Quantum Monte Carlo techniques and applications for warm dense matter

Ethan Brown, Miguel A Morales, Carlo Pierleoni, and David Ceperley

Abstract The Quantum Monte Carlo (QMC) method is used to study physical problems which are analytically intractable due to many-body interactions and strong coupling strengths. This makes QMC a natural choice in the warm dense matter (WDM) regime where both the *Coulomb coupling parameter* $\Gamma \equiv e^2/(r_s k_B T)$ and the *electron degeneracy parameter* $\Theta \equiv T/T_F$ are close to unity. As a truly first-principles simulation method, it affords superior accuracy while still maintaining reasonable scaling, emphasizing its role as a benchmark tool.

Here we give an overview of QMC methods including diffusion MC, path integral MC, and coupled electron-ion MC. We then provide several examples of their use in the WDM regime, reviewing applications to the electron gas, hydrogen plasma, and first row elements. We conclude with a comparison of QMC to other existing methods, touching specifically on QMC's range of applicability.

Ethan Brown
Lawrence Livermore National Laboratory, 7000 East Ave., Livermore CA, 94550 and Department of Physics, University of Illinois Urbana-Champaign, 1110 W. Green St. , Urbana, IL 61801-3080, USA e-mail: brown122@illinois.edu

Miguel A. Morales
Lawrence Livermore National Laboratory, 7000 East Ave., Livermore CA, 94550 e-mail: moralessilva2@llnll.gov

Carlo Pierleoni
Dipartimento di Scienze Fisiche e Chimiche, Universita dell'Aquila and CNISM, UdR dell'Aquila, a V. Vetoio 10, Loc. Coppito, I-67100 L'Aquila, Italy e-mail: carlo.pierleoni@aquila.infn.it

David Ceperley
Department of Physics, University of Illinois Urbana-Champaign, 1110 W. Green St. , Urbana, IL 61801-3080, USA e-mail: ceperley@illinois.edu

1 Introduction

With the development of high performance computing over the last few decades, simulation has become a ubiquitous many-body physics tool, see, for example, some of the other contributions in this volume. Most simulations are based on a classical description of atoms and molecules and use either a molecular dynamics (MD) or a Monte Carlo (MC) algorithm.

However, the microscopic description of warm dense matter (WDM) poses a particular difficulty, since classical mechanics is far from adequate. It is not convenient to describe the system as a perturbation from the ground state, that is as a sum over electronic excitations, since there are so many states. At a sufficiently high temperature, one needs a method that treats electrons and other quantum particles as particles and not as delocalized wavefunctions. Imaginary time path integrals provide a particular fortuitous formalism, since they supply a direct mapping of the quantum system into a classical system and reduce to the classical limit at high temperature. A big conceptual difficulty preventing the straightforward use of the path integral method is the issue of how to map fermion statistics into a probability. As we discuss below, this has an "in principle" solution: the restricted path integral method which, however, requires an ansatz to treat general fermion systems.

In this short review, we briefly describe three quantum Monte Carlo (QMC) methods, including diffusion MC, path integral MC, and coupled electron-ion MC, and their application to several WDM systems. We then close with a summary of strengths and weakness of the QMC methods.

1.1 Diffusion Monte Carlo

Diffusion Monte Carlo (DMC), along with the simpler Variational Monte Carlo (VMC), are two of the most popular stochastic methods employed in ground-state calculations of the Schrödinger equation for continuous systems. Even though they are zero temperature methods, they can be used in conjunction with other methods to treat regimes not dominated by electronic excitations [1, 2].

1.1.1 Formalism

As the name suggests, Variational Monte Carlo (VMC) is based on the variational principle of quantum mechanics. It states that the expectation value of the Hamiltonian, with respect to any trial wave function Ψ_T, will be a minimum for the exact ground state wave function:

$$E[\Psi_T] = \frac{\int \Psi_T^* \hat{H} \Psi_T d\mathbf{r}}{\int |\Psi_T|^2 d\mathbf{r}} = \int \pi(\mathbf{r}) E_L(\mathbf{r}) d\mathbf{r},$$

$$\pi(\mathbf{r}) = \frac{|\Psi_T(\mathbf{r})|^2}{\int |\Psi_T|^2 d\mathbf{r}}, \quad E_L(\mathbf{r}) = \frac{\hat{H}\Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})}, \tag{1}$$

where $\mathbf{r}$ is a $3N_e$-dimensional vector. The functional, $E[\Psi_T]$, provides an upper bound to the ground state energy of the system. The basic problem is finding flexible trial wave functions that are good approximations to the ground state. This typically requires some intuition and knowledge of the system.[1]

The Diffusion Monte Carlo (DMC) method [3], on the other hand, is based on the connection between quantum mechanics and classical diffusion. The time-dependent Schrödinger equation in imaginary time, β, becomes:

$$\frac{\partial \Psi(\mathbf{r},\beta)}{\partial \beta} = \lambda \nabla^2 \Psi(\mathbf{r},\beta) - V(\mathbf{r})\Psi(\mathbf{r},\beta). \tag{2}$$

Expanding $\Psi(\mathbf{r},\beta)$ in the eigenstates of the Hamiltonian it is easy to see that:

$$\begin{aligned} \Psi(\mathbf{r},\beta) &= \sum_i a_i \phi_i(\mathbf{r}) e^{-\beta \varepsilon_i}, \\ &\propto a_0\phi_0(\mathbf{r}) + \sum_i a_i \phi_i(\mathbf{r}) e^{-\beta(\varepsilon_i - \varepsilon_0)}. \end{aligned} \tag{3}$$

As β gets large, all exited states are exponentially suppressed. Imaginary time propagation projects the trial wave function to the ground state of the system.

The standard DMC algorithm is obtained by considering the mixed distribution $f(\mathbf{r},\beta) = \Psi(\mathbf{r},\beta)\Psi_T(\mathbf{r})$, where Ψ_T is any given trial wave function. If Eq. 2 is multiplied by Ψ_T and a few simple manipulations are performed, the partial differential equation for $f(\mathbf{r},\beta)$ is given by:

$$\begin{aligned} \frac{\partial f(\mathbf{r},\beta)}{\partial \beta} &= \hat{L} f(\mathbf{r},\beta), \\ &= \lambda \nabla^2 f(\mathbf{r},\beta) - \lambda \nabla \cdot [f(\mathbf{r},\beta)\mathbf{F}(\mathbf{r})] + [E_T - E_L(\mathbf{r})] f(\mathbf{r},\beta), \end{aligned} \tag{4}$$

where we introduced an overall shift in the energy. $\mathbf{F}$ is the quantum force defined by $\mathbf{F} = \nabla \ln |\Psi|^2$ and $E_L(\mathbf{r})$ is the local energy. Eq. 4 can be mapped to a drift-diffusion random walk with a branching (death/birth) process. The standard DMC algorithm combines a propagation of a population of 'walkers' using a Langevin equation with a branching step. After each Langevin step, each walker gains a weight in the population given by:

$$\omega_i(t) = e^{-\frac{\tau}{2}(E_L(\mathbf{r}_i(t)) + E_L(\mathbf{r}_i(t-\tau)) - 2E_T)}, \tag{5}$$

so that at time step k, the total weight of walker i is given by:

$$\omega_i(t_k) = \prod_{j=1}^{k} \omega_i(t_j), \tag{6}$$

[1] See §1.4 for more details on wave functions and optimization methods.

and averages are obtained from weighted sums:

$$< O >= \frac{\sum_{i=1}^{N}\sum_{p=1}^{N_{pop}} \omega_p(t_i) O(\mathbf{r}_p(t_i))}{\sum_{i=1}^{N}\sum_{p=1}^{N_{pop}} \omega_p(t_i)}. \tag{7}$$

This algorithm is typically called *pure* DMC.

Since the multiplicative nature of the weights lead to very large fluctuations and potential instabilities in the time propagation, a branching scheme is introduced where instead of storing weights during the simulation, walkers are stochastically replicated (or killed) using

$$N_{add}(t) = INT\left[e^{-\frac{\tau}{2}(E_L(\mathbf{r}_i(t))+E_L(\mathbf{r}_i(t-\tau))-2E_T)} + \gamma\right] \tag{8}$$

where γ is a uniformly distributed random number in $[0,1]$ and INT is the integer part of its argument. The role of E_T is now apparent. By appropriately choosing E_T, the population of walkers can be stabilized. In modern implementations, a combination of both weights and branching is used to improve efficiency [4]. Weights are maintained for a few steps, but their fluctuations are controlled by branching when they get outside of specified bounds.

1.1.2 Fermions

As mentioned previously, we must take special care to treat the antisymmetric nature of wave functions for fermions. Since the electronic Hamiltonian is invariant to particle exchange, the methods described above can be applied to fermions as long as we restrict them to the antisymmetric sector of the Hilbert space. The ground state wave function corresponds to the lowest antisymmetric eigenstate of the electronic Hamiltonian, which is not the global minimum of the entire Hilbert space. This means that any method that does not explicitly limit solutions to antisymmetric functions will converge to the incorrect result (e.g. to a bosonic state).

An antisymmetric wave function will contain both positive and negative regions in configuration space (assuming a real wavefunction). Since, a probability distribution has to be non-negative, we must include the signs in the estimators. This causes a severe problem because the normalization of the resulting average will be proportional to the difference between the positive and negative volumes in configuration space. Particle symmetry implies that those volumes must be equal and, since the sampling is bosonic, the normalization will go to zero as the simulation time increases.

