

14 High-Pressure Phases of Hydrogen

Markus Holzmann

Univ. Grenoble Alpes, CNRS

LPMMC, 38000 Grenoble, France

Contents

1	Introduction	2
2	Computational methods – overview	4
2.1	Path-integral Monte Carlo calculations	4
2.2	Adiabatic approximation	5
2.3	Quantum Monte Carlo calculations at zero temperature	7
2.4	Finite size effects	14
2.5	Monte Carlo calculations with noisy action	16
3	Coupled electron-ion Monte Carlo calculations	18
4	Hydrogen under pressure: Some CEIMC results	20
4.1	Electronic band gaps	20
4.2	Approaching metallic solid hydrogen	22
4.3	Liquid-liquid phase transition	23
4.4	Hugoniot adiabatic	24

1 Introduction

The single hydrogen atom has a special significance in quantum mechanics, as exact analytical solutions to the Schrödinger equation can be presented already in elementary introductory courses. Small, but important deviations between theory and experiment have also played a central role in the development of quantum field theory and quantum electrodynamics. However, under normal conditions on earth, atomic hydrogen, H, is unstable forming molecules, H₂. The theoretical description of molecular hydrogen is immediately getting more challenging than the atom, as the quantization of the nuclear motion must be taken into account. Separation of electronic and nuclear degrees of freedom can be achieved within the adiabatic or Born-Oppenheimer approximation [1], based on the high mass ratio $m_p/m_e \approx 1836$ between protons and electrons of mass m_p and m_e , respectively. Vibrational excitation energies of the isolated H₂ molecule are around ~ 6000 K, whereas rotational energies are much smaller ~ 200 K.

Since the nuclei of the molecule are identical, the nuclear spin has to be taken into account as well, leading to two different molecular species. Para (ortho) H₂ is characterized by even (odd) rotational quantum numbers. A distinction between these molecular species becomes important around and below room temperature, whereas nuclear quantum effects may extend up to several thousands of Kelvin.

At zero pressure and low temperatures, hydrogen is in the solid state forming a molecular crystal with almost freely rotating molecules, translationally localized around crystalline lattice sites but with large zero point motion [2]. Up to at least six different solid phases are experimentally known [3]. A sketch of the phase diagram for temperatures and pressures reached by static diamond anvil cell (DAC) experiments is shown in Fig. 1. Most of the transition lines are inferred from changes in infrared or Raman spectra whereas direct information on the crystal structures is almost entirely missing. Only in phase I, measurements of the Bragg peaks by X-ray spectroscopy have recently confirmed m-HCP at 300 K up to 250 GPa [4, 5].

Increasing pressure an insulator to metal (IM) transition is expected. Originally predicted around 25 GPa [6], the search for metallic hydrogen has been one of the main driving forces in high-pressure physics. Evidence for a semi-metallic state at 350 GPa below 200 K has been reported [7], a discontinuous change of the direct gap observed near 425 GPa [8] indicates the transition to a "good" metal. Earlier reports of metallic hydrogen at 495 GPa [9] has received strong criticism [10, 11]. Whether the IM transition occurs within the molecular solid or whether it coincides together with a molecular dissociation transition to atomic hydrogen remains still an open question.

Disappearance of lattice modes of Raman spectroscopy signalling melting of the crystalline state [12] with a maximum around 1000 K showing a reentrant behavior of the liquid phase at higher pressure. This experimental line coincides astonishingly well with theoretical predictions based on Born-Oppenheimer molecular dynamics (BOMD) calculations with DFT-PBE electronic energies using classical nuclei [13]. However, more recent calculations for quantum nuclei using Path-Integral molecular dynamics with a machine learned force-field trained on QMC data indicate melting at considerable higher temperatures [14].

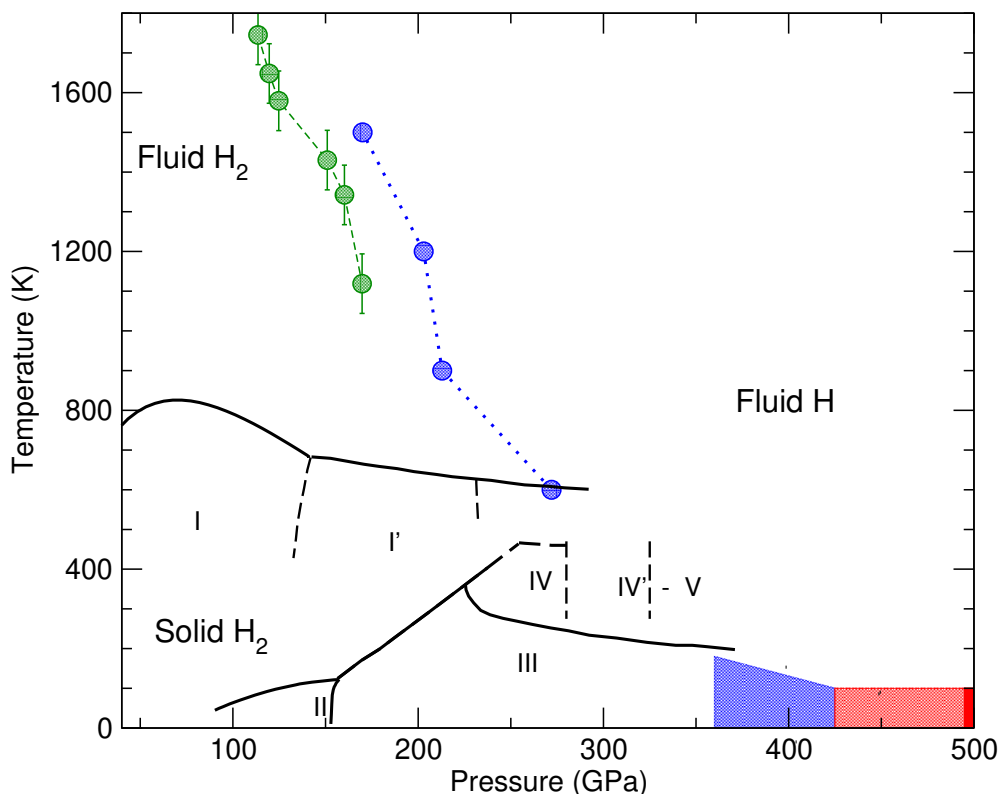


Fig. 1: Phase diagram of solid hydrogen as inferred from static compression experiments (solid and dashed black lines). The solid black line shows the melting of the crystal as identified by the disappearance of Raman active lattice modes [12]. The blue area indicates a semimetallic state [7], with the closure of the direct gap [8] signalling the transition to metallic solid hydrogen (red area). A reflective sample has been reported at 495 GPa [9–11]. Green circles report signatures of a liquid-liquid phase transition from diamond anvil cell (DAC) experiments [15], blue circles are CEIMC predictions [18] from the molecular (insulating) to the atomic (metallic) liquid (Figure courtesy of C. Pierleoni).

Above melting, the fluid may either be in the molecular or atomic state with a possible first order phase transition and a critical point at higher temperatures. Static compression experiments using DAC [15] and dynamic shock wave experiments [16, 17] report evidence for this liquid-liquid phase transition, supported by Coupled-Electron-Ion Monte Carlo (CEIMC) calculations [18–20] and molecular dynamics calculations based on electronic variational Monte Carlo energies [21, 22].

A proper discussion of the phase diagram of hydrogen can be found in Refs [2, 3, 23] and references therein. The above discussion mainly serves to illustrate some of the main difficulties and challenges theoretical and computational physics has to face when studying hydrogen. In the following, I will give an overview over various Monte Carlo methods and the approximations underlying CEIMC calculations, before briefly discussing some of the results obtained with.

2 Computational methods – overview

In the region of phase diagram discussed above, the properties of hydrogen are up to very high accuracy described within the following non-relativistic Hamiltonian

$$H = T + V \quad (1)$$

$$T = T_N + T_e \quad \text{with} \quad T_N = -\frac{\hbar^2}{2m_p} \sum_I \nabla_I^2 \quad \text{and} \quad T_e = -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 \quad (2)$$

$$V(\mathbf{r}, \mathbf{R}) = \sum_{i<j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{I<J} \frac{e^2}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{i,I} \frac{e^2}{|\mathbf{r}_i - \mathbf{R}_I|} \quad (3)$$

where \mathbf{r} and \mathbf{R} denotes the set of electronic and nuclear positions, \mathbf{r}_i (\mathbf{R}_I) labels the individual coordinate of electron i (nucleon I) and summations over i (I) extend over all N_e electrons (N_n nuclei). Electrons and nucleons interact via the Coulomb interactions between themselves and each other, e is the electron charge. The physical problem is then set up by the Schrödinger equation for the many-body wave function, imposing appropriate boundary conditions and symmetries due to particle statistics.

We are interested in the description of a macroscopic system characterized by the electronic density $n = N_e/V$, equal to the atomic number density due to charge neutrality, and temperature T . Density is frequently parametrized by the Wigner-Seitz parameter $r_s = a/a_B$ where $n = 3/(4\pi a^3)$ and a_B is the Bohr radius. At zero pressure and 4.2 K, the density corresponds to $r_s \simeq 1.768$ for p-H₂. Periodic boundary conditions on the density distribution in the simulation cell are used to eliminate surface effects, and the Coulomb $1/r$ interaction in Eq. (3) has to be replaced by the appropriate expressions, e.g., using Ewald's expressions [24–26].

Despite the simplicity of the Hamiltonian, Eq. (1), the extremely rich phase diagram indicates to us that a quantitative description/prediction will be challenging. In fact, straightforward approaches to directly determine hydrogen properties in a large part of the phase diagram are beyond our actual capacities, and we will introduce several approximations. These approximations are in general uncontrolled, in the sense that we will not be able to improve them systematically until a given precision is reached. However, our aim should be to estimate and quantify the systematic uncertainty stemming from the various approximation, if necessary. Experimental data should not be used to validate these approximations, rather, deviations with experiment should trigger reexaminations, as well as questioning the underlying description and modelling of experimental setups and their interpretation.

2.1 Path-integral Monte Carlo calculations

Path-Integral Monte Carlo (PIMC) methods give access to sample the elements of the thermal density matrix

$$\rho(\mathbf{R}'\mathbf{r}', \mathbf{R}\mathbf{r}; \beta) = \langle \mathbf{R}'\mathbf{r}' | e^{-\beta H} | \mathbf{R}\mathbf{r} \rangle \quad (4)$$

from which we can calculate static expectation values of any operator at inverse temperature $\beta = 1/k_B T$.

However, since functions of operators are in general defined via their spectral representation, a direct evaluation of Eq. (4) requires knowledge of all energy eigenfunctions. Based on the Trotter formula [27]

$$e^{-\beta H} = \lim_{M \rightarrow \infty} (e^{-\tau T} e^{-\tau V})^M, \quad \tau = \beta/M \quad (5)$$

the path integral formulation [28] circumvents explicit diagonalizations and the density matrix can be written as

$$\rho(\mathbf{R}'\mathbf{r}', \mathbf{R}\mathbf{r}; \beta) = \lim_{M \rightarrow \infty} \sum_{\mathcal{P}} (\pm 1)^{|\mathcal{P}|} \int_{\mathbf{R}\mathbf{r}}^{\mathcal{P}(\mathbf{R}'\mathbf{r}')} D\mathbf{R}[\tau] D\mathbf{r}[\tau] \left(\frac{\sqrt{m_e m_p}}{2\pi \hbar^2 \beta/M} \right)^{3M} e^{-S_M^p(\mathbf{R}[\tau]\mathbf{r}[\tau]; \beta)} \quad (6)$$

with $\int D\mathbf{R}[\tau] D\mathbf{r}[\tau] = \prod_{m=1}^{M-1} \int d\mathbf{R}^m \int d\mathbf{r}^m$ and

$$S_M^p(\mathbf{R}[\tau]\mathbf{r}[\tau]; \beta) = \sum_{m=0}^{M-1} \left(\frac{m_p (\mathbf{R}^{m+1} - \mathbf{R}^m)^2}{2\hbar^2 (\beta/M)} + \frac{m_e (\mathbf{r}^{m+1} - \mathbf{r}^m)^2}{2\hbar^2 (\beta/M)} + (\beta/M) V(\mathbf{r}^m, \mathbf{R}^m) \right) \quad (7)$$

setting $(\mathbf{R}^0, \mathbf{r}^0) = (\mathbf{R}, \mathbf{r})$ and $(\mathbf{R}^M, \mathbf{r}^M) = \mathcal{P}(\mathbf{R}', \mathbf{r}')$ and the summation extends over all possible permutations \mathcal{P} for correct (anti-)symmetrization. Integrals involved in the discretized path integral as well as the summation over permutations can then be sampled by Monte Carlo methods [29, 30].

