

An Overview of Quantum Monte Carlo Methods

David M. Ceperley

Department of Physics and National Center for Supercomputing Applications
University of Illinois Urbana-Champaign
Urbana, Illinois 61801

In this brief article, I introduce the various types of quantum Monte Carlo (QMC) methods, in particular, those that are applicable to systems in extreme regimes of temperature and pressure. We give references to longer articles where detailed discussion of applications and algorithms appear.

MOTIVATION

One does QMC for the same reason as one does classical simulations; there is no other method able to treat exactly the quantum many-body problem aside from the direct simulation method where electrons and ions are directly represented as particles, instead of as a “fluid” as is done in mean-field based methods. However, quantum systems are more difficult than classical systems because one does not know the distribution to be sampled, it must be solved for. In fact, it is not known today which quantum problems can be “solved” with simulation on a classical computer in a reasonable amount of computer time. One knows that certain systems, such as most quantum many-body systems in 1D and most bosonic systems are amenable to solution with Monte Carlo methods, however, the “sign problem” prevents making the same statement for systems with electrons in 3D. Some limitations or approximations are needed in practice. On the other hand, in contrast to simulation of classical systems, one does know the Hamiltonian exactly: namely charged particles interacting with a Coulomb potential. Given sufficient computer resources, the results can be of quite high quality and for systems where there is little reliable experimental data. For this reason, QMC methods, though more expensive, are useful to benchmark and validate results from other methods.

The two main applications that we will discuss in this review are the “electronic structure problem”; computing the energy of the an interacting systems of electrons and fixed ions, and the problem of quantum effects of the nuclei.

RANDOM WALK AND MARKOV CHAINS

Most QMC algorithms are based on random walks, known mathematically as Markov chains. This is a general class of algorithms, introduced in the famous work by Metropolis, the Rosenbluths and the Tellers (1953) for sampling any probability distribution. Let us denote the desired distribution as $\Pi(S)$, where the variable “S” represents the state of the walk, for example the coordinates of all of the particles. In a random walk, one invents some moving rules (transition probabilities) for changing the state of the system so that the probability of being in state S_{n+1} in the (n+1) step, given that the random walk was S_n is given by $T(S_n ; S_{n+1})$. By

enforcing the detailed balance condition, and the ergodicity of random walk, one can guarantee that the asymptotic distribution will be $\Pi(S)$ (Kalos and Whitlock, 1986).

A typical Monte Carlo simulation starts out the state of the system at some reasonable configuration; for example an appropriate number of electrons near each ion. The random walk algorithm is then applied many times. Runs of 10^6 to 10^{12} steps are typical. During the random walk, system averages are blocked together, for example the energy of system during a thousand steps. These block averages are saved for later analysis. During the analysis one examines the various properties to see, whether after an initial “warm-up,” the properties stabilize about some mean value, and fluctuate about it, hopefully in a normal fashion. According to the central limit theorem, the mean value of the simulation will converge to the exact value: i.e. $\langle E \rangle = \int dS \Pi(S) E(S)$ within an “error bar” which scales as $M^{-1/2}$ where M is the number of steps of the random walk. The estimated error is determined by the fluctuations in the block averages from the overall mean, and their autocorrelation. All we need to do to get another decimal place of accuracy is to run the system one hundred times longer, or to borrow another 99 processors from your colleagues. The steady progress of computer technology and algorithms has made calculations presented here feasible and even routine.