This can be seen if we assume that we sample the states according to the absolute value of the original distribution. In this case, the estimator for the energy, for example, becomes:

$$< E >= \frac{\sum_i \omega(\mathbf{r}_i) E_L(\mathbf{r}_i)}{\sum_i \omega(\mathbf{r}_i)}, \tag{9}$$

where $\omega(\mathbf{r}_i)$ is $\{+1,-1\}$ depending on the sign of the wave function. As the number of samples increases, both numerator and denominator approach zero, and the signal/noise ratio decreases. It can be shown that the efficiency of this direct approach decreases exponentially with the number of particles [5]. This exponential scaling with the number of particles is known as the *Fermion Sign Problem*.

There is currently no exact solution to the sign problem for general systems. The standard approach to perform stable simulations is known as the fixed-node approximation [6].[2] In this approximation, the simulation is restricted to a nodal pocket of the trial wave function, and walkers are not allowed to cross nodes. The resulting energy represents an upper bound to the exact ground state energy, becoming exact when the nodes of the trial wave function coincide with those of the exact solution. The fixed node method combined with a projection technique, such as DMC, produces the best energy consistent with the chosen nodal structure. This makes results, in principle, sensitive to the nodal structure. In practice, results are found to be very accurate.

1.2 Path Integral Monte Carlo

Next we briefly describe path integral Monte Carlo (PIMC), which is similar to DMC but can treat systems at non-zero temperatures; a many-body density matrix replaces the trial wave function. For a more complete overview of the method and its application to fermion systems, see Refs. [5] and [7] respectively.

1.2.1 Path Integrals

To begin, we define the solution to the Bloch equation $d\rho/d\beta = -\hat{H}\rho$, the many particle density matrix, to be

$$\rho(R,R';\beta) = \langle R \mid e^{-\beta\hat{H}} \mid R' \rangle \tag{10}$$

where $R \equiv (r^{(1)},\ldots,r^{(N)})$ with $r^{(i)}$ specifying the spacial coordinates of the i^{th} of N particles, and $\beta \equiv 1/k_BT$, the inverse temperature. The quantum statistical partition function is defined as the trace of the density matrix,

$$Z(\beta) = Tr(\rho) = \int dR \langle R \mid e^{-\beta\hat{H}} \mid R \rangle = \int dR \rho(R,R;\beta). \tag{11}$$

The expectation value of any observable may be computed from this definition as

$$\langle \hat{\mathcal{O}} \rangle = Tr(\hat{\mathcal{O}}\rho)/Z = Tr(\hat{\mathcal{O}}\rho)/Tr(\rho). \tag{12}$$

[2] Note that this assumes real wave functions. If otherwise, the fixed-phase approximation may be used. [4]

Using the product property of the density matrix M times, such that $\beta = M\tau$, we discretize the partition function integral. The partition function becomes

$$Z(\beta) = \prod_{i=0}^{M-1} \int d^d R_i \rho(R_0, R_1; \tau)\rho(R_1, R_2; \tau)\ldots\rho(R_{M-1}, R_0; \tau). \tag{13}$$

Thus, we have effectively reduced the problem of finding a low temperature density matrix to one of finding the product of many high temperature density matrices. For each, we define an action

$$S(R_i, R_j; \tau) \equiv -ln[\rho(R_i, R_j; \tau)] \tag{14}$$

which may be broken into kinetic and potential parts. For systems with a long-range interaction, a variant of the Ewald summation technique is used [8]. Finally, in order to properly account for the particle statistics of the simulated system, we must sum over permutations $\mathscr{P}$, giving

$$Z(\beta) = \frac{1}{N!}\sum_{\mathscr{P}} (\pm 1)^{\mathscr{P}} \int_{R\to\mathscr{P}R} dR_t e^{-S[R_t]} \tag{15}$$

where the action S represents the path starting at R and ending at $\mathscr{P}R$.

1.2.2 Restricted Paths

Once again, with fermions negative terms enter in this sum, leading to a sign problem. As was done in the previous discussion of DMC, one way to circumvent this issue is to impose a nodal constraint [6]. We define the *nodal surface* $\Upsilon_{R_\star\beta}$ for a given point $R_\star$ and inverse temperature β to be

$$\Upsilon_{R_\star\beta} = \{R \mid \rho(R, R_\star; \beta) = 0\} \tag{16}$$

which is a $(dN-1)$-dimensional manifold in dN-dimensional configuration space. Here, $R_\star$ is dubbed the *reference point*, as it is needed to define the nodal surfaces. Inside a nodal cell, by definition the sign of the density matrix is uniform. Using Dirichlet boundary conditions, we may solve the Bloch equation within each nodal cell. We define the *reach* $\Gamma_\beta(R)$ as the set of all continuous paths R_τ, for which $\rho(R_\tau, R_\star, \beta) \neq 0$ for all intermediate τ ($0 < \tau \leq \beta$), i.e. a node-avoiding path.

$$\Gamma_\beta(R) = \{\gamma : R \to R' \mid \rho(R, R_\tau; \beta) \neq 0\}. \tag{17}$$

It is clear then that all paths contributing to the Bloch equation solution must belong to this reach. For all diagonal contributions, odd permutations must cross a node an odd number of times and thus are not allowed by this constraint.[3] In fact, they are exactly cancelled by all paths of node-crossing even permutations. This

[3] Note, for non-diagonal elements the sum over odd permutations must be retained.

leaves us with the following expression for the density matrix,

$$\rho(R,R;\beta) = \frac{1}{N!}\sum_{\mathscr{P},even}\int_{\gamma:R\to\mathscr{P}R}^{\gamma\in\Gamma_\beta(R)} \mathscr{D}R_\tau e^{-\mathscr{S}[R_\tau]/\hbar}. \tag{18}$$

We have thus turned the sign-full expression for the density matrix into one which includes only terms of a single sign, allowing efficient computation. However, because ρ appears on both sides of Eq. 18, this requires a priori knowledge of the density matrix nodal structure, which is generally unknown. To escape this self-consistency issue, an ansatz density matrix that approximates the actual nodal structure, is introduced. This ansatz is a solution to the Bloch equation inside the trial nodal cells and obeys the correct initial conditions, providing an exact solution only when its nodes are the true nodes of the density matrix. This method is often called *restricted* PIMC (RPIMC).

For spinless fermions, it is often simplest to write down the anti-symmetrized density matrix as the Slater determinant of single-particle distinguishable density matrices, $\rho(R,R_\star;\beta) = \frac{1}{N!}\det\rho_{ij_\star}$ where

$$\rho_{ij_\star} = (4\pi\lambda\beta)^{-d/2} e^{-\frac{(r_i - r_{j\star})^2}{4\lambda\beta}}. \tag{19}$$

Generally, we expect this approximation to be best at high temperature and when correlation effects are weak. Furthermore, due to the constraint of translational invariance, free particle nodes are quite reasonable for homogeneous systems, specifically in the WDM regime. As a simple illustration, Fig. 1 shows the exact, the ground state, and the free particle nodal structures for a system of three particles in two dimensions. One can see that in the classical limit, the exact nodes match those of free particles, while in the $T \to 0$ limit, the nodes are that of the ground state.

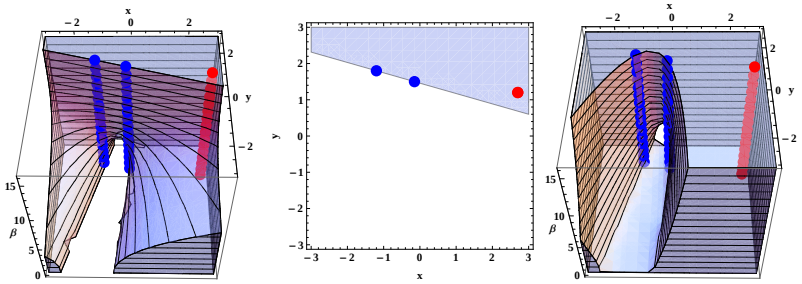

Fig. 1 An example nodal surface for three particles in two dimensions. The nodal line cuts through two of the particles (blue), defining the reach of the third (red). Here all time slices occupy the same point in space, showing how the exact nodal structure (a) is the precise extrapolation from the ground state nodal structure (b) to those of free particles (c). The third axis in (a) and (c) represents β, the inverse temperature.

1.2.3 Possible Sources of Error

We now discuss some of the sources of error. The largest is caused by the finite-size of the simulation cell. This error may be controlled either by performing multiple-sized simulations and extrapolating to the thermodynamic limit or by using an analytic correction. For homogeneous systems with long-range interactions, we can use the finite temperature extension of the analytic correction given in Ref. [9] as we discuss in §1.4.

The second largest error comes from the discretization of imaginary time by τ. While the naive Trotter breakup of $\exp[-\hat{S}] \approx \exp[-\hat{T}]\exp[-\hat{V}]$ has error $\sim \tau^2$, one can often do better. Specifically, we employ the *pair product approximation* for the action which has error $\sim \lambda\tau^3$ [5]. For fermions a time-step error arises from the nodal constraint when paths cross and recross a node within τ. Though these contributions may be difficult to separate, one can always extrapolate to the $\tau \to 0$ limit by adding more time slices. The third controllable error is due to the statistical fluctuations of the Monte Carlo algorithm itself. However, as in DMC, one may run the simulation longer to provide smaller errors on measured quantities.