The negative sign occurring for odd permutations of fermions will strongly affect the signal to noise ratio of direct Monte Carlo evaluations [23]. Such quantum statistical effects will be important below the Fermi temperature. For an ideal gas of electrons, we have $T_F^e \simeq 581454 r_s^{-2}$ Kelvin, such that $T/T_F^e \lesssim 1\%$ in the our region of interest, Fig. 1. The Fermi temperature of free protons, T_F^p , is considerably lower, $T_F^p = (m_e/m_p)T_F^e \simeq T_F^e/1836$, so that inter-molecular exchange effects are expected to be negligible for establishing the phase diagram in the central region of Fig. 1. However, proper intra-molecular exchanges for para and ortho H_2 may need to be taken into account, certainly at the lower temperatures around and below the rotational energy excitations ~ 200 K.

2.2 Adiabatic approximation

The path-integral expression above, Eq. (6) with Eq. (7), is based on the so-called primitive approximation of the short time (or high temperature) density matrix

$$\rho(\mathbf{R}'\mathbf{r}', \mathbf{R}\mathbf{r}; \tau) \approx \langle \mathbf{R}'\mathbf{r}' | e^{-\tau T} e^{-\tau V} | \mathbf{R}\mathbf{r} \rangle \sim e^{-S_1^p(\mathbf{R}'\mathbf{r}', \mathbf{R}\mathbf{r}; \tau)} \quad (8)$$

choosing τ sufficiently small, such that residual effects due to the commutator $[T, V]$ can be neglected. Different short-time approximations can be chosen which may allow us to reach the same precision within larger time steps, τ , so that less discretizations M of the path are needed for our computations [29].

From physical considerations we expect that the Born-Oppenheimer approximation should provide an excellent description in our region of interest. We therefore choose a different short time approximation, $\rho(\mathbf{R}'\mathbf{r}', \mathbf{R}\mathbf{r}; \tau) \approx \rho_{BO}(\mathbf{R}'\mathbf{r}', \mathbf{R}\mathbf{r}; \tau)$ with

$$\rho_{BO}(\mathbf{R}'\mathbf{r}', \mathbf{R}\mathbf{r}; \tau) = \langle \mathbf{R}'\mathbf{r}' | e^{-\tau T_N} e^{-\tau H_e} | \mathbf{R}\mathbf{r} \rangle = \langle \mathbf{R}'\mathbf{r}' | e^{-\tau T_N} | \mathbf{R}\mathbf{r}' \rangle \langle \mathbf{R}\mathbf{r}' | e^{-\tau(T_e+V)} | \mathbf{R}\mathbf{r} \rangle \quad (9)$$

where matrix elements of $e^{-\tau T_N}$ are independent of electronic degrees of freedom and those of $e^{-\tau(T_e+V)}$ are now diagonal in the nuclear coordinates. Denoting $\Psi_n(\mathbf{r}|\mathbf{R})$ the eigenfunctions of energy $E_n(\mathbf{R})$ of the Born-Oppenheimer Hamiltonian $H_e = T_e + V$, the dependence on \mathbf{R} enters only parametrically, we get

$$\rho_{BO}(\mathbf{R}'\mathbf{r}', \mathbf{R}\mathbf{r}; \tau) \sim \left(\frac{m_p}{2\pi\hbar^2\tau} \right)^{3/2} \sum_n e^{-m(\mathbf{R}-\mathbf{R}')^2/2\hbar^2\tau - \tau E_n(\mathbf{R})} \Psi_n^*(\mathbf{r}'|\mathbf{R}) \Psi_n(\mathbf{r}|\mathbf{R}). \quad (10)$$

In the resulting discretized path integral, we can then integrate out the electronic degrees of freedom (setting $\mathbf{r}^M = \mathbf{r}^0$ and integrating over \mathbf{r}^0)

$$\begin{aligned} \rho(\mathbf{R}', \mathbf{R}; \beta) &= \lim_{M \rightarrow \infty} \sum_{\mathcal{P}} (\pm 1)^{|\mathcal{P}|} \int_{\mathbf{R}}^{\mathcal{P}(\mathbf{R}')} D\mathbf{R}[\tau] D\mathbf{r}[\tau] \prod_m \rho_{BO}(\mathbf{R}^m \mathbf{r}^m, \mathbf{R}^{m+1} \mathbf{r}^{m+1}; \tau) \\ &= \lim_{M \rightarrow \infty} \sum_{\mathcal{P}} (\pm 1)^{|\mathcal{P}|} \int_{\mathbf{R}}^{\mathcal{P}(\mathbf{R}')} D\mathbf{R}[\tau] \left(\frac{m_p}{2\pi\hbar^2\beta/M} \right)^{3M/2} \sum_n e^{-S_{Mn}^{BO}(\mathbf{R}[\tau]; \beta)} \end{aligned} \quad (11)$$

with

$$S_{Mn}^{BO}(\mathbf{R}[\tau]; \beta) = \sum_{m=0}^{M-1} \left(i(\mathbf{R}^{m+1} - \mathbf{R}^m) \cdot \mathbf{A}_n(\mathbf{R}^{m+1}, \mathbf{R}^m) + \frac{m_p(\mathbf{R}^{m+1} - \mathbf{R}^m)^2}{2\hbar^2(\beta/M)} + (\beta/M) E_n(\mathbf{R}^m) \right) \quad (12)$$

where the (real-valued) vector potential \mathbf{A}_n is defined by the electronic overlap integrals

$$\int d\mathbf{r} \Psi_n^*(\mathbf{r}|\mathbf{R}') \Psi_n(\mathbf{r}|\mathbf{R}) \simeq e^{-i(\mathbf{R}' - \mathbf{R}) \cdot \mathbf{A}_n(\mathbf{R}', \mathbf{R})} \quad (13)$$

for $\mathbf{R}' \rightarrow \mathbf{R}$ and we have neglected non-adiabatic transitions between different electronic states. In general, we should have kept terms of order $(\mathbf{R}' - \mathbf{R})^2 \sim \hbar^2\beta/Mm_p$ to recover the exact path-integral in the limit $M \rightarrow \infty$. Neglecting these terms which are suppressed by $1/m_p$ corresponds to the Born-Oppenheimer approximation.

In the following, we will focus on the diagonal part of the density matrix, $\rho(\mathbf{R}, \mathbf{R}; \beta)$ which determines most of the basic thermodynamic observables. Further, we will assume sufficiently high temperatures to neglect proton exchanges, and all proton paths are closed. Then, the phase term in Eq. (12) only contributes when a Berry phase is acquired during the trajectory in imaginary time. In general, this requires exceptionally high symmetry configurations, so that this adiabatic phase is frequently dropped in the Born-Oppenheimer sampling of solids and liquids.

For $T \ll T_F^e$, we may further neglect any contributions from electronic excitations, $n > 0$, and restrict to configurations within the ground state Born-Oppenheimer surface $E_0(\mathbf{R})$. However, we still have to resolve the electronic ground state problem to obtain $E_0(\mathbf{R})$ for given nuclear positions.

Ab-initio or first-principles molecular dynamics methods [31] are frequently based on electronic structure calculations using density functional theory (DFT). How to decide which DFT functional should be used? Frequently, comparisons with experiments gauges the choice of DFT functional. Since we want to avoid this criterium, we need some other strategies to calculate the Born-Oppenheimer electronic energies.

2.3 Quantum Monte Carlo calculations at zero temperature

2.3.1 Variational principle

Quantum Monte Carlo methods at zero temperature are based on the variational principle

$$E_0 \leq E_T \equiv \frac{\langle \Psi_T | H_e | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \quad (14)$$

where Ψ_T is any trial wave function obeying the same boundary conditions as our true (Born-Oppenheimer) ground state wave function Ψ_0 . Here, and in the following we suppress the parametric dependence on the nuclear coordinates.

The simplest anti-symmetric trial wave function for N electrons is $\det_n \varphi_n(\mathbf{r}_i)$, a Slater determinant, for spin-polarized electrons, whereas it reduces to a product of two determinants, one for each spin-component, for a gas of unpolarized electrons. Optimizing the variational energy obtained by a Slater determinant with respect to the orbitals, $\varphi_n(\mathbf{r})$ gives the Hartree-Fock energy. To go beyond the Hartree-Fock approximation, we can add explicit pair correlations [32–35] to our wave function

$$\Psi_{SJ}(\mathbf{r}) = \det_{ni} \varphi_n(\mathbf{r}_i) e^{-U(\mathbf{r})}. \quad (15)$$

The frequently called Jastrow factor $U(\mathbf{r})$ is symmetric under electron exchanges, e.g. $U = \sum_{i < j} u_{ee}(|\mathbf{r}_i - \mathbf{r}_j|) + \sum_{i,J} u_{ep}(|\mathbf{r}_i - \mathbf{R}_J|)$ provides an explicit, size consistent and translational invariant form containing electron-electron and electron-proton correlations. In addition to the orbitals $\varphi_n(\mathbf{r})$, the one dimensional functions $u_{ee}(r)$ and $u_{ep}(r)$ need also to be determined by minimizing the resulting energy expectation value.

Backflow wave functions introduce an explicit dependence on all electron coordinates \mathbf{r} into the Slater determinant [36–38]

$$\Psi_{bf}(\mathbf{r}) = \det_{ni} \varphi_n(\mathbf{q}_i(\mathbf{r})) e^{-U(\mathbf{r})} \quad (16)$$

by the use of backflow coordinates $\mathbf{q}_i(\mathbf{r})$ which need to be symmetric with respect to electron exchange holding i fixed. A simple form for liquid and solid hydrogen [39] is

$$\mathbf{q}_i = \mathbf{r}_i + \sum_{j \neq i} (\mathbf{r}_i - \mathbf{r}_j) b_{ee}(|\mathbf{r}_i - \mathbf{r}_j|) + \sum_J (\mathbf{r}_i - \mathbf{R}_J) b_{ep}(|\mathbf{r}_i - \mathbf{R}_J|)$$

with one dimensional functions $b_{ee}(r)$ and $b_{ep}(r)$, ultimately determined by energy minimization. Systematic improvement for many-body correlations either explicitly [40] or implicitly via an iterative (deep) structure [41–44] are possible, as well as combined with neural network representations [45–50]. In addition to providing lower energy expectation values, better trial wave functions also lower the variance of the energy,

$$\sigma_T^2 = \frac{\langle \Psi_T | H^2 | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} - \left(\frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \right)^2 \quad (17)$$

which vanishes in case Ψ_T coincides with an exact energy eigenstate. Calculations with different classes of trial wave functions can be used for heuristic extrapolations of E_T to zero variance, thus providing an estimate for the residual energy error of the best trial wave function [41–43].

2.3.2 Variational Monte Carlo calculations

All of the wave functions described before have in common that the evaluation for a given configuration \mathbf{r} is sufficiently fast for systems ranging from around ten to thousands of electrons. However, we still have to calculate and optimize E_T . Whereas the Hartree-Fock energy can still be calculated by deterministic quadrature, the integrals involved in the calculations based on correlated wave functions, e.g., the Slater-Jastrow and backflow forms, involve integrations over $3N$ dimensions

$$E_T = \frac{\int d\mathbf{r} |\Psi_T(\mathbf{r})|^2 E_L(\mathbf{r})}{\int d\mathbf{r} |\Psi_T(\mathbf{r})|^2} \quad (18)$$

where we have introduced the so-called local energy

$$E_L(\mathbf{r}) \equiv \frac{\langle \mathbf{r} | H | \Psi_T \rangle}{\langle \mathbf{r} | \Psi_T \rangle} = \frac{H \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} \quad (19)$$

in the position representation $\Psi_T(\mathbf{r}) \equiv \langle \mathbf{r} | \Psi_T \rangle$.