VARIATIONAL MONTE CARLO

Variational Monte Carlo (VMC) is a simple generalization of classical Monte Carlo, introduced by McMillan(1965) for liquid helium. Ceperley (1977, 1978) has a detailed description of VMC for fermions and for the pure electron gas. In classical MC for an interaction $V(R)$, one samples the coordinates, $R = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ from the Boltzmann distribution: $\exp(-\beta V(R)) / Z$ where $\beta = 1/(k_B T)$. In VMC, we sample the electron coordinates from the squared modulus of a trial function; $|\phi_T(R)|^2$. Since electrons are fermions, one needs to use an antisymmetric function for $\phi_T(R)$; without correlation one would take a single Slater determinant. Typically there are $N/2$ spatial orbitals in the spin up determinant and the same number in the spin down determinant. If the wavefunction were a simple determinant, then we would have a Hartree-Fock result, and there would be no need to use a Monte Carlo procedure to determine the energy. The advantage of MC simulation is that one can introduce correlation directly by multiplying by a correlation factor, for example the well-known “Jastrow” form: $\phi_T(R) = \exp(-U(R)) \text{Det}(R)$. The Jastrow function $U(R) = \sum u(r_{ij})$ is usually a sum of two-body and three-body terms, it has a form similar to the potential energy function. The “cusp condition” gives the exact value of the wavefunction derivative as two charged particles approach each other. The case of two spin-unlike electrons is especially important, since the Slater determinant does not correlate electrons in different spin states. Typically, the function: $u(r) = a/(b+r)$ is used to correlate electrons, where (a, b) are variational parameters, or can be set by analytic conditions. But if we set $du(r_i - r_j)/dr = -1/2$ we restore the correct value near the co-incident point $r_i \sim r_j$ and thereby recover most of the correlation energy missing in the independent electron method. In this equation, and in what follows, we use atomic units.

We call this variational Monte Carlo, because one can vary parameters in the trial function to minimize the variational energy, defined as the expectation of the Hamiltonian, H , over the trial function. By the usual theorem, this is always an upper bound to the exact energy and gives the exact energy when the trial function is an exact eigenstate. The variational energy is computed

as $E_V = \langle E_L(\mathbf{R}) \rangle$: we average the “local energy” over the VMC random walk and $E_L(\mathbf{R}) = \phi_T(\mathbf{R})^{-1} H \phi_T(\mathbf{R})$ is the “local energy” of the trial function. One important aspect of QMC is the zero variance principle (Ceperley, 1977): as the trial function tends to an exact eigenstate of H , the fluctuations in the local energy tend to zero, and hence the number of MC steps needed to achieve a given error bar decreases to zero. As a result of this principle, QMC calculations can achieve much higher accuracy than classical MC calculations. Of course in electronic structure problems, one needs much higher accuracy.

The key problem in VMC is to come up with increasingly accurate trial wavefunctions and to efficiently optimize the parameters in them. There has been recent progress on both of these issues. Concerning better trial functions one can incorporate pair (geminal) (Casula and Sorella, 2003) and backflow transformation of orbitals (Holzmann et al., 2003) Optimization methods can now treat thousands of free parameters (Umrigar et al., 2007) within the context of a VMC random walk.

Another issue is how to treat pseudopotentials, particularly non-local ones, because QMC methods scale quite badly with the atomic number. We need to get rid of the core states. In VMC one uses an auxiliary integration in the core; this can be generalized for the other quantum method, e.g. as in Mitas et al. (1991). However, the use of pseudopotentials is a key approximation in Quantum Monte Carlo because typical pseudopotentials have not been constructed for or tested for correlated calculations.

The VMC methods are quite efficient, and scale with the number of electrons as N^3 for bulk properties and N^4 for energy gaps. Methods have been introduced to reduce this scaling, using correlated sampling for example, or for localized systems such as insulators with large band gaps, sparse matrix algorithms (Williamson et al., 2001). The difficulty with VMC is that it favors simple, well-understood phases over less well-characterized phases, and this can bias computation of something like a pressure-induced phase transition. However, it is a very valuable method, because the trial function that comes out of VMC is necessary for the algorithm that we discuss next.

DIFFUSION MONTE CARLO

The primary zero temperature QMC method is projector MC, where a function of the Hamiltonian takes the trial function and projects out the ground state. The most common implementation of this method is diffusion MC where one computes a modified trial function defined as $\phi(\mathbf{R},t) = \exp(-Ht) \phi_T(\mathbf{R})$ where the trial function $\phi_T(\mathbf{R})$ is also the initial state. For large projection time “ t ” $\phi(\mathbf{R},t)$ will converge to the lowest state that has an overlap with $\phi_T(\mathbf{R})$. The first question is how to implement the projection numerically. There are two different methods commonly used.