The final source of error, the nodal error, is the most problematic since it is uncontrollable. As mentioned, there is no a priori way to know the exact nodal structure. PIMC's variational principle is through the free energy, as opposed to the internal energy in DMC. Thus one possible solution is parameterize the nodal ansatz, and then minimize the free energy by varying the parameters. However, the free energy is not usually easily obtainable; it may require a thermodynamic integration. Nevertheless, in the systems mentioned below, estimates suggest that the nodal error arising from the free-particle ansatz to be smaller than the statistical error.

1.3 Coupled Electron Ion Monte Carlo

In applying QMC methods to WDM, systems of nuclei and electrons must be considered. The large nucleon-electron mass ratio implies a wide separation of time scales. Because of this, first principles methods based on Density Functional Theory (DFT), find the adiabatic, or Born-Oppenheimer (BO), approximation to be very useful even for the lightest element, hydrogen. Conversely, the QMC methods described so far ignore this separation and treat nuclei and electrons on equal footing, causing difficulties. The imaginary time step of the path integral representation (both in DMC and PIMC) is imposed by the light electron mass. In DMC this means that nuclear "dynamics" (the speed of sampling configuration space) is much slower than electron "dynamics" requiring very long (and time consuming) trajectories. In PIMC the separation of time scales presents itself as a separation in the regions where thermal effects are relevant: in high pressure hydrogen for instance nuclear quantum effects becomes relevant below $\sim 1000K - 5000K$ where electrons are, to a very good approximation, in their ground state. Performing PIMC in this region of temperatures requires very long electronic paths causing a slowing down of the

exploration of configuration space and effectively limiting the ability of PIMC to perform accurate calculations at low temperatures.

To study this interesting temperature region, a QMC method based on the BO approximation, the Coupled Electron-Ion Monte Carlo (CEIMC) has been developed [1]. In CEIMC a Monte Carlo calculation for finite temperature nuclei (either classical or quantum represented by path integrals) is performed using the Metropolis method with the BO energy obtained by a separate QMC calculation for ground state electrons. CEIMC has been extensively reviewed in refs. [1, 10]. Here, we only briefly report the main technical features of the method.

1.3.1 Penalty method

In CEIMC the difference of BO energies of two nearby nuclear configurations in a MC attempted step, as obtained by an electronic QMC run, is affected by statistical noise which, if ignored, results in a biased nuclear sampling. To cope with this situation either the statistical noise needs to be reduced to a negligible value by long electronic calculations (very inefficient), or the Metropolis acceptance/rejection scheme needs modifications to cope with noisy energy differences. The latter strategy is implemented in the Penalty Method [11] which enforces detailed balance to hold on average over the noise distribution. The presence of statistical noise causes an extra rejection with respect to the noiseless situation. An extra "penalty" defined as the variance of the energy difference over the square of the physical temperature is added to energy differences. Therefore running at lower temperatures requires a reduced variance to keep an acceptable efficiency of the nuclear sampling. Small variances can be obtained if correlated sampling is used to compute the energy of the two competing nuclear configurations. In an attempted nuclear MC step, a single electronic run is performed with a trial wave function which is a linear combination of the wave functions of the two nuclear configurations considered. The BO energy of the two nuclear configurations is obtained by a reweighting procedure which provides energy differences with a much reduced variance with respect to performing two independent electronic runs if the "distance" between the two nuclear configurations is limited (i.e. the overlap between the trial wave functions of the two configurations is large) [12]. This strategy allows an efficient sampling of nuclear configuration space for high pressure hydrogen and helium down to temperature as low as $\sim$100-200K.

1.3.2 PIMC

When nuclear quantum effects are included using a path integral representation (see §1.2), the relevant inverse temperature in the penalty method is the imaginary time discretization step τ, so that no loss of efficiency is experienced when lowering the temperature (i.e. taking longer paths). For quantum protons in high pressure hydrogen, CEIMC can be used to efficiently study systems at temperatures as low

as $\sim$50K. In the present implementation of nuclear quantum effects in CEIMC, we introduce an effective pair potential between nuclei and use the pair density matrix (see §1.2). The residual difference between the effective pair potential and the BO energy is considered at the primitive approximation level of the Trotter break-up of the proton propagator [1]. In high pressure hydrogen ($r_s = 1.40$) it is found that with this strategy, $\tau^{-1} \simeq$4500K is enough to reach convergence of the thermodynamics properties, which allows to study systems at low temperature with a limited number of time slices ($\leq$100).

1.3.3 VMC vs RQMC

The main ingredient of CEIMC is the electronic QMC engine used to compute the BO energy. As mentioned a very important aspect for the efficiency of CEIMC is the noise level which is related to the variance of the local energy. In ground state QMC (see §1.1) the "zero variance principle" applies: if the trial wave function is an eigenfunction of the Hamiltonian, the local energy is no longer a function of the electronic coordinates and a single calculation provides the exact corresponding eigenvalues. Therefore by improving the trial wave function and approaching the exact ground state, the variance of the local energy decreases to zero. In connection with CEIMC, this is important not only for the accuracy of the BO energy but also for the efficiency of the nuclear sampling since the extra rejection due to the noise is reduced for a more accurate trial wave function. For high pressure hydrogen we have devised a quite accurate trial function of the Slater-Jastrow, single determinant, form. The Slater determinant is built with single electron orbitals, obtained by a self-consistent DFT solution, expressed in terms of quasi-particle coordinates with a backflow transformation. The Jastrow part has an electron-proton and electron-electron RPA Jastrow plus two-body and three-body numerical terms [13, 14].

To go beyond VMC accuracy in CEIMC we have implemented the so called Reptation QMC method (RQMC) [15, 1]. The theoretical basis of RQMC and DMC are the same, however their implementations are quite different. In DMC a generalized drift-diffusion-source equation is solved which represents the Schroedinger equation in imaginary time. An initial population of replicas of the system is propagated in imaginary time, and the projected wave function at a given time is represented by the population of replicas at that time. In RQMC a single replica is considered, and the entire evolution of the system is represented by an imaginary time path which, as in PIMC, is kept in the memory of the computer. The problem is therefore transformed into sampling the configuration space of such paths which can be accomplished using Metropolis Monte Carlo. Expectations of the ground state wave function are calculated in the middle of the path [1]. Since RQMC uses an explicit representation of the statistical weight of each path, the reweighting procedure needed for estimating energy differences in CEIMC can be more easily applied in RQMC than in DMC; this is the reason for our choice of RQMC in CEIMC.

1.4 Improved wavefunctions and density matrices

The simplest and most common trial wave function used in QMC calculations is the Slater-Jastrow form, $\Psi_T(\mathbf{r}) = D^{\uparrow}D^{\downarrow}e^{-J}$, where $D^{\uparrow}$ is a Slater determinant of single particle states and J is a Jastrow factor. The orbitals in the determinant are typically obtained from a mean-field method, while the Jastrow term incorporates correlation directly and typically includes terms involving electron-ion, electron-electron, and electron-electron-ion terms. Over the last decade, many wave function forms have been explored. Some of these include: Pfaffians [16], generalized valence bond (GVB) [17], antisymmetrized geminal power (AGP) [18], backflow transformation [19], and multi-determinant expansions [20, 21, 22, 23]. In particular, the multideterminant form has shown great promise in molecular calculations, leading to wave functions with controlled accuracy capable of reaching chemically accurate results, and competing with the best methods in quantum chemistry. This wave function has the form: $\Psi_T(\mathbf{r}) = e^{-J}\sum_k c_k D_k^{\uparrow}D_k^{\downarrow}$, the linear parameters, c_k, as well as all the variational parameters in the Jastrow are optimized using VMC.

Over the last decade, robust and efficient wave function optimization methods have been developed [24, 25]. Among the most notable is the linear optimization method of Umrigar and co-workers [26]. In this case, the Hamiltonian is diagonalized in the sub-space obtained by linearizing the wave function derivatives with respect to its variational parameters. For linear parameters, as in the multi-determinant expansion, the resulting sub-space is complete, and the solution to the eigenvalue equations is the lowest energy solution. Rescaling of the lowest energy eigenvector is allowed when non-linear parameters are included, either by a line minimization, controlling the normalization, or any other criterion, to speed the convergence to the minimum. Typically, fewer than ten iterations are required, each using an increasing number of Monte Carlo steps. The continuous development of robust wave function optimization methods with the implementation and use of flexible trial wave function forms has lead to a large improvement in the accuracy of QMC methods over the last several years.

Density matrices may be improved in the same fashion, though all proper breakups of the action will converge to the exact density matrix given small enough time step τ. Several approximations which include higher-order electronic correlations have been proposed, though generally the pair product approximation is chosen as the best compromise between accuracy and efficiency [5]. For fermions, however, there is still a need to devise better nodal density matrices. At temperatures above zero, QMC methods obey a variational principle in the free energy (as opposed to the internal energy at zero-temperature). To this end, there have been some recent efforts to improve the nodal ansatz used in PIMC by minimizing the free energy [27]. The inclusion of backflow effects has also been proposed, though not tested, as a possible way forward. The literature discusses several recent attempts to improve nodal surfaces [28, 27].