In variational Monte Carlo (VMC) calculations [51–53], the standard Metropolis algorithm is used to sample the $3N$ dimensional configuration space \mathbf{r} according to a weight $\sim |\Psi_T(\mathbf{r})|^2$ via Markov chains. An estimate of the trial energy is then obtained by averaging the local energy over the sampled configurations

$$E_T = \mathbb{E}_{\mathbf{r} \sim \Psi_T^2} [E_L(\mathbf{r})]. \quad (20)$$

This estimate is unbiased, but affected by a stochastic error of order $\sim \sqrt{\sigma_T^2 / N_{MC}}$, controlled by the number of independent Monte Carlo samples, N_{MC} . Better wave functions, not only lower the energy expectation value, but also reduce the statistical error of the VMC calculations when reducing the variance towards an exact eigenstate (zero variance principle).

2.3.3 Stochastic optimization

The variational principle is of fundamental importance, since it allows us to compare the "quality" of different trial wave functions and select our "best" one, based on an objective criterium, without invoking comparison with experiment. Within VMC we can evaluate the energy of broad classes of wave functions, those which we can evaluate efficiently on our computer. Above, we have described generic forms of trial wave functions, all of them containing one or several unspecified functions. Any of these (one dimensional) functions can be expanded in a basis set providing a parametrization of the wave function.

Our trial wave function $\Psi_T(\mathbf{r}|\alpha)$ thus contains potentially a very large number of parameters α . The variational principle allows us to fix them by searching for the minimum of the energy expectation value $E_T(\alpha)$. This is almost a standard minimization problem, non-linear in α and intrinsically affected by stochastic noise, since $E_T(\alpha)$ is obtained via Monte Carlo integration. Ignoring the problem of noise for the moment, one would expand $E_T(\alpha + \delta\alpha)$ up to second order in small changes $\delta\alpha$ of our parameters and use some iterative method based on Newton's method

$$\alpha \leftarrow \alpha - (\partial_\alpha^2 E_T(\alpha))^{-1} \partial_\alpha E_T(\alpha) \quad (21)$$

where $\partial_\alpha E_T(\alpha)$ denotes the gradient vector and $\partial_\alpha^2 E_T(\alpha)$ the second derivative (Hessian) matrix, or some variant of it [54]. Apart that the Hessian might be quite costly to calculate for large number of parameters, the inversion is likely to be corrupted by the stochastic noise. Instead of trying to reduce the noise as much as possible by using long MC runs, a simple gradient descent

$$\alpha \leftarrow \alpha - \epsilon \partial_\alpha E_T(\alpha) \quad (22)$$

with learning rate ϵ can be more efficient [55].

For a real trial wave function, the gradient vector of the trial energy is written as

$$\begin{aligned} \partial_\alpha E_T(\alpha) &= 2 \frac{\int d\mathbf{r} \Psi_T(\mathbf{r}|\alpha) H \partial_\alpha \Psi_T(\mathbf{r}|\alpha)}{\int d\mathbf{r} \Psi_T^2(\mathbf{r}|\alpha)} - 2 \frac{\int d\mathbf{r} \Psi_T(\mathbf{r}|\alpha) H \Psi_T(\mathbf{r}|\alpha)}{\int d\mathbf{r} \Psi_T^2(\mathbf{r}|\alpha)} \frac{\int d\mathbf{r} \Psi_T(\mathbf{r}|\alpha) \partial_\alpha \Psi_T(\mathbf{r}|\alpha)}{\int d\mathbf{r} \Psi_T^2(\mathbf{r}|\alpha)} \\ &= 2 \mathbb{E}_{\mathbf{r} \sim \Psi_T^2} [(E_L(\mathbf{r}|\alpha) - E_T(\alpha)) \partial_\alpha \log \Psi_T(\mathbf{r}|\alpha)] \end{aligned} \quad (23)$$

where we have used that the Hamiltonian is hermitian to obtain an unbiased estimator in the last line. The resulting stochastic gradient descent based on the Monte Carlo estimation of $\partial_\alpha E_T$ is guaranteed to converge to the minimal energy [55], even in the case of large stochastic noise. Strategies based on stochastic gradient descent, also known as stochastic reconfiguration [56], have been successfully used for minimizing very large set of variational parameters, and have been combined with machine learning approaches more recently [57–59].

2.3.4 Projector Monte Carlo methods

Propagating a wave function in imaginary time $\Psi_t(\mathbf{r}) \sim \langle \mathbf{r} | e^{-tH} | \Psi_T \rangle$, the energy expectation value $E_t = \langle \Psi_t | H | \Psi_t \rangle / \langle \Psi_t | \Psi_t \rangle$ approaches the ground state energy of the same symmetry sector exponentially fast in t . Imaginary time projection can be performed stochastically based on the path-integral representation of the propagator $\rho_e(\mathbf{r}', \mathbf{r}; t) = \langle \mathbf{r}' | e^{-tH_e} | \mathbf{r} \rangle$ along the lines discussed above. In the primitive approximation for $\rho_e(\mathbf{r}', \mathbf{r}; t)$ we get the variational path-integral expression (VPIMC or PIGS for path-integral ground state techniques)

$$\Psi_t^2(\mathbf{r}) \sim \frac{\int d\mathbf{r}^0 \int d\mathbf{r}^{2M} \int_{\mathbf{r}^0}^{\mathbf{r}^{2M}} D\mathbf{r}[\tau] \pi(\mathbf{r}[\tau]) \delta(\mathbf{r} - \mathbf{r}^M)}{\int d\mathbf{r}^0 \int d\mathbf{r}^{2M} \int_{\mathbf{r}^0}^{\mathbf{r}^{2M}} D\mathbf{r}[\tau] \pi(\mathbf{r}[\tau])} \quad (24)$$

with

$$\pi(\mathbf{r}[\tau]) = \Psi_T(\mathbf{r}^0) e^{-S_{2M}^p(\mathbf{r}[\tau]; 2t)} \Psi_T(\mathbf{r}^{2M}). \quad (25)$$

In contrast to PIMC for the diagonal density matrix, Eqs (6) and (7), where the paths are periodic in imaginary time ($\mathbf{r}^0 = \mathbf{r}^M$), in VPIMC, Eq. (24), the paths are open with their ends weighted by the trial wave function, $\Psi_T(\mathbf{r}^0)$ and $\Psi_T(\mathbf{r}^{2M})$, and the ground state distribution is represented in the central slice \mathbf{r}^M of the path.

The primitive approximation is rarely used in this context. As ground state projector Monte Carlo methods are build on top of an optimized trial wave function, one usually wants to take advantage, using it as a guiding wave function for importance sampling. For this we introduce a similarity transformation of our Hamiltonian in the position representation

$$\tilde{H}_e = e^{U_T(\mathbf{r})} H_e e^{-U_T(\mathbf{r})} = T_e + \frac{\hbar^2}{m_e} \sum_i [\nabla_i U_T(\mathbf{r})] \cdot \nabla_i + E_L(\mathbf{r}) \quad (26)$$

where $\Psi_T(\mathbf{r}) = e^{-U_T(\mathbf{r})}$ is used as a guiding wave function.

The propagator corresponding to \tilde{H} can then be written as

$$\tilde{\rho}_T(\mathbf{r}', \mathbf{r}; \tau) = \sum_n \frac{\langle \mathbf{r}' | (-\tau \tilde{H}_e)^n | \mathbf{r} \rangle}{n!} = \sum_n \frac{e^{U_T(\mathbf{r}')} \langle \mathbf{r}' | (-\tau H_e)^n | \mathbf{r} \rangle e^{-U_T(\mathbf{r})}}{n!} = e^{U_T(\mathbf{r}')} \rho(\mathbf{r}', \mathbf{r}; \tau) e^{-U_T(\mathbf{r})} \quad (27)$$

which shows that $\tilde{\rho}_T$ and ρ are also related by a similarity transformation.

Let us now apply Trotter's approximation to $\tilde{\rho}_T$ for small τ

$$\begin{aligned} \tilde{\rho}_T(\mathbf{r}', \mathbf{r}; \tau) &\approx \langle \mathbf{r}' | e^{-\tau T_e} e^{-i\tau \sum_j (\hbar/m_e) [\nabla_j U_T] \cdot \hat{\mathbf{p}}_j} e^{-\tau E_L} | \mathbf{r} \rangle \\ &\approx \langle \mathbf{r}' | e^{-\tau T_e} | \mathbf{r} - \tau \mathbf{f} \rangle e^{-\tau E_L(\mathbf{r})} \approx e^{-\tau E_L(\mathbf{r}')} \langle \mathbf{r}' + \tau \mathbf{f}' | e^{-\tau T_e} | \mathbf{r} \rangle \\ &\sim e^{-m_e(\mathbf{r} - \mathbf{r}' - \tau \mathbf{f}')^2 / 2\hbar^2\tau} e^{-\tau[E_L(\mathbf{r}') + E_L(\mathbf{r})]/2} \equiv \tilde{\rho}_T^D(\mathbf{r}' \rightarrow \mathbf{r}; \tau) e^{-\tau[E_L(\mathbf{r}') + E_L(\mathbf{r})]/2} \end{aligned} \quad (28)$$

where $\mathbf{f}_i = (\hbar^2/m_e)\nabla_i \log \Psi_T(\mathbf{r})$ and we have used that $e^{-ip\delta/\hbar}|x\rangle = |x+\delta\rangle$. We have split the propagator into a drifted random walk, $\tilde{\rho}_T^D$, and a weight according to the local energy of the configurations. However, our approximate expression involving $\tilde{\rho}_T^D$ violates the exact relation $\tilde{\rho}_T(\mathbf{r}', \mathbf{r}; \tau)/\tilde{\rho}_T(\mathbf{r}, \mathbf{r}'; \tau) = \Psi_T^2(\mathbf{r})/\Psi_T^2(\mathbf{r}')$. It is important to restore this symmetry, e.g. using

$$\tilde{\rho}_T^{DMC}(\mathbf{r}' \rightarrow \mathbf{r}; \tau) = \tilde{\rho}_T^D(\mathbf{r}' \rightarrow \mathbf{r}; \tau) \min \left[1, \frac{\Psi_T^2(\mathbf{r}) \tilde{\rho}_T^D(\mathbf{r} \rightarrow \mathbf{r}'; \tau)}{\Psi_T^2(\mathbf{r}') \tilde{\rho}_T^D(\mathbf{r}' \rightarrow \mathbf{r}; \tau)} \right] \quad (29)$$

which coincides with Eq. (28) up to higher order terms in τ . We can now replace the weight of the path, Eq. (25), involved in the projection with

$$\pi(\mathbf{r}[\tau]) \approx \psi_T^2(\mathbf{r}^0) \prod_{m=0}^{2M-1} \tilde{\rho}_T^{DMC}(\mathbf{r}^m \rightarrow \mathbf{r}^{m+1}; \tau) e^{-\tau[E_L(\mathbf{r}^m) + E_L(\mathbf{r}^{m+1})]/2} \quad (30)$$

Reptation Monte Carlo (RMC) [60] adds moves where new configurations are proposed by a global shift in imaginary time, $\mathbf{r}^{m+1} \leftarrow \mathbf{r}^m$ (and, with equal probability, $\mathbf{r}^{m-1} \leftarrow \mathbf{r}^m$), dropping the configuration \mathbf{r}^m with $m > 2M$ ($m < 0$), and creating a new configuration for the freed place \mathbf{r}^0 (\mathbf{r}^{2M}). This proposition is then accepted or refused following Metropolis' rule according to the change of the weight $\pi(\mathbf{r}[\tau])$. The autocorrelation can be further reduced with minimal modifications implementing an algorithm similar to a directed loop algorithm [61].

PIMC and RMC methods store the (discretized) path $\mathbf{r}[\tau]$, ground state properties are directly accessible in the middle of the path, $\mathbf{r}(t/2)$, Eq. (24), in the limit $t \rightarrow \infty$, $\tau = t/2M \rightarrow 0$. Both limits must be numerically extrapolated.

