressure p^* (right panel). Notice that the effect on the Hugoniot pressure is reduced because of error cancellations in Eq.(44). The Hugoniot in the figure was obtained with the VMC-MLP and a DFT-PBE baseline.

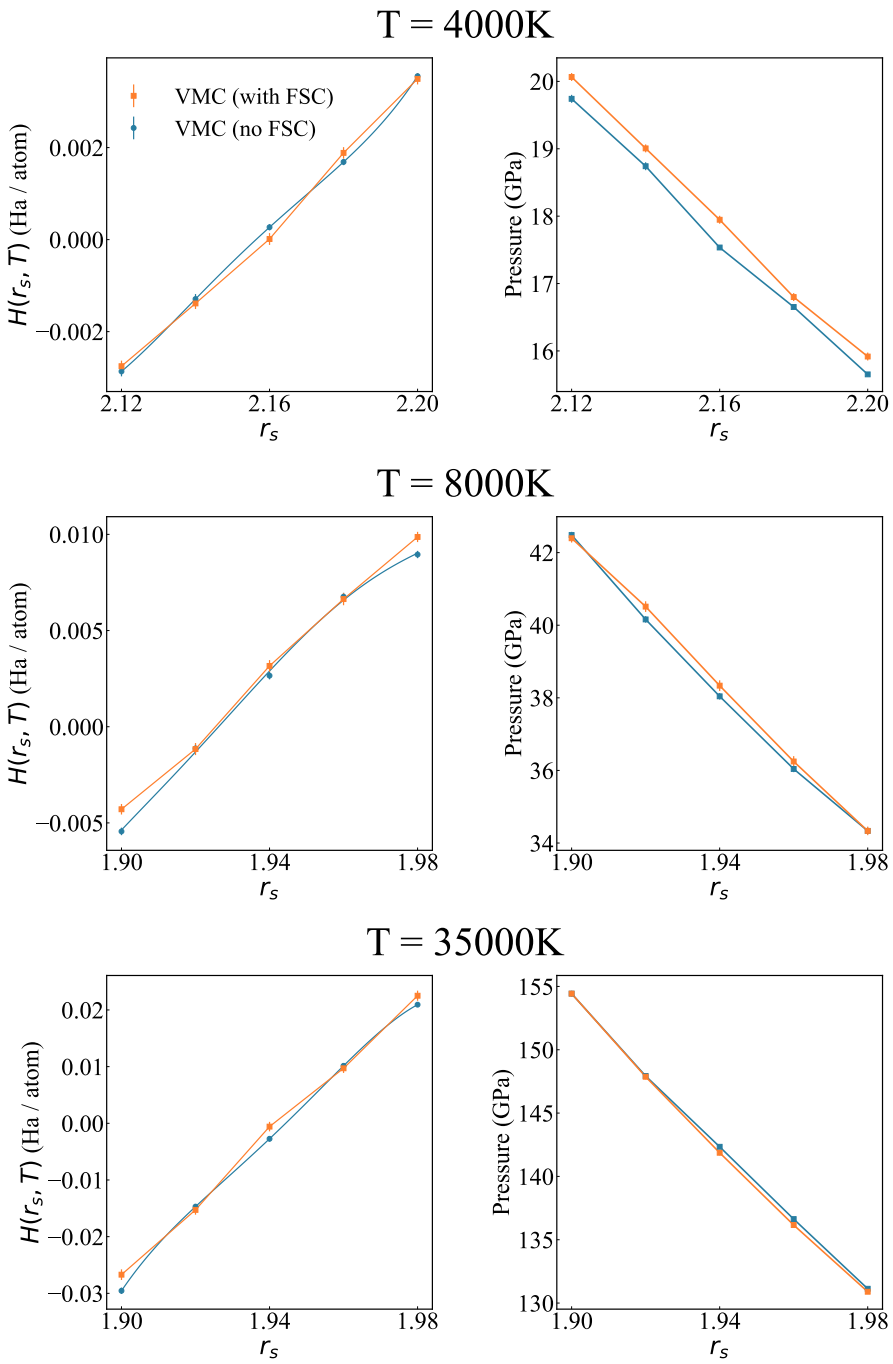

Figure 18: Comparison of the Hugoniot curve obtained with two models trained on VMC reference data points, with and without finite size corrections for T = 4000 K, 8000 K, and 35000 K. The FSC were estimated by adding to energies, forces, and pressures the difference between the values obtained with the LDA functional and those calculated with the KZK functional [216] at the corresponding size (N = 128 in our case).

4.9 WRAP-UP

In this Chapter, we studied the deuterium Hugoniot curve, using MLPs trained on QMC data points, constructed using a combination of the Δ-learning approach and KRR. We employed an opt-JSD WF, for which we optimized both the Jastrow factor and the antisymmetric part. This not only increases the accuracy of the results, but also cures the self-consistency error discussed in Ch. 2. In our Hugoniot determination, we again exploited the Δ-learning approach and utilized a FT-DFT baseline to approximate the effect of finite temperature electrons, which are relevant for temperatures above $T = 10000$ K. Both our models, trained on VMC and LRDMC reference calculations, respectively, give results that are in good agreement with recent experiments [181] and previous QMC calculations [200]. The slightly higher compressibility of our Hugoniot curve with respect to the experimental data at higher temperatures can stimulate further investigations on this part of the phase diagram.

5 BIAS CORRECTION IN THE HUGONIOT DATASET

The possibility to construct accurate machine learning potentials crucially depends on the availability of extended datasets. Techniques such as Δ-learning can be used to reduce the number of necessary configurations and obtain reliable models even with relatively small training sets. However, the generation of thousands of configurations within expensive electronic structure methods, such as VMC, is still a formidable task, usually requiring large computational resources. The efficiency of VMC is largely due to the complexity of the WF employed in the calculations. One of the simplest WFs is the JSD, where only the Jastrow factor is optimized and the SD part is taken directly from DFT calculations (see Sec. 2.2.3). Ideally, the JSD is an appealing candidate for VMC-based machine learning application, providing an excellent tradeoff between accuracy and computational cost. Indeed, the Jastrow optimization is much less demanding than the "full" optimization involving also the determinant. However, as discussed in Sec. 2.3, forces and pressure computed with the JSD WF can have a bias due to the self-consistency error. Machine learning potentials are by definition consistent, given that energy derivatives are obtained using Eq. (35), but a biased training set can nevertheless spoil the final accuracy of the model. A solution would be to use only the energies for the training; this approach is often impractical because it usually increases by orders of magnitude the number of configurations necessary to reach a given accuracy. As shown in Sec. 2.3, the self-consistency error can be removed by applying a suitable correction, adding a relatively cheap computational overhead. The objective of this Chapter is to demonstrate how this technique can be employed to generate unbiased datasets with a JSD WF and also to compare their quality with more refined ones obtained from a fully optimized ansatz (opt-JSD). In particular, we will apply this approach to a subset of the configurations we used in Ch. 4 to train the MLPs for studying the Hugoniot curve. The calculation of the bias correction on the entire dataset is currently ongoing and will be the topic of future work.

The self-consistency error for the Hugoniot configurations will be analyzed in Sec. 5.1. A comparison between MLPs trained on datasets with and without bias corrections and on datasets employing an opt-JSD WF, taken as reference, is reported in Sec. 5.2. Finally, we summarize the contents of this Chapter in Sec. 5.3.

5.1 SELF-CONSISTENCY ERROR FOR THE HUGONIOT CONFIGURATIONS

We will start our discussion by assessing the entity of the self-consistency error in the Hugoniot dataset. To do this, we used a JSD WF with a basis set of [4s2p1d] and [2s2p1d] GTOs for the antisymmetric part and Jastrow, respectively, the same employed in Ch. 4. We initialized the Slater determinant by running a DFT calculation with the LDA functional at the Baldereschi point. Contrary to what we did previously (see Sec. 4.7) here we did not use GEOs and utilized the full primitive basis. Finally, the Jastrow part was optimized using the SR method. We dub the dataset obtained in this way as the "unoptimized" dataset, to distinguish it from the "optimized" one employed in Sec. 4.1 and discussed in Sec. 4.7. The latter employs the opt-JSD, where both the nodal surface and the Jastrow factor are optimized.

We assessed the consistency of the JSD forces and pressures by calculating $\mathbf{F}_i^{\mathrm{VMC}}$ and P^{VMC} via the combination of two approaches: (i) using Eq.(25) (along with the analogous expression for pressure) and (ii) applying the finite-differences method (FDM) by fitting the PES. The results are shown in Fig. 19 for one of the force components and the virial pressure.

We can observe a significant bias for both quantities, with a discrepancy $\sim 10\%$ of their value. This clearly highlights the importance of curing the self-consistency error for the unoptimized dataset. For this reason, we applied the correction of Eq. (26) to both forces and pressures. For forces, we computed the numerical derivatives of the parameters $\frac{d\lambda^{\mathbf{R}}_{\mu,\nu}}{d\mathbf{R}}$ with the FDM (i. e., using Eq. (27)), with a displacement $\Delta R_i^u = \pm 0.003$ Å. The derivatives with respect to the volume, entering in the pressure correction, were obtained with the FDM using relative volume variations of $\pm 0.1\%$. In Fig. 19 we can observe that the corrected forces and pressure are perfectly compatible with their values computed with the FDM. Fig. 20 shows the values of both the biased and corrected forces components (pressures) versus the forces (pressures) estimated from the PES, for several configurations. Notice how the pressure correction acts almost like a rigid shift of ~ 1 GPa. The root mean squared error (RMSE) between biased and corrected quantities is also shown, further demonstrating the effectiveness of the correction in removing the self-consistency error.

5.2 MLPS WITH CORRECTED TRAINING SETS

In this Section, we will present some preliminary results illustrating how the correction can be applied to generate an unbiased JSD dataset for constructing MLPs. We selected 245 among the 561 configurations used for training the model in Ch. 4 and applied the correction as

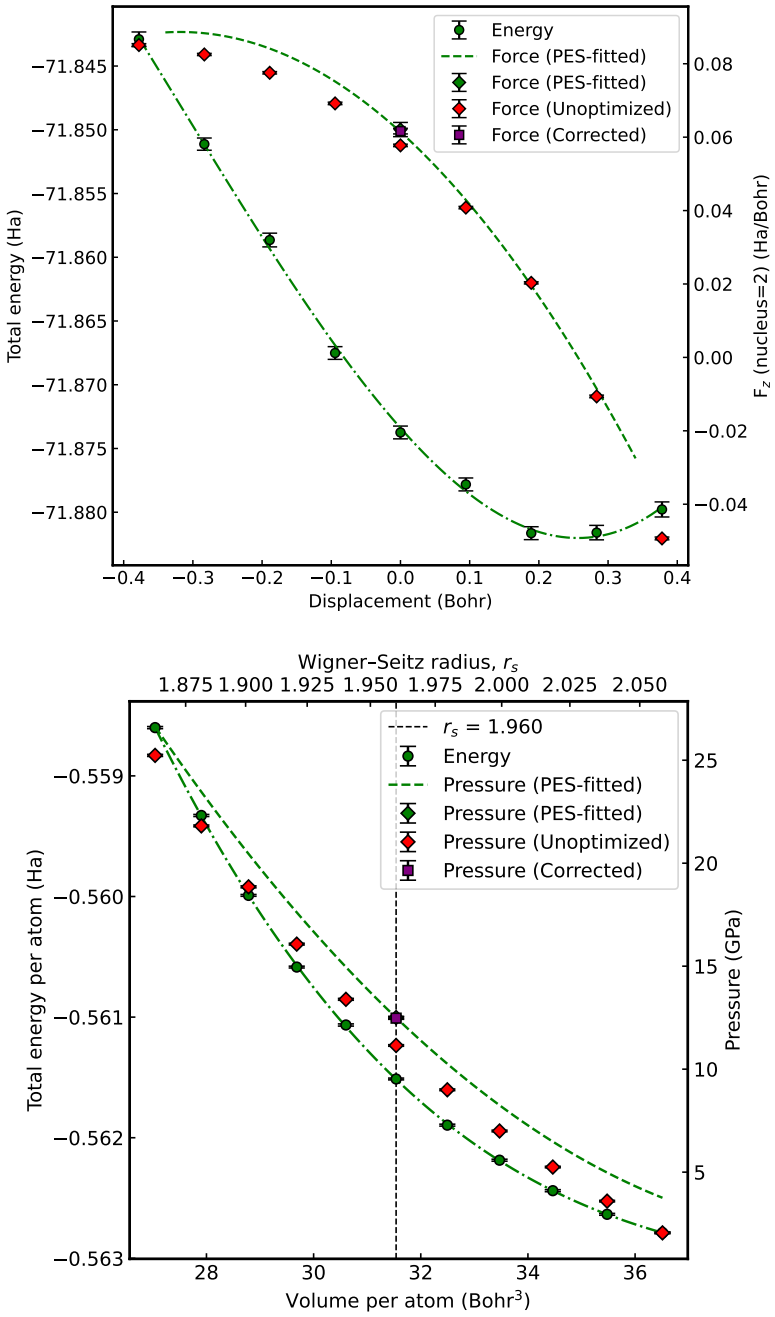

Figure 19: (Top panel) Comparison between the biased force evaluated with Eq. (25) (red diamonds), the numerical force calculated by fitting the PES (dashed green line and green diamond), and the corrected force obtained by applying Eq. (26) (violet square). The PES of the system along the atomic displacement is also shown (green dots and dash-dotted line). (Bottom panel) Comparison between the biased pressure (red diamonds), the numerical pressure obtained by fitting the PES (dashed green line and green diamond), and the corrected pressure (violet square). The PES of the system is also shown (green dots and dash-dotted line). The calculations have been performed on a system of 128 atoms.

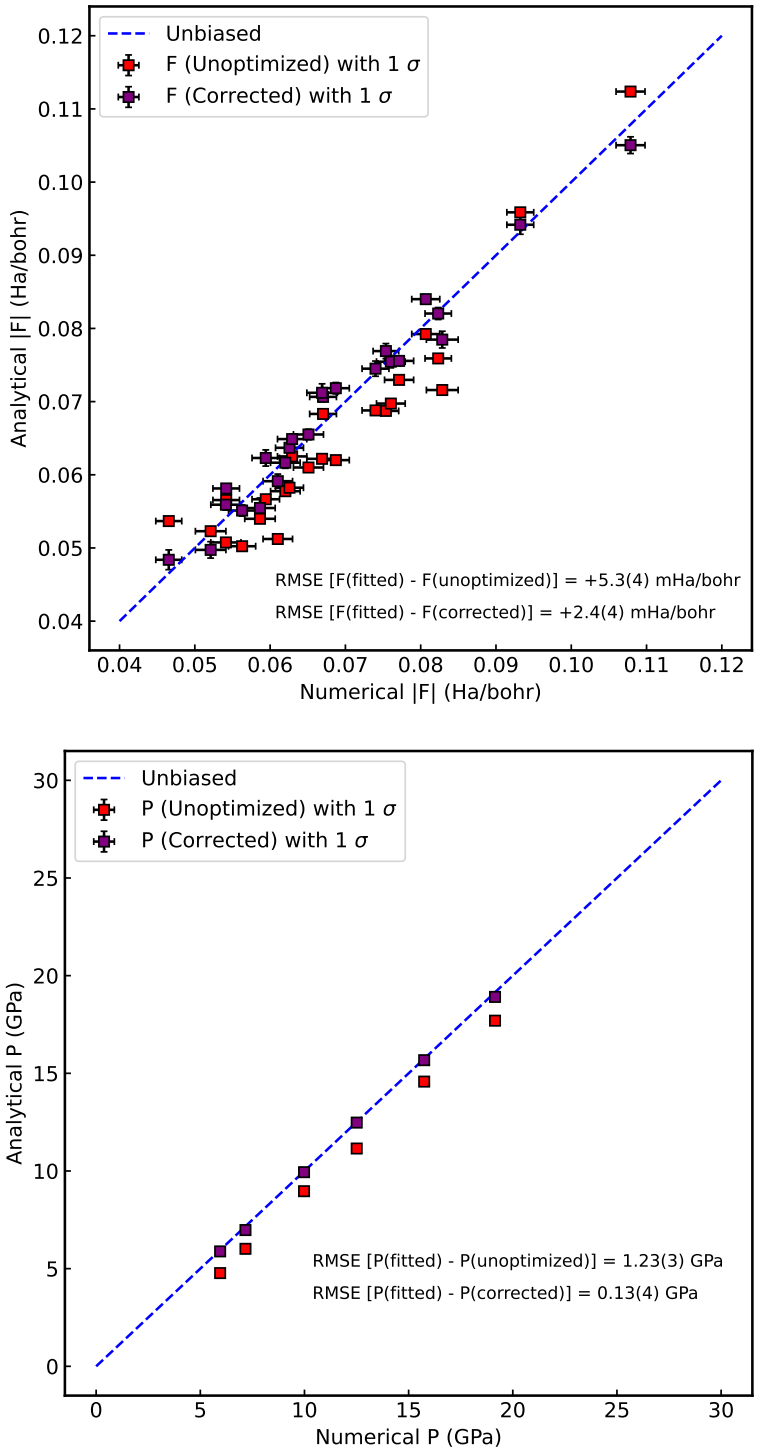

Figure 20: (Top panel) Values of the biased (red markers) and corrected (violet markers) force components as a function of the numerical force estimated by the PES. The dashed line indicates perfect consistency. (Bottom panel) Value of the biased (red markers) and corrected (violet markers) pressure for 6 different 128-atom configurations as a function of the numerical pressure estimated from the derivative of the PES with respect to the volume.

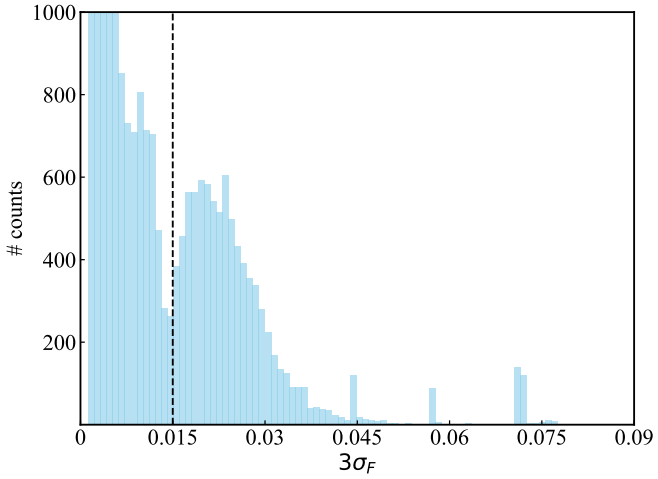

Figure 21: Distribution of the standard deviations for the corrected forces. The vertical line indicates the 3σ threshold chosen in our application.

previously described. We noticed that for some force components and pressures, the correction is sometimes accompanied by a large error bar. We identified the cause of this phenomenon in our current implementation, which estimates the parameters derivatives with the FDM by running several DFT calculations (see Eq. (27)) for the displaced coordinates. When using a localized basis in DFT, a cutoff is normally employed to reduce the redundancy of the basis set, by cutting elements corresponding to eigenvalues of the overlap matrix [129] smaller than the cutoff. When the ions are displaced, the removed elements change, thus introducing a dependence of the basis on the atomic displacement within the FDM. In the future, we plan to replace this part with more reliable and faster approaches (e. g., using linear response theory [118]). To deal with these erratic components, for the forces we kept the biased value whenever the standard deviation was above a given threshold, here set to 0.015 Ha/Bohr. As shown in Fig. 21, this corresponds to a "minimum" in the standard deviation distribution for the corrected forces, which we can interpret as the onset of the deviation from normality. For pressure, a similar strategy was applied. In this case, since the correction is nearly constant for all the configurations, whenever the noise was larger than 3×10^{-5} a.u., we shifted the biased value by the average correction $\Delta p \approx 1$ GPa.

The "corrected" dataset was used to train MLPs using KRR and the Δ-learning approach with a DFT PBE baseline potential. The hyperparameters of these models are equal to the ones described in Sec. 4.8. Among the 245 configurations of the dataset, we selected 45 configurations for testing. We investigated the effect of varying the training weights in the loss function (Eq. (34)) on the RMSE computed on the test set for energy, forces, and pressure. To assess the effect of the cor-

rection on the models' performance, we also trained MLPs using the biased "unoptimized" dataset and the "optimized" one introduced in Ch. 4, based on the JSD and opt-JSD WFs, respectively. Also in this case, only 245 configurations (split into training and test set) were considered for consistency.

When training with different target methods, the value of the RMSE may not be the best metric to assess the relative accuracy of a model with respect to the others. To do this we defined the relative improvement of the model for each system property ($X = E, \mathbf{F}, P$) as:

$$\Delta_X = \frac{\mathrm{RMSE}(X_{\mathrm{test}} - X_{\mathrm{pred}}) - \mathrm{RMSE}(X_{\mathrm{test}} - X_{\mathrm{dummy}})}{\mathrm{RMSE}(X_{\mathrm{test}} - X_{\mathrm{dummy}})}, \tag{49}$$

where $E_{\mathrm{dummy}} = \frac{1}{N_{\mathrm{test}}} \sum_{i \in \mathrm{test}} E_i$, $\mathbf{F}_{\mathrm{dummy}} = 0$, and $P_{\mathrm{dummy}} = 0$. In other words, Δ_X measures the relative improvement of the model prediction on the quantity X, with respect to a "dummy" model, corresponding to a perfectly flat PES equal to the average energy on the test set for all configurations. Notice how $\Delta_X = -1$ indicates that the model has learned exactly the quantity X, while a value $\Delta_X \geqslant 0$ means that the model has an error equal to, or larger than, a flat model.