1.5 Finite size effects and extrapolations

One of the main approximations in simulations is to reach the thermodynamic limit by doing simulations with a microscopic system. Simulations are limited today to using fewer than several thousand electrons; typically this is used to model a physical system which contains enormously more electrons. There has been significant progress on this during the past decade. For WDM these methods are particularly appropriate since particle correlations are, for a large part, local, and if the correct boundary conditions are applied, a supercell can mimic an infinite system. To minimize the finite size errors, periodic boundary conditions are used. The dependence of the energy on the size of the simulation cell is often the largest systematic error in QMC calculations. Finite size effects are larger in QMC than in mean field methods because electrons are explicitly represented.[4] Fortunately, they can be minimized by several methods. We will discuss a recent analysis of finite system effects, first of the kinetic energy, and second of the potential energy.

If calculations are done with periodic boundary conditions, the largest effect for systems with a Fermi surface at low temperture comes from the kinetic energy, arising from the discontinuous filling of the shells in k-space. In the original calculations of the HEG, calculations were done with a sequence of closed shells, and Fermi Liquid theory was used to extrapolate to the thermodynamic limit [29]. Although this works well for homogeneous systems with small unit cells, it does not work as well for more complex systems, having larger cells, or in a liquid. General boundary conditions on the phase are:

$$\Psi(\mathbf{r}_1+\mathbf{L},\mathbf{r}_2,\ldots)=e^{i\theta}\Psi(\mathbf{r}_1,\mathbf{r}_2,\ldots) \tag{20}$$

where $\mathbf{L}$ is a lattice vector of the supercell. If the twist angle θ is averaged over, (twist averaged boundary conditions or TABC), most single-particle finite-size effects, arising from shell effects in filling the plane wave orbitals, are eliminated [30]. In fact, within the grand canonical ensemble, there are no single-particle kinetic finite size effects. One can understand this in terms of the momentum distribution since the kinetic energy per electron may be written as

$$T_N/N=\frac{(2\pi)^3}{2\Omega}\sum_{\mathbf{k}}k^2\rho_N(k) \qquad T_\infty=\frac{1}{2}\int d\mathbf{k}k^2\rho_\infty(k) \tag{21}$$

where ρ_N and ρ_∞ are the momentum distribution for an N electron and infinite systems respectively. Twist averaging gets rid of the largest approximation, namely the "sampling" of $\mathbf{k}$-space, all the while still using $\rho_N(\mathbf{k})$.

The potential energy per electron is the integral over the two particle correlation function, conveniently expressed in k-space in terms of the structure factor (S_k):

$$V_N=\sum_{\mathbf{k}}V_N(\mathbf{k})S_N(k) \qquad V_\infty=\int d\mathbf{k}V_\infty(\mathbf{k})S_\infty(k) \tag{22}$$

[4] All exact methods that work with finite systems share this difficulty.

Here $V_N(k) = V_\infty(k) = 4\pi k^{-2}$ is the Fourier transform of the Coulomb interaction, and we have assumed that the Ewald image procedure is used to treat the long-range interaction in a periodic system. As with the expression for the kinetic energy, we assume that $S_N(\mathbf{k}) = S_\infty(\mathbf{k})$. However, twist averaging does not convert the sum in Eq. 22 to an integral since $S(k)$ is a two-particle quantity; the values of $\mathbf{k}$ that can be computed are all in the reciprocal lattice of the supercell no matter what the twist angle. The dominant finite size error comes from the contribution near $\mathbf{k} = 0$. For small values of $\mathbf{k}$ we have that $S(\mathbf{k}) = ck^2$. Integrating over the sphere $k < \pi/L$ (assuming a cubic supercell) one gets a correction $4\pi k_0^3 c$. This procedure was introduced in ref. [31], and thorough analysis and comparison with other methods was performed in ref. [32].

2 Applications

With the above methods, the QMC community has been able to study several systems in WDM. Here provide recent calculations of the electron gas, hydrogen, helium, carbon, and water, and address their importance to WDM research.

2.1 Electron Gas

The one-component plasma (OCP), a fundamental many-body model, consists of a single species of charged particles immersed in a rigid neutralizing background. For electrons, the OCP is a model of simple metals and is often called the homogeneous electron gas (HEG) or *jellium*. It is customary to define the natural length scale $r_s a_B \equiv (3/4\pi n)^{1/3}$ and energy scale $Ry = e^2/2a_B$, where n is the system density. When r_s, the *Wigner-Seitz radius*, is small (high density) ($r_s \to 0$), the kinetic energy term dominates and the system becomes qualitatively similar to a non-interacting gas. At low density ($r_s \to \infty$), the potential energy dominates and the system is predicted to form a Wigner crystal [33]. In 3D at intermediate densities, a partially polarized state is predicted to emerge [34, 35].

Over the past few decades QMC calculations of the ground state HEG examined each of these phases [3, 36]. In addition to determining phase boundaries, the results of these studies have proven invaluable in the rigorous parameterization of functionals in ground state density functional theory (DFT) [37]. In the opposite limit, the classical OCP, whose equation of state depends only on the *Coulomb coupling parameter* $\Gamma \equiv e^2/(r_s k_B T)$, has been studied extensively [38, 39, 40]. However, the accuracy of these results quickly deteriorates as the temperature is lowered and quantum correlations play a greater role [41]. This breakdown is most apparent in the WDM regime where both Γ and the *electron degeneracy parameter* $\Theta \equiv T/T_F$ are close to unity.

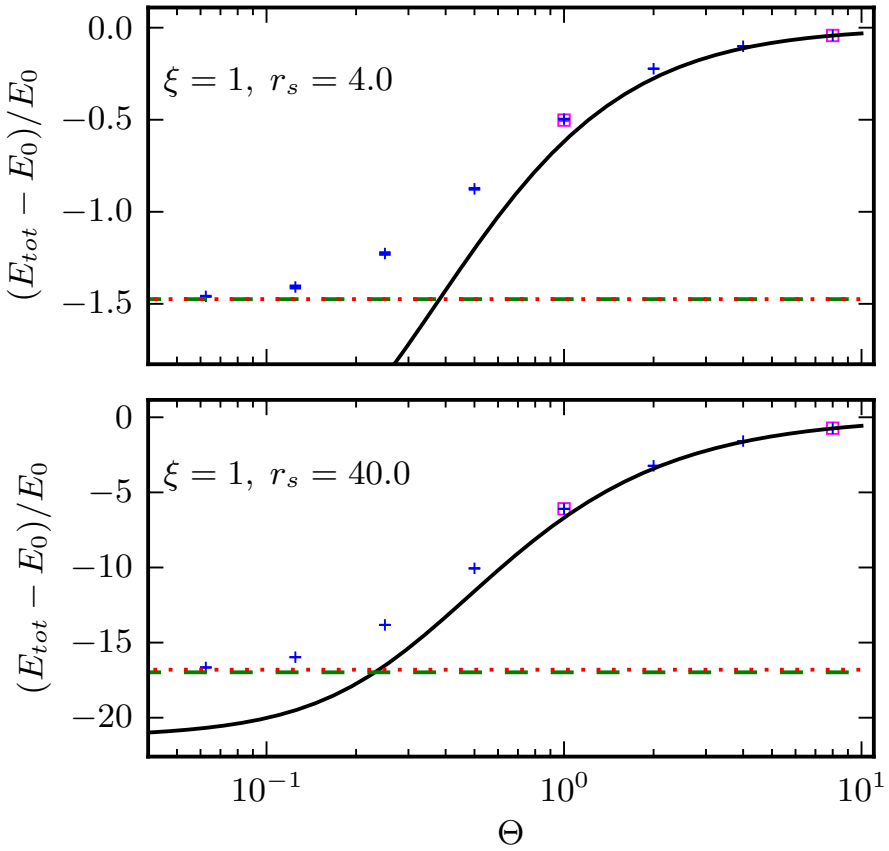

Fig. 2 (color online) Excess energies for $r_s = 4.0$ (top) and $r_s = 40.0$ (bottom) for the polarized state. For both densities, the high temperature results fall smoothly on top of previous Monte Carlo energies for the classical electron gas [38] (solid line). Differences from the classical coulomb gas occur for $\Theta < 2.0$ for $r_s = 4.0$ and $\Theta < 4.0$ for $r_s = 40.0$. Simulations with the Fermion sign (squares) confirm the fixed-node results at $\Theta = 1.0$ and 8.0. The zero temperature limit (dotted line) smoothly extrapolates to the ground-state QMC results of Ceperley-Alder [3] (dashed line).

There have been several attempts to tackle this strongly interacting regime through both analytics (diagrammatic techniques, perturbative expansions, RPA, etc.) and numerics (DFT, hypernetted-chain calculations, etc.). However, without a reliable first principles benchmark, these efforts have gone unverified, except in the weakly-interacting limits. Ref. [42], using the RPIMC method (see 1.2), provides such accurate, first-principles thermodynamic data for the 3*D* HEG throughout the WDM regime for both the fully spin-polarized $\xi = 1$ and unpolarized $\xi = 0$ systems. This reference makes connections to both previous semi-classical and ground-state studies, as well as matches on to *exact*, signful calulations where possible. In the proceeding, we briefly touch on those results, as well as provide a new comparison to existing finite temperature energy parameterizations.