Let us now regard the mixed distribution $\Psi_t(\mathbf{r})\Psi_T(\mathbf{r})$ at the external ends of the path which we write

$$f_t(\mathbf{r}) = \Psi_t(\mathbf{r})\Psi_0(\mathbf{r}) \sim \frac{\int d\mathbf{r}^0 \int d\mathbf{r}^M \int_{\mathbf{r}^0}^{\mathbf{r}^M} D\mathbf{r}[\tau] \pi_M(\mathbf{r}[\tau]) \delta(\mathbf{r} - \mathbf{r}^M)}{\int d\mathbf{r}^0 \int d\mathbf{r}^M \int_{\mathbf{r}^0}^{\mathbf{r}^M} D\mathbf{r}[\tau] \pi_M(\mathbf{r}[\tau])} \quad (31)$$

$$\pi_M(\mathbf{r}[\tau]) = \psi_T^2(\mathbf{r}^0) \prod_{m=0}^{M-1} \tilde{\rho}_T^{DMC}(\mathbf{r}^m \rightarrow \mathbf{r}^{m+1}; \tau) e^{\tau\{E_T - [E_L(\mathbf{r}^m) - E_L(\mathbf{r}^{m+1})]/2\}} \quad (32)$$

where we have introduced the trial energy E_T , a so-far arbitrary constant corresponding to a global weight, for later use. By noting that

$$\pi_{M+1}(\mathbf{r}[\tau]) = \pi_M(\mathbf{r}[\tau]) \tilde{\rho}_T^{DMC}(\mathbf{r}^M \rightarrow \mathbf{r}^{M+1}; \tau) e^{\tau\{E_T - [E_L(\mathbf{r}^{M+1}) + E_L(\mathbf{r}^M)]/2\}} \quad (33)$$

we can grow the projection time t based on

$$f_{t+\tau}(\mathbf{r}) \sim \mathbb{E}_{\mathbf{r}' \sim f_t(\mathbf{r}')} [\tilde{\rho}_T^{DMC}(\mathbf{r}' \rightarrow \mathbf{r}; \tau) e^{\tau\{E_T - [E_L(\mathbf{r}) + E_L(\mathbf{r}')]/2\}}]. \quad (34)$$

Diffusion Monte Carlo (DMC) calculations directly sample the limiting distribution $t \rightarrow \infty$, iterating Eq. (34). Starting from an initial VMC distribution of N_w "walkers" $\mathbf{r}' \sim f_0(\mathbf{r}') \equiv \Psi_T^2(\mathbf{r}')$, the walkers are displaced by a drifted random walk according to $\tilde{\rho}_T^{DMC}(\mathbf{r}' \rightarrow \mathbf{r}; \tau)$ acquiring a weight $w(\mathbf{r}, \mathbf{r}') = e^{\tau\{E_T - [E_L(\mathbf{r}) + E_L(\mathbf{r}')]/2\}}$. A branching process is usually added to take this weight into account by keeping on average $w(\mathbf{r}, \mathbf{r}')$ copies of a propagated walker. The trial energy E_T must be chosen (and eventually adapted) to keep the population of walkers asymptotically stable. The grows process of Eq. (34) will reach a stationary non-vanishing distribution, $f_\infty(\mathbf{r}) \sim \Psi_T(\mathbf{r})\Psi_0(\mathbf{r})$ for a trial energy coinciding with the true ground state energy E_0 [62, 63, 56].

Since DMC can be implemented on top of VMC with very few modifications, it is by far the most applied zero temperature projection method. However, the stationary distribution $f_\infty(\mathbf{r})$ does not correspond to the ground state density unless $\Psi_T \sim \Psi_0$. Thus, general ground state observables are not directly accessed, apart from the energy where

$$E_0 = \lim_{t \rightarrow \infty} \frac{\langle \Psi_{t/2} | H | \Psi_{t/2} \rangle}{\langle \Psi_{t/2} | \Psi_{t/2} \rangle} = \lim_{t \rightarrow \infty} \frac{\langle \Psi_T | H | \Psi_t \rangle}{\langle \Psi_T | \Psi_t \rangle} = \mathbb{E}_{\mathbf{r} \sim f_\infty} E_L(\mathbf{r}) \quad (35)$$

provides an unbiased estimator of the true ground state energy.

DMC calculations need to be extrapolated to the limit of vanishing time step, $\tau \rightarrow 0$, and an infinitely large population of walkers, $N_w \rightarrow \infty$. In particular, the bias due to the finite size of the population will eventually grow strongly with system size. Depending on the quality of the trial wave function, the scalability of DMC to converge to the ground state energy of large systems may be questionable [64].

2.3.5 Fixed-node approximation

Quite generally, the overall ground state of any (regular) Hamiltonian, is nodeless and symmetric with respect to particle exchange. Thus, all projection Monte Carlo methods described above can be directly applied to obtain the ground state of a system containing N bosons. This is not the case for Fermions, since the ground state wave function of a Fermi system must be anti-symmetric,

$$\Psi_F(\dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots) = -\Psi_F(\dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots), \quad \text{for any } i, j \quad (36)$$

with nodes where $\Psi(\mathbf{r}) = 0$, e.g., when $\mathbf{r}_i = \mathbf{r}_j$. Thus, in general, the ground state of fermions is never the lowest eigenstate of the Hamiltonian of the system. Only in particular situations

which we do not address here, e.g., for some particular Hamiltonian in one spatial dimension, it can be degenerate with the bosonic ground state.

It is possible to extend the variational principle for the energy to some excited states, $|\Psi_m\rangle$ with $E_m > E_0$, imposing orthogonality of the trial wave function to all lower eigenfunctions, $\langle\Psi_T|\Psi_n\rangle = 0$ for all n with $E_n < E_m$. Since the fermionic ground state wave function is the lowest eigenfunction in the space of anti-symmetric wave functions, orthogonality to states with lower energy is guaranteed by symmetry. Since VMC based methods sample $|\Psi_T(\mathbf{r})|^2 \geq 0$, they can be directly applied to Fermions by using anti-symmetric trial wave functions which obey Eq. (36). All of the previously discussed trial wave functions are constructed manifestly anti-symmetric based on determinants.

In contrast to VMC, Projection Monte Carlo methods stochastically sample $\Psi_0(\mathbf{r})$ which now contains negative regions where the wave function cannot be interpreted as probability. Let us try to represent a fermionic wave function, starting with

$$\Psi_T^\pm = \frac{1}{2}(|\Psi_T| \pm \Psi_T) \quad \text{so that} \quad \Psi_T = \Psi_T^+ - \Psi_T^-, \quad \Psi_T^+ \geq 0, \quad \Psi_T^- \geq 0 \quad (37)$$

and diffuse $\Psi_T^\pm \geq 0$ separately. Using $|\Psi_T|$ for importance sampling, e.g. in DMC, we obtain the mixed distribution

$$f_t^\pm(\mathbf{r}) = |\Psi_T(\mathbf{r})| \Psi_t^\pm(\mathbf{r}). \quad (38)$$

However, both Ψ_T^\pm do have some overlap with the bosonic ground state, Ψ_B , of energy E_B ,

$$\Psi_T^\pm = \frac{1}{2} \left(c_B e^{-tE_B} \Psi_B(\mathbf{r}) \pm c_F e^{-tE_F} \Psi_F(\mathbf{r}) + \dots \right) \quad \text{with} \quad c_{B/F} = \int d\mathbf{r} \Psi_{B/F}(\mathbf{r}) |\Psi_T(\mathbf{r})|. \quad (39)$$

We can now calculate the expectation value of some operator for the fermionic state as

$$\langle O \rangle = \frac{\int d\mathbf{r} s(\mathbf{r}) O(\mathbf{r}) (f_t^+ - f_t^-)}{\int d\mathbf{r} s(\mathbf{r}) (f_t^+ - f_t^-)} = \frac{1}{\bar{s}} \frac{\int d\mathbf{r} s(\mathbf{r}) O(\mathbf{r}) (f_t^+ - f_t^-)}{\int d\mathbf{r} (f_t^+ + f_t^-)} \quad (40)$$

where $s(\mathbf{r}) \equiv \Psi_T(\mathbf{r})/|\Psi_T(\mathbf{r})| = \pm 1$ and

$$\bar{s} = \frac{\int d\mathbf{r} s(\mathbf{r}) (f_t^+ - f_t^-)}{\int d\mathbf{r} (f_t^+ + f_t^-)} = \frac{c_F^2 e^{-tE_F} + \dots}{c_B^2 e^{-tE_B} + \dots}. \quad (41)$$

Therefore, we have

$$\bar{s} \sim \exp(-Nt\Delta\mathcal{E}) \quad (42)$$

where $\Delta\mathcal{E} = (E_F - E_B)/N$ is independent of N for large systems. The mean sign \bar{s} in general enters in the normalization of expectation values for physical observables. Assuming a finite gap, $\mathcal{E} > 0$, between the fermionic and bosonic ground state energy per particle, \bar{s} vanishes exponentially in $Nt \gg 1$. Since $\langle s^2 \rangle = 1$, the variance approaches one in this limit

$$\sigma_s^2 = \overline{s^2} - \bar{s}^2 \approx 1 \gg \bar{s}^2 \sim \exp(-2Nt\Delta\mathcal{E}). \quad (43)$$

In order to get the error bar of the sign $\sim \sqrt{\sigma_s^2/N_{MC}}$ sufficiently small to resolve the value of the average sign, we roughly need

$$N_{MC} \sim \frac{1}{\bar{s}^2} \sim \exp(2Nt\Delta\mathcal{E}) \quad (44)$$

independent samples, increasing exponentially with N and t . This is the fermion sign problem. Can we circumvent this exponential signal to noise ratio? Importance sampling DMC is based on the overlap

$$f(\mathbf{r}) = \Psi_T(\mathbf{r}) \Psi_0(\mathbf{r}). \quad (45)$$

Let us consider that we have found a trial wave function with exactly the same positive and negative regions as the fermionic ground state we are looking for, so that $f(\mathbf{r}) \geq 0$ in the full configuration space. In this case, if we impose $f(\mathbf{r}) \geq 0$ during the time evolution in DMC, we expect that DMC converges to the exact fermionic ground state. What happens? Looking at the drifted random walk created by the importance sampling, imposing $f(\mathbf{r}) \geq 0$ for all \mathbf{r} , we reject any move $\mathbf{r} \rightarrow \mathbf{r}'$ with $\Psi_T(\mathbf{r})\Psi_T(\mathbf{r}') < 0$. Our population of walkers can be separated into two sets, positive walkers at \mathbf{r}^+ which satisfy $\Psi_T(\mathbf{r}^+) \geq 0$, and negative walkers at \mathbf{r}^- with $\Psi_T(\mathbf{r}^-) < 0$. Positive and negative walkers are separated by the nodal surface \mathbf{s} where $\Psi_T(\mathbf{r}=\mathbf{s}) \equiv 0$, and it is enough to know the exact nodal surface. Note that the nodal surface \mathbf{r} is a hypersphere in $Nd-1$ dimensions where d is the spatial dimension. For any sufficiently regular anti-symmetric trial function, applying the permutation operator to any positive configuration, we obtain a negative walker, and vice versa. It is therefore sufficient to sample only the positive space as long as we are only interested in physical observables which commute with the permutation operator.

Everything above is fine, but we still do not know the nodal surface for almost all fermion problems we are interested in, so let us search for the best approximation we can do. In the fixed-node approximation, we simply impose the nodes of a given trial wave function. Once started with positive walkers, our fixed-node DMC algorithm will converge to an eigenfunction of the Hamiltonian

$$H\Psi_{FN}(\mathbf{r}) = E_{FN}\Psi_{FN}(\mathbf{r}), \quad \text{for all } \mathbf{r} \text{ in } \mathbf{r}^+, \text{ the positive region with } \Psi_T(\mathbf{r}) \geq 0. \quad (46)$$

On the nodes \mathbf{s} of Ψ_T we also have $\Psi_{FN}(\mathbf{s}) = 0$, and we can continue the wave function to the negative regions, \mathbf{r}^- using permutations $\Psi_{FN}(\mathbf{r}) = (-1)^{|P|} \Psi_{FN}(P\mathbf{r})$, where the permutation P can be determined from solving $\Psi_T(\mathbf{r}) = (-1)^{|P|} \Psi_T(P\mathbf{r})$ for P . As long as Ψ_T is a sufficiently regular fermionic trial wave function, we can reach all configuration space by this procedure, and the continued Ψ_{FN} is a continuous anti-symmetric wave function [65].