Figs. 22 and 23 show the results obtained for Δ_E, Δ_F and Δ_P using MLPs trained with different values of the training weights w_E, w_F, and w_P. We can first analyze what happens by progressively increasing the weight on the forces for the different models without explicitly training on pressures (i. e., $w_P = 0$). The force bias in the unoptimized dataset is evident from the larger Δ_F of the corresponding model compared to that of both the optimized and corrected datasets (see Fig. 22b). Notice how, for small values of w_F/w_E, Δ_F is positive in the unoptimized case, which in the Δ-lerning framework means that the correction on forces is detrimental. As the weight on forces is increased, all models reach a plateau in Δ_F, while simultaneously losing some accuracy on the other quantities. For the energy, this loss of accuracy is significantly larger in the models trained on the unoptimized dataset, further demonstrating the effect of the bias. We observed very similar performances between the optimized and corrected models. For pressure, the models including bias corrections show the best performances, presenting a minimum at $w_F/w_E \sim 10^{-2}$. Since the loss does not include the pressure in this case, this is another indication of consistency. Fig. 23 shows the results obtained by varying the weight on pressure, keeping $w_E = 1$ and $w_F = 3/128$ fixed. Remarkably, for the corrected model, including the pressure in the training step does not significantly improve Δ_p. This suggests that the information from energy and consistent forces alone is sufficient to "reconstruct" the relative virial pressure.

Finally, we ran MD simulation at densities and temperatures corresponding to the Hugoniot curve position using a MLP trained on the

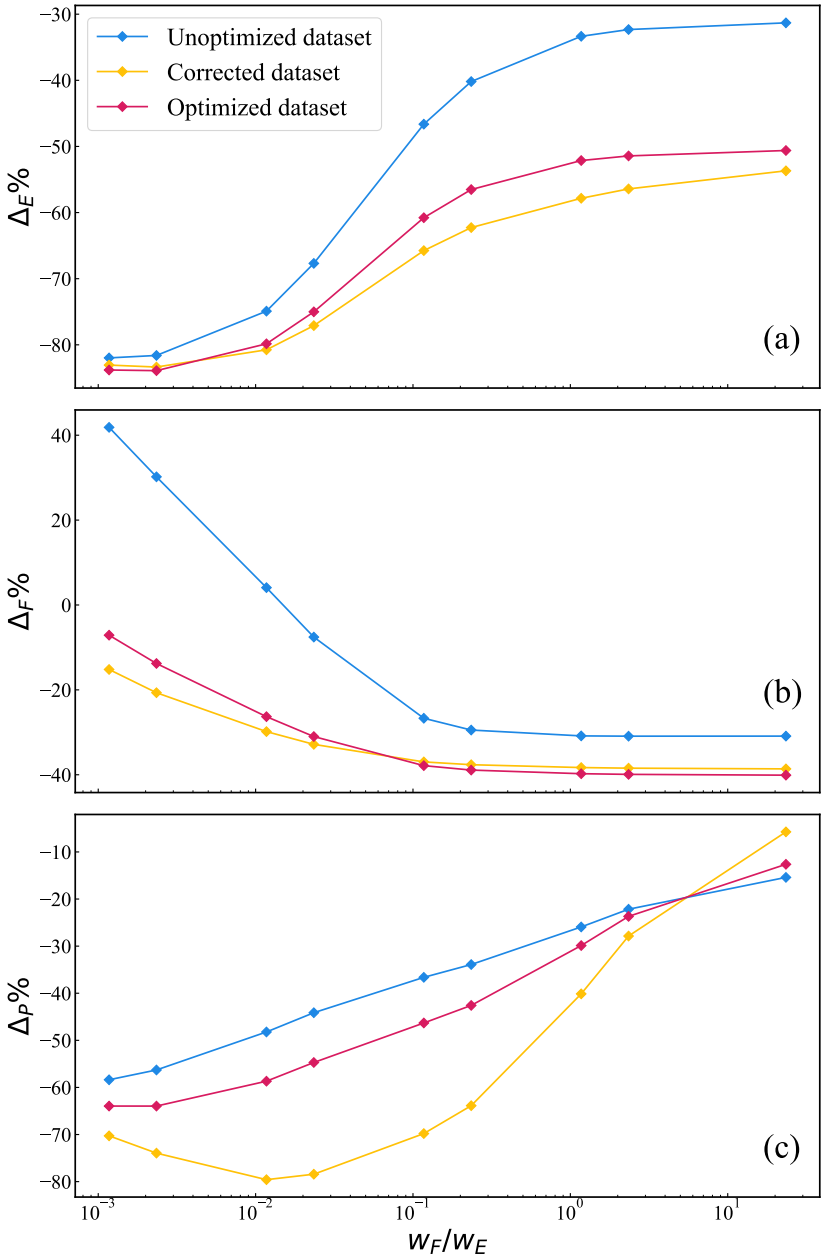

Figure 22: Relative RMSE variation Δ_X (defined in the text) for (a) energy, (b) forces, and (c) pressure, as a function of the force/energy weight ratio in the loss function used for training (expressed in the appropriate inverse atomic units). All datasets (optimized, unoptimized, and corrected) comprise 245 configurations with 128 atoms each, 200 of which are used for training. The weight on the pressure was set to zero.

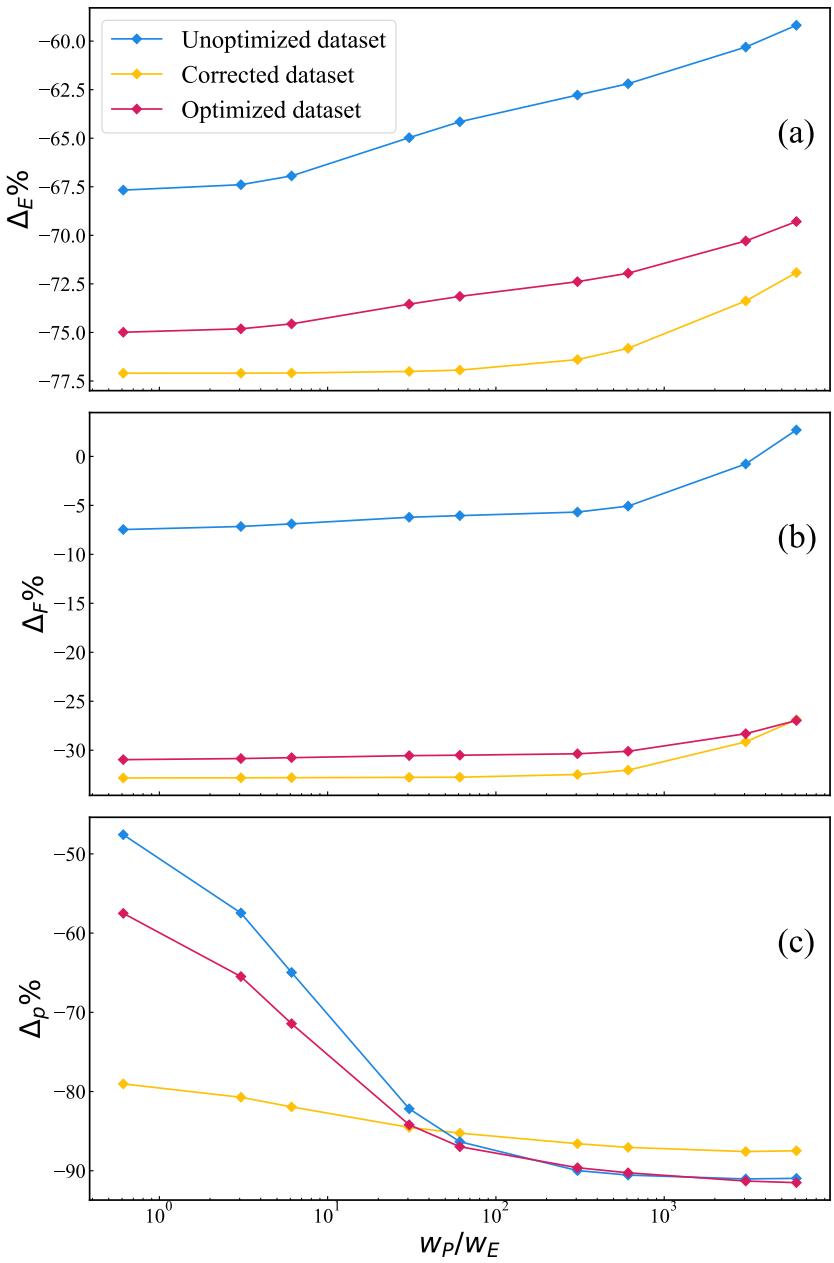

Figure 23: Δ_X (defined in the text) for (a) energy, (b) forces, and (c) pressure, as a function of the pressure/energy weight ratio in the loss function used for training (expressed in the appropriate inverse atomic units), for the different models. All datasets (optimized, unoptimized, and corrected) comprise 245 configurations with 128 atoms each, 200 of which are used for training. The weight on energy and forces was set to 1 and 3/128, respectively.

	P_{opt} (GPa)	P_{unopt} (GPa)	P_{corr} (GPa)
$r_s = 2.16$, $T = 4000$ K	17.5(1)	16.2(1)	16.7(1)
$r_s = 2.02$, $T = 6000$ K	28.9(2)	25.6(2)	27.5(2)
$r_s = 1.92$, $T = 8000$ K	40.2(2)	37.4(2)	39.1(3)

Table 6: Average pressure from MD simulations obtained with different MLPs trained on the (full) optimized dataset, the (full) unoptimized dataset, and a corrected dataset containing 200 configurations. All models employ the Δ-learning technique with a DFT-PBE baseline.

corrected dataset. For this model, the loss weights were set to $w_E = 1$, $w_F = 3/128$ and $w_P = 10^6/128^2$, based on the previous analysis. We compared the results with those of Ch. 4 and those obtained using the biased JSD dataset. For these simulations, both the optimized and unoptimized models were trained on the full dataset described in Sec. 4.6. Moreover, for each model, we selected the training weights to have a good tradeoff between the RMSE values on energy, forces, and pressure. A comparison of the equilibrium pressure during the dynamics for three different temperatures is reported in Tab. 6. Notice that, in principle, the optimized dataset and the corrected one will give different results. Indeed, the optimization of the antisymmetric part of the WF not only improves the consistency of forces and pressure, but also modifies the PES of the system. In fact, as already shown in Fig. 15 the energy significantly changes with the optimization. Nevertheless, if compared with the unoptimized model, the MLP trained on the corrected data gives results that are consistently closer to the optimized ones. This suggests that, with the bias correction, the physical description of the system improves considerably. Similar conclusions can be reached by looking at the radial distribution function $g(r)$ (Fig. 24). In particular, the $g(r)$ of the corrected model displays a better agreement with the one obtained in Ref. [41] using the optimized model.

5.3 WRAP-UP

In this Chapter, we analyzed the impact of the self-consistency error (described in Sec. 2.3) on the construction of QMC-based MLPs. We compared the relative accuracy and physical predictions of several models, all trained on configurations extracted from the Hugoniot dataset (introduced in Ch. 4) but with energy, forces, and pressures computed with three different approaches. In particular, we considered:

1. an "unoptimized" dataset, using a JSD WF with forces and pressures evaluated with the standard VMC formula (e. g., Eq. (25));

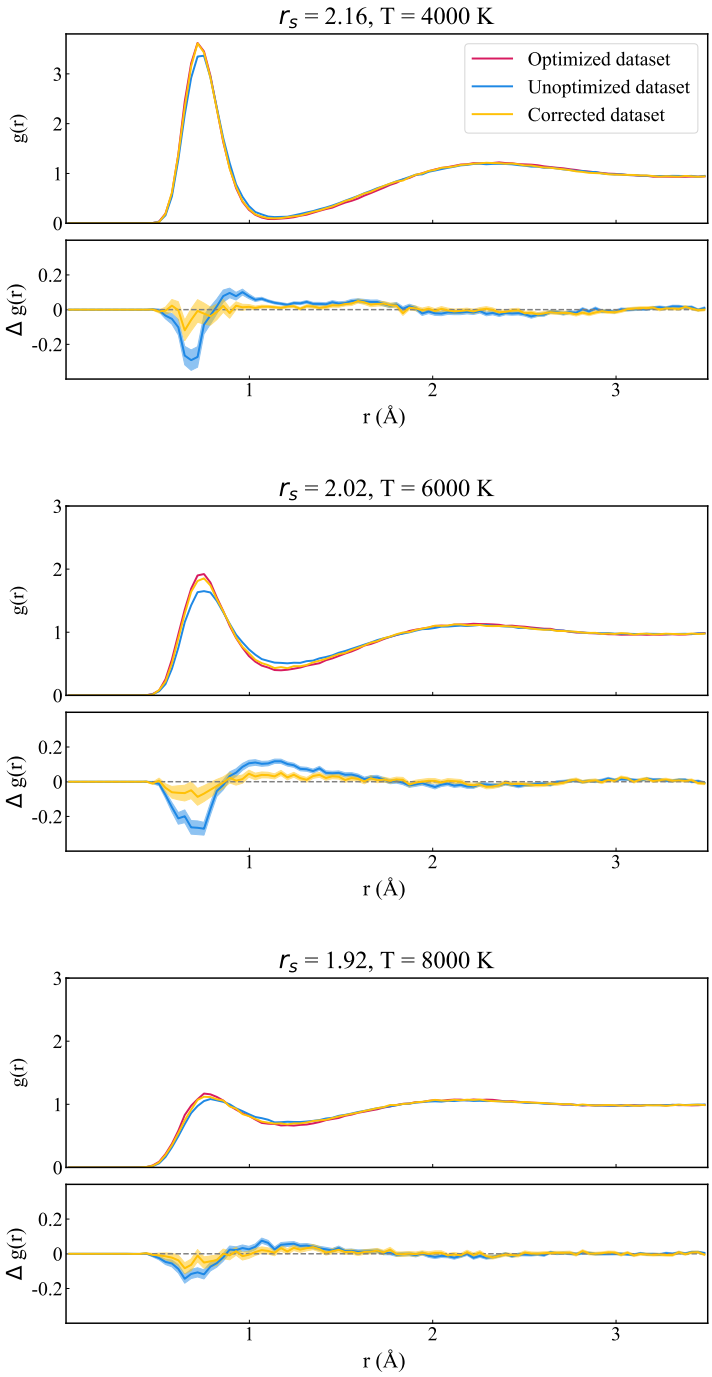

Figure 24: Radial distribution function $g(r)$ for densities and temperatures close to the Hugoniot position, obtained with MLPs trained on different datasets. The difference $\Delta g(r)$ with respect to the model trained on the optimized dataset is also shown for both the "unoptimized" and "corrected" models. The shaded area indicates the uncertainty in $\Delta g(r)$.

2. an "optimized" dataset, employing the opt-JSD WF described in Sec. 4.7, where the WF nodal surface is optimized to reduce the self-consistency error;

3. a "corrected" dataset, utilizing a JSD WF where both forces and pressure are corrected using the additional term introduced in Sec. 2.3.1 (see also Eq. (26)).

The comparison highlights the detrimental effect of the self-consistency error, which is manifested by a large dependence on the training weights in the loss function for the MLPs using the unoptimized dataset. On the contrary, the models trained on the corrected dataset show performances comparable to, if not better than, the "optimized" ones, as far as the consistency between energy and forces/pressure components is concerned. The stark improvement brought by the correction is also apparent when we compare results for average thermodynamic quantities obtained from MD, such as the pressure and radial distribution function. Taking the results yielded by the opt-JSD datasets as a reference demonstrates that the correction also improves the physical description of the system. This opens the possibility of using JSD datasets for ML applications, which are much cheaper to generate compared to ones employing fully optimized WFs.

6

LIQUID LIQUID PHASE TRANSITION IN HIGH PRESSURE HYDROGEN

While the realization of metallic solid hydrogen is still highly debated [27], a broader consensus exists regarding the transition from a molecular insulating state to a metallic atomic one in the liquid part of the phase diagram. This liquid-liquid phase transition (LLPT) has been extensively studied experimentally [218–228] and by means of numerical simulations [64, 65, 68, 69, 113, 145, 191, 192, 229–242] given its relevance for planetary science, in particular for understanding the interior of giant gas planets [243] and their magnetic fields [244].

As is often the case for hydrogen, experiments based on static and dynamic compression give contrasting results, with the latter predicting a larger transition pressure. Moreover, a definitive experimental answer on the order of the LLPT is currently missing. A similar uncertainty is also present in the numerical simulations. Results obtained with AIMD using DFT show a large variability with respect to the choice of the exchange-correlation functional. For example, the transition pressure can vary by 200 GPa when including long-range Van der Walls corrections [245]. The characterization of the LLPT is also a very debated topic, with many first-principles simulations results suggesting a "weak" first-order character [65, 68, 69, 229, 232, 238] based on the observation of "kinks" in the EOS below a critical temperature T_c in the 1500-2000 K range. Given the large autocorrelation time expected near the transition, the results obtained with DFT or QMC have been questioned because of their short time-scale covered and/or the small size of the systems considered. In particular, recent large-scale simulations with a DFT MLP [64] have suggested that the LLPT is a smooth crossover, even though the accuracy of the model has been criticized [65, 66].

In this Chapter, we will present results obtained with two different MLPs, applied to study the LLPT. In Sec. 6.1 we will discuss a KRR model trained on VMC reference data. Details about the dataset construction and model training will be presented in Sec. 6.1.1, followed by a discussion of the results for the LLPT (Sec. 6.1.2).

In Sec. 6.2 we will consider a new MLP obtained with MACE and trained on DFT data. Sec. 6.2.1 describes the dataset used to train the model, with the focus on a modified loss function we employed to improve the description of the molecular phase. In Sec. 6.2.2 we will study the order of the LLPT as a function of both temperature and system size. Results for the Widom line will be shown in Sec. 6.2.3. Finally, in Sec. 6.2.4 we report a comparison between structural quan-

tities obtained with the MACE MLP and AIMD, in order to validate our results. Sec. 6.3 contains a summary of the contents of this Chapter.

6.1 VMC-BASED MLP FOR THE LLPT

As discussed previously, the location of the hydrogen LLPT strongly depends on the underlying approximations of the simulation method, e. g., in DFT, on the choice of the exchange-correlation functional. Therefore, if one wants to obtain quantitatively accurate predictions for this transition, more advanced methods, such as QMC, must be employed. Similarly to what we did in Chap.4, here we will show how the techniques described in Sec.3.4 can be applied to train models on QMC data, allowing for QMC-quality simulations at more tractable AIMD cost. This application was originally published in Ref. [39].

6.1.1 *KRR model dataset and training*

We used the Δ-learning approach to construct a KRR model trained on the difference between VMC and DFT reference calculations.

Our initial dataset comprised 510 liquid configurations with 128 atoms each, extracted from the dataset of Ref. [64] containing dynamic snapshots from PBE-MD simulations at T ranging from 800 K to 1500 K, and with $1.26 \leqslant r_s \leqslant 1.60$. In particular, we selected configurations to have the same r_s distribution as the one of the original dataset. We then added another set of configurations according to the active learning scheme described in Sec. 4.6, for a total of 684 data points. We then split the dataset into training and test sets, and "sparsified" the former using the local FPS method described in Sec.3.4.3.

DFT energies, forces, and pressures were calculated using the Quantum Espresso package [202–204] with the PBE exchange-correlation functional [75]. We employed a PAW pseudopotential [205] with a plane wave cut-off of 60 Ry, and a $4 \times 4 \times 4$ Monkhorst-Pack $\mathbf{k}$-points grid.

For the VMC reference calculations, we used a JSD WF with a [3s] basis set for both the Jastrow and antisymmetric parts (see Appendix A). The Jastrow factor was optimized with the SR method for 500 energy minimization steps. To minimize size effects we used twisted average boundary conditions on a $4 \times 4 \times 4$ Monkhorst-Pack mesh of the Brillouin zone.

The final KRR model was trained using a cutoff $r_c = 3.5$ Bohr and hyperparameters $2\sigma^2 = 0.5$ Bohr^2, $n = \eta = 2$ (see Eqs. (37)-(39)-(40)). The weights for training in Eq.(34) were $w_E = 0.1 \times 128^2$, $w_F = 3 \times 128$, and $w_P = 10^2$ (in the appropriate inverse atomic units). For further details on the model performances, see Appendix B.

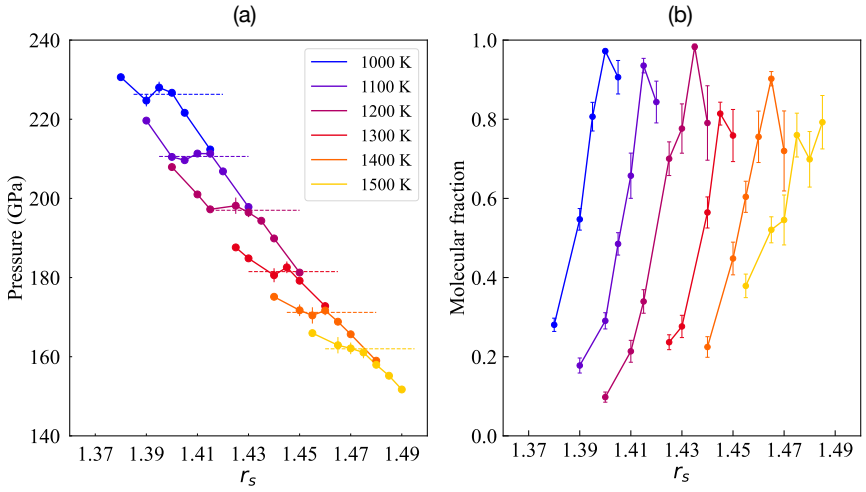

Figure 25: (a) Equation of state of hydrogen at different values of the temperature T, as obtained by our VMC-MLP for a system of N = 256 atoms. The dashed lines at each T indicate the estimated transition pressure. (b) Value of the molecular fraction (defined in the main text) as a function of r_s, for each temperature.