With importance sampling, one can show that one can interpret the projection as a diffusing, branching, random walk and is known as Diffusion MC (DMC). This method, with “importance sampling” was first used in the work of Ceperley and Alder (1980) and is described in detail for molecular systems in Reynolds et al. (1982) or for solid state systems by Foulkes (2001).

A more recent implementation stores the whole evolution as a path in phase space. Computer time is then used to wiggle this path. This is known under several names, such as Variational Path Integrals (Ceperley, 1995), or Reptation Quantum Monte Carlo (Baroni and Moroni, 1999). The path integral approach has the advantage that all static properties can be computed without bias, while in the DMC approach, only the energy is unbiased by the trial function. Also, it is a more convenient approach to estimating energy differences (Pierleoni and Ceperley, 2006).

In either case, the algorithm can be interpreted as a stochastic process for fermions only with an approximation, because the underlying integrand must be made positive to avoid the “fermion sign problem.” The most common approach is called the fixed-node approximation. Here again, the trial function enters. One searches for the lowest energy state ϕ with the restriction that $\phi(\mathbf{R}) \geq 0$ for all possible coordinate values \mathbf{R} . To implement this, one simply requires that the random walk does not change the sign of $\phi(\mathbf{R})$. The resulting upper bound is much lower than the starting variational energy. In systems such as the electron gas, or for a simple metal, one can obtain very accurate results, because the wavefunction nodes are highly constrained by symmetry. Estimates of the fixed-node error are obtained either from expensive “nodal-release” calculations (Ceperley and Alder, 1980) or from comparison to experiment or results from other theoretical approaches. Systems for which the fixed-node approach are less successful are molecular systems with a ground state having contributions from several Slater determinants, such as the molecule C_2 . Of course if the trial function includes the needed determinants, then the fixed-node answer can be arbitrarily accurate.

In general, one would like to use complex wavefunctions, both because this allows one to work with the full symmetry of the problem and because this allows a faster extrapolation to the infinite system, with the so-called Twist Average Boundary Condition (Lin et al., 2001). In this method, aimed at metallic systems, one integrates over the Brillouin zone of the simulation cell (or supercell), thereby reducing finite-size effects. The complex wavefunctions can be used either in VMC and DMC, using the so-called fixed-phase method, (Ortiz et al., 1993). This method reduces to the fixed-node method for real trial functions, but for arbitrary phases gives an upper bound which is exact if the phase of the trial function is exact. Further progress on the reducing finite size effects from the Coulomb interaction are obtained by consideration of properties of the static structure factor for long wavelengths (Chiesa et al. 2006).

PATH INTEGRAL MC

The two previous methods we discussed, were methods for zero temperature, single many-body eigenstates. We now discuss a method, useful for quantum calculations at non-zero temperatures. Path Integral MC (PIMC) is based on Feynman’s (1953) imaginary time path integral description of liquid ^4He . In this method, the quantum statistical mechanics of distinguishable particle or bosonic systems maps into a classical statistical problem of interacting polymers. Hence the probability distribution to sample is:

$$\Pi(S) = C \exp \left\{ - \sum (R_t - R_{t-1})^2 / 4\lambda\tau - \tau V(R_t) \right\}$$

where R_t is the path and it is cyclic in the time variable $0 \leq t \leq \beta = 1/(k_B T)$ so that $R_0 = R_\beta$. The PIMC procedure consists of sampling the path with the random walk algorithm using this action. Here C is a constant and $\lambda = \hbar^2/2m$. For details of this method see Ceperley (1995).

For bosons, since they are indistinguishable, one has to consider paths which are cyclic up to a permutation of path labels, so that $R_0 = P R_\beta$ where P is a permutation operator acting on particle labels. Figure (1) on the left, illustrates paths coming from distinguishable particle statistics (i.e. $P=1$) and, on the right, bosonic statistics with exchange. Both the path R_t and the permutation P are sampled in a random walk.