Unfortunately, the partial derivatives of the constructed FN wavefunction with respect to \mathbf{r}_i are in general not continuous at the nodal surface. The variational principle does not directly apply, since the FN wave function is outside the variational space of wave functions (continuous wave functions with continuous first derivatives). However, we can smear out our wave function at a distance ϵ close to the nodes to make them sufficiently smooth to apply the variational theorem, so that the smoothed function provides an upper bound for the energy. This smoothing will increase the absolute value of the curvature $\sim \epsilon^{-1}$ close to the node and the Laplacian of the kinetic energy will produce large absolute values, $\sim \epsilon^{-1}$. However, since the wave function vanishes as ϵ , the kinetic energy contribution of the smoothed wave function close to the nodal region $\sim \int_{\epsilon} \psi \nabla^2 \psi \sim \epsilon$ vanishes. Therefore, the energy of our fixed-node wave

function provides a true upper bound to the fermion ground state energy, E_F , [66]

$$E_F \leq E_{FN} = \frac{\int d\mathbf{r} \Psi_{FN}(\mathbf{r}) E_L(\mathbf{r}) \Psi_{FN}(\mathbf{r})}{\int d\mathbf{r} \Psi_{FN}(\mathbf{r}) \Psi_{FN}(\mathbf{r})}. \quad (47)$$

For many-body fermion problems, the fixed-node energies are the most accurate variational values and routinely used in electronic structure DMC calculations [67].

As a generalization of the fixed-node approach, the fixed-phase approximation [68] is based on a complex trial wave function

$$\Psi_T(\mathbf{r}) = A(\mathbf{r})e^{-i\varphi(\mathbf{r})}, \quad \text{with non negative amplitude } A(\mathbf{r}) \geq 0 \text{ and real phase } \varphi(\mathbf{r}). \quad (48)$$

For any given phase, we can then minimize the energy of the trial wave function for an explicitly given phase, $\varphi(\mathbf{r})$. However, since the phase is only well defined for non-vanishing amplitude, we also have to fix the nodes of the amplitude and make the wave function single valued. An argument similar to that above shows that the fixed-phase wave function provides also an upper bound for the ground state energy in the same symmetry class as Ψ_T . Fixed-phase methods are needed for treating twisted boundary conditions, see below, and for systems with broken time-reversal invariance, e.g., including magnetic field effects.

2.4 Finite size effects

With the methods described above, we will be able to simulate systems containing N particles (electrons or protons), typically $N \lesssim 10^3$, and one might ask how such small systems may faithfully reproduce material properties in the bulk. To eliminate surface effects, periodic boundary conditions are in general used. Still, residual effects of the underlying finite simulation cell remain and the extrapolation to the thermodynamic limit represents one of the major sources of bias. Heuristically, for classical particles, interacting via short-range forces exponential convergence may be expected once the size of the simulation cell exceeds the correlation length. Large systems are needed approaching phase transitions where the correlation length diverges, and care is needed to correctly describe ordered phases which are in general sensitive to boundary conditions. Methods to address finite size effects having their origin in structural formation have been developed in the context of classical molecular dynamics or Monte Carlo calculations [25, 26]. In the following we will focus on finite size effects of electronic origin.

As discussed previously, all temperatures shown in Fig. 1 are much lower than the Fermi temperature of electrons, a typical situation in condensed matter physics and material science. In contrast to the nuclei, electrons are in a strongly degenerate quantum state where the wave character dominates and the sensitivity to boundary conditions is strongly enhanced.

To illustrate finite size effects, let us consider a non-interacting gas of electrons at density $N/L^3 \sim k_F^3 \sim r_s^{-3}$ where k_F is the Fermi wave vector corresponding to the highest occupied single particle state. From the ratio of the Fermi wave length $\lambda_F = 2\pi/k_F$ to the size of the box, $\lambda_F/L \simeq 2.03N^{-1/3}$, we already see that size effects will be far from negligible even for large simulation cells containing thousands of electrons or more.

These size effects are rather well understood. They correspond to shell effects which are already present for an ideal gas with a Slater determinant composed of the first N plane wave orbitals of wave vectors \mathbf{k}_i on a discrete grid of spacing $2\pi/L$. The corresponding kinetic energy per particle, $\mathcal{T}_N^0 = T_N^0/N = N^{-1} \sum_{i=1}^N (\hbar k_i)^2/2m_e$, extrapolates irregularly to the thermodynamic limit due to the sharp Fermi surface. Imposing twisted boundary conditions on the wave function [69], $\Psi(\dots, \mathbf{r}_i+L, \dots) = e^{i\vartheta} \Psi(\dots, \mathbf{r}_i, \dots)$, the wave vectors of the plane wave orbitals are collectively displaced, corresponding to a shifted grid $\mathbf{k} + \vartheta$. Averaging the twist angle ϑ over N_ϑ twists on a dense grid in the Brillouin zone of the simulation cell then mimics the thermodynamic limit integration $\lim_{N_\vartheta \rightarrow \infty} N_\vartheta^{-1} \sum_{\vartheta} \sum_{\mathbf{k}} f(\mathbf{k} + \vartheta) = V/(2\pi)^3 \int d^3q f(\mathbf{q})$ for any function f . However, in a many-body calculation with fixed number of particles, twist averaged boundary conditions (TABC) do not necessarily restore a sharp Fermi surface, since exactly N orbitals of lowest single particle energies are occupied for each twist, ϑ . To reproduce the exact single particle energy with a sharp Fermi surface, the number of particles must be allowed to vary with ϑ . This is implemented in grand-canonical twist averaging (GC-TABC) [70, 71]. Further important size effects are due to the long range Coulomb interactions. Let us write down the potential energy per particle, considering only electron-electron interactions,

$$\mathcal{V}_N \equiv V_N/N = \frac{1}{2V} \sum_{\mathbf{k} \neq 0} v_k (S_N(k) - 1) \quad (49)$$

where $v_k \sim k^{-2}$ is the Fourier transform of the Coulomb potential, $S_N(k) = \langle \rho_{\mathbf{k}} \rho_{-\mathbf{k}} \rangle / N$ the static structure factor, and $\rho_{\mathbf{k}} = \sum_j e^{i\mathbf{k} \cdot \mathbf{r}_j}$ the Fourier transform of the density operator.

Since $S_N(k)$ is a local operator, we may expect fast convergence to the thermodynamic limit, $S_N(k) \simeq S_\infty(k)$, by analogy of $\int d\mathbf{r} \Psi_N^2(\mathbf{r})$ to the configuration integral of classical systems. Then, the dominate size error is given by the replacement of the discrete summation in reciprocal space by an integration is

$$\mathcal{V}_\infty - \mathcal{V}_N = \left(\int \frac{d\mathbf{k}}{(2\pi)^3} - \frac{1}{V} \sum_{\mathbf{k} \neq 0} \right) \frac{v_k}{2} (S_\infty(k) - 1). \quad (50)$$

Nonanalytical behavior of the integrand will dominate the quadrature error. Since v_k diverges at the origin, leading order corrections can be determined by focusing on the integration around $k \rightarrow 0$. The dominating term is actually the Madelung constant, e.g., the contribution of the interaction of one particle with all its periodic images,

$$v_M = - \left(\int \frac{d\mathbf{k}}{(2\pi)^3} - \frac{1}{V} \sum_{\mathbf{k} \neq 0} \right) \frac{v_k}{2} \sim \sum_{\mathbf{n} \neq 0} \frac{1}{|\mathbf{n}|L} \sim N^{-1/3} \quad (51)$$

and Ewald's method should be used for evaluation [24–26].

Since $S(k) \sim k^2$ for charged systems, with the prefactor fixed by sum rules, the next order term corresponds to the missing term with $k = 0$ in the summation on the rhs of Eq. (50), $\lim_{k \rightarrow 0} v_k S(k)/V \sim N^{-1}$. Similar to the Madelung constant, this term depends only on the shape of the simulation cell. Both terms are therefore easily corrected for [70].

Seemingly innocent, the quadratic behavior of $S(k)$ around the origin is due to charge density fluctuations, the plasmons. The size correction of the potential energy $\sim N^{-1}$ corresponds to including half of the zero point energy of the long wavelength plasmons which do not fit inside the simulation cell. Kinetic energy corrections will add the missing other half of the plasmon energies [70].

For neutral quantum particles, these size corrections decay slightly faster $\sim N^{-4/3}$, since the energy of the long wave length phonon modes vanish linearly in $|\mathbf{k}|$, and one can show that $S(k) \sim k$ in this case.

Understanding size effects can be a powerful tool. Using all information of calculations at a single system size, allows us to make reliable predictions of thermodynamic limit values. This is particularly important for calculations on hydrogen discussed later, where several calculations varying system sizes are hardly affordable.

We have outlined above, that finite size effects on the electronic ground state energy are intrinsically connected to non-analytical behavior of the wave function. Beyond shell effects, the behavior of the structure factor $S(k) \sim k^\alpha$ for $k \rightarrow 0$ determines the exponent of the leading order power law $\mathcal{E}_\infty - \mathcal{E}_N \sim N^{-(\alpha+1)/3}$ for the total energy per particle, $\mathcal{E}_N \equiv E/N$ in the case of Coulomb interactions. Although the exponent as well as the prefactor of $S(k)$ can be determined via general considerations, inaccuracies in the trial wave function, either due to limitations of the functional form or due to insufficient optimization, might lead to deviations which then propagate to size effects.

Of course, such predictions depend crucially on the underlying assumptions, as well as on the actual values for the asymptotics, e.g., extrapolating $\lim_{k \rightarrow 0} S(k)/k^2$ from our finite size data. Estimating the error of such procedures is a difficult task. Whenever affordable, numerical extrapolations of different system sizes provide important cross-checks.

Pure numerical extrapolation of size effects is delicate as our computations actually only guarantee upper bounds to the exact ground state energies. As the computational cost for optimization as well as for projection methods like DMC strongly increase with system size, deteriorations of the energies for large system sizes as compared to those predicted based on finite size corrections, may actually indicate convergence problems in the data.

Here, we have focused on the error of the ground state energy. Similar considerations apply for other observables, in particular for the pressure. Using the virial estimator for Coulomb systems, size effects on the pressure can be obtained from the separate information on kinetic and potential energy corrections.

2.5 Monte Carlo calculations with noisy action

Above we have discussed the main systematic errors of QMC calculations together with some strategies to estimate and reduce them. However, when using QMC energies in Born-Oppenheimer calculations for the nuclei, we also have to address the stochastic error. To be concrete, let us discuss Monte Carlo calculations for classical nuclei distributed by the Boltzmann weight $\sim e^{-\beta E_{\mathbf{R}}}$ according to the Born-Oppenheimer ground state energies $E_{\mathbf{R}} \equiv E_0(\mathbf{R})$.

Using QMC methods to determine electronic energies we do not have access to $E_{\mathbf{R}}$ with arbitrary precision. Our QMC calculations output $\varepsilon = E_{\mathbf{R}} + r$ where r is the error due to the stochastic noise. Assuming the error is Gaussian distributed, $g(\varepsilon|E_{\mathbf{R}}\sigma_{\mathbf{R}}^2) \sim e^{-(\varepsilon-E)^2/2\sigma^2}$ with variance $\sigma_{\mathbf{R}}^2$, the configuration integral can be written

$$Z = \int d\mathbf{R} e^{-\beta E_{\mathbf{R}}} = \int d\varepsilon \int d\mathbf{R} e^{-\beta\varepsilon - \beta^2\sigma_{\mathbf{R}}^2/2} g(\varepsilon|E_{\mathbf{R}}\sigma_{\mathbf{R}}^2) \quad (52)$$

where the last equality can be verified by explicit integration over ε . It is then straightforward to perform Monte Carlo calculations using the noisy energies ε . However, we need to replace the (unknown) exact values of the energies, $E_{\mathbf{R}}$, by $\varepsilon + \beta\sigma^2/2$ in the Metropolis acceptance ratio, where ε is our QMC estimate of known variance $\sigma^2 \equiv \sigma_{\mathbf{R}}^2$. We are thus able to obtain unbiased samples of \mathbf{R} according to the Boltzmann weight of the energies $E_{\mathbf{R}}$ without ever calculating them exactly!