6.1.2 *LLPT results analysis*

We performed classical NVT simulations for several values of the density and temperatures between 1000 K and 1500 K on two systems of N = 128 and N = 256 H atoms, respectively. At each MD step, we first computed DFT-PBE forces for the given configuration and then summed the correction given by the MLP. We used a time step of 10 atomic units (i.e. 0.242 fs) and controlled the temperature using either second-order Langevin dynamics [206] or the stochastic velocity rescaling method [136]. For each density-temperature point, we ran trajectories of length between 2 ps and 10 ps, over which we averaged the thermodynamic quantities of interest.

The resulting equations of state for N = 256 are reported in Fig. 25a. For all temperatures, the results suggest the presence of a pressure plateau, in agreement with previous works pointing to the first-order nature of the LLPT in this system [68, 69, 229, 232], even though the length of our simulations cannot exclude the possibility of a crossover, in particular for the highest temperatures considered. The transition pressures and corresponding densities are quantitatively different from the ones obtained with the underlying PBE baseline (see Fig.26), further supporting the importance of the VMC correction for accurately estimating these quantities. We also calculated the "stable molecular fraction" [233, 238, 246] corresponding to the densities close to the transition. In particular, we computed the average number of hydrogen pairs whose constituent atoms remain within a distance of 2 Bohr for at least a time $\tau \sim 70$ fs (2280 a.u.) along the trajectory (see Fig.25b). Across the transition, this quantity

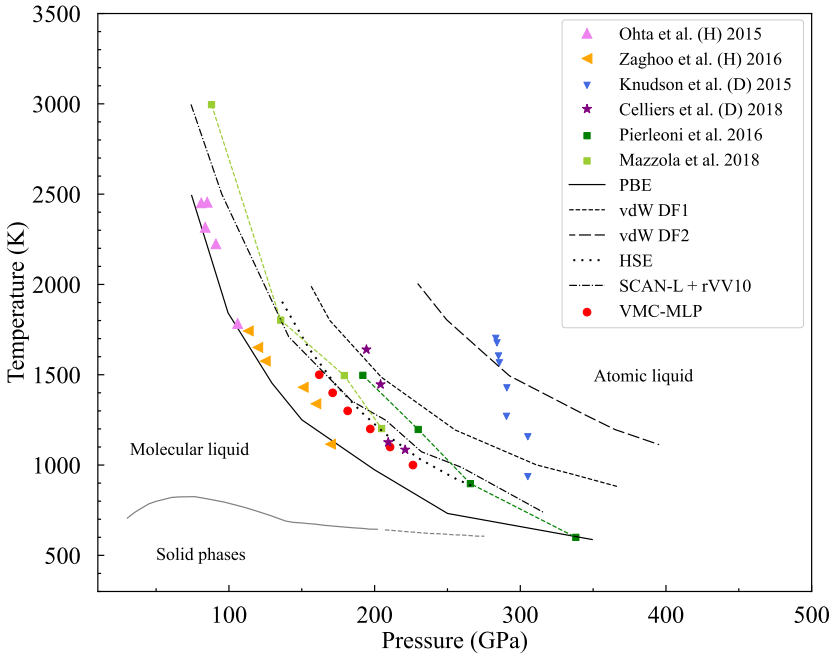

Figure 26: Proposed transition lines of the LLPT in hydrogen. Results obtained with our VMC-MLP (red circles). Previous (classical) QMC calculations taken from Refs. [236] and [69] (light green and dark green squares). DFT results for several functionals as reported in Ref.[69] (black lines). Recent classical results calculated in Ref. [65] with the PBE functional (black solid lines) and in Ref. [240] with the SCAN-L functional (dash-dotted black line) are also shown. DAC experimental data are taken from Refs. [221] and [224] (pink and orange triangles). Dynamical compression experiments are from Refs.[222] and [227] (blue triangles and violet stars, respectively).

jumps from small values, corresponding to an atomic liquid, to values closer to 1, typical of a molecular system.

In Fig. 26 we report the transition line computed with the VMC-MLP, alongside previous numerical and experimental results. As expected, our findings are in good agreement with the VMC results of Ref. [236], which employed the same type of wave function (i.e., the JSD). The minor discrepancies observed are likely attributable to the higher statistical noise (roughly an order of magnitude larger) affecting energies, forces, and pressures in their simulations. With respect to other QMC-based simulations, we notice that our transition line is approximately ~ 30 GPa lower than the ones predicted by the CEIMC method [69]. This can be explained by the different cell sizes used here (256 vs 54 and 128 in Ref. [69]), and by the different WF, which in their case included the backflow transformation [247]. Moreover, since we used the JSD WF, the self-consistency error is, in principle, present in our calculations of ionic forces and pressures (see Sec. 2.3). Our

results are in excellent agreement with the DFT ones obtained with the hybrid Heyd-Scuseria-Ernzerhof (HSE) [248] and the strongly constrained and appropriately normed (SCAN-L) [240] functionals, suggesting their accuracy in describing the LLPT.

As already mentioned, our results only considered classical ions. However, NQEs are significant in this range of temperature, shifting the transition line by 20-30 GPa towards lower pressures [69]. Taking this into account, our simulations are close to the DAC experiments of Refs [221] and [224].

Before proceeding further, we briefly comment on the first-order character of the LLPT suggested by our results. As previously discussed, the relatively short length of our simulations prevents a clear conclusion regarding the nature of the transition. In fact, the pressure plateaux visible in Fig. 25 are subject to large uncertainties, and distinguishing between a sharp crossover or a first-order process is not trivial (see, for example, the EOS at $T = 1500$K).

Even if our MLP correction is very cheap to evaluate, using DFT as a baseline limits the capability of our approach to perform simulations longer than 10 ps with reasonable resources. The simplest solution to this problem is to replace the *ab initio* DFT calculation with a cheaper method. Our experience suggests that this is not trivial, since (within a fixed training set) the accuracy of the resulting Δ-learning model rapidly decreases if the baseline is too approximate. To avoid this, the use of other MLPs trained on DFT data (for which it is possible to generate extended datasets) as baseline potentials is promising. This approach was recently applied to study the hydrogen melting line, using two NN trained on DFT and QMC [40]. Having an accurate DFT-based MLP is crucial in this case.

6.2 DFT MACE MODEL

We will now describe results obtained using a MACE model [132] trained on a DFT-PBE dataset in the LLPT region, which have not been published yet at the time of writing. This MLP allowed us to carefully study the character of the transition and how it is affected by the system size.

6.2.1 *MACE model dataset and training*

To train the final MACE model, we generated a dataset of 21812 configurations. We extracted ~ 17000 configurations of $N = 128$ hydrogen atoms from the VMC-MD simulations of Refs. [39] and [41] (see Fig. 27). An additional set of 3000 configurations with 128 atoms was selected from a series of AIMD simulations at lower temperatures, i. e., $T = 800$ K and $T = 900$ K. We also added 500 snapshots with a larger number of atoms, i. e., $N = 256$ and $N = 512$ extracted from MD

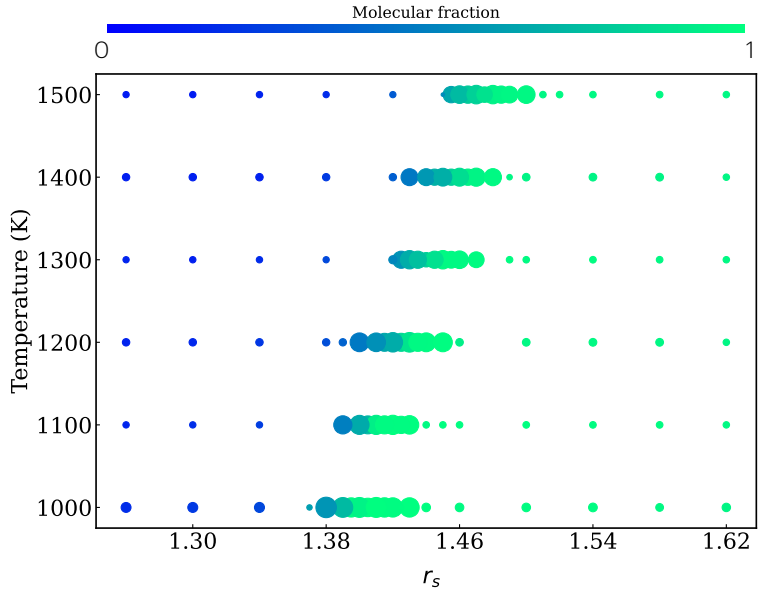

Figure 27: Distribution of configurations sampled from MD runs of Ref. [39] to generate part of the dataset of the MACE model. The dimension of the dot is proportional to the number of configurations extracted from each MD simulation, while the color indicates the stable molecular fraction along the trajectory (with $\tau \sim 20$ fs).

runs with early iterations of our MACE model. Finally, we included solid structures (see Ref. [54]) and ~ 100 low-temperature configurations from Ref. [40]. To ensure consistency across the dataset, we recomputed energies, forces, and pressure within DFT, using the PBE functional. A PAW pseudopotential together with a 60 Ry plane waves cutoff was used. A sufficiently dense **k**-point grid for each system size was employed to cure finite size effects. For instance, a $4 \times 4 \times 4$ grid was used for the configurations with 128 atoms.

We constructed the MACE model with the aim of reaching a good tradeoff between accuracy and speed. A cutoff radius of $r_c = 3$ Å was used; notice that within MACE the "effective" radius is twice this value, because of the message passing step. The model considered 128 equivariant messages and a correlation order of 3. MACE normally employs a "standard" loss function like the one of Eq.(34), which can include energy, forces, and virial stress terms. In the first iterations of the model, we noticed the appearance of rather stable molecular solid-like structures at high temperatures (T ~ 1200 K) during the dynamics. These structures are clearly unphysical because they are not observed in AIMD simulations. In particular, the models trained using the "standard" loss function energetically favor these configurations, even after their inclusion in the training set, as shown in Fig. 28. On the contrary, these models seem to correctly describe the atomic

phase. With the aim of increasing the accuracy in the molecular phase, we modified the loss to include a penalty

$$\Delta\mathcal{L}\left(\{\beta\}\right) = w_E \frac{\lambda}{N_{batch}} \left| \sum_{\mu\in mol} \frac{1}{N_\mu} \left(E^\mu_{pred}\left(\{\beta\}\right) - E^\mu_{ref} \right) \right|, \tag{50}$$

where λ is a parameter controlling the relative weight of the penalty with respect to the energy term of the standard loss function, and the sum only considers "molecular" configurations in a given training set batch of dimension N_{batch}. To perform this classification, we used a "static" criterion that, contrary to the "dynamical" one described in Sec. 6.1.2, only depends on the atomic positions of the given configuration. In particular, a configuration is classified as molecular if the first peak of the radial distribution function $g(r)$ (estimated by fitting the $g(r)$ with a Gaussian function for $r \in [0, 1.3]$Å) is larger than 1.8. We observed that the models trained using the "modified" loss $\mathcal{L} + \Delta\mathcal{L}$ have a much smaller energy error for the molecular solid structures (see Fig. 28), which also do not appear during the dynamics at high temperature. The penalty slightly alters the error distribution on the rest of the training set, but shows an overall smaller value of the energy RMSE and mean absolute error (MAE) (see Fig. 29).

A similar outcome may also be obtained by assigning a larger weight to these snapshots, even though we have not explored this possibility.

The training was performed using the Adam optimizer [249] with a $N_{batch} = 16$ batch size and initial learning rate of 0.01. We extracted 280 configurations for testing from the full dataset; among the remaining configurations, 95% of them were used for training, while 5% was taken as the validation set. This last group of configurations is used during the training to assess the performance of the model. The loss weights were taken equal to $w_E = 1$, $w_F = 100$, $w_{virial} = 100$ for the first 320 epochs (defined as N_t/N_{batch} optimization steps, with N_t the training set dimension) and $w_E = 200$, $w_F = 10$, $w_{virial} = 10$ for the remaining 130 epochs. In Eq. (50) we set $\lambda = 50$. At the end of the optimization, the best model is selected as the one minimizing the value of the loss on the validation set. The final accuracy of our MACE model, measured by the RMSE on energy per atom and forces computed on the training, validation, and test sets is reported in Tab. 7. The value of the RMSE on the virial pressure calculated on the test set was ~ 1 GPa. Compared to the previous NN MLP proposed in Ref. [64] our model has an error on energy $6-7$ times smaller and it is twice as accurate on forces.

6.2.2 *LLPT simulations: size scaling*

Thanks to our MACE model we were able to study the behavior of the LLPT as a function of both temperature and system size. To do this,

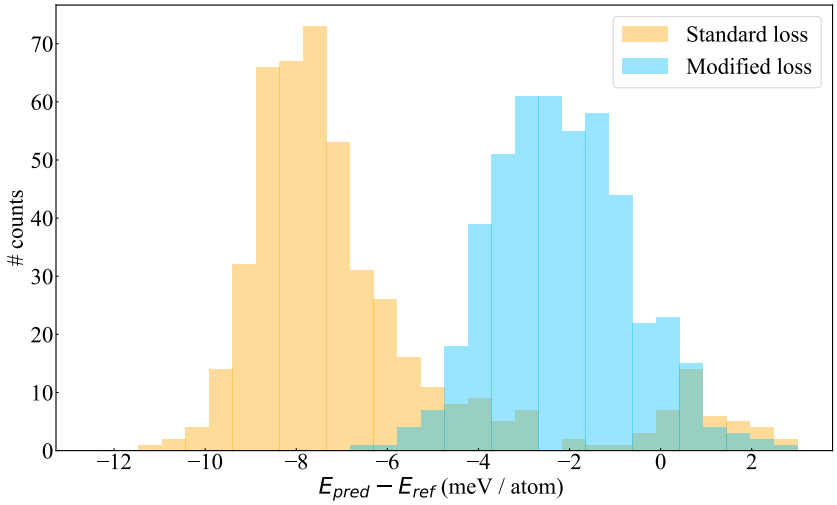

Figure 28: Histogram of the difference between the energy per atom predicted by the model and the reference PBE value for solid-like molecular configurations that appeared during the dynamics at high temperature. We reported in orange the results of the MACE model trained with the standard loss and in blue the ones obtained with a model trained with the modified loss described in the text.

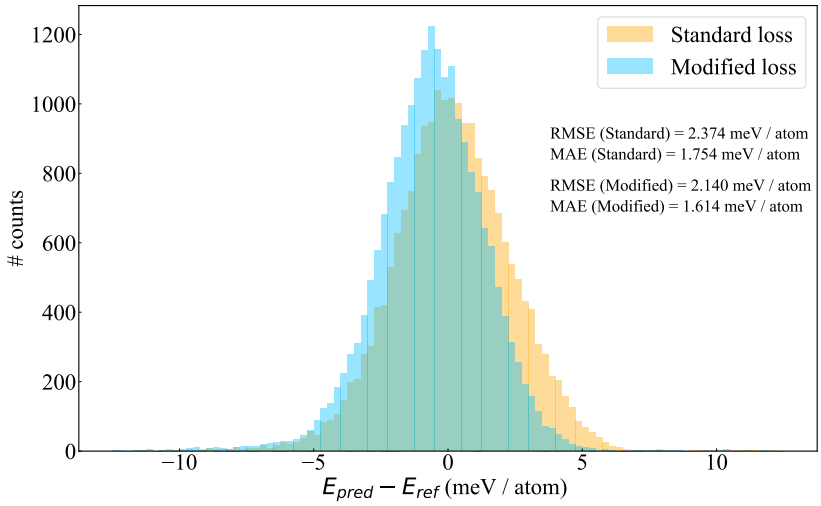

Figure 29: Histogram of the difference between the energy per atom predicted by the model and the reference PBE value for the configurations in the training set. We reported in orange the results of the MACE model trained with the standard loss and in blue the ones obtained with a model trained with the modified loss described in the text.

	$RMSE_E$ (meV / atom)	$RMSE_F$ (meV / Å)
Training	2.2	116
Validation	2.1	118
Test	2.0	116

Table 7: RMSE for the energy per atom ($RMSE_E$) and forces ($RMSE_F$) calculated on the training, validation, and test sets, respectively, using the final MACE model.

we ran MD simulations in the NVT ensemble using the LAMMPS code [250] interfaced with MACE. We considered different systems with $N = 128$, $N = 256$, $N = 512$, and $N = 2048$ hydrogen atoms and cubic supercells. To accurately study the order of the transition we used a very dense r_s grid in the vicinity of the LLPT, and performed simulations of length in between $0.1 - 0.3$ ns, about two orders of magnitude longer than our previous simulations [39] and what can be usually achieved with AIMD. In the dynamics, we used a time step of 0.2 fs and controlled the temperature via the stochastic velocity rescaling method [136], with a characteristic time $\tau = 0.1$ ps. The EOSs obtained with our simulations are shown in Figs. 30-33 for four different temperatures.

From these results, we can identify three distinct regimes for the LLPT in the temperature range between $T = 900$ K and $T = 1100$ K. At the lowest temperature, i. e.$T = 900$ K (Fig. 30) the model clearly predicts a first-order transition, signaled by the presence of hysteresis. This behavior is present for all the system sizes explored. For $T = 950$ K and $T = 1000$ K (Figs. 31 and 32), the EOS in the transition region has a strong dependence on the system size: the small systems (i. e., $N = 128$ and $N = 256$) again suggest a first-order transition, while for larger N the pressure plateau and hysteresis are missing. Finally at $T \geqslant 1100$ K (Fig. 33), the results indicate a smooth crossover between the atomic and molecular state, with a relatively small size dependence.

To investigate the LLPT, the sole observation of the EOS is not sufficient, since it is known that for liquid-liquid transitions the density is not the correct order parameter [251]. In particular, we computed the stable molecular fraction of the system, using the dynamic criterion with a cutoff of 1.05 Å and a H_2 lifetime of $\tau \sim 80$ fs. The results corresponding to values of r_s slightly before/after the transition are also shown in Figs.30-33. In all cases, the molecular fraction rapidly increases from values ~ 0.5 to values close to 1. Another quantity that we can use to analyze the structure and spatial correlations of the system near the transition is the structure factor $S(\mathbf{k})$, where, in our case, $\mathbf{k} = \frac{2\pi}{L}(n_1, n_2, n_3)$, with L being the side of the cubic simulation box and n_1, n_2, n_3 integer numbers. In particular, the

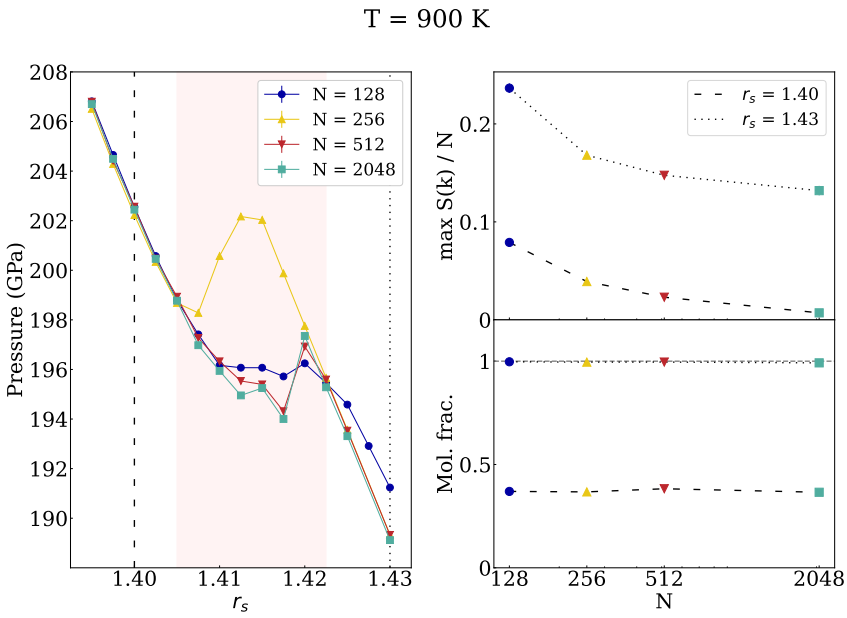

Figure 30: [Left panel] EOS at T = 900 K for different system sizes. The red-shaded region indicates the r_s range of the first-order transition predicted by the model, indicated by the presence of hysteresis. [Right panel] Values for the average $\max_{\mathbf{k}} S(\mathbf{k})$ [top] and for the molecular fraction [bottom] as a function of the number of atoms N for two different r_s values.