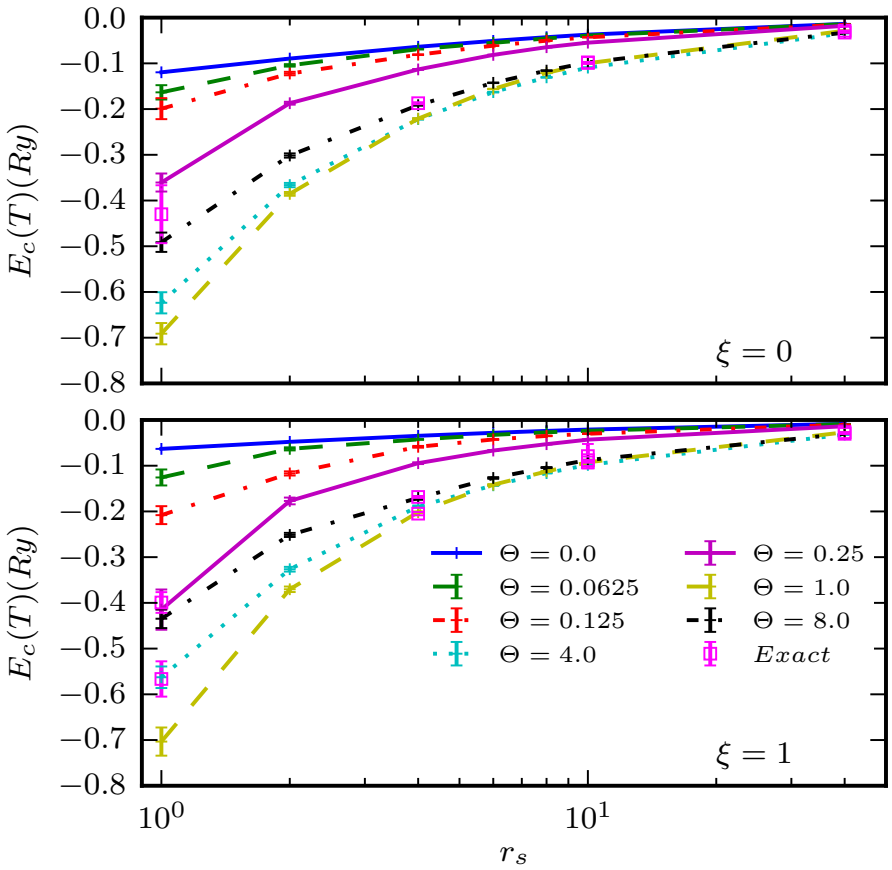

Fig. 3 (color online) Correlation energy $E_c(T)$ of the 3*D* HEG at several temperatures and densities for the unpolarized (top) and fully spin-polarized (bottom) states. *Exact* (signful) calculations (squares) confirm the fixed-node results where possible. For comparison, we plot the $\Theta = 0.0$ correlation energy used in LDA DFT calculations.

For densities ranging from $r_s = 1.0$ to 40.0 and temperatures ranging from $\Theta \equiv T/T_F = 0.0625$ to 8.0, energies, pair correlation functions, and structure factors were computed.[5] At each density, we observe a smooth convergence to previous semi-classical studies [38] at high temperature. In Fig. 2 we plot the total excess energy $(E_{tot} - E_0)/E_0)$ for the polarized system at all temperatures with $r_s = 4.0$ and 40.0. At the highest temperatures, the results match well with the purely classical Monte Carlo results of Ref. [38] (solid line). For a few select points, the much more time-consuming but more accurate, signful PIMC simulation (squares) was done. These points which are essentially exact, i.e. without possible nodal error, match well with fixed-node results. Finally, we know from Fermi liquid theory the

[5] All data can be found in a repository hosted at http://github.com/3dheg/3DHEG.

low-temperature gas should have a linear form for the heat capacity, and therefore a quadratic form for the internal energy. Thus for each density we fit the low-temperature points to a quadratic function and extrapolate to $0T$. Fig. 2 shows the extrapolated results (dotted line) match well with the zero temperature QMC results of Ceperley-Alder [3] (dashed line).

In Ref. [42], was found similar limiting agreement in both the calculated structure factors and pair correlation functions. Through this comparison against existing numerical and analytical data, we conclude the free-particle nodal approximation performs well for the densities studied. Further investigation is needed at even smaller values of r_s and lower temperatures in order to determine precisely where this approximation begins to fail. Such studies will necessarily require algorithmic improvements, however, because of difficulty in sampling paths at high density and low temperature [43].

Finally, we have evaluated the exchange-correlation energy E_{xc}, an essential quantity in any DFT formulation, defined

$$E_{xc}(T) \equiv E_{tot}(T) - E_0(T) \tag{23}$$

where E_0 is the kinetic energy of a free Fermi gas at temperature T. As is customary, we further break up E_{xc} into exchange and correlation parts,

$$E_{xc}(T) = E_x(T) + E_c(T) \tag{24}$$

where $E_x(T)$ is the Hartree-Fock exchange energy for a free Fermi gas at temperature T.

By calculating $E_{tot}(T)$ through RPIMC simulations $E_c(T)$ was determined at all studied densities and spin-polarizations. As one can see in Fig. 3, correlation effects increase both with density (smaller r_s) and temperature up to the Fermi temperature T_F. Above this temperature, the electron gas begins to be less correlated. This represents the point at which electron screening is a dominant effect, the interaction becomes effectively short-ranged, and the Debye-approximation becomes relatively accurate [44]. As the density increases, the relative temperature at which this occurs decreases. At $r_s = 1.0$ the maximal effect of interactions occurs very near T_F, $\Theta = 1$. Most notably, we see a departure from the $T = 0.0$ correlation energy used ubiquitously in both ground state and finite temperature LDA DFT calculations. This discrepancy is significant throughout the WDM regime, calling into question the use of ground state correlation functionals at such temperatures and densities.

We conclude with a brief comparison of these calculations to previous estimations of the $3D$ HEG correlation energy. Such parameterizations generally fall into two categories: those which extend down from the classical regime and those which assume some interpolation between the $T = 0$ and high-T regimes. From the former group, in Fig. 4, we plot E_c coming from Debye-Huckel (DH) theory which assumes short-ranged interactions and the classical Monte Carlo simulations of Hansen, *et. al.* of the full Coulomb system both with Wigner-Kirkwood corrections (H+WK) and without (H). Clearly these methods do not perform well in the quantum regime below the Fermi temperature since they lack quantum exchange.

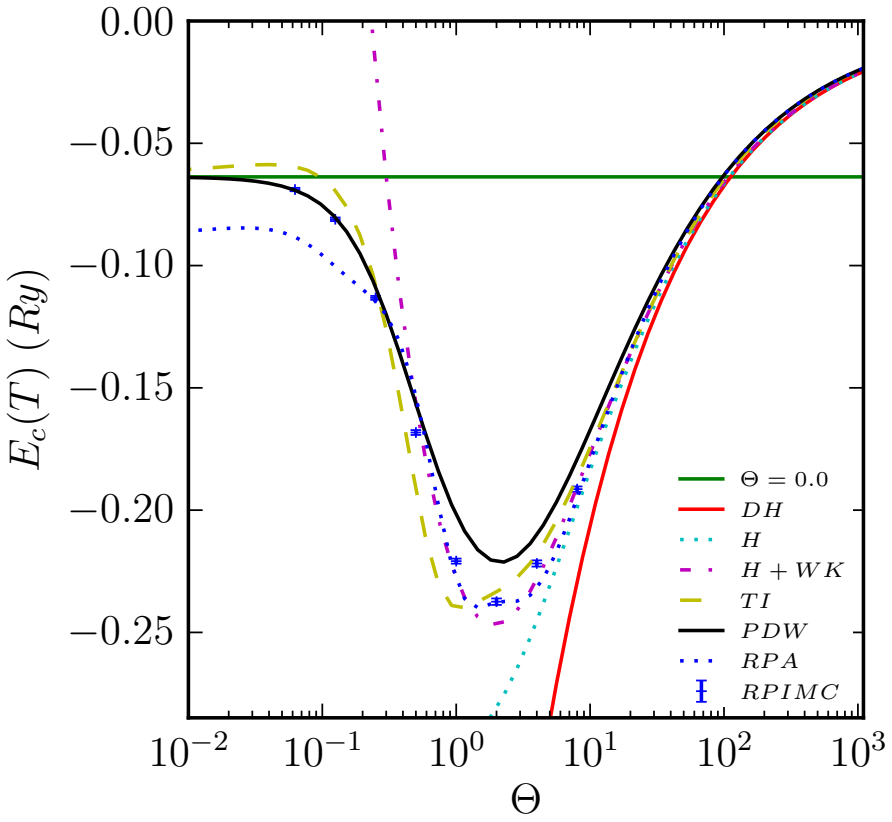

Fig. 4 (color online) Correlation energy $E_c(T)$ of the $3D$ HEG at $r_s = 4.0$ for the unpolarized $\xi = 0$ state from the above calculations (RPIMC) and several previous parameterizations. The latter include Debye-Huckel (DH), Hansen (H), Hansen+Wigner-Kirkwood (H+WK), random phase approximation (RPA), Tanaka and Ichimaru (TI), and Perrot and Dharma-wardana (PDW). We also include the ground state $\Theta = 0.0$ result for comparison.

The random phase approximation (RPA), from the latter group, is a reasonable approximation in the low-density, high-temperature limit (where it reduces to DH) and the low-temperature, high-density limit, since these are both weakly interacting regimes. Its failure, however, is most apparent in its estimation of the equilibrium, radial distribution function $g(r)$ which becomes negative for stronger coupling. Extensions of the RPA into intermediate densities and temperatures have largely focused on constructing local-field corrections (LFC) through interpolation since diagrammatic resummation techniques often become intractable in strongly-couplled regimes. Singwi, *et. al.* [45] introduced one such strategy relying on two assumptions. First, they use the static polarization-potential approximation allowing one to write the LFC $G(k,z) \equiv G(k,z=0) \equiv G(k)$. Next they assume the two-particle distribution function is a function of the $n(r)$ and $g(r)$ which allows a self-consistent solution for $G(k)$. Tanaka and Ichimaru [46] (TI) extended this method to finite

temperatures and provided the shown parameterization of the $3D$ HEG correlation energy. A similar method (not shown) by Dandrea, *et. al.* uses the Vashista-Singwi LFC [47] to interpolate between the high- and low-temperature limits. Both methods appear to perform marginally better than the RPA at all temperatures, though both still fail to produce a positive-definite $g(r)$ at values of $r_s > 2$.