It is instructive to formulate such an algorithm in more detail. Our Monte Carlo state vector is actually specified by $(\mathbf{R}, \varepsilon)$ with weight

$$\Pi(\mathbf{R}, \varepsilon) = e^{-\beta\varepsilon - \beta^2\sigma_{\mathbf{R}}^2/2} g(\varepsilon|E_{\mathbf{R}}\sigma_{\mathbf{R}}^2) \quad (53)$$

and our transition probability of the Markov chain is given by

$$T(\mathbf{R}\varepsilon \rightarrow \mathbf{R}'\varepsilon') = \mathcal{A}(\mathbf{R} \rightarrow \mathbf{R}') g(\varepsilon'|E_{\mathbf{R}'}\sigma_{\mathbf{R}'}^2) a(\mathbf{R}\varepsilon \rightarrow \mathbf{R}'\varepsilon') \quad (54)$$

where $\mathcal{A}(\mathbf{R} \rightarrow \mathbf{R}')$ is the a-priori probability for the proposed move and the acceptance probability is given by the Metropolis-Hastings rule

$$a(\mathbf{R}\varepsilon \rightarrow \mathbf{R}'\varepsilon') = \min \left[1, \frac{\Pi(\mathbf{R}', \varepsilon') \mathcal{A}(\mathbf{R}' \rightarrow \mathbf{R})}{\Pi(\mathbf{R}, \varepsilon) \mathcal{A}(\mathbf{R} \rightarrow \mathbf{R}')} \frac{g(\varepsilon|E_{\mathbf{R}}\sigma_{\mathbf{R}}^2)}{g(\varepsilon'|E_{\mathbf{R}'}\sigma_{\mathbf{R}'}^2)} \right] \quad (55)$$

$$= \min \left[1, e^{-(\beta(\varepsilon' - \varepsilon) + \beta^2(\sigma_{\mathbf{R}'}^2 - \sigma_{\mathbf{R}}^2)/2)} \frac{\mathcal{A}(\mathbf{R}' \rightarrow \mathbf{R})}{\mathcal{A}(\mathbf{R} \rightarrow \mathbf{R}')} \right] \quad (56)$$

satisfying detailed balance

$$\Pi(\mathbf{R}, \varepsilon) T(\mathbf{R}\varepsilon \rightarrow \mathbf{R}'\varepsilon') = \Pi(\mathbf{R}', \varepsilon') T(\mathbf{R}'\varepsilon' \rightarrow \mathbf{R}\varepsilon). \quad (57)$$

Therefore, $\Pi(\mathbf{R}, \varepsilon)$ will be a stationary distribution of our Markov chain, so that the correct unbiased Boltzmann distribution results via integration over ε , yielding unbiased expectation values of all observables involving \mathbf{R} without explicit dependence on $E_{\mathbf{R}}$ (or ε), e.g., nuclear structural properties.

The electronic energies, $E_{\mathbf{R}}$, as well as general electronic properties can be calculated by independent re-sampling of the Born-Oppenheimer wave function at the obtained nuclear configurations. The mere average of ε of the weight $\Pi(\mathbf{R}, \varepsilon)$ will differ from $E_{\mathbf{R}}$ by $\sim \beta\sigma^2$.

But why should we actually store the value of ε , and return to this value whenever our attempted move to a new configuration $(\mathbf{R}'\varepsilon')$ is rejected? After all ε is a gaussian variable which is integrated over stochastically.

We should rather try to construct a Markov chain where our uncertainty is only entering inside a transition probability which depends on the energy difference $\Delta\varepsilon = E_{\mathbf{R}'} - E_{\mathbf{R}} + r$ with gaussian errors r

$$T(\mathbf{R} \rightarrow \mathbf{R}') = \int d\Delta\varepsilon \tilde{T}(\mathbf{R} \rightarrow \mathbf{R}'; \Delta\varepsilon) \quad (58)$$

Let us plug Eq. (58) into the usual detailed balance condition

$$e^{-\beta E_{\mathbf{R}}} \int d\Delta\varepsilon \tilde{T}(\mathbf{R} \rightarrow \mathbf{R}', \Delta\varepsilon) = e^{-\beta E_{\mathbf{R}'}} \int d\Delta\varepsilon \tilde{T}(\mathbf{R}' \rightarrow \mathbf{R}; \Delta\varepsilon). \quad (59)$$

As long as \tilde{T} satisfies Eq. (59), our Markov chain based on the transition probability Eq. (58) will converge to the correct equilibrium distribution. Replacing the deterministic integration in the transition matrix, Eq. (58), with a stochastic evaluation of the integral over the noise cannot affect the asymptotic, stationary distribution of the Markov chain, and we will still converge to the correct, unbiased equilibrium distribution. Therefore, it is enough that our new transition matrix \tilde{T} is determined by Eq. (59) such that detailed balance for \tilde{T} needs only to be satisfied on average. In Ref. [72] it has been shown that

$$\tilde{T}(\mathbf{R} \rightarrow \mathbf{R}'; \Delta\varepsilon) = g(\Delta\varepsilon | E_{\mathbf{R}'} - E_{\mathbf{R}}, \sigma^2) \min\left(1, e^{-\beta\Delta\varepsilon - \beta^2\sigma^2/2}\right) \quad (60)$$

satisfies detailed balance on average, Eq. (59). Since the variance of the energy difference reduces the acceptance probability, the algorithm for the random walk with uncertain energies has been called penalty method. Uncertainty in the variance can be similarly taken into account, but modifies the acceptance ratio further [72].

3 Coupled electron-ion Monte Carlo calculations

Coupled Electron-Ion Monte Carlo (CEIMC) calculations [73, 74] sample the discretized path-integral representation of nuclei in the Born-Oppenheimer approximation, Eqs (11,12), using Monte Carlo methods, both for the nuclear degrees of freedom, and for the determination of the adiabatic electronic ground state energies $E_0(\mathbf{R})$. So far, calculations have been performed for discernible nuclei, neglecting ionic exchange effects as well as excitations to higher electronic energy surfaces and effects of a possible Berry phase.

Electronic energies are calculated by ground state QMC methods as described above, the penalty methods is used to provide unbiased sampling of the nuclear degrees of freedom despite the intrinsic stochastic error in the Born-Oppenheimer energy surface.

Accuracy of the electronic energies is of crucial importance to provide meaningful comparisons and predictions for experiment. This requires robust high quality trial wave functions for the electronic Born-Oppenheimer ground state, as well as elimination of dominating finite size effects on the fly for each set of nuclear positions, \mathbf{R} .

The trial wave functions are build out of orbitals obtained from a separate density-functional (DFT) calculation at each nuclear configuration, many-body correlations are included via analytical expressions for Jastrow, three-body, and backflow potentials with parametric depen-

dence on the nuclear coordinates [39, 75] augmented by simple functional forms which are optimized over representative samples of nuclear positions. This provides an accurate and smooth parametrization of the trial wave functions over all relevant nuclear configurations avoiding on-the-fly optimizations. Although the orbitals have been determined by DFT, the dependence on the underlying DFT functional is weakened by the presence of the explicit many-body correlations in the wave function. Ultimately, the selection of the DFT flavor is based on the variational principle, similar to the functional choice and parametrization of the Jastrow, three-body and backflow functions.

Electronic energies are systematically averaged over twisted boundary conditions (TABC) for each nuclear configuration to reduce the influence of the finite simulation cell. This is particularly important, because shell effects strongly bias nuclear configurations on an energy scale easily exceeding the nuclear temperature. The use of TABC in CEIMC is almost for free, since the averaging of independent calculations with different twists reduces the overall stochastic error and can be implemented massively parallel.

More details concerning the methodology of CEIMC calculations can be found in Ref. [74]. Most of the CEIMC calculations so far have used VMC energies for the Born-Oppenheimer energy surface. Additional projection in imaginary time using RMC or DMC provides an overall lowering of the energies, typically smooth and weakly depend on the nuclear constellation. Such effects can be quantified a-posteriori by reweighting selected configurations of CEIMC trajectories based on VMC energies. Similarly, the influence of different DFT functionals used for the orbitals in VMC, as well as optimization of more flexible functional forms for the many-body correlations can be estimated and quantified, if necessary. Reweighting methods are expected to work, as long as structural properties are not strongly modified, e.g., the energy landscape for the nuclei is not significantly changed.

Since CEIMC calculations are computationally expensive, it is in general not possible to explore a phase diagram like that of hydrogen in Fig. 1 fully from scratch. Rather, less expensive DFT calculations are typically used for structural optimization at fixed pressure either at zero or finite temperatures which provide the shape of the simulation cell used in fixed volume CEIMC calculations. Analyzing the subsequent CEIMC results may then provide indications to either confirm or question such a procedure.

Vice-versa, QMC energies of selected configurations can be used to benchmark DFT functionals [76] or to train a machine-learned effective potential [14, 77, 22] such to fully explore phase diagrams and afford molecular dynamics calculations on much bigger length and time scales. The resulting candidate phases can then be explored more precisely within CEIMC.

One of the most appealing features of CEIMC calculations is the possibility to improve existing results on a long term by separately addressing the various assumptions or simplifications used in practice and outlined in the computational methods section above. The underlying variational principle provides an unambiguous criterium to improve and judge the quality of CEIMC results. Use of Langevin dynamics based on QMC forces within the Born-Oppenheimer approximation has also been developed [78], providing comparable results to CEIMC [21, 22].

4 Hydrogen under pressure: Some CEIMC results

Predicting and exploring phases of high pressure hydrogen and deuterium with particular focus on solid phases and on the liquid-liquid phase transition has been one of the main motivations for developing CEIMC. Further, since the major fraction of hydrogen in the universe is in the fluid phase, establishing an accurate equation of state for hydrogen and hydrogen-helium mixtures, pressure as a function of temperature, density, and composition, is relevant for planetary models. Beyond pure structural properties, e.g., the crystal structure in the various solid phases, the existence and location of possible insulator to metal transitions, and their connection to a molecular to atomic transition is one of the longest outstanding issues of high pressure research.

Since CEIMC calculations are based on the many-body density matrix, electronic properties can be probed beyond the single particle approximation. In particular, off-diagonal matrix elements of the reduced single particle density matrix can be calculated and used to determine localized (insulating) or extended (metallic) behavior of the electrons. Furthermore, electronic excitation gaps can be calculated within QMC accuracy as discussed below, providing more direct information on the transition to metallic hydrogen.

4.1 Electronic band gaps

From a theoretical point of view, band insulators and semiconductors are characterized by a nonvanishing value of the fundamental gap [63]

$$\tilde{\Delta} = E_0(N_n+1) + E_0(N_n-1) - 2E_0(N_n) \quad (61)$$

where $E_0(N_e)$ denotes the electronic ground state energy of a system of N_e electrons where the number N_n and positions of nuclei are kept fixed, and charge neutrality is assured by applying a uniform background charge. Since Eq. (61) only involves electronic ground state energies, ground state QMC methods can be applied straightforwardly to obtain upper bounds for $E_0(N_n)$ as well as for $E_0(N_n \pm 1)$. Although this does not guarantee us an upper bound for the gap, we can still use the variational principle to judge the quality of all the quantities involved.

Above, we have discussed the importance of size effects on the electronic energies per particle, $E(N_e)/N_e$. In Eq. (61) this extensive part of the ground state energies will drop out, since the gap itself is an intensive quantity. Calculations of gaps may therefore require higher control of size effects. A detailed discussion is given in Ref. [79] where it is shown that the fundamental gap approaches slowly its value in the thermodynamic limit, inversely proportional to the linear extension of the simulation cell, $\sim (\epsilon L)^{-1}$. Leading order finite size corrections are given in terms of the dielectric constant, ϵ , and can be determined within calculations of the same system size. Similar to ground state energy calculations, predictions for the thermodynamic limit value of the gap can thus be made without the need to extrapolate numerically results for different system sizes.

For hydrogen, the definition of the fundamental gap, Eq. (61), needs to be modified to take into account nuclear quantum effects

$$\Delta = F(N_n+1) + F(N_n-1) - 2F(N_n) \quad (62)$$

where $F(N_e) = -T \log Z(N_e)$ denotes the free energy of the system containing N_e electrons and N_n nuclei at temperature T . Within the Born-Oppenheimer approximation, we have

$$Z(N_n \pm 1) = Z(N_n) \langle e^{-\beta[E_{\mathbf{R}}(N_n \pm 1) - E_{\mathbf{R}}(N_n)]} \rangle \geq e^{-\beta \langle E_{\mathbf{R}}(N_n \pm 1) - E_{\mathbf{R}}(N_n) \rangle} \quad (63)$$

where $\langle \dots \rangle \equiv \mathbb{E}_{\mathbf{R} \sim \pi}[\dots]$ denotes the averaging over the probability distribution of the nuclei according to the Born-Oppenheimer energies of $N_e = N_n$ electrons.