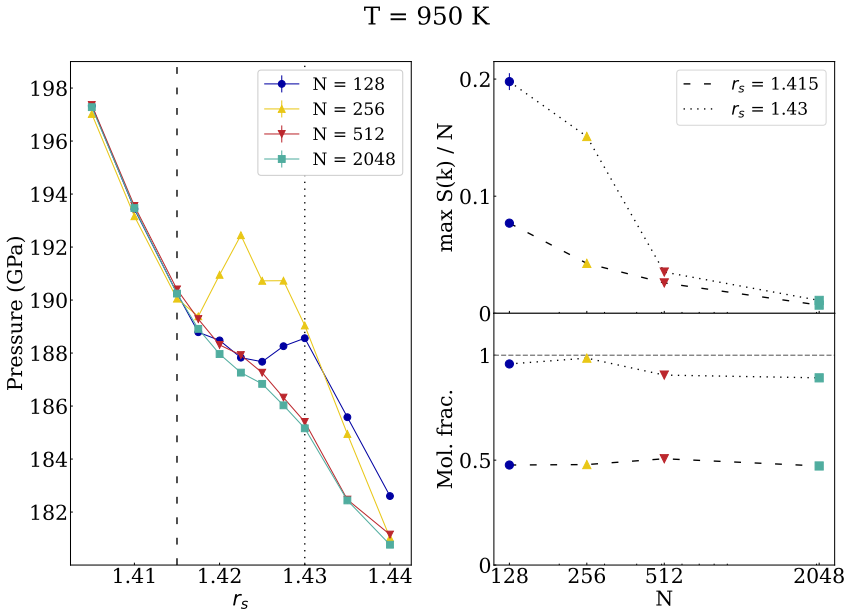

Figure 31: [Left panel] EOS at T = 950 K for different system sizes. [Right panel] Values for the average $\max_{\mathbf{k}} S(\mathbf{k})$ [top] and for the molecular fraction [bottom] as a function of the number of atoms N for two different r_s values.

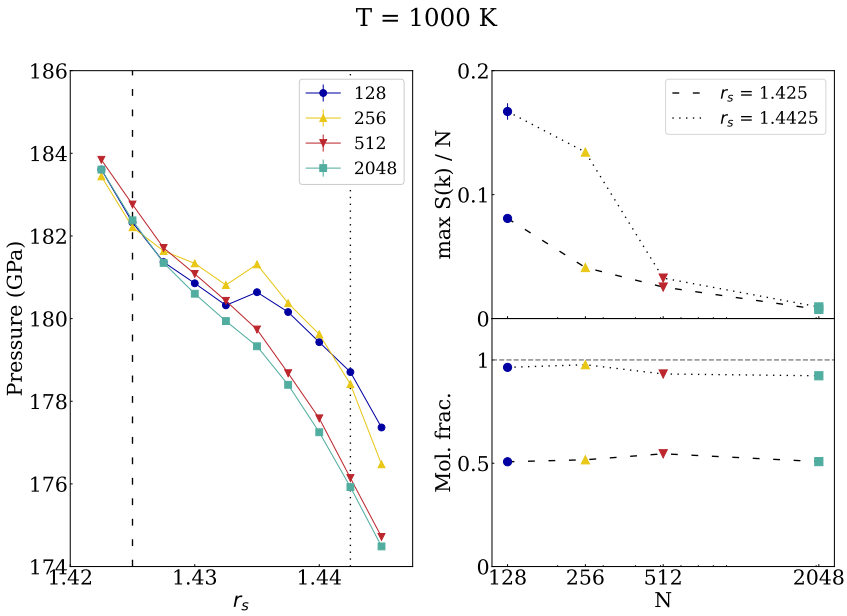

Figure 32: [Left panel] EOS at T = 1000 K for different system sizes. [Right panel] Values for the average $\max_{\mathbf{k}} S(\mathbf{k})$ [top] and for the molecular fraction [bottom] as a function of the number of atoms N for two different r_s values.

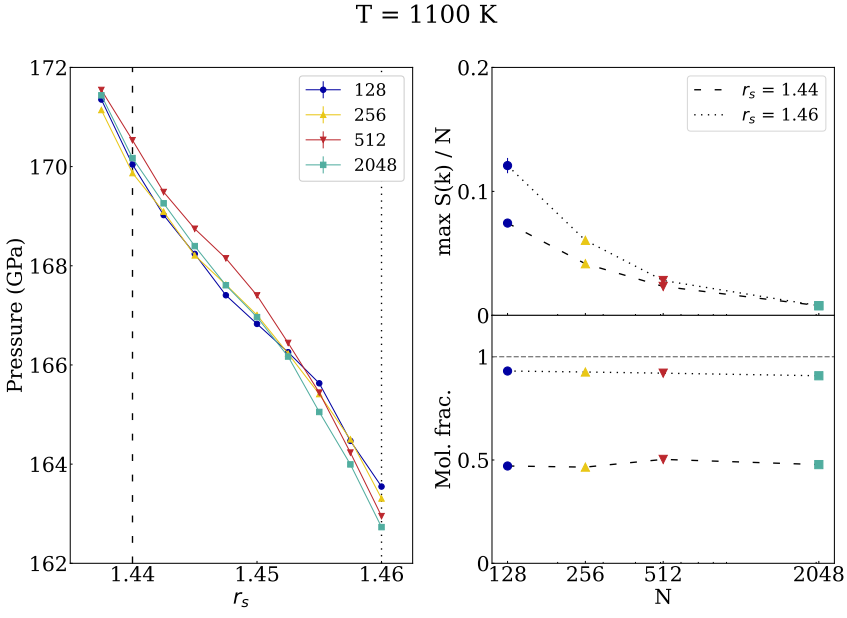

Figure 33: [Left panel] EOS at T = 1100 K for different system sizes. [Right panel] Values for the average $\max_{\mathbf{k}} S(\mathbf{k})$ [top] and for the molecular fraction [bottom] as a function of the number of atoms N for two different r_s values.

maximum value of $S(\mathbf{k})$ can be used as a proxy for the formation of crystalline structures. Indeed, in a solid system we expect only a discrete number of reciprocal $\mathbf{k}$ vectors to contribute to the structure factor, i. e., those corresponding to the particular crystal symmetry, so that $\max_{\mathbf{k}} S(\mathbf{k}) \propto N$ in the thermodynamic limit. On the contrary, in a liquid, the rotational invariance implies that the contribution of each $\mathbf{k}$ to the $S(\mathbf{k})$ will stay constant for $N \to \infty$. The time-averaged value of $\max_{\mathbf{k}} S(\mathbf{k})/N$ as a function of the number of particles is reported in the top right panels of Figs.30-33. Interestingly, here we can realize that the first-order transition observed at $T = 900$ K is accompanied by a non-vanishing value of $\max_{\mathbf{k}} S(\mathbf{k})/N$ in the thermodynamic limit, revealing a long-range spatial order of the molecular phase at this temperature. Moreover, for $T = 950$ K and $T = 1000$ K, a large value of $\max_{\mathbf{k}} S(\mathbf{k})/N$ is present for $N = 128$ and $N = 256$ at values of the r_s after the pressure plateau. This suggests that the first-order transition seen at these values of N is due to finite-size effects and it is between a molecular crystal and an atomic liquid. This explanation was already proposed in Ref. [64], even though the persistence of these solid structures with N and consequent transition at lower temperatures was not observed there, probably because of the lower accuracy of their MLP. For the larger systems, i. e., $N \geqslant 512$, the formation of the solid is not favored anymore. This is shown in Fig.34, where the behavior of $\max_{\mathbf{k}} S(\mathbf{k})/N$ is plotted as a function of the simulation time, for $T = 950$ K and $r_s = 1.43$. Notice how the system with $N = 512$ occasionally crystallizes, as indicated by the two "jumps" of the structure factor maximum at the beginning and the end of the MD, but it mostly remains liquid. On the contrary, the small system with $N = 256$ is in a solid-like state for the majority of the time. Fig.34 further suggests that ~ 100 ps long simulations are necessary in order to have truly converged results near the transition (10 ps are necessary to melt the crystal for $N = 512$). To further characterize the structure at the lowest temperature $T = 900$ K, we computed the mean squared displacement (MSD):

$$\mathrm{MSD}(t) = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{R}_i(t) - \mathbf{R}_i(0)|^2 . \tag{51}$$

The MSD computed for $T = 900$ K and $T = 1000$ K for different values of r_s is shown in Fig. 35. Notice how the "solid-like" structure (corresponding to $T = 900$ and $r_s = 1.43$) shows a non-zero diffusivity. In Ref. [65], this observation was used as an indication of the liquid phase. Here we obtained very similar values of the MSD (see the supplementary material of Ref. [65]), although our measure of the $\max_{\mathbf{k}} S(\mathbf{k})/N$ clearly indicates the presence of long-range spatial correlations. A snapshot taken from an MD simulation at $T = 900$ K and $r_s = 1.43$ for a 2048 atoms system is shown in Fig. 36. From this, we can notice the appearance of "planes" formed by the hydrogen

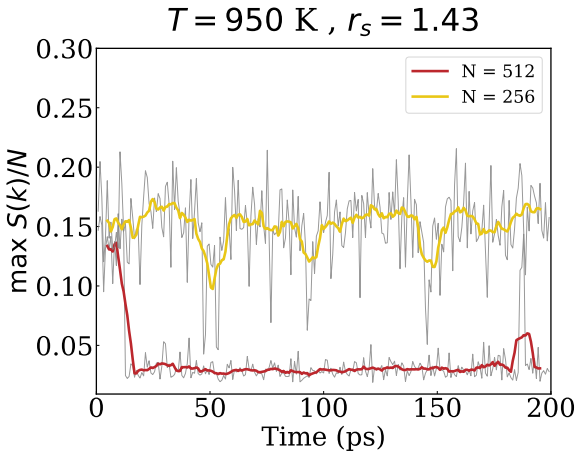

Figure 34: Value of $\max_{\mathbf{k}} S(\mathbf{k})/N$ as a function of the simulation time for two systems, with N = 256 and N = 512 atoms, respectively, at temperature T = 950 K and $r_s = 1.43$.

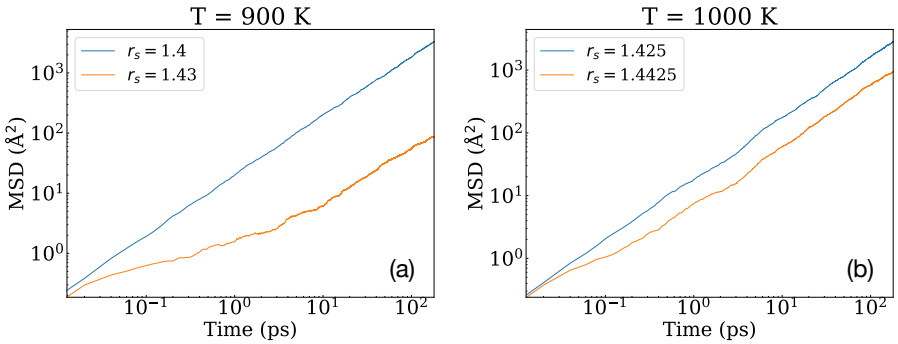

Figure 35: MSD as a function of time for (a) T = 900 K and $r_s = 1.40$, $r_s = 1.43$ (b) T = 1000 K and $r_s = 1.425$, $r_s = 1.4425$.

molecules. In summary, our results suggest that the first-order character observed for T = 900 K can be due to the formation of defective solid structures, favored by the vicinity of the hydrogen melting line. It is likely that a proper treatment of the melting, e. g., using non-cubic simulation boxes, might predict a proper solid-liquid transition for these temperatures. Regarding this, we noticed that our transition point at T = 900 K lies exactly on the PBE melting line recently obtained in Ref. [40] using a NN MLP, further confirming this hypothesis (see Fig. 40).

In the atomic phase, our simulations confirm that the system is a liquid at all temperatures. A genuine molecular liquid is also observed for all system sizes at temperature $T \geqslant 1100$ K. Our MACE model indicates that above the melting line for $T \geqslant 950$ K the LLPT is a continuous crossover in the thermodynamic limit (Figs. 32 and 33).

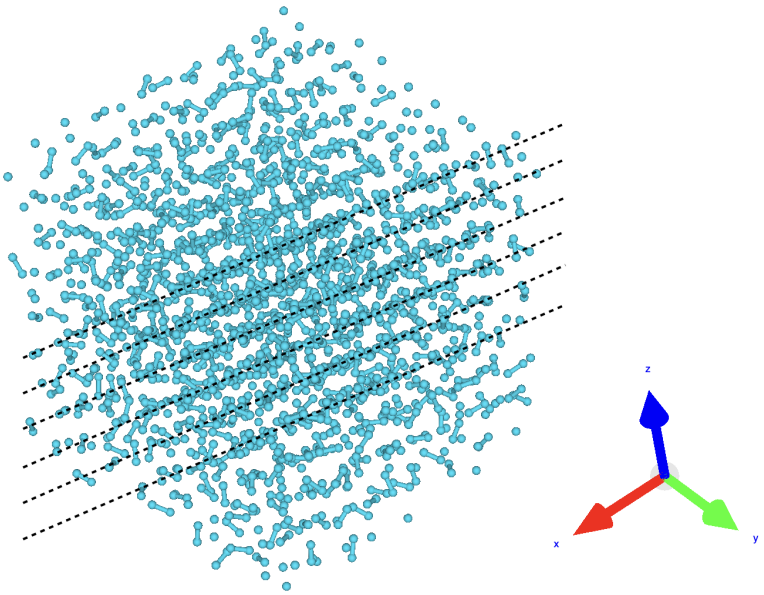

Figure 36: 2048 atoms structure extracted from an MD simulation at T = 900 K and r_s = 1.43. The dashed lines are a guide for the eye to highlight the intermolecular planes. The image was obtained using the VESTA visualization program [252].

At this point, a comparison between the physical picture given by the MACE model and previous AIMD results is necessary. For the PBE functional, many authors [65, 68, 229, 232, 235, 238] reported a first-order LLPT below a critical temperature T_c, based on the observation of "kinks" in the EOS. We have already noted in Sec. 6.1.2, how short simulations make the identification of the plateaux unclear. This is reflected by the large variability of the predicted T_c, which varies from T = 4000 K [235], to T = 1500 K [68, 232]. Recent AIMD PBE simulations reported in Ref. [238] brought down the estimated T_c to ~ 1250 K. Our model predicts a first-order transition below a temperature between 900 K and 950 K, consistently lower than these estimates. Therefore we cannot entirely attribute the observation of a first-order LLPT in these previous works to the formation of solid structures.

To analyze the possibility that the plateaux identified in AIMD originate from the short length of these simulations, we analyzed 0.2 ns long trajectories obtained with the MACE model for values of r_s close to the transition at different temperatures above T = 900 K. The results are shown in Fig. 37 for a system of N = 128 atoms at a temperature of T = 1400 K and a system of N = 512 atoms at T = 1000 K. Using the long trajectory, we computed the running average of the pressure using a variable size time window τ_{run}, corresponding to simulations times achievable within AIMD, e. g., τ_{run} ~ 10 ps. The variance of the running average gives us an idea of the variability of the estimated equilibrium pressure that one is expected to have

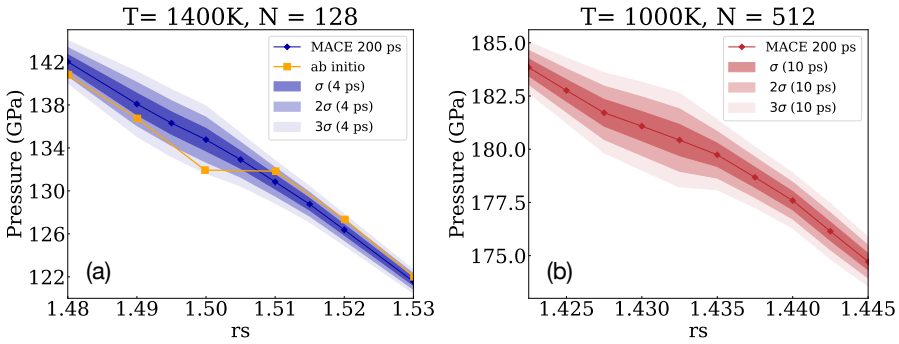

Figure 37: Results for 200 ps long simulations and corresponding confidence intervals, estimated from a running average with a τ_{run}-long time window. (a) Results for $N = 128$, $T = 1400$ K and $\tau_{run} = 4$ ps. The AIMD result reported in Ref. [39] is also shown. (b) Results for $N = 512$, $T = 1000$ K and $\tau_{run} = 10$ ps.

when using an MD simulation of length equal to τ_{run}. From Fig. 37, we can notice that the estimated variance grows in the vicinity of the transition, and that the presence of plateaux in the pressure can be understood as an artifact due to the lack of sampling of the phase space. For the small system (Fig. 37a), we also compared the results with the PBE AIMD ones reported in Ref. [39] at $T = 1400$ K for the same size. The pressure plateau here is within the 3σ uncertainty region estimated from a running average of 4 ps (slightly longer than the average length of the AIMD simulations). The results at lower temperature $T = 1000$ K for a larger system of 512 atoms (Fig. 37b) show that even a 10 ps long dynamics can, in principle, produce artificial kinks of the size reported in the literature [232, 238]. Even though, in principle, this estimation depends on the thermostat used in the simulations, we do not expect this to change our conclusions.

Before proceeding, we mention that we also ran PIOUD simulations on a system of $N = 512$ atoms to see if the inclusion of NQEs changes our physical picture. The preliminary results of these simulations again show solid-like molecular structures and a first-order transition at low temperatures and a continuous LLPT at higher temperatures. The transition is shifted toward lower pressures because of NQEs, consistently with previous results [68, 69, 222].

6.2.3 *Results for the Widom line*

Both the first-order transition and the crossover above the critical temperature T_c can be further characterized by studying the behavior of the specific heat c_v (i. e., the heat capacity per particle) as a function of temperature, density, and system dimension. The results for c_v obtained from simulations with $N = 2048$ hydrogen atoms at temperatures up to 1800 K are reported in Fig. 38. For each temperature, we

estimated the location and value of the maximum. For temperatures between 900 K and 1100 K, we also studied the size scaling of the c_v maximum, as shown in Fig. 39. In a first-order transition, the c_v peak presents a divergence $\sim N$ in the thermodynamic limit [253]. As shown in the figure, linear scaling is observed at $T = 900$ K, while the ratio c_v/N tends to zero for higher temperatures. This is consistent with the conclusion of the previous section, and further indicates the location of the critical point between $T = 900$ K and $T = 950$ K. The location of the c_v maximum at higher temperatures allows us to locate the Widom line for the supercritical region. The results are reported in Fig. 40. Our Widom line shows remarkable agreement with previous simulations of the LLPT obtained with the PBE functional. In particular, our results well reproduce the ones reported in Ref. [232] for temperatures up to 1400 K and those of Ref. [65] for higher temperatures, even though they do not agree on the LLPT first-order character for $T \geqslant 900$ K. Since fluctuations are large at the Widom line, this can further suggest that the EOS plateaux observed in previous works could be due to short simulation times. As mentioned earlier, the first-order transition observed at $T = 900$ K lies on the melting line estimated in Ref. [40], supporting the idea that, below this temperature, the transition could coincide with melting. This behavior found by our MACE model qualitatively agrees with the conclusions of Ref. [64], even though their model could not resolve the aforementioned low-temperature regime, and predicted a Widom line (estimated from the isobaric specif heat c_p maximum) far from the AIMD result (see Fig. 40).

6.2.4 *Comparison with AIMD*

We conclude this Chapter by directly comparing results obtained with our MACE model and AIMD simulations to validate our MLP results. Here we report NVT results obtained at values of density and temperature roughly corresponding to $p = 150$ GPa that we took from Ref. [65]. We performed simulations of length ~ 2 ps with both AIMD and the MACE model on a system of 512 atoms, and compared the structural properties of the system. The time-resolved value of $\max_{\mathbf{k}} S(\mathbf{k})$ for each density and temperature is shown in Fig. 41 for AIMD and Fig. 42 for MACE. The two methods show good agreement and both predict solid-like structures at $T = 1050$ K. Notice how this temperature matches the value of the melting line of Ref. [40] at $p = 150$ GPa. A comparison of the radial distribution function is reported in Fig. 43. At all temperatures, the MACE model shows a remarkable agreement with the AIMD result, both in the molecular and atomic phases. This is a further confirmation of the reliability of our MLP.

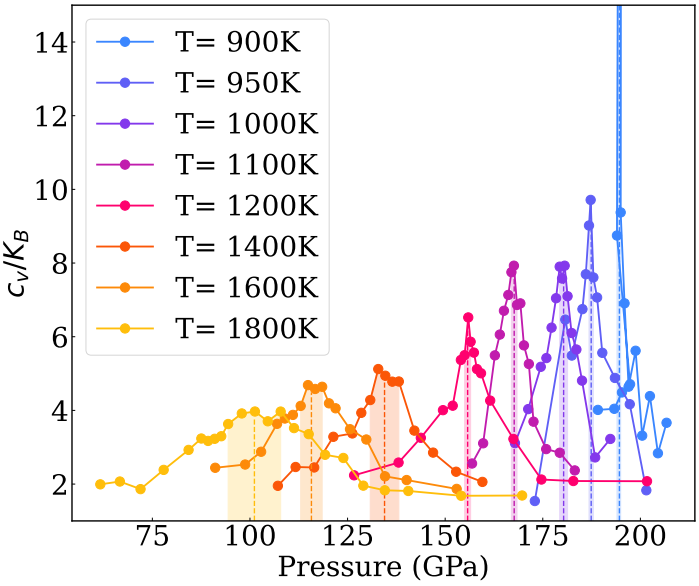

Figure 38: Specific heat per particle vs pressure along the isotherms, for a system of N = 2048 hydrogen atoms obtained with the MACE model. The shaded areas indicate the uncertainty in the peak position.