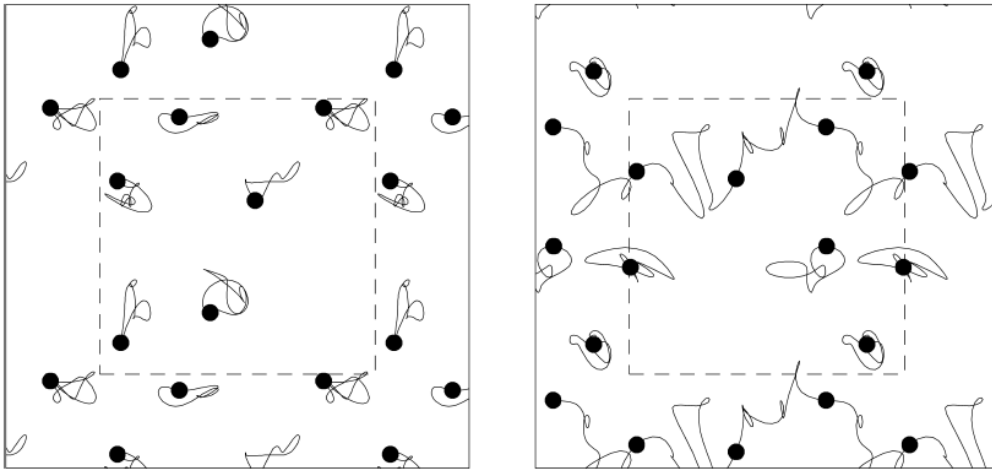


Figure 1: An example of a typical “imaginary-time path” for 6 hard spheres. The black dots represent the beginning of the path. Left: the identity permutation appropriate for distinguishable particles. Right: a permutation of 3 atoms winding across the periodic box, appropriate to systems involving bose condensation, represents superfluid flow in PIMC.

For fermions one has to include a minus sign for odd permutations so that those paths contribute negatively to any averages. At low enough temperature, all permutations are equally likely, so the sign becomes random, uncorrelated with the configuration, so all averages become zero over zero; in practice the signal to noise ratio decreases as $\exp(-2 N E_F/k_B T)$ where E_F is the Fermi energy (Ceperley, 1996). The fixed-node approximation can also be used with path integrals allowing one to treat electrons using path integrals (Ceperley, 1992). Such calculations have been carried out for hot dense hydrogen, as for example occurs in the interior of the giant planets. See for example Magro et al. (1996) and Militzer and Ceperley (2001). All quantum effects are included in such a calculation, including those of the proton quantum motion. However, we have not been successful in going to low temperatures with this approach; i.e. temperatures less than about 20% of the Fermi energy, because of the difficulty of sampling the paths.

Another very useful application for PIMC, besides its primary application for all types of bosonic systems, is to treat the effects of nuclear degrees of freedom, *e.g.* the effects of zero point motion in lithium (Filippi et al., 1998). The same type of approach is also employed for calculations of water that include zero point motion: a molecular dynamics calculation is done for both the electronic degrees of freedom and the imaginary time path of the protons.

COUPLED ELECTRON ION MC

In the field of Molecular Dynamics simulations of atomic and molecular systems, a great leap-forward was taken when Car and Parrinello (1985), introduced the *ab initio* MD method, replacing an *ad hoc* intermolecular potential with a potential computed using density functional theory (DFT) “on the fly.” We now describe a recent attempt to do something similar using potentials coming from QMC, in particular from reptation QMC. The Coupled Electron-Ion MC (CEIMC) assumes the Born-Oppenheimer potential and makes either a classical or a quantum (Path Integral) description of the ions (Pierleoni et al., 2006). Assuming one is doing a classical MC of the ions, after a move of the ionic coordinates from S to S' , in Metropolis MC one accepts the move with probability $\min[1, \exp(-V(S')-V(S)/k_B T)]$. With CEIMC one computes the electronic energy difference $\Delta V = V(S') - V(S)$ using the Projector MC technique described above. A possible difficulty is that this energy difference will be noisy, since it is determined by a MC procedure, and unless the noise is much smaller than $k_B T$, the results will be biased, since the acceptance formula is non-linear. We have shown (Ceperley and Dewing, 1999) that by adding a “penalty” of $\sigma^2/(2 T^2)$ in the above exponent, that this bias can be avoided. Here σ is the statistical error in ΔV . The basic idea of CEIMC is to compute, using correlated sampling QMC methods, accurate energy differences.