Due to the homogeneous background charge added to the electron/hole-doped systems with $N_e = N_n \pm 1$, the density of the additional electron/hole is necessarily smeared out over the whole simulation cell. Therefore, the modification of the Born-Oppenheimer energy surface of the doped system uniformly approaches the one of the neutral system up to corrections of $1/N_n$, so that the equality in Eq. (63) applies in the thermodynamic limit. We thus have

$$\Delta = \langle E_{\mathbf{R}}(N_n + 1) + E_{\mathbf{R}}(N_n - 1) - 2E_{\mathbf{R}}(N_n) \rangle \equiv \langle \tilde{\Delta}_{\mathbf{R}} \rangle \quad (64)$$

and the fundamental gap is obtained by averaging the Born-Oppenheimer gap, $\tilde{\Delta}_{\mathbf{R}}$, over the distribution of the nuclei, baptized "quantum average" [80].

Note, that the definition of the fundamental gap, Eq. (62), is given in terms of thermodynamic quantities and does in general not coincide with the spectroscopic gap entering linear response functions given by the minimal difference between two energy eigenstates of the full electron-proton system. In particular, in the semi-classical region the spectroscopic gap is given by the minimum of the Born-Oppenheimer gap, $\tilde{\Delta}_{\mathbf{R}}$ with respect to the nuclear configurations \mathbf{R} . Thus, the resulting semi-classical gap is in general smaller than the average value of the fundamental gap, Eq. (64). However, such a semi-classical description loses its meaning when nuclear quantum effects become important. At zero temperature, averaging over nuclear configurations corresponds to taking the nuclear ground state expectation value, and the spectroscopic gap for electron/hole excitations will coincide in that limit with the quantum averaged fundamental gap, Eq. (64).

The above discussion can be extended for neutral (optical) gaps of particle-hole excitations with same number of electrons and nuclei, $N_e = N_n$ [81, 82]. However, electron-hole excitations eventually localize and bind together. Size effects on the neutral gap will eventually differ from those of the fundamental gap when the simulation cell reaches the size of the localization length and the electronic wave function can account for excitonic effects. Further, spatial localization of excitons may then also affect the Born-Oppenheimer energy surface of the excitation, differing from the ground state energy surface in this localized region, an effect not necessarily negligible in the thermodynamic limit. Such a modification is neglected when averaging the neutral Born-Oppenheimer excitation energies over the nuclear configurations of the electronic ground state energy surface, so that the resulting averaged gaps represent in general only upper bounds. Approaching the metallic state, such localization effects will vanish and not affect the determination of the insulator to metal transition.

In Fig. 2 are shown the results of a detailed study of the electronic excitation energies of phase I hydrogen at room temperature [82]. The fundamental (or quasi-particle) gap, Δ_{qp} , determined

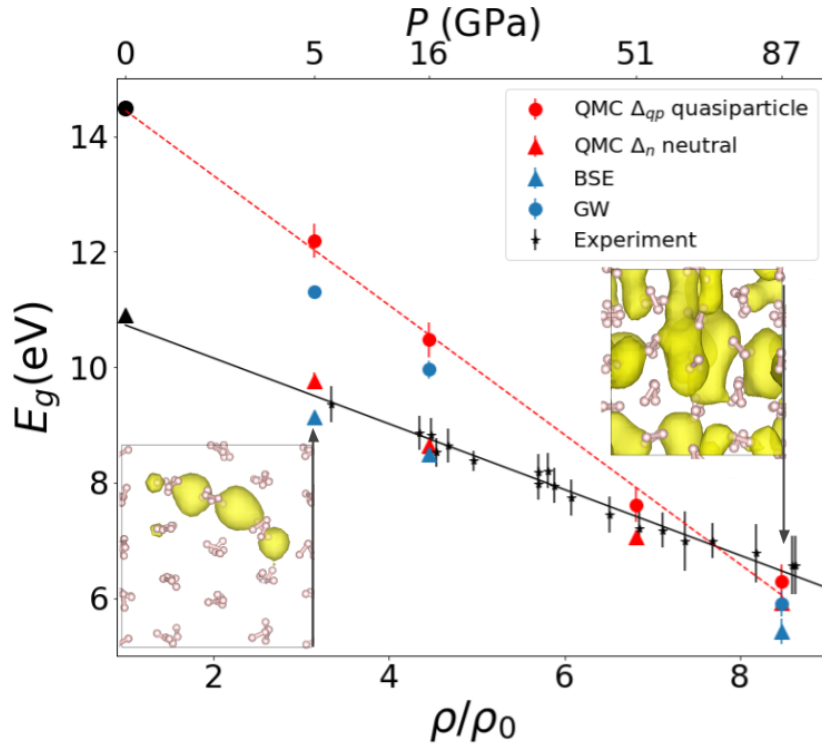


Fig. 2: Fundamental gap of quasi-particle (Δ_{qp}) and neutral (Δ_n) electronic excitations for phase I hydrogen at room temperature from QMC calculations and many-body perturbation theory (GW, BSE) [82]. Experimental values of the gap are determined from IXS spectra [83].

by Eq. (64), and the corresponding neutral gap Δ_n of particle-hole excitations from QMC calculations are compared with those obtained by many-body perturbation theory, charged excitations from GW, and neutral ones from solutions of the Bethe-Salpeter equation (BSE). In the pressure range between 5 and 90 GPa, the system changes from a wide-gap molecular insulator to a semiconductor. Differences between Δ_{qp} and Δ_n are due to excitonic effects and decrease with pressure, ~ 2 eV at 5 GPa down to ~ 0.5 eV at 90 GPa. Thermal and quantum nuclear motion reduces the gap by ~ 2 eV with respect to the ideal structure, roughly independent of pressure. Experimental values have been obtained by inelastic X-ray scattering (IXS) from the lower limit of the phonon energy-loss spectra [83].

4.2 Approaching metallic solid hydrogen

The metallization of crystalline hydrogen under pressure has been one of the driving forces of high pressure physics. Below 200 K, in phase III, experiments indicate a transition to a semimetallic state at 350 GPa [7], but synchrotron infrared spectroscopy measurements show that the direct gap remains open up to 425 GPa with an abrupt collapse attributed to metallization [8]. Structural information of phase III is indirect via vibrational spectroscopy. In Ref. [80], we have studied two candidate structures, Cmca-12 and C2/c-24, in CEIMC calculations at 200 K including nuclear thermal and quantum effects. The fundamental electronic gap has then been calculated for a subset of the sampled nuclear configurations using reptation

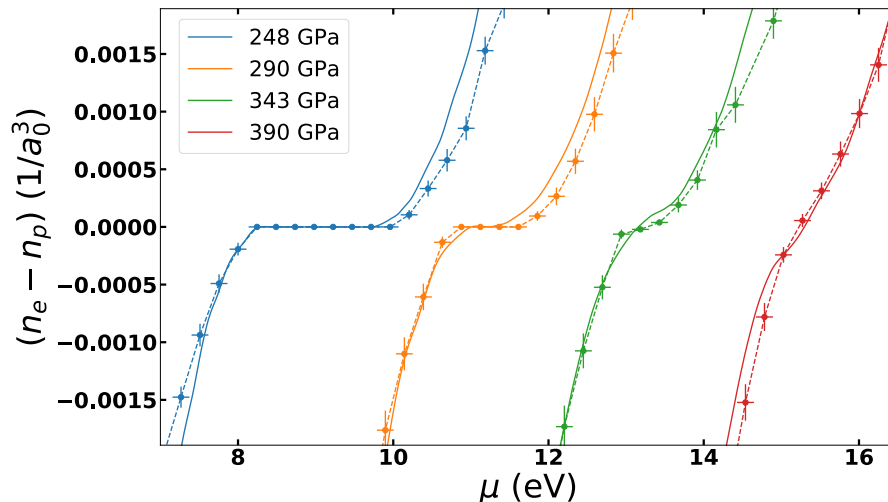


Fig. 3: Electronic excess density $n_e - n_p$ as a function of the chemical potential for solid hydrogen in the C2/c-24 structure (candidate for phase III) at 200 K from QMC using GC-TABC (points) compared to DFT-HSE (smooth lines) at various pressures [80]. The observed plateau when the electronic density coincides with the protonic one, $n_e = n_p$, at the lower pressures is a signature of the (indirect) gap. It is closing around $P \approx 370\text{--}380$ GPa.

Monte Carlo methods combined with grand-canonical twist averaging (GC-TABC) and leading and next-to-leading order finite size corrections.

Figure 3 shows the electronic excess density as a function of the electronic chemical potential in the C2/c-24 structure at 200 K. The incompressible behavior $\partial_\mu n_e = 0$ is characteristic for an insulator, the width of the plateau of the undoped system ($n_e = n_p$) corresponds to the fundamental gap.

Analysis of the electronic states at the edges of the gap shows that the corresponding fundamental gap is indirect [84]. The closure of the indirect gap occurs around 380 GPa for C2/c-24 (340 GPa for Cmca-12). Information on the direct gap can be obtained by unfolding the band-structure of the simulation cell. The direct gap remains open until ~ 450 GPa for C2/c-24 (500 GPa for Cmca-12). The calculations thus indicate a formation of a bad metal upon closure of the indirect gap, but the solid remains black (absorbing) until closure of the direct gap, a qualitative scenario supporting experimental observations.

4.3 Liquid-liquid phase transition

At higher temperatures, the solid melts to a molecular fluid at the low pressure side, whereas monatomic fluid hydrogen is expected at high pressure. Born-Oppenheimer molecular dynamics simulations using DFT and CEIMC calculations [85, 13, 86, 75] predict a first order liquid-liquid phase transition from the molecular to the atomic fluid below a critical temperature somewhere between 2000 and 3000 K.

The structural transition from the molecular to the atomic liquid is expected to coincide with an insulator to metal transition. Optical conductivity, reflectivity, and absorption can be computed within DFT methods [19], but a dependence on the underlying DFT functional remains.

Within CEIMC, the change of nature from the insulating to the metallic fluid can be detected in the reduced single-body density matrix of the many-electron ground state. Whereas the off-diagonal part, $n(r)$, decays exponentially for large distances r in the insulator, an algebraic decay $\sim r^{-3}$ indicates Fermi-liquid behavior of a metal with a sharp Fermi surface [18, 87]. Calculations of the fundamental electronic gap in the liquid [20] further confirm that molecular-atomic and insulator-metal transition occur together [87, 88].

Experiments probing the liquid-liquid phase transition have been performed with diamond anvil cells (DAC) [15] or by dynamic compression [16, 17], but temperature is not directly measured in the shock wave experiments and the interpretation of the data is sensitive to the adopted model. In Fig. 1, we show the transition from molecular to atomic hydrogen inferred from DAC experiments [15] in comparison with the CEIMC predictions [18].

However, recent calculations of the melting of the solid with a machine learned potential based on QMC Born-Oppenheimer energies [14] (not shown in Fig. 1) have questioned the experimental melting line inferred from Raman spectroscopy [12], predicting melting at considerably higher temperatures with the reentrant part getting very close to the liquid-liquid phase transition indicated by CEIMC. Further studies, theoretical and experimental, will be needed to firmly establish the phase diagram in that region.

4.4 Hugoniot adiabetic

Dynamic compression techniques can be used to experimentally probe the equation of state up to high pressures and temperatures [89]. Assuming that the shock is created by a planar surface, hydrodynamics relates the discontinuity in the energy, pressure, and density occurring at the shock by energy and momentum conservation by [90]

$$E(n, T) - E_0 + \frac{1}{2} (n^{-1} - n_0^{-1}) (P(n, T) + P_0) = 0 \quad (65)$$

where E_0 and n_0^{-1} are the initial values of the energy per atom and volume per atom at the equilibrium pressure P_0 and temperature T_0 before the shock. Only values of the equation of state which satisfy Eq. (65), the so-called shock or Hugoniot adiabetic, can be reached on the other side of the discontinuity surface with energy per particle $E(n, T)$ and pressure $P(n, T)$ at the volume per particle n^{-1} and temperature T .

Given the initial conditions, E_0 , n_0 , and P_0 , and knowledge of the equation of state in the fluid, $E(n, T)$ and $P(n, T)$, the Hugoniot adiabat then determines the line of P vs. n (or compression n/n_0) which can be compared to experimental measurements of P and n at the shock.

CEIMC calculations of the deuterium Hugoniot result in close agreement to measurements at low pressures [91], but predict a larger compression than experiments at higher pressure [91, 92]. The origin of the discrepancies remains still unclear as similar differences have been observed by independent calculations of QMC quality [93, 22].

Acknowledgments

I want to thank Carlo Pierleoni and David Ceperley for having introduced me to the physics of hydrogen more than twenty years ago and for ongoing collaborations since then.