6.3 WRAP-UP

To conclude this Chapter, we summarize the main physical outcomes of our simulations. Using the Δ-learning approach we trained a KRR model on the difference between VMC and DFT data to study the hydrogen liquid-liquid phase transition (LLPT). Our results are in good agreement with previous *ab initio* QMC simulations [236], and with results obtained using advanced DFT exchange-correlation functionals like HSE and SCAN-L. Shifting our transition line by 20 GPa toward lower pressures, roughly corresponding to the effect of NQEs estimated in Ref. [69], makes our predictions close to the static compression experiments of Refs. [221, 224]. Our VMC simulations indicate a first-order LLPT, although the limited length of our dynamics prevents a definitive conclusion regarding the nature of the transition. To answer this question, the efficiency of the Δ-learning scheme has to be improved, which can be done, for instance, by using a DFT-MLP as baseline.

With this in mind, we constructed a DFT MLP using MACE, thanks to which we studied the transition between molecular and atomic hydrogen for temperatures between 900 K and 1800 K, running MD simulations for different systems sizes. At the lowest temperature (T = 900 K) the model predicts a first-order transition in the thermodynamic limit. The structure factor analysis reveals that long-range

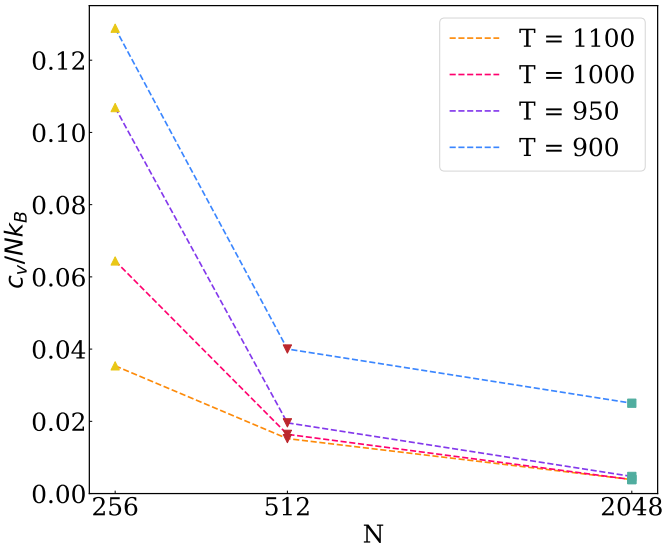

Figure 39: Scaling of $c_v/k_B N$ with the number of atoms in the system for four different temperatures. The scaling confirms a first-order transition for T = 900 K and a smooth crossover for higher temperatures in the thermodynamic limit.

spatial order is present, thus suggesting that this is a transition between a molecular solid-like system (frustrated by the cubic cell) and an atomic liquid. This is also in agreement with recent results for the PBE melting line [40]. For temperatures above T = 900 K, the MACE model predicts a proper LLPT in the thermodynamic limit ($N \geqslant 512$), which is a smooth crossover, similar to what was reported in Ref. [64]. We observed that up to T = 1000 K, the frustrated crystal structures appear for small system sizes, giving rise to a first-order transition that disappears as N is increased. For higher temperatures, all system sizes agree on the crossover nature of the LLPT. Our Widom line, estimated from the maximum of the specific heat c_v, shows an excellent agreement with the LLPT location reported by previous PBE AIMD results [65, 232]. This might explain why a first-order transition was predicted by these works: being fluctuations important close to the Widom line, short simulations can incorrectly identify density jumps/pressure plateaux in the EOS. Finally, preliminary results obtained with PIMD suggest that our conclusions remain the same even with the inclusion of NQEs.

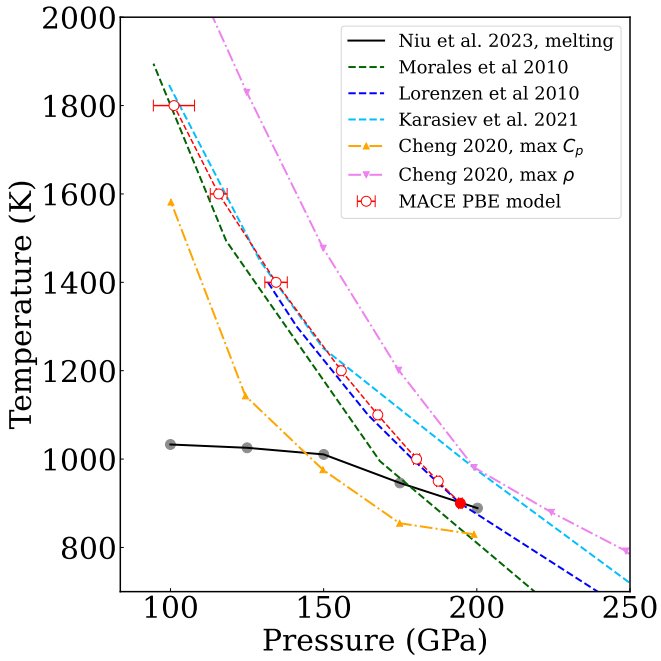

Figure 40: Classical PBE-LLPT location as computed with different methods. BOMD results by Ref. [232] (blue dashed line), Ref. [68] (green dashed line), Ref. [65] (light blue dashed line). The results for the molecular-to-atomic crossover obtained with an NN MLP by Ref. [64] are reported with orange and violet markers, corresponding to the maximum of the isobaric specific heat and density, respectively. The black markers and line indicate the recently proposed PBE melting line by Ref.[40], obtained with an NN MLP and the two-phase method. Our MACE model results are indicated with red markers. The filled point at T = 900 K indicates the first-order character of the transition, while the empty points correspond to the location of the Widom line given by the c_v maximum (see Fig. 38).

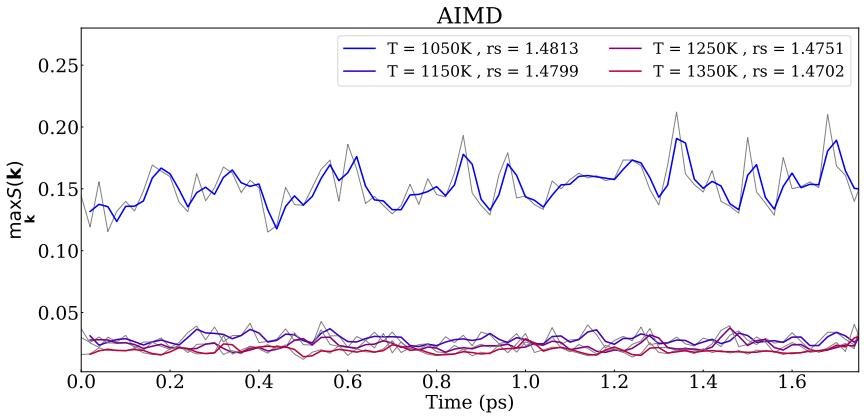

Figure 41: $\max_{\mathbf{k}} S(\mathbf{k})$ resolved in time for different values of r_s and temperatures, obtained from AIMD simulations for a system of 512 atoms.

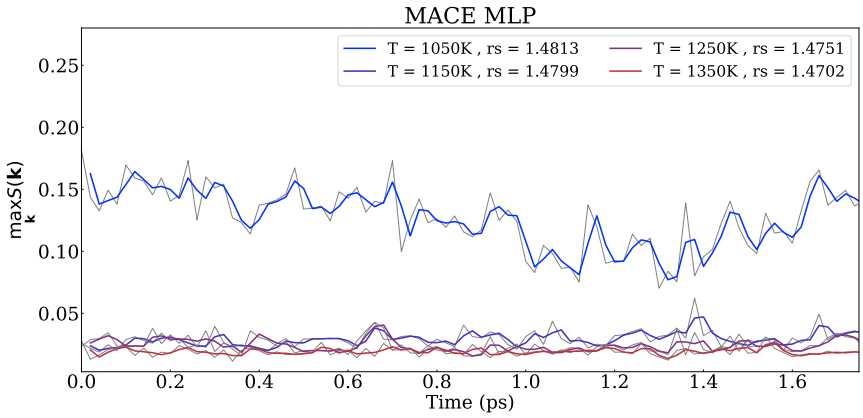

Figure 42: $\max_{\mathbf{k}} S(\mathbf{k})$ resolved in time for different values of r_s and temperatures, obtained from MD simulations using our MACE model for a system of 512 atoms.

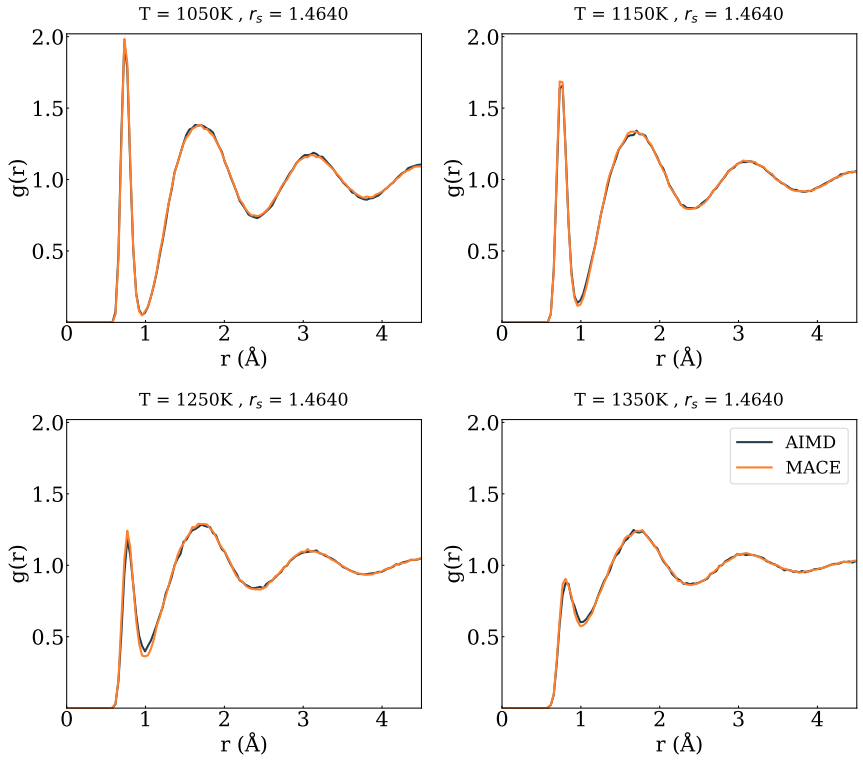

Figure 43: Comparison between AIMD and MACE for the radial distribution function $g(r)$ at different values of r_s and temperature.

7

CONCLUSIONS AND PERSPECTIVES

Quantum Monte Carlo (QMC) methods provide accurate and reliable results for the properties of high-pressure hydrogen, but their high computational cost limits their applicability in molecular dynamics (MD) simulations. In this thesis, we discussed how, thanks to the application of machine learning techniques, one can perform QMC-based MD simulations at a fraction of their original cost.

The construction of these machine learning potentials (MLPs) using QMC data points is difficult for multiple reasons. First, generating extended QMC datasets requires a large amount of resources, orders of magnitude larger than those necessary within density functional theory (DFT), at least for its simplest local or semi-local exchange-correlation functional approximations. In our work, we solved this problem by employing the Δ-learning approach. Thanks to this technique we obtained accurate models trained on both variational Monte Carlo (VMC) and lattice regularized diffusion Monte Carlo (LRDMC) with small and affordable datasets, which we used to study the liquid part of the hydrogen phase diagram, namely the Hugoniot curve (Ch. 4) and the liquid-liquid phase transition (LLPT) (Ch. 6). The downside of the Δ-learning method resides in the efficiency of the final dynamics, which is not only due to the MLP correction but also to the cost of evaluating the baseline potential. In the applications presented here, we always employed a DFT baseline, meaning that our simulations have the same efficiency as AIMD. Even though this approach is much faster than computing QMC energies and forces at each MD step, it is still quite limited in terms of simulation time. Very accurate DFT-MLP like the MACE model presented in Ch. 6, may represent the ideal candidate for replacing the *ab initio* baseline. A natural follow-up of the work presented in this thesis is to verify whether the physical outcome on the LLPT order found with the DFT-MLP is changed by increasing the level of theory from DFT to QMC. The combination of the Δ-learning approach with an efficient and accurate MACE DFT baseline allows such a study, an application we will explore in the near future.

The other aspect that makes the training of QMC-MLP challenging is that, within QMC, the calculation of energy derivatives is not trivial, as we have extensively discussed when talking about the self-consistency error. In Ch.2 we developed an approach to remove the bias in the VMC forces and pressures originating from a partially optimized wave function (WF), such as the Jastrow correlated Slater determinant (JSD) with a frozen antisymmetric part taken from DFT. The

availability of this correction is crucial, because it allows the utilization of the relatively cheap-to-optimize JSD WF for generating large unbiased training sets for machine learning applications, as demonstrated in Ch. 5.

In conclusion, QMC-based machine learning potentials represent one of the most promising approaches for the numerical simulation of high-pressure hydrogen and potentially other complex systems. This thesis has introduced and validated a robust set of techniques for constructing these models reliably and demonstrated their practical applications. Looking forward, fast and efficient QMC-MLPs capable of exploring extended phase diagrams, from solid structures to liquid phases, appear to be within closer reach.

****

APPENDIX

A

LOCALIZED BASIS SETS FOR QMC CALCULATIONS

In the following sections, we describe the localized basis sets employed in the QMC calculations of this thesis.

A.1 HUGONIOT DATASET

The localized basis employed in the QMC calculations for the Hugoniot dataset discussed in Sec. 4.7 is reported in Tab. 8. For the Slater determinant (SD) part we computed the exponents by explicitly optimizing them for a few selected configurations at the conditions relevant to the Hugoniot curve, starting with a cc-pVTZ basis. We verified that the optimal values do not significantly depend on the density and temperature in the range explored. For the d-type orbital, we kept the initial exponent of the cc-pVTZ GEO, since this decreases the linear dependency of the basis and does not impact the final energy. We optimized the Jastrow basis exponents for every configuration in the dataset. In Tab. 8 we thus report only their starting values. The GTOs labeled s^* and p^* have the corresponding (Gaussian) radial part multiplied by r^2 and r, respectively.

A.2 LLPT DATASET

The VMC calculations on the dataset we used for training the MLP presented in Sec. 6.1 employed a basis of [3s] GTOs for both the antisymmetric and Jastrow parts. The exponents of the Gaussian functions are reported in Tab. 9.

	Orbital type	GTO exponents
SD	s	2.9727
		0.8465
		0.2667
		0.1027
	p	0.2701
		0.1530
	d	1.057
Jastrow	s	0.2497
	s*	0.4543
	p	0.5878
	p*	0.5179
	d	0.2953

Table 8: Basis set exponents employed for the QMC calculations on the Hugoniot dataset. The values relative to the Jastrow factor were optimized for each configuration, starting from the values reported here.

	Orbital type	GTO exponents
SD	s	1.962
		0.4446
		0.122
Jastrow	s	1.962
		0.4446
		0.122

Table 9: GTOs exponents employed for the QMC calculations on the LLPT dataset, for both the Slater determinant part and Jastrow factor.

B

VALIDATION OF THE LLPT VMC MODEL

Here we will give further details on the performances and validation of the model described in Sec. 6.1.

The RMSE of our MLP measured on the test set is 0.39 mHa / atom and 2.47 mHa / Bohr for energy and forces respectively. We notice a significant bias in the VMC pressure in our dataset, caused by the self-consistency error. This does not spoil much the accuracy of the model, since the relative weight w_P we used in the loss function is very small compared to that of energy and forces ($w_P/w_E \approx 0.06$). Notice how a much larger value was employed for the Hugoniot model (see Sec. 4.8), i. e., $w_P/w_E \approx 30$ for the MLP trained on VMC data.

To further validate our model, we adopted a similar strategy to the one shown in Sec. 4.8 for the MLP trained on the Hugoniot dataset. In particular, we trained a second model on the difference between VMC and DFT data with the BLYP functional (the MLP used for the simulations employed the PBE functional). We compare the EOS at $T = 1000$ K obtained with the two models for a system of $N = 128$ hydrogen atoms in Fig. 44. The equations of state predicted by the two MLPs show good consistency in the entire range of r_s explored. By looking at their maximum discrepancy at $r_s \sim 1.4$, we can assign an uncertainty of ~ 5 GPa to the transition pressure predicted by our model.

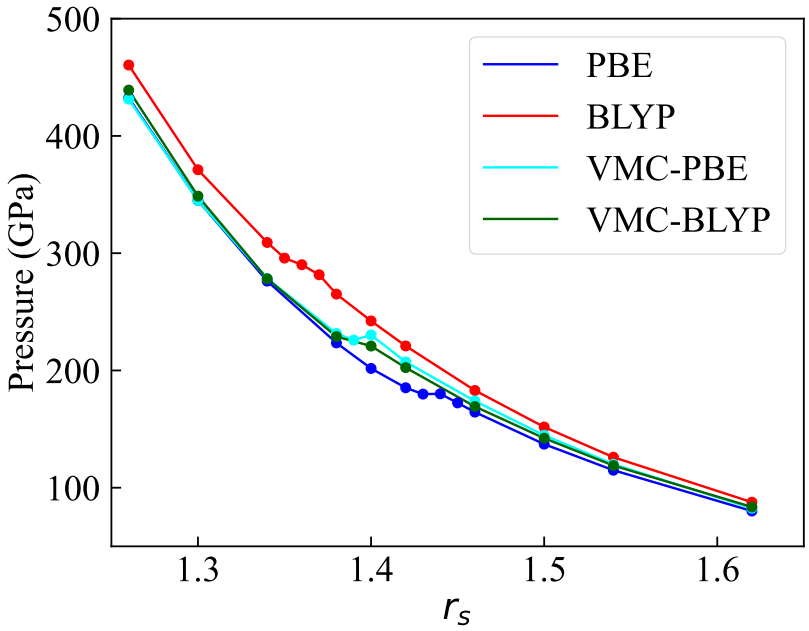

Figure 44: Equation of state obtained with two VMC-MLPs employing two different baseline potentials, using the PBE and BLYP functionals, respectively. The corresponding DFT EOSs obtained with AIMD are also reported.

BIBLIOGRAPHY

[1] Ralph H Fowler. "On dense matter." In: *Monthly Notices of the Royal Astronomical Society, Vol. 87, p. 114-122* 87 (1926), pp. 114–122. DOI: 10.1093/mnras/87.2.114.

[2] E. Wigner and H. B. Huntington. "On the Possibility of a Metallic Modification of Hydrogen." In: *The Journal of Chemical Physics* 3.12 (1935), pp. 764–770. DOI: 10.1063/1.1749590.

[3] David J. Stevenson. "Jupiter's Interior as Revealed by Juno." In: *Annual Review of Earth and Planetary Sciences* 48.Volume 48, 2020 (2020), pp. 465–489. DOI: 10.1146/annurev-earth-081619-052855.

[4] Burkhard Militzer et al. "Juno Spacecraft Measurements of Jupiter's Gravity Imply a Dilute Core." In: *The Planetary Science Journal* 3.8 (2022), p. 185. DOI: 10.3847/PSJ/ac7ec8.

[5] Ravit Helled, David J. Stevenson, Jonathan I. Lunine, Scott J. Bolton, Nadine Nettelmann, Sushil Atreya, Tristan Guillot, Burkhard Militzer, Yamila Miguel, and William B. Hubbard. "Revelations on Jupiter's formation, evolution and interior: Challenges from Juno results." In: *Icarus* 378 (2022), p. 114937. DOI: 10.1016/j.icarus.2022.114937.

[6] Howard, S. et al. "Jupiter's interior from Juno: Equation-of-state uncertainties and dilute core extent." In: *A&A* 672 (2023), A33. DOI: 10.1051/0004-6361/202245625.

[7] L. Iess et al. "Measurement and implications of Saturn's gravity field and ring mass." In: *Science* 364.6445 (2019), eaat2965. DOI: 10.1126/science.aat2965.

[8] Ni, Dongdong. "Understanding Saturn's interior from the Cassini Grand Finale gravity measurements." In: *A&A* 639 (2020), A10. DOI: 10.1051/0004-6361/202038267.

[9] Andrew P Ingersoll. "Cassini Exploration of the Planet Saturn: A Comprehensive Review." In: *Space Science Reviews* 216.8 (2020), p. 122. DOI: 10.1007/s11214-020-00751-1.

[10] Miguel, Y., Guillot, T., and Fayon, L. "Jupiter internal structure: the effect of different equations of state." In: *A&A* 596 (2016), A114. DOI: 10.1051/0004-6361/201629732.

[11] Mazevet, S., Licari, A., and Soubiran, F. "Benchmarking the ab initio hydrogen equation of state for the interior structure of Jupiter." In: *A&A* 664 (2022), A112. DOI: 10.1051/0004-6361/201935764.

[12] S. X. Hu, B. Militzer, V. N. Goncharov, and S. Skupsky. "First-principles equation-of-state table of deuterium for inertial confinement fusion applications." In: *Phys. Rev. B* 84 (22 2011), p. 224109. DOI: 10.1103/PhysRevB.84.224109.

[13] R. S. Craxton et al. "Direct-drive inertial confinement fusion: A review." In: *Physics of Plasmas* 22.11 (2015), p. 110501. DOI: 10.1063/1.4934714.