A third, more recent approach introduced by Perrot and Dharma-wardana (PDW) [48] relies on a classical mapping where the distribution functions of a classical system at temperature T_{cf}, solved for through the hypernetted-chain equation, reproduce those for the quantum system at temperature T. In a previous work, PDW showed such a temperature T_q existed for the classical system to reproduce the correlation energy of the quantum sytem at $T = 0$ [49]. To extend this work to finite temperature quantum systems, they use the simple interpolation formula $T_{cf} = \sqrt{T^2 + T_q^2}$. This interpolation is clearly valid in the low-T limit where Fermi liquid theory gives the quadratic dependence of the energy on T. Further in the high-T regime, T dominates over T_q as the system becomes increasingly classical. The PDW line in Fig. 4 clearly matches well with the RPIMC results in these two limits. It is not surprising, however, that in the intermediate temperature regime, where correlation effects are greatest, the quadratic interpolation fails. A contemporary, but similar approach by Dutta and Dufty [50] uses the same classical mapping as PDW which relies on matching the $T = 0$ pair correlation function instead of the correlation energy. While we expect this to give more accurate results near $T = 0$, we would still expect a breakdown of the assumed Fermi liquid behavior near the Fermi temperature.

Future electron gas work will include creating a new parameterization of the correlation energy which uses the above data directly. In doing so, simulations at higher densities and both lower and higher temperatures may be necessary in order to complete the interpolation between the ground state and classical limits.

2.2 Hydrogen

Hydrogen is the simplest element of the periodic table and also the most abundant element in the Universe. Together with its isotopes, deuterium and tritium, it occupies a special place in WDM since it is particularly relevant in astrophysics, planetary physics and nuclear energy applications.

Because of its simple electronic structure, it has been instrumental in the development of quantum mechanics and remains important for developing ideas and theoretical methods. QMC is not an exception, and a DMC investigation of the ground state of hydrogen across the pressure-induced molecular dissociation regime appeared as early as 1987 [29], followed by more recent and accurate studies [51, 52]. Few years later RPIMC was developed to investigate the WDM regime of hydrogen [53, 54, 55] and to study the primary and secondary Hugoniots [56, 57]. RPIMC predictions for the principal Hugoniot of deuterium were first in disagreement with pulsed laser-produced shock compression experiments [58, 59, 60], but

were later confirmed by magnetically generated shock compression experiments at the Z-pinch machine [61, 62, 63, 64, 65, 66] and by converging explosive-driven shock waves techniques [67, 68]. A general agreement between RPIMC and First Principle Molecular Dynamics (FPMD) predictions is also observed [10] except at the lowest temperatures that could be reached by RPIMC ($\sim 10000K$). There the Hugoniot from FPMD is slightly softer than from RPIMC (see Fig. 13 of ref. [10]).[6] Lower temperatures cannot be easily reached by RPIMC without reducing the level of accuracy. However, most of the interesting phenomena like molecular dissociation under pressure, metallization, solid-fluid transition, a possible liquid-liquid phase transition and its interplay with melting, the various crystalline phases and the transition to the atomic phases [10], occur at lower temperature out of the reach of RPIMC. Investigating this regime by QMC methods has been the main motivation in developing CEIMC. The other motivation is to benchmark the much more developed (and less demanding) alternative theoretical method, namely FPMD based on DFT. Indeed the numerical implementation of DFT is based on approximations (the exchange-correlation functional) the accuracy of which can only be established against experiments or, better, against more accurate theories. As mentioned earlier, QMC is based on the variational principle and therefore has an internal measure of accuracy.

CEIMC has been applied to investigate the WDM regime and benchmark FPMD [69, 70, 71] and to investigate the Liquid-Liquid phase transition (LLPT) region in hydrogen [72, 73]. The emerging picture is that a weak first-order phase transition occurs in hydrogen between a molecular-insulating fluid and a metallic-mostly monoatomic fluid. At higher temperature, molecular dissociation and metallization occur continuously. However the precise location of the transition line and the critical point are still matter of debate since several levels of the theory provide different locations. Within FPMD-DFT the location of the transition line depends strongly on the exchange-correlation functional employed and on whether classical or quantum protons are considered [74]. Transition lines from the PBE and vdW-DF2 approximations differ by roughly 200-250GPa, the PBE one being located at lower pressure. The PBE melting line with quantum protons is not in agreement with experiments, which highlights the failure of the PBE approximation when employed together with the quantum description of the nuclei. On the other hand, optical properties for the vdW-DF2 approximation are in agreement with experiments supporting the use of this functional for hydrogen in the WDM regime. The LLPT line from CEIMC lies in between the lines from PBE and vdW-DF2 functionals [72, 73]. However, those calculations were limited to classical protons and nuclear quantum effects are still to be considered.

Note that at higher temperatures and pressures, when molecules are entirely dissociated, the system is in the metallic state and protons can be considered classical particles, PBE and CEIMC have been found to be in very good agreement [71].

Our current understanding of the metallic state and metallization phenomena in high pressure hydrogen is mostly based on single electron theory [75]. To the best of

[6] Note that both theories are compatible with experiments because of the large uncertainty of the latter.

our knowledge there has only been one attempt to compute the DC conductivity of hydrogen in the WDM regime by a QMC based methods [76]. Protonic configurations at finite temperature are generated by CEIMC while the Correlation Function QMC [77, 78] is employed to compute the low energy many-body excitation spectrum of the liquid. Combining the excitation energies with the Green-Kubo formula, the electrical conductivity of hydrogen is obtained based entirely on QMC; these calculations do not suffer from self-interaction errors but suffer from other limitations, notably the numerical difficulty in obtaining accurate properties of excited states because of the QMC sign problem, and large finite cell size effects. However, good agreement with the limited data from shock experiments measurements was obtained [79].

Much more effort is needed in developing QMC method for WDM, in particular to explore the sensitivity of CEIMC to orbitals from DFT with different functionals and to further develop the method for transport and dynamical properties.

2.3 Helium, Carbon, and Water

While the application of quantum Monte Carlo methods in the WDM regime has been focused mainly on hydrogen, there has been recent interest in the extension of these methods to other materials in the periodic table. The first application of PIMC to elements beyond hydrogen involved the study of the phase diagram and equation of state of helium [80] and helium-hydrogen mixtures [81]. Similar to the case of pure hydrogen, the properties of hydrogen-helium mixtures are critical to the development of interior models for giant planets. Combining DFT and PIMC calculations, B. Militzer [82] was able to generate an equation of state for helium purely based on first-principles calculations, covering many orders of magnitude in density and temperature. This work represents a big step forward in the understanding of this material at extreme conditions and is an important piece in the development of next-generation planetary models.

The PIMC method was recently extended to other first row elements in the periodic table, including carbon and oxygen [83]. These calculations employed all-electron simulations of carbon and water at extreme conditions and covered temperature ranges from $10^4 - 10^9$ K. These represent the first PIMC calculations including core electrons and realistic materials. While some precautions must be taken when including core electrons, the work of Driver, *et al.* shows a marked improvement in the description of these materials at extreme temperatures and leads the way for the extension of the method to other interesting materials. This work also shows the need to generalize and improve the nodal constraint used in path integral simulations at low temperatures, particularly when bound states appear. With the development of optimization techniques for density matrices and the use of more complicated ansatz for the nodal constraint, the method can become the standard simulation technique at finite temperatures, applicable across the entire periodic table.

3 Discussion

We conclude this review with a quick comparison of QMC to other existing methods followed by a brief summary of some of the positive and negative aspects of QMC methods for the simulation of WDM.

3.1 Comparison of Methods

As discussed throughout this chapter, several simulation techniques exist that can be used in the study of materials in the WDM regime. Since most methods are based on very different approximations, which work better in different regimes, it is important to understand the regimes of applicability of methods as well as the expected accuracy as a function of both temperature and density. For example, methods based on density functional theory (DFT) are typically accurate at low temperatures, but are limited to temperatures below $10^4 - 10^5$ K due to computational difficulties above this regime. Path integral methods, on the other hand, are very accurate at high temperatures, but the accuracy is limited as the temperature is lowered. Coupled Electron-Ion Monte Carlo (CEIMC), while being accurate also at low temperatures, becomes inefficient in the limit of very low temperatures. A combination of all three methods can then be used to cover very wide ranges in density and temperature, and at the same time producing accurate properties in an efficient way. As a consequence, considerable work has been performed in order to compare the accuracy and limitations of these methods in the regimes where they overlap.

In the case of helium, a recent study compared the EOS between DFT-based BOMD and PIMC methods [82]. Excellent agreement between the two methods is obtained at temperatures above 80,000 K. This work is particularly interesting because it establishes the importance of the electronic temperature in DFT-MD calculations. In the case of hydrogen at Mbar pressures, several comparisons between DFT-MD, CEIMC, and PIMC have been made over the years [84, 2, 85]. In particular, CEIMC calculations have been used to benchmark DFT-MD simulations at high pressure in the development of equation of state models for hydrogen from first-principles [85, 23].