CEIMC should be especially useful for systems where DFT functionals are less reliable, such as close to a metal-insulator transition. Most of our work on this has focused on warm dense hydrogen. It is possible to construct quite accurate many-body trial functions for hydrogen, avoiding the use of pseudopotentials. Using this method we have recently examined the equation of state of liquid hydrogen (Morales et al., 2009) the atomic-molecular and melting transition in hydrogen and the demixing transition in hydrogen-helium mixtures. One is able to achieve simulations at quite low temperatures (say 500K) and use high accuracy trial wavefunctions.

Some of the advantages of CEIMC, compared with AIMD, are that the MC procedure is more robust than M, the energy functional includes fully correlated electrons without problems such as self-interaction arising in DFT, pseudopotentials are not needed for light atoms and finally, one can average over other degrees of freedom without slowing down the computation. Two examples of this last point are in the average over twisted boundary conditions and in doing quantum versus classical protons. Using this method on massively parallel computers, with hundreds of thousands weakly coupled processors, one can perform a simulation of higher accuracy than AIMD, to verify its predictions. The development of QMC to calculate forces, so that one can do a dynamical simulation, should allow prediction of dynamical properties as well.

AUXILLARY FIELD MC

The Auxiliary-Field Monte Carlo (AFQMC) method (Zhang et al., 2003) is a significantly different alternative to DMC. Its random walks have a state space of (single particle) Slater determinants that are subject to a fluctuating external potential. The sign problem is eliminated by requiring that the phase of the determinant remain close to the phase of the trial determinant. Applications show that this often reduces the reliance of the solution on the quality of the trial

wave function. Because one represents the random walkers in a single-particle basis, similar to DFT and other standard electronic structure methods in chemistry and physics, much of the technology developed from decades of effort in standard one-particle methods can be directly used for correlated calculations. In complex systems such as transition metal oxides, the use of an effective basis in which to carry out the random walks could lead to significant gains in efficiency, though in practice the calculations are more time consuming than DMC calculations. Further, the AF method allows us to do QMC calculations on both "simple" and realistic model Hamiltonians in which higher and irrelevant energy scales can be systematically removed. To date, the new method has been applied to close to 100 systems, several recent calculations have treated solids (Kwee et al 2008). The method demonstrated excellent accuracy, better than 0.1eV in most cases, and was consistently able to correct errors in the mean-field trial wave function.

PROSPECTIVES

QMC methods are already very valuable for calculations of high pressure bulk properties such as the equation of state of materials. Path integral Monte Carlo is useful for high temperature electronic structure calculations and calculation of quantum properties involving light ions. Coupled electron Ion Monte Carlo is good for low temperature properties, where one needs to go beyond methods based on Density Functional Theory particularly for disordered systems. There have been many important algorithmic developments in the last decade which allow calculations on realistic material systems. That, combined with access to computer systems with millions of hours of computation, have resulted in a growing number of QMC calculations relevant to geophysics and astrophysics. As methods and computers develop, the QMC algorithms can take advantage of massive parallelism to deliver results of high quality, and to benchmark less expensive methods, *e.g.* to decide which DFT functional is appropriate for a given problem.