[14] S. X. Hu, V. N. Goncharov, T. R. Boehly, R. L. McCrory, S. Skupsky, L. A. Collins, J. D. Kress, and B. Militzer. "Impact of first-principles properties of deuterium–tritium on inertial confinement fusion target designs." In: *Physics of Plasmas* 22.5 (2015), p. 056304. DOI: 10.1063/1.4917477.

[15] Gaetano Squadrito, Gaetano Maggio, and Agatino Nicita. "The green hydrogen revolution." In: *Renewable Energy* 216 (2023), p. 119041. DOI: 10.1016/j.renene.2023.119041.

[16] A P Drozdov, M I Eremets, I A Troyan, V Ksenofontov, and S I Shylin. "Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system." In: *Nature* 525.7567 (2015), pp. 73–76. DOI: 10.1038/nature14964.

[17] Maddury Somayazulu, Muhtar Ahart, Ajay K. Mishra, Zachary M. Geballe, Maria Baldini, Yue Meng, Viktor V. Struzhkin, and Russell J. Hemley. "Evidence for Superconductivity above 260 K in Lanthanum Superhydride at Megabar Pressures." In: *Phys. Rev. Lett.* 122 (2 2019), p. 027001. DOI: 10.1103/PhysRevLett.122.027001.

[18] M I Eremets et al. "High-Temperature Superconductivity in Hydrides: Experimental Evidence and Details." In: *Journal of Superconductivity and Novel Magnetism* 35.4 (2022), pp. 965–977. DOI: 10.1007/s10948-022-06148-1.

[19] N. W. Ashcroft. "Metallic Hydrogen: A High-Temperature Superconductor?" In: *Phys. Rev. Lett.* 21 (26 1968), pp. 1748–1749. DOI: 10.1103/PhysRevLett.21.1748.

[20] Alexander F. Goncharov and Russell J. Hemley. "Probing hydrogen-rich molecular systems at high pressures and temperatures." In: *Chem. Soc. Rev.* 35 (10 2006), pp. 899–907. DOI: 10.1039/B607523C.

[21] William A. Bassett. "Diamond anvil cell, 50th birthday." In: *High Pressure Research* 29.2 (2009), pp. 163–186. DOI: 10.1080/08957950802597239.

[22] Agnès Dewaele, Paul Loubeyre, Florent Occelli, Olivier Marie, and Mohamed Mezouar. "Toroidal diamond anvil cell for detailed measurements under extreme static pressures." In: *Nature Communications* 9.1 (2018), p. 2913. DOI: 10.1038/s41467-018-05294-2.

[23] W J Nellis. "Dynamic compression of materials: metallization of fluid hydrogen at high pressures." In: *Rep. Prog. Phys.* 69.5 (2006), pp. 1479–1580. DOI: 10.1088/0034-4885/69/5/r05.

[24] Ranga P. Dias and Isaac F. Silvera. "Observation of the Wigner-Huntington transition to metallic hydrogen." In: *Science* 355.6326 (2017), pp. 715–718. DOI: 10.1126/science.aal1579.

[25] Mikhail I Eremets, Alexander P Drozdov, PP Kong, and Hongbo Wang. "Semimetallic molecular hydrogen at pressure above 350 GPa." In: *Nature Physics* 15.12 (2019), pp. 1246–1249. DOI: 10.1038/s41567-019-0646-x.

[26] Paul Loubeyre, Florent Occelli, and Paul Dumas. "Synchrotron infrared spectroscopic evidence of the probable transition to metal hydrogen." In: *Nature* 577.7792 (2020), pp. 631–635. DOI: 10.1038/s41586-019-1927-3.

[27] Eugene Gregoryanz, Cheng Ji, Philip Dalladay-Simpson, Bing Li, Ross T. Howie, and Ho-Kwang Mao. "Everything you always wanted to know about metallic hydrogen but were afraid to ask." In: *Matter and Radiation at Extremes* 5.3 (2020), p. 038101. DOI: 10.1063/5.0002104.

[28] Michael Bonitz et al. "Toward first principles-based simulations of dense hydrogen." In: *Physics of Plasmas* 31.11 (2024), p. 110501. DOI: 10.1063/5.0219405.

[29] Gerald Irwin Kerley. "A Model for the Calculation of Thermodynamic Properties of a Fluid." In: *Molecular-Based Study of Fluids*. Washington DC: ACS, 1983. Chap. 5, pp. 107–138. DOI: 10.1021/ba-1983-0204.ch005.

[30] W. Ebeling and W. Richert. "Plasma phase transition in hydrogen." In: *Physics Letters A* 108.2 (1985), pp. 80–82. DOI: 10.1016/0375-9601(85)90521-3.

[31] Didier Saumon, Gilles Chabrier, and Hugh M van Horn. "An equation of state for low-mass stars and giant planets." In: *Astrophys. J., Suppl. Ser.* 99 (1995), p. 713. DOI: 10.1086/192204.

[32] Marvin Ross. "Linear-mixing model for shock-compressed liquid deuterium." In: *Phys. Rev. B* 58 (2 1998), pp. 669–677. DOI: 10.1103/PhysRevB.58.669.

[33] Gerald Kerley. *Equations of state for hydrogen and deuterium.* 2003. DOI: 10.2172/917468.

[34] Richard M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004.

[35] Jeffrey M. McMahon, Miguel A. Morales, Carlo Pierleoni, and David M. Ceperley. "The properties of hydrogen and helium under extreme conditions." In: *Rev. Mod. Phys.* 84 (4 2012), pp. 1607–1653. DOI: 10.1103/RevModPhys.84.1607.

[36] Federico Becca and Sandro Sorella. *Quantum Monte Carlo approaches for correlated systems*. Cambridge University Press, 2017.

[37] Jörg Behler. "Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations." In: *Phys. Chem. Chem. Phys.* 13 (40 2011), pp. 17930–17955. DOI: 10.1039/C1CP21668F.

[38] Jörg Behler. "Perspective: Machine learning potentials for atomistic simulations." In: *The Journal of Chemical Physics* 145.17 (2016), p. 170901. DOI: 10.1063/1.4966192.

[39] Andrea Tirelli, Giacomo Tenti, Kousuke Nakano, and Sandro Sorella. "High-pressure hydrogen by machine learning and quantum Monte Carlo." In: *Phys. Rev. B* 106 (4 2022), p. L041105. DOI: 10.1103/PhysRevB.106.L041105.

[40] Hongwei Niu, Yubo Yang, Scott Jensen, Markus Holzmann, Carlo Pierleoni, and David M. Ceperley. "Stable Solid Molecular Hydrogen above 900 K from a Machine-Learned Potential Trained with Diffusion Quantum Monte Carlo." In: *Phys. Rev. Lett.* 130 (7 2023), p. 076102. DOI: 10.1103/PhysRevLett.130.076102.

[41] Giacomo Tenti, Kousuke Nakano, Andrea Tirelli, Sandro Sorella, and Michele Casula. "Principal deuterium Hugoniot via quantum Monte Carlo and Δ-learning." In: *Phys. Rev. B* 110 (4 2024), p. L041107. DOI: 10.1103/PhysRevB.110.L041107.

[42] David M Ceperley, Scott Jensen, Yubo Yang, Hongwei Niu, Carlo Pierleoni, and Markus Holzmann. "Training models using forces computed by stochastic electronic structure methods." In: *Electronic Structure* 6.1 (2024), p. 015011. DOI: 10.1088/2516-1075/ad2eb0.

[43] Kevin K. Ly and David M. Ceperley. *Melting of atomic hydrogen and deuterium with path-integral Monte Carlo*. 2024. DOI: 10.48550/arXiv.2409.19484.

[44] Shubhang Goswami, Scott Jensen, Yubo Yang, Markus Holzmann, Carlo Pierleoni, and David M. Ceperley. *High temperature melting of dense molecular hydrogen from machine-learning interatomic potentials trained on quantum Monte Carlo*. 2024. DOI: 10.48550/arXiv.2411.15665.

[45] Kousuke Nakano, Michele Casula, and Giacomo Tenti. "Efficient calculation of unbiased atomic forces in ab initio variational Monte Carlo." In: *Phys. Rev. B* 109 (20 2024), p. 205151. DOI: 10.1103/PhysRevB.109.205151.

[46] Catherine E Housecroft and Alan G Sharpe. *Inorganic chemistry*. Vol. 1. Pearson Education, 2008.

[47] David Halliday, Robert Resnick, and Jearl Walker. *Fundamentals of physics*. John Wiley & Sons, 2013.

[48] Cheng Ji et al. "Ultrahigh-pressure isostructural electronic transitions in hydrogen." In: *Nature* 573.7775 (2019), pp. 558–562. DOI: 10.1038/s41586-019-1565-9.

[49] Cheng Ji et al. "Crystallography of low Z material at ultrahigh pressure: Case study on solid hydrogen." In: *Matter and Radiation at Extremes* 5.3 (2020), p. 038401. DOI: 10.1063/5.0003288.

[50] Chris J Pickard and Richard J Needs. "Structure of phase III of solid hydrogen." In: *Nature Physics* 3.7 (2007), pp. 473–476. DOI: 10.1038/nphys625.

[51] Jorge Kohanoff, Sandro Scandolo, Stefano de Gironcoli, and Erio Tosatti. "Dipole-Quadrupole Interactions and the Nature of Phase III of Compressed Hydrogen." In: *Phys. Rev. Lett.* 83 (20 1999), pp. 4097–4100. DOI: 10.1103/PhysRevLett.83.4097.

[52] Sam Azadi and Thomas D. Kühne. "Unconventional phase III of high-pressure solid hydrogen." In: *Phys. Rev. B* 100 (15 2019), p. 155103. DOI: 10.1103/PhysRevB.100.155103.

[53] Maria Hellgren, Damian Contant, Thomas Pitts, and Michele Casula. "High-pressure II-III phase transition in solid hydrogen: Insights from state-of-the-art ab initio calculations." In: *Phys. Rev. Res.* 4 (4 2022), p. L042009. DOI: 10.1103/PhysRevResearch.4.L042009.

[54] Lorenzo Monacelli, Michele Casula, Kousuke Nakano, Sandro Sorella, and Francesco Mauri. "Quantum phase diagram of high-pressure hydrogen." In: *Nature Physics* 19.6 (2023), pp. 845–850. DOI: 10.1038/s41567-023-01960-5.

[55] Alexander F. Goncharov, I. I. Mazin, Jon H. Eggert, Russell J. Hemley, and Ho-kwang Mao. "Invariant Points and Phase Transitions in Deuterium at Megabar Pressures." In: *Phys. Rev. Lett.* 75 (13 1995), pp. 2514–2517. DOI: 10.1103/PhysRevLett.75.2514.

[56] Bruce J. Baer, William J. Evans, and Choong-Shik Yoo. "Coherent Anti-Stokes Raman Spectroscopy of Highly Compressed Solid Deuterium at 300 K: Evidence for a New Phase and Implications for the Band Gap." In: *Phys. Rev. Lett.* 98 (23 2007), p. 235503. DOI: 10.1103/PhysRevLett.98.235503.

[57] Alexander F. Goncharov, Russell J. Hemley, and Ho-kwang Mao. "Vibron frequencies of solid H2 and D2 to 200 GPa and implications for the P–T phase diagram." In: *The Journal of Chemical Physics* 134.17 (2011), p. 174501. DOI: 10.1063/1.3574009.

[58] Ross T. Howie, Thomas Scheler, Christophe L. Guillaume, and Eugene Gregoryanz. "Proton tunneling in phase IV of hydrogen and deuterium." In: *Phys. Rev. B* 86 (21 2012), p. 214104. DOI: 10.1103/PhysRevB.86.214104.

[59] Ross T. Howie, Christophe L. Guillaume, Thomas Scheler, Alexander F. Goncharov, and Eugene Gregoryanz. "Mixed Molecular and Atomic Phase of Dense Hydrogen." In: *Phys. Rev. Lett.* 108 (12 2012), p. 125501. DOI: 10.1103/PhysRevLett.108.125501.

[60] Tommaso Morresi, Rodolphe Vuilleumier, and Michele Casula. "Hydrogen phase-IV characterization by full account of quantum anharmonicity." In: *Phys. Rev. B* 106 (5 2022), p. 054109. DOI: 10.1103/PhysRevB.106.054109.

[61] Philip Dalladay-Simpson, Ross T Howie, and Eugene Gregoryanz. "Evidence for a new phase of dense hydrogen above 325 gigapascals." In: *Nature* 529.7584 (2016), pp. 63–67. DOI: 10.1038/nature16164.

[62] Bartomeu Monserrat, Neil D. Drummond, Philip Dalladay-Simpson, Ross T. Howie, Pablo López Ríos, Eugene Gregoryanz, Chris J. Pickard, and Richard J. Needs. "Structure and Metallicity of Phase V of Hydrogen." In: *Phys. Rev. Lett.* 120 (25 2018), p. 255701. DOI: 10.1103/PhysRevLett.120.255701.

[63] Chang-sheng Zha, Hanyu Liu, John S. Tse, and Russell J. Hemley. "Melting and High P—T Transitions of Hydrogen up to 300 GPa." In: *Phys. Rev. Lett.* 119 (7 2017), p. 075302. DOI: 10.1103/PhysRevLett.119.075302.

[64] Bingqing Cheng, Guglielmo Mazzola, Chris J Pickard, and Michele Ceriotti. "Evidence for supercritical behaviour of high-pressure liquid hydrogen." In: *Nature* 585.7824 (2020), pp. 217–220. DOI: 10.1038/s41586-020-2677-y.

[65] Valentin V Karasiev, Joshua Hinz, SX Hu, and SB Trickey. "On the liquid–liquid phase transition of dense hydrogen." In: *Nature* 600.7889 (2021), E12–E14. DOI: 10.1038/s41586-021-04078-x.

[66] Bingqing Cheng, Guglielmo Mazzola, Chris J Pickard, and Michele Ceriotti. "Reply to: On the liquid–liquid phase transition of dense hydrogen." In: *Nature* 600.7889 (2021), E15–E16. DOI: 10.1038/s41586-021-04079-w.

[67] Ross T Howie, Philip Dalladay-Simpson, and Eugene Gregoryanz. "Raman spectroscopy of hot hydrogen above 200 GPa." In: *Nature materials* 14.5 (2015), pp. 495–499. DOI: 10.1038/nmat4213.

[68] Miguel A. Morales, Carlo Pierleoni, Eric Schwegler, and D. M. Ceperley. "Evidence for a first-order liquid-liquid transition in high-pressure hydrogen from ab initio simulations." In: *Proceedings of the National Academy of Sciences* 107.29 (2010), pp. 12799–12803. DOI: 10.1073/pnas.1007309107.

[69] Carlo Pierleoni, Miguel A. Morales, Giovanni Rillo, Markus Holzmann, and David M. Ceperley. "Liquid–liquid phase transition in hydrogen by coupled electron–ion Monte Carlo simulations." In: *Proceedings of the National Academy of Sciences* 113.18 (2016), pp. 4953–4957. DOI: 10.1073/pnas.1603853113.

[70] Carlo Pierleoni, Markus Holzmann, and David M. Ceperley. "Local structure in dense hydrogen at the liquid–liquid phase transition by coupled electron–ion Monte Carlo." In: *Contributions to Plasma Physics* 58.2-3 (2018), pp. 99–106. DOI: 10.1002/ctpp.201700184.

[71] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln." In: *Annalen der Physik* 389.20 (1927), pp. 457–484. DOI: 10.1002/andp.19273892002.

[72] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas." In: *Phys. Rev.* 136 (3B 1964), B864–B871. DOI: 10.1103/PhysRev.136.B864.

[73] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects." In: *Phys. Rev.* 140 (4A 1965), A1133–A1138. DOI: 10.1103/PhysRev.140.A1133.

[74] J. P. Perdew and Alex Zunger. "Self-interaction correction to density-functional approximations for many-electron systems." In: *Phys. Rev. B* 23 (10 1981), pp. 5048–5079. DOI: 10.1103/PhysRevB.23.5048.

[75] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized Gradient Approximation Made Simple." In: *Phys. Rev. Lett.* 77 (18 1996), pp. 3865–3868. DOI: 10.1103/PhysRevLett.77.3865.

[76] Roberto Dovesi et al. "Quantum-mechanical condensed matter simulations with CRYSTAL." In: *WIREs Computational Molecular Science* 8.4 (2018), e1360. DOI: 10.1002/wcms.1360.

[77] N. David Mermin. "Thermal Properties of the Inhomogeneous Electron Gas." In: *Phys. Rev.* 137 (5A 1965), A1441–A1443. DOI: 10.1103/PhysRev.137.A1441.

[78] Valentin V. Karasiev, Travis Sjostrom, James Dufty, and S. B. Trickey. "Accurate Homogeneous Electron Gas Exchange-Correlation Free Energy for Local Spin-Density Calculations." In: *Phys. Rev. Lett.* 112 (7 2014), p. 076403. DOI: 10.1103/PhysRevLett.112.076403.

[79] Valentin V. Karasiev, James W. Dufty, and S. B. Trickey. "Nonempirical Semilocal Free-Energy Density Functional for Matter under Extreme Conditions." In: *Phys. Rev. Lett.* 120 (7 2018), p. 076401. DOI: 10.1103/PhysRevLett.120.076401.

[80] Kushal Ramakrishna, Tobias Dornheim, and Jan Vorberger. "Influence of finite temperature exchange-correlation effects in hydrogen." In: *Phys. Rev. B* 101 (19 2020), p. 195129. DOI: 10.1103/PhysRevB.101.195129.

[81] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. "Quantum Monte Carlo simulations of solids." In: *Rev. Mod. Phys.* 73 (1 2001), pp. 33–83. DOI: 10.1103/RevModPhys.73.33.

[82] Kousuke Nakano et al. "TurboRVB: A many-body toolkit for ab initio electronic simulations by quantum Monte Carlo." In: *The Journal of Chemical Physics* 152.20 (2020), p. 204121. DOI: 10.1063/5.0005037.

[83] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. "Equation of State Calculations by Fast Computing Machines." In: *The Journal of Chemical Physics* 21.6 (1953), pp. 1087–1092. DOI: 10.1063/1.1699114.

[84] W. K. Hastings. "Monte Carlo sampling methods using Markov chains and their applications." In: *Biometrika* 57.1 (1970), pp. 97–109. DOI: 10.1093/biomet/57.1.97.

[85] Sandro Sorella and Luca Capriotti. "Algorithmic differentiation and the calculation of forces by quantum Monte Carlo." In: *The Journal of Chemical Physics* 133.23 (2010), p. 234111. DOI: 10.1063/1.3516208.

[86] Sandro Sorella. "Green Function Monte Carlo with Stochastic Reconfiguration." In: *Phys. Rev. Lett.* 80 (20 1998), pp. 4558–4561. DOI: 10.1103/PhysRevLett.80.4558.

[87] Sandro Sorella, Michele Casula, and Dario Rocca. "Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods." In: *The Journal of Chemical Physics* 127.1 (2007), p. 014105. DOI: 10.1063/1.2746035.

[88] Guglielmo Mazzola, Andrea Zen, and Sandro Sorella. "Finite-temperature electronic simulations without the Born-Oppenheimer constraint." In: *The Journal of Chemical Physics* 137.13 (2012), p. 134112. DOI: 10.1063/1.4755992.

[89] Sandro Sorella. "Wave function optimization in the variational Monte Carlo method." In: *Phys. Rev. B* 71 (24 2005), p. 241103. DOI: 10.1103/PhysRevB.71.241103.

[90] C. J. Umrigar, Julien Toulouse, Claudia Filippi, S. Sorella, and R. G. Hennig. "Alleviation of the Fermion-Sign Problem by Optimization of Many-Body Wave Functions." In: *Phys. Rev. Lett.* 98 (11 2007), p. 110201. DOI: 10.1103/PhysRevLett.98.110201.

[91] Julien Toulouse and C. J. Umrigar. "Optimization of quantum Monte Carlo wave functions by energy minimization." In: *The Journal of Chemical Physics* 126.8 (2007), p. 084102. DOI: 10.1063/1.2437215.

[92] Léon Bottou. "On-line Learning and Stochastic Approximations." In: *On-Line Learning in Neural Networks*. Ed. by DavidEditor Saad. Publications of the Newton Institute. Cambridge University Press, 1999, pp. 9–42.

[93] Tosio Kato. "On the eigenfunctions of many-particle systems in quantum mechanics." In: *Communications on Pure and Applied Mathematics* 10.2 (1957), pp. 151–177. DOI: 10.1002/cpa.3160100201.

[94] Michele Casula and Sandro Sorella. "Geminal wave functions with Jastrow correlation: A first application to atoms." In: *The Journal of Chemical Physics* 119.13 (2003), pp. 6500–6511. DOI: 10.1063/1.1604379.

[95] M. Bajdich, L. Mitas, L. K. Wagner, and K. E. Schmidt. "Pfaffian pairing and backflow wavefunctions for electronic structure quantum Monte Carlo methods." In: *Phys. Rev. B* 77 (11 2008), p. 115112. DOI: 10.1103/PhysRevB.77.115112.

[96] Mariapia Marchi, Sam Azadi, Michele Casula, and Sandro Sorella. "Resonating valence bond wave function with molecular orbitals: Application to first-row molecules." In: *The Journal of Chemical Physics* 131.15 (2009), p. 154116. DOI: 10.1063/1.3249966.

[97] S. Sorella, N. Devaux, M. Dagrada, G. Mazzola, and M. Casula. "Geminal embedding scheme for optimal atomic basis set construction in correlated calculations." In: *J. Chem. Phys.* 143.24 (2015), p. 244112. DOI: 10.1063/1.4938089.