References

- [1] M. Born and R. Oppenheimer, *Ann. Phys.* **389**, 457 (1927)
- [2] I.F. Silvera, *Rev. Mod. Phys.* **52**, 393 (1980)
- [3] J.M. McMahon, M.A. Morales, C. Pierleoni, and D.M. Ceperley, *Rev. Mod. Phys.* **84**, 1607 (2012)
- [4] J. Cheng *et al.*, *Nature* **573**, 558 (2019)
- [5] J. Cheng *et al.*, *Matter and Radiation at Extremes* **5**, 038401 (2020)
- [6] E. Wigner and H.B. Huntington, *J. Chem. Phys.* **3**, 764 (1935)
- [7] M.I. Eremets, A.P. Drozdov, P.P. Kong, and H. Wang, *Nat. Phys.* **15**, 1246 (2019)
- [8] P. Loubeyre, F. Occelli, and P. Dumas, *Nature* **577**, 631 (2020)
- [9] R.P. Dias and I.F. Silvera, *Science* **355**, 715 (2017)
- [10] P. Loubeyre, F. Occelli, and P. Dumas, arXiv:1702.07192
- [11] M.I. Eremets and A.P. Drozdov, arXiv:1601.04479
- [12] C.-S. Zha, H. Liu, J.S. Tse, and R.J. Hemley, *Phys. Rev. Lett.* **119**, 075302 (2017)
- [13] M.A. Morales, C. Pierleoni, E. Schwegler, and D.M. Ceperley, *Proc. Natl. Acad. Sci. USA* **107**, 12799 (2010)
- [14] H. Niu, Y. Yang, S. Jensen, M Holzmann, C. Pierleoni, and D.M. Ceperley, *Phys. Rev. Lett.* **130**, 76102 (2023)
- [15] M. Zaghoo, A. Salamat, and I.F. Silvera, *Phys. Rev. B* **93**, 155128 (2016)
- [16] 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, *Science* **348**, 1455 (2015)
- [17] P.M. Celliers, M. Millot, S. Brygoo, R.S. McWilliams, D.E. Fratanduono, J.R. Rygg, A.F. Goncharov, P. Loubeyre, J.H. Eggert, J.L. Peterson, N.B. Meezan, S. Le Pape, G.W. Collins, R. Jeanloz, and R.J. Hemley, *Science* **361**, 677 (2018)
- [18] C. Pierleoni, M.A. Morales, G. Rillo, M. Holzmann, and D.M. Ceperley, *Proc. Natl. Acad. Sci. USA* **113**, 4954 (2016)
- [19] G. Rillo, M.A. Morales, D.M. Ceperley, and C. Pierleoni, *Proc. Natl. Acad. Sci. USA* **116**, 9770 (2019)
- [20] V. Gorelov, D.M. Ceperley, M. Holzmann, and C. Pierleoni, *Phys. Rev. B* **102**, 195133 (2020)
- [21] G. Mazzola, R. Helled, and S. Sorella, *Phys. Rev. Lett.* **120**, 025701 (2018)

- [22] G. Tenti, K. Nakano, A. Tirelli, S. Sorella, M. Casula, Phys. Rev. B **110**, L041107 (2024)
- [23] M. Bonitz *et al.*, arXiv 2405.106238
- [24] P.P. Ewald, Ann. Phys. **64** (1921) 253.
- [25] D. Frenkel and B. Smit: *Understanding molecular simulation: from algorithms to applications* (Academic Press, San Diego, 2002)
- [26] M.P. Allen and D.J. Tildesley: *Computer Simulation of Liquids* (Oxford Univ. Press, 2017)
- [27] H. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959)
- [28] R. Feynman: *Statistical Mechanics: A Set of Lectures* (Benjamin, New York, 1972)
- [29] D.M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995)
- [30] M. Boninsegni, N.V. Prokofev, and B.V. Svistunov, Phys. Rev. E **74**, 036701 (2005)
- [31] D. Marx and J. Hutter, *Ab initio Molecular Dynamics: Basic Theory and Advanced Methods*, (Cambridge University Press, 2009)
- [32] A. Bijl, Physica **7**, 869 (1940)
- [33] R.B. Dingle, Phil. Mag. **40**, 573 (1949)
- [34] R. Jastrow, Phys. Rev. **98**, 1479 (1955)
- [35] E. Feenberg, *Theory of Quantum Fluids* (Academic Press, New York, 1969)
- [36] R.P. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956)
- [37] V.R. Pandharipande and N. Itoh, Phys. Rev. A **8**, 2564 (1973)
- [38] K.E. Schmidt and V.R. Pandharipande, Phys. Rev. B **19**, 2504 (1979)
- [39] M. Holzmann, D.M. Ceperley, C. Pierleoni, and K. Esler, Phys. Rev. E **68**, 046707 (2003)
- [40] M. Holzmann, B. Bernu, and D.M. Ceperley, Phys. Rev. B **74**, 104510 (2006)
- [41] M. Taddei, M. Ruggeri, S. Moroni, and M. Holzmann, Phys. Rev. B **91**, 115106 (2015)
- [42] M. Ruggeri, S. Moroni, and M. Holzmann, Phys. Rev. Lett. **120**, 205302 (2018)
- [43] M. Holzmann and S. Moroni, Phys. Rev. B **99**, 085121 (2019)
- [44] M. Holzmann and S. Moroni, Phys. Rev. Lett. **124**, 206404 (2020)
- [45] G. Pescia, J. Han, A. Lovato, J. Lu, and G. Carleo, space, Phys. Rev. Res. **4**, 023138 (2022)
- [46] X. Li, Z. Li, and J. Chen, Nat Commun. **13**, 7895 (2022)
- [47] M. Wilson, S. Moroni, M. Holzmann, N. Gao, F. Wudarski, T. Vegge, and A. Bhowmik, Phys. Rev. B **107**, 235139 (2023)

- [48] G. Cassella, H. Sutterud, S. Azadi, N.D. Drummond, D. Pfau, J.S. Spencer, and W.M.C. Foulkes, *Phys. Rev. Lett.* **130**, 036401 (2023)
- [49] H. Xie, Z.-H. Li, H. Wang, L. Zhang, and L. Wang, *Phys. Rev. Lett.* **131**, 126501 (2023)
- [50] G. Pescia, J. Nys, J. Kim, A. Lovato, and G. Carleo, arXiv:2305.07240
- [51] W.L. McMillan, *Phys. Rev.* **138**, A442 (1965)
- [52] D. Ceperley, G.V. Chester, and M.H. Kalos, *Phys. Rev. B* **16**, 3081 (1977)
- [53] D. Ceperley, *Phys. Rev. B* **18**, 3126 (1978)
- [54] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery: *Numerical Recipes* (Cambridge University Press, 1992)
- [55] H. Robbins and S. Monro, *Ann. Math. Stat.* **22**, 400 (1951)
- [56] F. Becca and S. Sorella, *Quantum Monte Carlo approaches for correlated systems* (Cambridge University Press, 2017)
- [57] G. Carleo and M. Troyer, *Science* **355**, 602 (2017)
- [58] D. Pfau, J.S. Spencer, A.G.D.G. Matthews, and W.M.C. Foulkes, *Phys. Rev. Res.* **2**, 033429 (2020)
- [59] J. Hermann, Z. Schätzle, and F. Noé, *Nat. Chem.* **12**, 891 (2020)
- [60] S. Baroni and S. Moroni, *Phys. Rev. Lett.* **82**, 4745 (1999)
- [61] C. Pierleoni and D. Ceperley, *Chem. Phys. Chem.* **6**, 1872 (2005)
- [62] B.L. Hammond, W.A. Lester Jr., and P.J. Reynolds: *Monte Carlo Methods in ab initio Quantum Chemistry* (World Scientific, Singapore, 1994)
- [63] R.M. Martin, L. Reining, and D.M. Ceperley, *Interacting Electrons* (Cambridge University Press, 2016)
- [64] M. Boninsegni and S. Moroni, *Phys. Rev. E* **86**, 056712 (2012)
- [65] D.M. Ceperley, *J. Stat. Phys.* **63**, 1237 (1991)
- [66] P.J. Reynolds, D.M. Ceperley, B.J. Alder, W.A. Lester, Jr., *J. Chem. Phys.* **77**, 5593 (1982)
- [67] W.M. Foulkes, L. Mitas, R.J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001)
- [68] G. Ortiz, D.M. Ceperley, and R.M. Martin, *Phys. Rev. Lett.* **71**, 2777 (1993)
- [69] C. Lin, F.-H. Zong, and D.M. Ceperley, *Phys. Rev. E* **64**, 016702 (2001)
- [70] S. Chiesa, D.M. Ceperley, R.M. Martin, M. Holzmann, *Phys. Rev. Lett.* **97**, 076404 (2006)
- [71] M. Holzmann, R.C. Clay, III, M.A. Morales, N.M. Tubman, D.M. Ceperley, and C. Pierleoni, *Phys. Rev. B* **94**, 035126 (2016)

- [72] M. Dewing and D.M. Ceperley, *J. Chem. Phys.* **110**, 9812 (1999)
- [73] C. Pierleoni, D.M. Ceperley, and M. Holzmann, *Phys. Rev. Lett.* **93**, 146402 (2004)
- [74] C. Pierleoni and D.M. Ceperley: *The Coupled Electron-Ion Monte Carlo Method* Lect. Notes Phys. **703**, 641 (2006)
- [75] C. Pierleoni, K.T. Delaney, M.A. Morales, D.M. Ceperley, and M. Holzmann, *Comp. Phys. Commun.* **179**, 89 (2008)
- [76] R.C. Clay, M. Holzmann, D.M. Ceperley, M.A. Morales, *Phys. Rev. B* **93**, 035121 (2016)
- [77] A. Tirelli, G. Tenti, K. Nakano, and S. Sorella, *Phys. Rev. B* **106**, L041105 (2022)
- [78] C. Attaccalite and S. Sorella, *Phys. Rev. Lett.* **100**, 114501 (2008)
- [79] Y. Yang, V. Gorelov, C. Pierleoni, D.M. Ceperley, and M. Holzmann, *Phys. Rev. B* **101**, 085115 (2020)
- [80] V. Gorelov, M. Holzmann, D.M. Ceperley, and C. Pierleoni, *Phys. Rev. Lett.* **124**, 116401 (2020)
- [81] V. Gorelov, Y. Yang, M. Ruggeri, D.M. Ceperley, C. Pierleoni, and M. Holzmann, *Condens. Matter Phys.* **26**, 33701 (2023)
- [82] V. Gorelov, M. Holzmann, D.M. Ceperley, C. Pierleoni, *Phys. Rev. B* **109**, L241111 (2024)
- [83] B. Li, Y. Ding, D.Y. Kim, L. Wang, T.-C. Weng, W. Yang, Z. Yu, C. Ji, J. Wang, J. Shu, J. Chen, K. Yang, Y. Xiao, P. Chow, G. Shen, W.L. Mao, and H.-K. Mao, *Phys. Rev. Lett.* **126**, 036402 (2021)
- [84] V. Gorelov, D.M. Ceperley, M. Holzmann, and C. Pierleoni, *J. Chem. Phys.* **153**, 234117 (2020)
- [85] S. Scandolo, *Proc. Natl. Acad. Sci. USA* **100**, 3051 (2003)
- [86] W. Lorenzen, B. Holst, and R. Redmer, *Phys. Rev. B* **82**, 195107 (2010)
- [87] C. Pierleoni, G. Rillo, D.M. Ceperley, and M. Holzmann, *J. Phys.: Conf. Ser.* **1136**, 012005 (2018)
- [88] C. Pierleoni, M. Holzmann, and D.M. Ceperley, *Contrib. Plasma Phys.* **58**, 99 (2018)
- [89] W. Nellis, *Rep. Prog. Phys.* **69**, 1479 (2006)
- [90] L.D. Landau and E.M. Lifshitz: *Fluid Mechanics* (Pergamon, Oxford, 1987)
- [91] M. Ruggeri, M. Holzmann, D.M. Ceperley, C. Pierleoni, *Phys. Rev. B* **102**, 144108 (2020)
- [92] N.M. Tubman, E. Liberatore, C. Pierleoni, M. Holzmann, and D.M. Ceperley, *Phys. Rev. Lett.* **115**, 045301 (2015)
- [93] R.C. Clay III, M.P. Desjarlais, and L. Shulenburger, *Phys. Rev. B* **100**, 075103 (2019)