The recent work of Driver, *et al.* [86] on all-electron simulations of water and carbon using path integral methods also offers a comparison to DFT-MD over an extended temperature range (see Fig. 5). Similar to the case of hydrogen and helium, a reasonable agreement between the two methods is observed at temperatures above 200,000 K. This result serves to emphasize PIMC's role as an intermediary between DFT and classical methods such as hyper-netted chain (HNC) or Debye-Huckel theory.

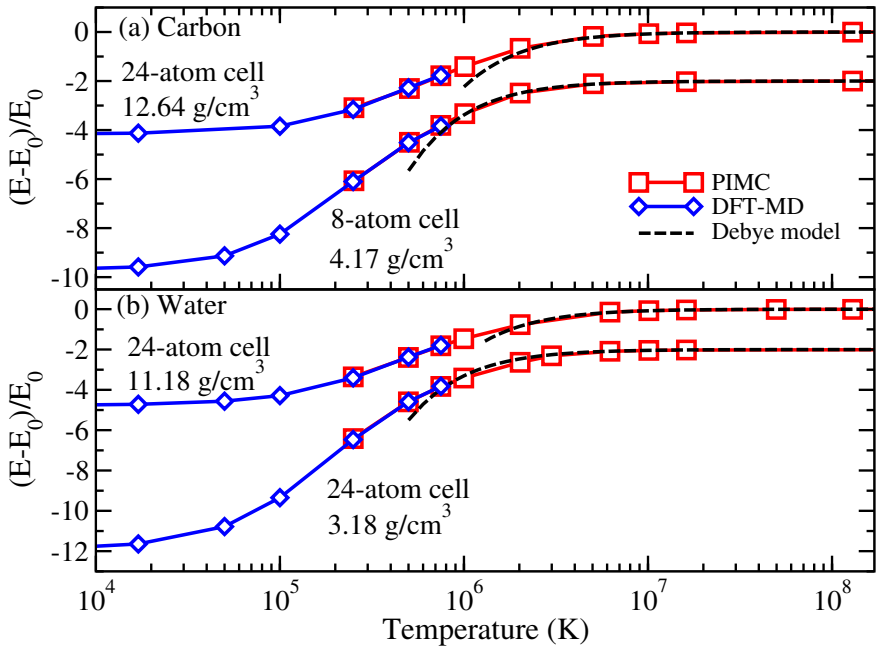

Fig. 5 A direct comparison of the excess energy in (a) Carbon and (b) Water, calculated through DFT-MD, PIMC, and Debye-Huckel theory. From this plot, one can conclude that PIMC may act as a bridge between the two quicker computational methods. [86]

3.2 Positives

In the spirit of this chapter being used a reference, we conclude by listing the specific areas where QMC performs well and where QMC still needs improvement. We begin with the things QMC is known to do well:

- Strong correlation is not a problem for QMC, in fact strong correlation is when QMC works best, e.g. in a Wigner crystal. Also there is no problem with the self-interaction, e.g. the formation of hydrogen molecules is exactly captured. For the singlet state of an isolated molecule, there is also no fermion sign problem.
- Because of its high accuracy, QMC usually serves as a benchmark for other methods. As an example, QMC provides the exchange-correlation energy of the electron gas at zero temperature. And, as discussed in §2.1, QMC can now providing high temperature correlation energy for the electron gas. Other properties besides the energy are possible to compute, e.g. correlation functions and static response functions.
- QMC is the most general method of simulating quantum systems. For example, one can add nuclear zero point motion of the ions, or include other quantum particles such as positrons, study bose condensation, all within the same framework.

- Adding additional randomness typically does not cause a large increase in computer resources. Adding randomness simply increases the dimensionality of the integrals by a small amount, while the computational efficiency depends on the variance and not on the number of integration variables. For other methods, where the quantum mechanics is computed deterministically, this is not true. This means for problems where randomness is part of the problem, QMC methods are not as expensive. Some examples of this averaging are twist average boundary conditions to reduce finite size effects (see §1.5), a liquid as treated with CEIMC (see §2.2), and proton zero point motion (see §2.2).
- QMC has reasonable scaling versus number of particles. The expensive part of scaling is the fermion determinant that takes N^3 operations if the matrix is full. However at high temperature the matrix elements have a range of the thermal deBroglie wavelength. For example, at the Fermi temperature, each row and column have only on the order of 10 elements with an value greater than 0.01. In this case $O(N)$ or $O(N^2)$ scaling should be possible.
- Perhaps most importantly, imaginary time path integrals give new understanding of quantum statistical mechanics. A notable example is the winding number relation to superfluidity [5]. There is a lot of work that needs to be be done in the direction to find new relations between properties of restricted paths and quantum observables.

3.3 Negatives

Finally, we list the issues that are currently holding up the applications of QMC to problems in WDM:

- There are a number of convergence issues: the algorithms are mathematically well-posed, but become inefficient as the needed limit is taken. For example, there are finite size effects (taking the simulation box size to infinity) (see §1.5). We do not view this problem as critical for simple systems, since the correlation length in WDM is not that large, but it becomes more difficult for mixtures. Another example of a convergence issue is the *Trotter error* or time step error (see §1.2). This also is not a practical limitation, but more efficient actions would increase the overall efficiency. Most important are ergodic problems of the Markov chain for restricted paths at low temperatures, aptly dubbed *reference point freezing* [87]. We do not know a solution to this problem, but also, we do not know if it is a fundamental problem.
- Imaginary time path intergrals can calculate imaginary time Green's functions. There are several problems in using this to find out dynamical response properties. First of all, this is equilibrium response only, true many-body dynamics is yet to be developed. Some linear response is possible, but it is not known precisely what are its limitations. Even within equilibrium linear response theory, there are two serious problems. First, the restricted path method for solving the sign problem changes the dynamics, even with the correct nodes. Second, there

is the issue of imaginary to real time conversion, though for disordered systems in WDM, this should not be the biggest problem.
- For ground state problems, methods to remove core electrons using non-local pseudopotentials have been developed. However with path integrals, only all electron calculations have been done, see §2.3. Core electrons have problems for non-local pseudopotentials with restricted paths that have not been resolved. It is not known yet whether this is a serious problem or if pseudopotentials developed at $T = 0$ would work in the WDM regime.
- The final, but most fundamental problem is the *fermion sign problem*. This is a well-studied problem with as yet no known general solution (in fact it has been shown that there exist quantum glass states that are NP-complete [88]). Above the Fermi temperature the nodes should be dominated by the free particle nodes, and perturbation theory should take us to lower temperatures. Also, as mentioned, a variational principle should be able to guide us to lower temperatures as it does at zero temperature. Little work has been done along these lines in practice.

The problems listed above are opportunities to find solutions at the intersection of applied mathematics and many-body theory. None of these issues, beyond the sign problem in a specific spin system, have been shown to be fundamental and hence insurmountable. Because of this, we foresee QMC continuing to be a central tool as WDM research moves forward.

Acknowledgements DC is supported by DOE DE-NA0001789. EB is supported by DOE DE-AC52-07NA27344, LDRD 10-ERD-058 and the LLNL Lawrence Scholar program. CP is supported by the Italian Institute of Technology (IIT) under the SEED project grant number 259 SIMBEDD Advanced Computational Methods for Biophysics, Drug Design and Energy Research. This work was performed in part under the auspices of the US DOE by LLNL under Contract DE-AC52-07NA27344.

References

1. C. Pierleoni, D.M. Ceperley, Lecture Notes in Physics **714**, 641 (2006)
2. M.A. Morales, C. Pierleoni, D.M. Ceperley, Phys. Rev. E **81**, 021202 (2010)
3. D.M. Ceperley, B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980)
4. J. Kolorenč, L. Mitas, Reports on Progress in Physics **74**(2), 026502 (2011)
5. D.M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995)
6. D.M. Ceperley, Journal of Statistical Physics **63**, 1237 (1991). 10.1007/BF01030009
7. W. Selke, Journal of Statistical Physics **87**, 959 (1997). 10.1007/BF02181259
8. V. Natoli, D.M. Ceperley, Journal of Computational Physics **117**(1), 171 (1995)
9. S. Chiesa, D.M. Ceperley, R.M. Martin, M. Holzmann, Phys. Rev. Lett. **97**, 076404 (2006)
10. J. McMahon, M. Morales, C. Pierleoni, D.M. Ceperley, Review of Modern Physics **84**, 1607 (2012)
11. D.M. Ceperley, M. Dewing, J. Chem. Phys. **110**, 9812 (1999)
12. D. Ceperley, M. Dewing, C. Pierleoni, Lecture Notes in Physics **605**, 473 (2002)
13. M. Holzmann, D.M. Ceperley, C. Pierleoni, K. Esler, Phys. Rev. E **68**, 046707:1 (2003)
14. C. Pierleoni, K. Delaney, M. Morales, D. Ceperley, M. Holzmann, Computer physics communications **179**(1-3), 89 (2008)