[98] Michele Casula, Claudia Filippi, and Sandro Sorella. "Diffusion Monte Carlo Method with Lattice Regularization." In: *Phys. Rev. Lett.* 95 (10 2005), p. 100201. DOI: 10.1103/PhysRevLett.95.100201.

[99] D. M. Ceperley and M. H. Kalos. "Quantum Many-Body Problems." In: *Monte Carlo Methods in Statistical Physics*. Ed. by Kurt Binder. Berlin, Heidelberg: Springer Berlin Heidelberg, 1986, pp. 145–194. DOI: 10.1007/978-3-642-82803-4_4.

[100] Nandini Trivedi and D. M. Ceperley. "Ground-state correlations of quantum antiferromagnets: A Green-function Monte Carlo study." In: *Phys. Rev. B* 41 (7 1990), pp. 4552–4569. DOI: 10.1103/PhysRevB.41.4552.

[101] Matteo Calandra Buonaura and Sandro Sorella. "Numerical study of the two-dimensional Heisenberg model using a Green function Monte Carlo technique with a fixed number of walkers." In: *Phys. Rev. B* 57 (18 1998), pp. 11446–11456. DOI: 10.1103/PhysRevB.57.11446.

[102] D. F. B. ten Haaf, H. J. M. van Bemmel, J. M. J. van Leeuwen, W. van Saarloos, and D. M. Ceperley. "Proof for an upper bound in fixed-node Monte Carlo for lattice fermions." In: *Phys. Rev. B* 51 (19 1995), pp. 13039–13045. DOI: 10.1103/PhysRevB.51.13039.

[103] P. J. Reynolds, R. N. Barnett, B. L. Hammond, R. M. Grimes, and W. A. Lester Jr. "Quantum chemistry by quantum monte carlo: Beyond ground-state energy calculations." In: *International Journal of Quantum Chemistry* 29.4 (1986), pp. 589–596. DOI: 10.1002/qua.560290403.

[104] Roland Assaraf and Michel Caffarel. "Computing forces with quantum Monte Carlo." In: *The Journal of Chemical Physics* 113.10 (2000), pp. 4028–4034. DOI: 10.1063/1.1286598.

[105] Claudia Filippi and C. J. Umrigar. "Correlated sampling in quantum Monte Carlo: A route to forces." In: *Phys. Rev. B* 61 (24 2000), R16291–R16294. DOI: 10.1103/PhysRevB.61.R16291.

[106] Simone Chiesa, D. M. Ceperley, and Shiwei Zhang. "Accurate, Efficient, and Simple Forces Computed with Quantum Monte Carlo Methods." In: *Phys. Rev. Lett.* 94 (3 2005), p. 036404. DOI: 10.1103/PhysRevLett.94.036404.

[107] A. Badinski and R. J. Needs. "Total forces in the diffusion Monte Carlo method with nonlocal pseudopotentials." In: *Phys. Rev. B* 78 (3 2008), p. 035134. DOI: 10.1103/PhysRevB.78.035134.

[108] Roland Assaraf, Michel Caffarel, and A. C. Kollias. "Chaotic versus Nonchaotic Stochastic Dynamics in Monte Carlo Simulations: A Route for Accurate Energy Differences in N-Body Systems." In: *Phys. Rev. Lett.* 106 (15 2011), p. 150601. DOI: 10.1103/PhysRevLett.106.150601.

[109] S. Moroni, S. Saccani, and C. Filippi. "Practical Schemes for Accurate Forces in Quantum Monte Carlo." In: *Journal of Chemical Theory and Computation* 10.11 (2014). PMID: 26584369, pp. 4823–4829. DOI: 10.1021/ct500780r.

[110] Jesse van Rhijn, Claudia Filippi, Stefania De Palo, and Saverio Moroni. "Energy Derivatives in Real-Space Diffusion Monte Carlo." In: *Journal of Chemical Theory and Computation* 18.1 (2022). PMID: 34930005, pp. 118–123. DOI: 10.1021/acs.jctc.1c00496.

[111] Roland Assaraf and Michel Caffarel. "Zero-variance zero-bias principle for observables in quantum Monte Carlo: Application to forces." In: *The Journal of Chemical Physics* 119.20 (2003), pp. 10536–10552. DOI: 10.1063/1.1621615.

[112] C. J. Umrigar. "Two aspects of quantum monte carlo: Determination of accurate wavefunctions and determination of potential energy surfaces of molecules." In: *International Journal of Quantum Chemistry* 36.S23 (1989), pp. 217–230. DOI: 10.1002/qua.560360826.

[113] Claudio Attaccalite and Sandro Sorella. "Stable Liquid Hydrogen at High Pressure by a Novel Ab Initio Molecular-Dynamics Calculation." In: *Phys. Rev. Lett.* 100 (11 2008), p. 114501. DOI: 10.1103/PhysRevLett.100.114501.

[114] Claudia Filippi, Roland Assaraf, and Saverio Moroni. "Simple formalism for efficient derivatives and multi-determinant expansions in quantum Monte Carlo." In: *The Journal of Chemical Physics* 144.19 (2016), p. 194105. DOI: 10.1063/1.4948778.

[115] Kousuke Nakano, Abhishek Raghav, and Sandro Sorella. "Space-warp coordinate transformation for efficient ionic force calculations in quantum Monte Carlo." In: *The Journal of Chemical Physics* 156.3 (2022), p. 034101. DOI: 10.1063/5.0076302.

[116] Juha Tiihonen, III Clay Raymond C., and Jaron T. Krogel. "Toward quantum Monte Carlo forces on heavier ions: Scaling properties." In: *The Journal of Chemical Physics* 154.20 (2021), p. 204111. DOI: 10.1063/5.0052266.

[117] Osvaldo Goscinski. "Antisymmetrized geminal power (AGP) wave-functions in finite and extended systems." In: *International Journal of Quantum Chemistry* 22.S16 (1982), pp. 591–603. DOI: 10.1002/qua.560220851.

[118] Julien Toulouse. *Introduction to the calculation of molecular properties by response theory*. https://hal.science/hal-03934866/document. 2018.

[119] Frank Jensen. *Introduction to computational chemistry*. John wiley & sons, 2017.

[120] M. Chandler Bennett, Cody A. Melton, Abdulgani Annaberdiyev, Guangming Wang, Luke Shulenburger, and Lubos Mitas. "A new generation of effective core potentials for correlated calculations." In: *The Journal of Chemical Physics* 147.22 (2017), p. 224106. DOI: 10.1063/1.4995643.

[121] M. Chandler Bennett, Guangming Wang, Abdulgani Annaberdiyev, Cody A. Melton, Luke Shulenburger, and Lubos Mitas. "A new generation of effective core potentials from correlated calculations: 2nd row elements." In: *The Journal of Chemical Physics* 149.10 (2018), p. 104108. DOI: 10.1063/1.5038135.

[122] Abdulgani Annaberdiyev, Guangming Wang, Cody A. Melton, M. Chandler Bennett, Luke Shulenburger, and Lubos Mitas. "A new generation of effective core potentials from correlated calculations: 3d transition metal series." In: *The Journal of Chemical Physics* 149.13 (2018), p. 134108. DOI: 10.1063/1.5040472.

[123] Guangming Wang, Abdulgani Annaberdiyev, Cody A. Melton, M. Chandler Bennett, Luke Shulenburger, and Lubos Mitas. "A new generation of effective core potentials from correlated calculations: 4s and 4p main group elements and first row additions." In: *The Journal of Chemical Physics* 151.14 (2019), p. 144110. DOI: 10.1063/1.5121006.

[124] Qiming Sun et al. "PySCF: the Python-based simulations of chemistry framework." In: *WIREs Computational Molecular Science* 8.1 (2018), e1340. DOI: 10.1002/wcms.1340.

[125] Qiming Sun et al. "Recent developments in the PySCF program package." In: *The Journal of Chemical Physics* 153.2 (2020), p. 024109. DOI: 10.1063/5.0006074.

[126] Kousuke Nakano, Oto Kohulák, Abhishek Raghav, Michele Casula, and Sandro Sorella. "TurboGenius: Python suite for high-throughput calculations of ab initio quantum Monte Carlo methods." In: *The Journal of Chemical Physics* 159.22 (2023), p. 224801. DOI: 10.1063/5.0179003.

[127] Evgeny Posenitskiy et al. "TREXIO: A file format and library for quantum chemistry." In: *The Journal of Chemical Physics* 158.17 (2023), p. 174801. DOI: 10.1063/5.0148161.

[128] Klaus-Peter Huber. *Molecular spectra and molecular structure: IV. Constants of diatomic molecules*. Springer Science & Business Media, 2013. DOI: 10.1119/1.1932852.

[129] Kousuke Nakano, Tommaso Morresi, Michele Casula, Ryo Maezono, and Sandro Sorella. "Atomic forces by quantum Monte Carlo: Application to phonon dispersion calculations." In: *Phys. Rev. B* 103 (12 2021), p. L121110. DOI: 10.1103/PhysRevB.103.L121110.

[130] Pascal Vinet, John R. Smith, John Ferrante, and James H. Rose. "Temperature effects on the universal equation of state of solids." In: *Phys. Rev. B* 35 (4 1987), pp. 1945–1953. DOI: 10.1103/PhysRevB.35.1945.

[131] C. Pierleoni and D.M. Ceperley. "The Coupled Electron-Ion Monte Carlo Method." In: *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*. Ed. by Mauro Ferrario, Giovanni Ciccotti, and Kurt Binder. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 641–683. DOI: 10.1007/3-540-35273-2_18.

[132] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gabor Csanyi. "MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields." In: *Advances in Neural Information Processing Systems*. Ed. by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh. Vol. 35. Curran Associates, Inc., 2022, pp. 11423–11436. DOI: 10.48550/arXiv.2206.07697.

[133] Shuichi Nosé. "A unified formulation of the constant temperature molecular dynamics methods." In: *The Journal of Chemical Physics* 81.1 (1984), pp. 511–519. DOI: 10.1063/1.447334.

[134] Shūichi Nosé. "A molecular dynamics method for simulations in the canonical ensemble." In: *Molecular Physics* 52.2 (1984), pp. 255–268. DOI: 10.1080/00268978400101201.

[135] William G. Hoover. "Canonical dynamics: Equilibrium phase-space distributions." In: *Phys. Rev. A* 31 (3 1985), pp. 1695–1697. DOI: 10.1103/PhysRevA.31.1695.

[136] Giovanni Bussi, Davide Donadio, and Michele Parrinello. "Canonical sampling through velocity rescaling." In: *The Journal of Chemical Physics* 126.1 (2007), p. 014101. DOI: 10.1063/1.2408420.

[137] Michele Parrinello and Aneesur Rahman. "Polymorphic transitions in single crystals: A new molecular dynamics method." In: *Journal of Applied physics* 52.12 (1981), pp. 7182–7190. DOI: 10.1063/1.328693.

[138] T. Schneider and E. Stoll. "Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions." In: *Phys. Rev. B* 17 (3 1978), pp. 1302–1322. DOI: 10.1103/PhysRevB.17.1302.

[139] SA Adelman and CL Brooks III. "Generalized Langevin models and condensed-phase chemical reaction dynamics." In: *The Journal of Physical Chemistry* 86.9 (1982), pp. 1511–1524. DOI: 10.1021/j100206a011.

[140] Rep Kubo. "The fluctuation-dissipation theorem." In: *Reports on progress in physics* 29.1 (1966), p. 255. DOI: 10.1088/0034-4885/29/1/306.

[141] Hans Hellmann. *Einführung in die quantenchemie*. F. Deuticke, 1937.

[142] R. P. Feynman. "Forces in Molecules." In: *Phys. Rev.* 56 (4 1939), pp. 340–343. DOI: 10.1103/PhysRev.56.340.

[143] Dominik Marx and Jürg Hutter. *Ab initio molecular dynamics: basic theory and advanced methods*. Cambridge University Press, 2009.

[144] R. Car and M. Parrinello. "Unified Approach for Molecular Dynamics and Density-Functional Theory." In: *Phys. Rev. Lett.* 55 (22 1985), pp. 2471–2474. DOI: 10.1103/PhysRevLett.55.2471.

[145] Guglielmo Mazzola, Seiji Yunoki, and Sandro Sorella. "Unexpectedly high pressure for molecular dissociation in liquid hydrogen by electronic simulation." In: *Nature Communications* 5.1 (2014), p. 3487. DOI: 10.1038/ncomms4487.

[146] Andrea Zen, Ye Luo, Guglielmo Mazzola, Leonardo Guidoni, and Sandro Sorella. "Ab initio molecular dynamics simulation of liquid water by quantum Monte Carlo." In: *The Journal of Chemical Physics* 142.14 (2015), p. 144111. DOI: 10.1063/1.4917171.

[147] Michele Ceriotti, Michele Parrinello, Thomas E. Markland, and David E. Manolopoulos. "Efficient stochastic thermostatting of path integral molecular dynamics." In: *The Journal of Chemical Physics* 133.12 (2010), p. 124104. DOI: 10.1063/1.3489925.

[148] Michele Ceriotti, Giovanni Bussi, and Michele Parrinello. "Colored-Noise Thermostats à la Carte." In: *Journal of Chemical Theory and Computation* 6.4 (2010), pp. 1170–1180. DOI: 10.1021/ct900563s.

[149] Michele Ceriotti, David E. Manolopoulos, and Michele Parrinello. "Accelerating the convergence of path integral dynamics with a generalized Langevin equation." In: *The Journal of Chemical Physics* 134.8 (2011), p. 084104. DOI: 10.1063/1.3556661.

[150] Félix Mouhat, Sandro Sorella, Rodolphe Vuilleumier, Antonino Marco Saitta, and Michele Casula. "Fully Quantum Description of the Zundel Ion: Combining Variational Quantum Monte Carlo with Path Integral Langevin Dynamics." In: *J. Chem. Theory Comput.* 13.6 (2017), pp. 2400–2417. DOI: 10.1021/acs.jctc.7b00017.

[151] Chris M Handley and Paul L A Popelier. "Potential Energy Surfaces Fitted by Artificial Neural Networks." In: *The Journal of Physical Chemistry A* 114.10 (2010), pp. 3371–3383. DOI: 10.1021/jp9105585.

[152] Jörg Behler. "Four Generations of High-Dimensional Neural Network Potentials." In: *Chemical Reviews* 121.16 (2021), pp. 10037–10072. DOI: 10.1021/acs.chemrev.0c00868.

[153] Jörg Behler and Michele Parrinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces." In: *Phys. Rev. Lett.* 98 (14 2007), p. 146401. DOI: 10.1103/PhysRevLett.98.146401.

[154] Albert P. Bartók, Mike C. Payne, Risi Kondor, and Gábor Csányi. "Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons." In: *Phys. Rev. Lett.* 104 (13 2010), p. 136403. DOI: 10.1103/PhysRevLett.104.136403.

[155] Albert P. Bartók and Gábor Csányi. "Gaussian approximation potentials: A brief tutorial introduction." In: *International Journal of Quantum Chemistry* 115.16 (2015), pp. 1051–1057. DOI: 10.1002/qua.24927.

[156] Matthias Rupp, Alexandre Tkatchenko, Klaus-Robert Müller, and O. Anatole von Lilienfeld. "Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning." In: *Phys. Rev. Lett.* 108 (5 2012), p. 058301. DOI: 10.1103/PhysRevLett.108.058301.

[157] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. "Neural message passing for Quantum chemistry." In: *Proceedings of the 34th International Conference on Machine Learning - Volume 70*. ICML'17. Sydney, NSW, Australia: JMLR.org, 2017, pp. 1263–1272.

[158] Albert P. Bartók, Risi Kondor, and Gábor Csányi. "On representing chemical environments." In: *Phys. Rev. B* 87 (18 2013), p. 184115. DOI: 10.1103/PhysRevB.87.184115.

[159] Ralf Drautz. "Atomic cluster expansion for accurate and transferable interatomic potentials." In: *Phys. Rev. B* 99 (1 2019), p. 014104. DOI: 10.1103/PhysRevB.99.014104.

[160] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. "Learning representations by back-propagating errors." In: *Nature* 323.6088 (1986), pp. 533–536. DOI: 10.1038/323533a0.

[161] Christoph Schran, Krystof Brezina, and Ondrej Marsalek. "Committee neural network potentials control generalization errors and enable active learning." In: *The Journal of Chemical Physics* 153.10 (2020), p. 104105. DOI: 10.1063/5.0016004.

[162] Raghunathan Ramakrishnan, Pavlo O Dral, Matthias Rupp, and O Anatole von Lilienfeld. "Big Data Meets Quantum Chemistry Approximations: The Δ-Machine Learning Approach." In: *Journal of Chemical Theory and Computation* 11.5 (2015), pp. 2087–2096. DOI: 10.1021/acs.jctc.5b00099.

[163] Sandip De, Albert P Bartók, Gábor Csányi, and Michele Ceriotti. "Comparing molecules and solids across structural and alchemical space." In: *Phys. Chem. Chem. Phys.* 18.20 (2016), pp. 13754–13769. DOI: 10.1039/C6CP00415F.

[164] Luboš Mitáš, Eric L. Shirley, and David M. Ceperley. "Nonlocal pseudopotentials and diffusion Monte Carlo." In: *The Journal of Chemical Physics* 95.5 (1991), pp. 3467–3475. DOI: 10.1063/1.460849.

[165] Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. "Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges." In: *arXiv e-prints*, arXiv:2104.13478 (2021), arXiv:2104.13478. DOI: 10.48550/arXiv.2104.13478.

[166] Geneviève Dusson, Markus Bachmayr, Gábor Csányi, Ralf Drautz, Simon Etter, Cas van der Oord, and Christoph Ortner. "Atomic cluster expansion: Completeness, efficiency and stability." In: *Journal of Computational Physics* 454 (2022), p. 110946. DOI: 10.1016/j.jcp.2022.110946.

[167] M. van Thiel, M. Ross, B. L. Hord, A. C. Mitchell, W. H. Gust, M. J. D'Addario, R. N. Keeler, and K. Boutwell. "Shock-Wave Compression of Liquid Deuterium to 0.9 Mbar." In: *Phys. Rev. Lett.* 31 (16 1973), pp. 979–982. DOI: 10.1103/PhysRevLett.31.979.

[168] W. J. Nellis, A. C. Mitchell, M. van Thiel, G. J. Devine, R. J. Trainor, and N. Brown. "Equation-of-state data for molecular hydrogen and deuterium at shock pressures in the range 2–76 GPa (20–760 kbar)." In: *J. Chem. Phys.* 79.3 (1983), pp. 1480–1486. DOI: 10.1063/1.445938.

[169] N. C. Holmes, M. Ross, and W. J. Nellis. "Temperature measurements and dissociation of shock-compressed liquid deuterium and hydrogen." In: *Phys. Rev. B* 52 (22 1995), pp. 15835–15845. DOI: 10.1103/PhysRevB.52.15835.

[170] L. B. Da Silva et al. "Absolute Equation of State Measurements on Shocked Liquid Deuterium up to 200 GPa (2 Mbar)." In: *Phys. Rev. Lett.* 78 (3 1997), pp. 483–486. DOI: 10.1103/PhysRevLett.78.483.

[171] G. W. Collins, L. B. Da Silva, P. Celliers, D. M. Gold, M. E. Foord, R. J. Wallace, A. Ng, S. V. Weber, K. S. Budil, and R. Cauble. "Measurements of the Equation of State of Deuterium at the Fluid Insulator-Metal Transition." In: *Science* 281.5380 (1998), pp. 1178–1181. DOI: 10.1126/science.281.5380.1178.

[172] G. W. Collins et al. "Equation of state measurements of hydrogen isotopes on Nova." In: *Physics of Plasmas* 5.5 (1998), pp. 1864–1869. DOI: 10.1063/1.872857.

[173] T. R. Boehly et al. "Properties of fluid deuterium under double-shock compression to several Mbar." In: *Physics of Plasmas* 11.9 (2004), pp. L49–L52. DOI: 10.1063/1.1778164.

[174] D. G. Hicks, T. R. Boehly, P. M. Celliers, J. H. Eggert, S. J. Moon, D. D. Meyerhofer, and G. W. Collins. "Laser-driven single shock compression of fluid deuterium from 45 to 220 GPa." In: *Phys. Rev. B* 79 (1 2009), p. 014112. DOI: 10.1103/PhysRevB.79.014112.

[175] T. Sano et al. "Laser-shock compression and Hugoniot measurements of liquid hydrogen to 55 GPa." In: *Phys. Rev. B* 83 (5 2011), p. 054117. DOI: 10.1103/PhysRevB.83.054117.

[176] P. Loubeyre, S. Brygoo, J. Eggert, P. M. Celliers, D. K. Spaulding, J. R. Rygg, T. R. Boehly, G. W. Collins, and R. Jeanloz. "Extended data set for the equation of state of warm dense hydrogen isotopes." In: *Phys. Rev. B* 86 (14 2012), p. 144115. DOI: 10.1103/PhysRevB.86.144115.

[177] A. Fernandez-Pañella et al. "Shock Compression of Liquid Deuterium up to 1 TPa." In: *Phys. Rev. Lett.* 122 (25 2019), p. 255702. DOI: 10.1103/PhysRevLett.122.255702.

[178] Zhiyu He et al. "High-precision equation of state benchmark for cryogenic liquid deuterium at ultrahigh pressure." In: *Phys. Rev. B* 103 (13 2021), p. 134107. DOI: 10.1103/PhysRevB.103.134107.

[179] M. D. Knudson, D. L. Hanson, J. E. Bailey, C. A. Hall, J. R. Asay, and W. W. Anderson. "Equation of State Measurements in Liquid Deuterium to 70 GPa." In: *Phys. Rev. Lett.* 87 (22 2001), p. 225501. DOI: 10.1103/PhysRevLett.87.225501.

[180] M. D. Knudson, D. L. Hanson, J. E. Bailey, C. A. Hall, J. R. Asay, and C. Deeney. "Principal Hugoniot, reverberating wave, and mechanical reshock measurements of liquid deuterium to 400 GPa using plate impact techniques." In: *Phys. Rev. B* 69 (14 2004), p. 144209. DOI: 10.1103/PhysRevB.69.144209.

[181] M. D. Knudson and M. P. Desjarlais. "High-Precision Shock Wave Measurements of Deuterium: Evaluation of Exchange-Correlation Functionals at the Molecular-to-Atomic Transition." In: *Phys. Rev. Lett.* 118.3 (2017), pp. 1–5. DOI: 10.1103/PhysRevLett.118.035501.

[182] R. D. Dick and G. I. Kerley. "Shock compression data for liquids. II. Condensed hydrogen and deuterium." In: *The Journal of Chemical Physics* 73.10 (1980), pp. 5264–5271. DOI: 10.1063/1.439955.

[183] S I Belov et al. "Shock compression of solid deuterium." In: *Journal of Experimental and Theoretical Physics Letters* 76.7 (2002), pp. 433–435. DOI: 10.1134/1.1528696.

[184] GV Boriskov, AI Bykov, RI Il'Kaev, VD Selemir, GV Simakov, RF Trunin, VD Urlin, VE Fortov, and AN Shuĭkin. "Shock-Wave Compression of Solid Deuterium at a Pressure of 120 GPa." In: *Doklady Physics*. Vol. 48. 10. 2003.

[185] S K Grishechkin et al. "Experimental measurements of the compressibility, temperature, and light absorption in dense shock-compressed gaseous deuterium." In: *Journal of Experimental and Theoretical Physics Letters* 80.6 (2004), pp. 398–404. DOI: 10.1134/1.1830656.

[186] G. V. Boriskov, A. I. Bykov, R. I. Il'kaev, V. D. Selemir, G. V. Simakov, R. F. Trunin, V. D. Urlin, A. N. Shuikin, and W. J. Nellis. "Shock compression of liquid deuterium up to 109 GPa." In: *Phys. Rev. B* 71 (9 2005), p. 092104. DOI: 10.1103/PhysRevB.71.092104.

[187] G. E. Duvall and R. A. Graham. "Phase transitions under shock-wave loading." In: *Rev. Mod. Phys.* 49 (3 1977), pp. 523–579. DOI: 10.1103/RevModPhys.49.523.

[188] Thomas J. Lenosky, Joel D. Kress, and Lee A. Collins. "Molecular-dynamics modeling of the Hugoniot of shocked liquid deuterium." In: *Phys. Rev. B* 56 (9 1997), pp. 5164–5169. DOI: 10.1103/PhysRevB.56.5164.

[189] Thomas J. Lenosky, Scott R. Bickham, Joel D. Kress, and Lee A. Collins. "Density-functional calculation of the Hugoniot of shocked liquid deuterium." In: *Phys. Rev. B* 61 (1 2000), pp. 1–4. DOI: 10.1103/PhysRevB.61.1.

[190] Michael P. Desjarlais. "Density-functional calculations of the liquid deuterium Hugoniot, reshock, and reverberation timing." In: *Phys. Rev. B* 68 (6 2003), p. 064204. DOI: 10.1103/PhysRevB.68.064204.

[191] Stanimir A. Bonev, Burkhard Militzer, and Giulia Galli. "Ab initio simulations of dense liquid deuterium: Comparison with gas-gun shock-wave experiments." In: *Phys. Rev. B* 69 (1 2004), p. 014101. DOI: 10.1103/PhysRevB.69.014101.

[192] Bastian Holst, Ronald Redmer, and Michael P. Desjarlais. "Thermophysical properties of warm dense hydrogen using quantum molecular dynamics simulations." In: *Phys. Rev. B* 77 (18 2008), p. 184201. DOI: 10.1103/PhysRevB.77.184201.

[193] L. Caillabet, S. Mazevet, and P. Loubeyre. "Multiphase equation of state of hydrogen from ab initio calculations in the range 0.2 to 5 g/cc up to 10 eV." In: *Phys. Rev. B* 83 (9 2011), p. 094101. DOI: 10.1103/PhysRevB.83.094101.

[194] Cong Wang and Ping Zhang. "Wide range equation of state for fluid hydrogen from density functional theory." In: *Physics of Plasmas* 20.9 (2013), p. 092703. DOI: 10.1063/1.4821839.

[195] V. V. Karasiev, S. X. Hu, M. Zaghoo, and T. R. Boehly. "Exchange-correlation thermal effects in shocked deuterium: Softening the principal Hugoniot and thermophysical properties." In: *Phys. Rev. B* 99.21 (2019), pp. 1–9. DOI: 10.1103/PhysRevB.99.214110.

[196] D. I. Mihaylov, V. V. Karasiev, S. X. Hu, J. R. Rygg, V. N. Goncharov, and G. W. Collins. "Improved first-principles equation-of-state table of deuterium for high-energy-density applications." In: *Phys. Rev. B* 104 (14 2021), p. 144104. DOI: 10.1103/PhysRevB.104.144104.

[197] Yaroslav Lavrinenko, Pavel R. Levashov, Dmitry V. Minakov, Igor V. Morozov, and Ilya A. Valuev. "Equilibrium properties of warm dense deuterium calculated by the wave packet molecular dynamics and density functional theory method." In: *Phys. Rev. E* 104 (4 2021), p. 045304. DOI: 10.1103/PhysRevE.104.045304.

[198] B. Militzer and D. M. Ceperley. "Path Integral Monte Carlo Calculation of the Deuterium Hugoniot." In: *Phys. Rev. Lett.* 85 (9 2000), pp. 1890–1893. DOI: 10.1103/PhysRevLett.85.1890.

[199] Norm M. Tubman, Elisa Liberatore, Carlo Pierleoni, Markus Holzmann, and David M. Ceperley. "Molecular-Atomic Transition along the Deuterium Hugoniot Curve with Coupled Electron-Ion Monte Carlo Simulations." In: *Phys. Rev. Lett.* 115 (4 2015), p. 045301. DOI: 10.1103/PhysRevLett.115.045301.

[200] Michele Ruggeri, Markus Holzmann, David M. Ceperley, and Carlo Pierleoni. "Quantum Monte Carlo determination of the principal Hugoniot of deuterium." In: *Phys. Rev. B* 102.14 (2020), p. 144108. DOI: 10.1103/PhysRevB.102.144108.

[201] W. J. Nellis. "Shock Compression of Deuterium near 100 GPa Pressures." In: *Phys. Rev. Lett.* 89 (16 2002), p. 165502. DOI: 10.1103/PhysRevLett.89.165502.

[202] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials." In: *J. Phys.: Condens.Matter* 21.39 (2009), p. 395502. DOI: 10.1088/0953-8984/21/39/395502.

[203] P Giannozzi et al. "Advanced capabilities for materials modelling with Quantum ESPRESSO." In: *J. Phys.: Condens.Matter* 29.46 (2017), p. 465901. DOI: 10.1088/1361-648x/aa8f79.

[204] Paolo Giannozzi et al. "Quantum ESPRESSO toward the exascale." In: *J. Chem. Phys.* 152.15 (2020), p. 154105. DOI: 10.1063/5.0005082.

[205] H.PBEK-JPAW_PSL.1.0.0.UPF and H.PZ-KJPAW_PSL.1.0.0.UPF pseudopotentials available at http://pseudopotentials.quantum-espresso.org/legacy_tables/ps-library/h.

[206] Andrea Ricci and Giovanni Ciccotti. "Algorithms for Brownian dynamics." In: *Mol. Phys.* 101 (2003), pp. 1927–1931. DOI: 10.1080/0026897031000108113.

[207] M. Keith Matzen et al. "Pulsed-power-driven high energy density physics and inertial confinement fusion researcha)." In: *Physics of Plasmas* 12.5 (2005), p. 055503. DOI: 10.1063/1.1891746.

[208] Valentin V. Karasiev, D. I. Mihaylov, and S. X. Hu. "Meta-GGA exchange-correlation free energy density functional to increase the accuracy of warm dense matter simulations." In: *Phys. Rev. B* 105 (8 2022), p. L081109. DOI: 10.1103/PhysRevB.105.L081109.

[209] Chiheb Ben Mahmoud, Federico Grasselli, and Michele Ceriotti. "Predicting hot-electron free energies from ground-state data." In: *Phys. Rev. B* 106 (12 2022), p. L121116. DOI: 10.1103/PhysRevB.106.L121116.

[210] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. "Van der Waals Density Functional for General Geometries." In: *Phys. Rev. Lett.* 92 (24 2004), p. 246401. DOI: 10.1103/PhysRevLett.92.246401.

[211] Kristian Berland, Valentino R Cooper, Kyuho Lee, Elsebeth Schröder, T Thonhauser, Per Hyldgaard, and Bengt I Lundqvist. "van der Waals forces in density functional theory: a review of the vdW-DF method." In: *Rep. Prog. Phys.* 78.6 (2015), p. 066501. DOI: 10.1088/0034-4885/78/6/066501.

[212] Raymond C. Clay, Michael P. Desjarlais, and Luke Shulenburger. "Deuterium Hugoniot: Pitfalls of thermodynamic sampling beyond density functional theory." In: *Phys. Rev. B* 100.7 (2019), p. 75103. DOI: 10.1103/PhysRevB.100.075103.

[213] Chengteh Lee, Weitao Yang, and Robert G. Parr. "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density." In: *Phys. Rev. B* 37 (2 1988), pp. 785–789. DOI: 10.1103/PhysRevB.37.785.

[214] A. Baldereschi. "Mean-Value Point in the Brillouin Zone." In: *Phys. Rev. B* 7 (12 1973), pp. 5212–5215. DOI: 10.1103/PhysRevB.7.5212.

[215] Shivesh Pathak and Lucas K Wagner. "A light weight regularization for wave function parameter gradients in quantum Monte Carlo." In: *AIP Adv.* 10.8 (2020), p. 085213. DOI: 10.1063/5.0004008.

[216] Hendra Kwee, Shiwei Zhang, and Henry Krakauer. "Finite-Size Correction in Many-Body Electronic Structure Calculations." In: *Phys. Rev. Lett.* 100 (12 2008), p. 126404. DOI: 10.1103/PhysRevLett.100.126404.

[217] Stephen R Xie, Matthias Rupp, and Richard G Hennig. "Ultra-fast interpretable machine-learning potentials." In: *npj Computational Materials* 9.1 (2023), p. 162. DOI: 10.1038/s41524-023-01092-7.

[218] S. T. Weir, A. C. Mitchell, and W. J. Nellis. "Metallization of Fluid Molecular Hydrogen at 140 GPa (1.4 Mbar)." In: *Phys. Rev. Lett.* 76 (11 1996), pp. 1860–1863. DOI: 10.1103/PhysRevLett.76.1860.

[219] V. E. Fortov et al. "Phase Transition in a Strongly Nonideal Deuterium Plasma Generated by Quasi-Isentropical Compression at Megabar Pressures." In: *Phys. Rev. Lett.* 99 (18 2007), p. 185001. DOI: 10.1103/PhysRevLett.99.185001.

[220] Vasily Dzyabura, Mohamed Zaghoo, and Isaac F. Silvera. "Evidence of a liquid–liquid phase transition in hot dense hydrogen." In: *Proceedings of the National Academy of Sciences* 110.20 (2013), pp. 8040–8044. DOI: 10.1073/pnas.1300718110.

[221] Kenji Ohta, Kota Ichimaru, Mari Einaga, Sho Kawaguchi, Katsuya Shimizu, Takahiro Matsuoka, Naohisa Hirao, and Yasuo Ohishi. "Phase boundary of hot dense fluid hydrogen." In: *Scientific Reports* 5.1 (2015), p. 16560. DOI: 10.1038/srep16560.

[222] M. D. Knudson, M. P. Desjarlais, A. Becker, R. W. Lemke, K. R. Cochrane, M. E. Savage, D. E. Bliss, T. R. Mattsson, and R. Redmer. "Direct observation of an abrupt insulator-to-metal transition in dense liquid deuterium." In: *Science* 348.6242 (2015), pp. 1455–1460. DOI: 10.1126/science.aaa7471.

[223] R. Stewart McWilliams, D. Allen Dalton, Mohammad F. Mahmood, and Alexander F. Goncharov. "Optical Properties of Fluid Hydrogen at the Transition to a Conducting State." In: *Phys. Rev. Lett.* 116 (25 2016), p. 255501. DOI: 10.1103/PhysRevLett.116.255501.

[224] Mohamed Zaghoo, Ashkan Salamat, and Isaac F. Silvera. "Evidence of a first-order phase transition to metallic hydrogen." In: *Phys. Rev. B* 93 (15 2016), p. 155128. DOI: 10.1103/PhysRevB.93.155128.

[225] Mohamed Zaghoo and Isaac F. Silvera. "Conductivity and dissociation in liquid metallic hydrogen and implications for planetary interiors." In: *Proceedings of the National Academy of Sciences* 114.45 (2017), pp. 11873–11877. DOI: 10.1073/pnas.1707918114.

[226] Mohamed Zaghoo, Rachel J. Husband, and Isaac F. Silvera. "Striking isotope effect on the metallization phase lines of liquid hydrogen and deuterium." In: *Phys. Rev. B* 98 (10 2018), p. 104102. DOI: 10.1103/PhysRevB.98.104102.

[227] Peter M. Celliers et al. "Insulator-metal transition in dense fluid deuterium." In: *Science* 361.6403 (2018), pp. 677–682. DOI: 10.1126/science.aat0970.

[228] Shuqing Jiang, Nicholas Holtgrewe, Zachary M. Geballe, Sergey S. Lobanov, Mohammad F. Mahmood, R. Stewart McWilliams, and Alexander F. Goncharov. "A Spectroscopic Study of the Insulator–Metal Transition in Liquid Hydrogen and Deuterium." In: *Advanced Science* 7.2 (2020), p. 1901668. DOI: 10.1002/advs.201901668.

[229] Sandro Scandolo. "Liquid–liquid phase transition in compressed hydrogen from first-principles simulations." In: *Proceedings of the National Academy of Sciences* 100.6 (2003), pp. 3051–3053. DOI: 10.1073/pnas.0038012100.

[230] Kris T. Delaney, Carlo Pierleoni, and D. M. Ceperley. "Quantum Monte Carlo Simulation of the High-Pressure Molecular-Atomic Crossover in Fluid Hydrogen." In: *Phys. Rev. Lett.* 97 (23 2006), p. 235702. DOI: 10.1103/PhysRevLett.97.235702.

[231] J. Vorberger, I. Tamblyn, B. Militzer, and S. A. Bonev. "Hydrogen-helium mixtures in the interiors of giant planets." In: *Phys. Rev. B* 75 (2 2007), p. 024206. DOI: 10.1103/PhysRevB.75.024206.

[232] Winfried Lorenzen, Bastian Holst, and Ronald Redmer. "First-order liquid-liquid phase transition in dense hydrogen." In: *Phys. Rev. B* 82 (19 2010), p. 195107. DOI: 10.1103/PhysRevB.82.195107.

[233] Isaac Tamblyn and Stanimir A. Bonev. "Structure and Phase Boundaries of Compressed Liquid Hydrogen." In: *Phys. Rev. Lett.* 104 (6 2010), p. 065702. DOI: 10.1103/PhysRevLett.104.065702.

[234] J. Yang, Ch. L. Tian, F. Sh. Liu, L. C. Cai, H. K. Yuan, M. M. Zhong, and F. Xiao. "A new evidence of first-order phase transition for hydrogen at 3000 K." In: *Europhysics Letters* 109.3 (2015), p. 36003. DOI: 10.1209/0295-5075/109/36003.

[235] G E Norman and I M Saitov. "Critical point and mechanism of the fluid–fluid phase transition in warm dense hydrogen." In: *Doklady Physics* 62.6 (2017), pp. 294–298. DOI: 10.1134/S1028335817060088.

[236] Guglielmo Mazzola, Ravit Helled, and Sandro Sorella. "Phase Diagram of Hydrogen and a Hydrogen-Helium Mixture at Planetary Conditions by Quantum Monte Carlo Simulations." In: *Phys. Rev. Lett.* 120 (2 2018), p. 025701. DOI: 10.1103/PhysRevLett.120.025701.

[237] Chunling Tian, Fusheng Liu, Hongkuan Yuan, Hong Chen, and Anlong Kuan. "First-order liquid-liquid phase transition in compressed hydrogen and critical point." In: *The Journal of Chemical Physics* 150.20 (2019), p. 204114. DOI: 10.1063/1.5096400.

[238] Hua Y. Geng, Q. Wu, Miriam Marqués, and Graeme J. Ackland. "Thermodynamic anomalies and three distinct liquid-liquid transitions in warm dense liquid hydrogen." In: *Phys. Rev. B* 100 (13 2019), p. 134109. DOI: 10.1103/PhysRevB.100.134109.

[239] Giovanni Rillo, Miguel A. Morales, David M. Ceperley, and Carlo Pierleoni. "Optical properties of high-pressure fluid hydrogen across molecular dissociation." In: *Proceedings of the National Academy of Sciences* 116.20 (2019), pp. 9770–9774. DOI: 10.1073/pnas.1818897116.

[240] Joshua Hinz, Valentin V. Karasiev, S. X. Hu, Mohamed Zaghoo, Daniel Mejía-Rodríguez, S. B. Trickey, and L. Calderín. "Fully consistent density functional theory determination of the insulator-metal transition boundary in warm dense hydrogen." In: *Phys. Rev. Res.* 2 (3 2020), p. 032065. DOI: 10.1103/PhysRevResearch.2.032065.

[241] Taras Bryk, Carlo Pierleoni, Giancarlo Ruocco, and Ari Paavo Seitsonen. "Characterization of molecular-atomic transformation in fluid hydrogen under pressure via long-wavelength asymptote of charge density fluctuations." In: *Journal of Molecular Liquids* 312 (2020), p. 113274. DOI: 10.1016/j.molliq.2020.113274.

[242] Armin Bergermann, Lucas Kleindienst, and Ronald Redmer. "Nonmetal-to-metal transition in liquid hydrogen using density functional theory and the Heyd–Scuseria–Ernzerhof exchange-correlation functional." In: *The Journal of Chemical Physics* 161.23 (2024), p. 234303. DOI: 10.1063/5.0241111.

[243] Ravit Helled, Guglielmo Mazzola, and Ronald Redmer. "Understanding dense hydrogen at planetary conditions." In: *Nature Reviews Physics* 2.10 (2020), pp. 562–574. DOI: 10.1038/s42254-020-0223-3.

[244] J E P Connerney et al. "The Juno Magnetic Field Investigation." In: *Space Science Reviews* 213.1 (2017), pp. 39–138. DOI: 10.1007/s11214-017-0334-z.

[245] Miguel A. Morales, Raymond Clay, Carlo Pierleoni, and David M. Ceperley. "First Principles Methods: A Perspective from Quantum Monte Carlo." In: *Entropy* 16.1 (2014), pp. 287–321. DOI: 10.3390/e16010287.

[246] Thomas Bischoff, Bastian Jäckl, and Matthias Rupp. "Hydrogen under Pressure as a Benchmark for Machine-Learning Interatomic Potentials." In: *arXiv e-prints*, arXiv:2409.13390 (2024), arXiv:2409.13390. DOI: 10.48550/arXiv.2409.13390.

[247] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs. "Inhomogeneous backflow transformations in quantum Monte Carlo calculations." In: *Phys. Rev. E* 74 (6 2006), p. 066701. DOI: 10.1103/PhysRevE.74.066701.

[248] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. "Hybrid functionals based on a screened Coulomb potential." In: *The Journal of Chemical Physics* 118.18 (2003), pp. 8207–8215. DOI: 10.1063/1.1564060.

[249] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization." In: *arXiv e-prints*, arXiv:1412.6980 (2014), arXiv:1412.6980. DOI: 10.48550/arXiv.1412.6980.

[250] Aidan P. Thompson et al. "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales." In: *Computer Physics Communications* 271 (2022), p. 108171. DOI: 10.1016/j.cpc.2021.108171.

[251] Hajime Tanaka. "Liquid–liquid transition and polyamorphism." In: *The Journal of Chemical Physics* 153.13 (2020), p. 130901. DOI: 10.1063/5.0021045.

[252] Koichi Momma and Fujio Izumi. "*VESTA*: a three-dimensional visualization system for electronic and structural analysis." In: *Journal of Applied Crystallography* 41.3 (2008), pp. 653–658. DOI: 10.1107/S0021889808012016.

[253] K Binder. "Theory of first-order phase transitions." In: *Reports on Progress in Physics* 50.7 (1987), p. 783. DOI: 10.1088/0034-4885/50/7/001.

[254] Ankit Rohatgi. *WebPlotDigitizer*. Version 5.2.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The majority of the plots reported in this thesis have been generated using the Python package matplotlib. The literature data were digitalized using WEBPLOTDIGITIZER [254].

Final Version as of January 17, 2025 (`classicthesis`).