15. S. Baroni, S. Moroni, Phys. Rev. Lett. **82**(24), 4745 (1999)
16. M. Bajdich, L. Mitas, G. Drobný, L.K. Wagner, K.E. Schmidt, Phys. Rev. Lett. **96**, 130201 (2006)
17. A.G. Anderson, W.A. Goddard, The Journal of Chemical Physics **132**(16), 164110 (2010)
18. M. Casula, C. Attaccalite, S. Sorella, The Journal of Chemical Physics **121**(15), 7110 (2004)
19. Y. Kwon, D.M. Ceperley, R.M. Martin, Phys. Rev. B **58**, 6800 (1998)
20. M. Bajdich, M.L. Tiago, R.Q. Hood, P.R.C. Kent, F.A. Reboredo, Phys. Rev. Lett. **104**, 193001 (2010)
21. P. Seth, P.L. Rios, R.J. Needs, The Journal of Chemical Physics **134**(8), 084105 (2011)
22. B.K. Clark, M.A. Morales, J. McMinis, J. Kim, G.E. Scuseria, The Journal of Chemical Physics **135**(24), 244105 (2011)
23. M.A. Morales, J. McMinis, B.K. Clark, J. Kim, G.E. Scuseria, Journal of Chemical Theory and Computation **8**(7), 2181 (2012)
24. C.J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, R.G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007)
25. J. Toulouse, C.J. Umrigar, The Journal of Chemical Physics **128**(17), 174101 (2008)
26. J. Toulouse, C.J. Umrigar, The Journal of Chemical Physics **126**(8), 084102 (2007)
27. S.A. Khairallah, J. Shumway, E.W. Draeger, ArXiv e-prints (2011)
28. B. Militzer, E.L. Pollock, Phys. Rev. E **61**, 3470 (2000)
29. D.M. Ceperley, B.J. Alder, Phys. Rev. B **36**, 2092 (1987)
30. C. Lin, F.H. Zong, D.M. Ceperley, Phys. Rev. E **64**, 016702 (2001)
31. S. Chiesa, D.M. Ceperley, R.M. Martin, M. Holzmann, Phys. Rev. Lett. **97**, 076404 (2006)
32. N.D. Drummond, R.J. Needs, A. Sorouri, W.M.C. Foulkes, Physical Review B (Condensed Matter and Materials Physics) **78**(12), 125106 (2008)
33. E. Wigner, Phys. Rev. **46**, 1002 (1934)
34. F. Bloch, Z. Physik **57**, 545 (1929)
35. F.H. Zong, C. Lin, D.M. Ceperley, Phys. Rev. E **66**, 036703 (2002)
36. D. Ceperley, Phys. Rev. B **18**, 3126 (1978)
37. S. Kümmel, L. Kronik, Rev. Mod. Phys. **80**, 3 (2008)
38. E.L. Pollock, J.P. Hansen, Phys. Rev. A **8**, 3110 (1973)
39. B. Jancovici, Physica A Statistical Mechanics and its Applications **91**, 152 (1978)
40. J.P. Hansen, P. Vieillefosse, Physics Letters A **53**, 187 (1975)
41. M.D. Jones, D.M. Ceperley, Phys. Rev. Lett. **76**, 4572 (1996)
42. E.W. Brown, B.K. Clark, J.L. DuBois, D.M. Ceperley, Phys. Rev. Lett. **In press.** (2012)
43. D.M. Ceperley, Phys. Rev. Lett. **69**, 331 (1992)
44. J.P. Hansen, Phys. Rev. A **8**, 3096 (1973)
45. K.S. Singwi, M.P. Tosi, R.H. Land, A. Sjölander, Phys. Rev. **176**, 589 (1968)
46. S. Tanaka, S. Ichimaru, Journal of the Physical Society of Japan **55**, 2278 (1986)
47. R.G. Dandrea, N.W. Ashcroft, A.E. Carlsson, Phys. Rev. B **34**, 2097 (1986)
48. F.m.c. Perrot, M.W.C. Dharma-wardana, Phys. Rev. B **62**, 16536 (2000)
49. M.W.C. Dharma-wardana, F. Perrot, Phys. Rev. Lett. **84**, 959 (2000)
50. S. Dutta, J. Dufty, ArXiv e-prints (2012)
51. V. Natoli, R.M. Martin, D.M. Ceperley, Phys. Rev. Lett. **70**, 1952 (1993)
52. V. Natoli, R.M. Martin, D.M. Ceperley, Phys. Rev. Lett. **74**, 1601 (1995)
53. C. Pierleoni, D. Ceperley, B. Bernu, W. Magro, Phys. Rev. Lett. **73**(16), 2145 (1994)
54. W. Magro, D. Ceperley, C. Pierleoni, B. Bernu, Phys. Rev. Lett. **76**(8), 1240 (1996)
55. C. Pierleoni, W.R.Magro, D. Ceperley, B. Bernu, in "Physics of Strongly Coupled Plasmas" W.D. Kraeft and M. Schlanges eds., pp. 11–26 (World Scientific, 1996)
56. B. Militzer, D. Ceperley, Phys. Rev. Lett. **85**(9), 1890 (2000)
57. B. Militzer, D. Ceperley, J. Kress, J. Johnson, L. Collins, S. Mazevet, Phys. Rev. Lett. **87**(27), 275502 (2001)
58. L. DaSilva, P. Celliers, G. Collins, K. Budil, N. Holmes, T. Barbie Jr, B. Hammel, J. Kilkenny, R. Wallace, M. Ross, R. Cauble, Phys. Rev. Lett. **78**(3), 483 (1997)
59. G. Collins, L.D. Silva, P. Celliers, D. Gold, M. Foord, R. Wallace, A. Ng, S. Weber, K. Budil, R. Cauble, Science **281**(5380), 1178 (1998)

60. P. Celliers, G. Collins, L.D. Silva, D. Gold, R. Cauble, R. Wallace, M. Foord, B. Hammel, Phys. Rev. Lett. **84**(24), 5564 (2000)
61. M. Knudson, D. Hanson, J. Bailey, C. Hall, J. Asay, W. Anderson, Phys. Rev. Lett. **87**(22), 225501 (2001)
62. M. Knudson, D. Hanson, J. Bailey, C. Hall, J. Asay, Phys. Rev. Lett. **90**(3), 35505 (2003)
63. M. Knudson, D. Hanson, J. Bailey, C. Hall, J. Asay, C. Deeney, Physical Review B **69**(14), 144209 (2004)
64. J. Bailey, M. Knudson, A. Carlson, G. Dunham, M. Desjarlais, D. Hanson, J. Asay, Physical Review B **78**(14), 144107 (2008)
65. D. Hicks, T. Boegly, P. Celliers, J. Eggert, S. Moon, D. Meyerhofer, G. Collins, Physical Review B **79**, 014112 (2009)
66. M. Knudson, M. Desjarlais, Phys. Rev. Lett. **103**(22), 225501 (2009)
67. G. Boriskov, A. Bykov, R. Il'Kaev, V. Selemir, G. Simakov, R. Trunin, V. Urlin, A. Shuikin, W. Nellis, Physical Review B **71**(9), 092104 (2005)
68. S. Grishechkin, S. Gruzdev, V. Gryaznov, M. Zhernokletov, R. Il'Kaev, I. Iosilevskii, G. Kashintseva, S. Kirshanov, S. Manachkin, V. Mintsev, JETP letters **80**(6), 398 (2004)
69. C. Pierleoni, D.M. Ceperley, M. Holzmann, Phys. Rev. Letts. **93**, 146402:1 (2004)
70. C. Pierleoni, K. Delaney, M. Morales, D. Ceperley, M. Holzmann, Computer Physics Communications **179**(1-3), 89 (2008)
71. M.A. Morales, C. Pierleoni, D.M. Ceperley, Phys. Rev. E **81**(2), 021202 (2010)
72. M.A. Morales, C. Pierleoni, E. Schwegler, D.M. Ceperley, Proceedings of the National Academy of Sciences **107**(29), 12799 (2010)
73. E. Liberatore, M.A. Morales, D.M. Ceperley, C. Pierleoni, Molecular Physics pp. 1–8 (2011)
74. M.A. Morales, J.M. McMahon, C. Pierleoni, D.M. Ceperley, Phys. Rev. Lett. **In press.** (2013)
75. R.M. Martin, *Electronic Structure: Basic theory and practical methods* (Cambridge University Press, 2004)
76. F. Lin, M. Morales, K. Delaney, C. Pierleoni, Phys. Rev. Lett. **103**, 256401 (2009)
77. D.M. Ceperley, B. Bernu, J. Chem. Phys. **89**, 6316 (1988)
78. B. Bernu, D.M. Ceperley, W.A. Lester, J. Chem. Phys. **93**, 552 (1990)
79. S. Weir, A. Mitchell, W. Nellis, Phys. Rev. Lett. **76**(11), 1860 (1996)
80. B. Militzer, Phys. Rev. Lett. **97**, 175501 (2006)
81. B. Militzer, Journal of Low Temperature Physics **139**, 739 (2005)
82. B. Militzer, Phys. Rev. B **79**, 155105 (2009)
83. K.P. Driver, B. Militzer, Phys. Rev. Lett. **108**, 115502 (2012)
84. C. Pierleoni, D.M. Ceperley, M. Holzmann, Phys. Rev. Lett. **93**, 146402 (2004)
85. L. Caillabet, S. Mazevet, P. Loubeyre, Phys. Rev. B **83**, 094101 (2011)
86. K.P. Driver, B. Militzer, Phys. Rev. Lett. **108**, 115502 (2012)
87. B. Militzer, D.M. Ceperley, Phys. Rev. Lett. **85**, 1890 (2000)
88. M. Troyer, U.J. Wiese, Phys. Rev. Lett. **94**(17), 170201 (2005)