REFERENCES

Baroni S, Moroni S (1999) Reptation Quantum Monte Carlo: A Method for Unbiased Ground-State Averages Imaginary-Time Correlations. *Phys Rev Lett* 82: 4745

Car R, Parrinello M (1985) Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys Rev Lett* 55:2471-2474

Casula M, Sorella S (2003) Geminal wave functions with Jastrow correlation: A first application to atoms. *J Chem Phys* 119: 6500-6511

Ceperley D, Chester GV, Kalos MH (1977) Monte Carlo Simulation of a Many-Fermion System. *Phys Rev B* 16: 3081

Ceperley D (1978) Ground State of the Fermion One-Component Plasma: A Monte Carlo Study in Two Three Dimensions, *Phys Rev B* 18: 3126.

Ceperley DM, Alder BJ (1980) Ground State of the Electron Gas by a Stochastic Method. *Phys Rev Lett* 45: 566

Ceperley DM (1992) Path Integral Calculations of Normal Liquid ^3He . Phys Rev Lett 69:331

Ceperley DM (1995) Path Integrals in the Theory of Condensed Helium. Rev Mod. Phys 67: 279

Ceperley D (1996) Path Integral Monte Carlo Methods for Fermions in Monte Carlo Molecular Dynamics of Condensed Matter Systems. edited by K. Binder G. Ciccotti, Editrice Compositori, Bologna

Ceperley DM Dewing M (1999) The Penalty Method for Random Walks with Uncertain Energies. J Chem Phys 110: 9812-9820

Chiesa S, Ceperley DM, Martin RM, Holzmann M (2006) Finite Size Error in Many-body Simulations with Long-Ranged Interactions. Phys Rev Lett 97 076404:1-4

Feynman RP (1953) Atomic Theory of the λ Transition in Helium. Phys Rev 91: 1291-1301

Filippi C, Ceperley DM (1998) Path Integral Monte Carlo Calculation of the kinetic energy of condensed Lithium. Phys Rev B 57: 252

Foulkes WMC, Mitas L, Needs RJ, Rajagopal G (2001) Quantum Monte Carlo Simulations of Solids. Rev Mod Phys 73: 33-84

Holzmann M, Ceperley DM, Pierleoni C, Esler K (2003) Backflow Correlations for the Electron Gas Metallic Hydrogen. Phys Rev E 68: 046707:1-15

Kalos MH, Whitlock PA (1986) Monte Carlo Methods. Wiley

Kwee H, Zhang S, Krakauer H (2008) Finite-Size Correction in Many-Body Electronic Structure Calculations. Phys Rev Lett 100: 126404

Lin C, Zong FH, Ceperley DM (2001) Twist-averaged boundary conditions in continuum Quantum Monte Carlo algorithms. Phys Rev E 64: 016702

Magro WR, Ceperley DM, Pierleoni C, Bernu B (1996) Molecular Dissociation in Hot, Dense Hydrogen. Phys Rev Lett 76: 1240

McMillan WL (1965) Ground State of Liquid ^4He , Phys Rev 138: A442

Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E (1953) Equation of State Calculations by Fast Computing Machines. J Chem Phys 21: 1087

Militzer B, Ceperley DM (2001) Path Integral Monte Carlo Simulation of the Low-Density Hydrogen Plasma. Phys Rev E63: 66404

Morales MA, Pierleoni C, Ceperley DM, (2009) Equation of state of metallic hydrogen from Coupled Electron-Ion Monte Carlo simulations. submitted to Phys Rev E,; arXiv:0906.1594

Mitas L, Shirley EL, Ceperley DM (1991) Nonlocal Pseudopotentials in Diffusion Monte Carlo. J Chem Phys 95: 3467

Ortiz G, Ceperley DM, Martin RM (1993) New Stochastic Method for Systems with Broken Time-Reversal Symmetry; 2-D Fermions in a Magnetic Field: Phys Rev Lett 71: 2777

Pierleoni C, Bernu B, Ceperley DM, Magro WR (1994) Equation of State of the hydrogen plasma by path integral Monte Carlo simulation. Phys Rev Lett 73: 2145

Pierleoni C, Ceperley DM (2006) The Coupled Electron-Ion Monte Carlo Method. Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology , eds. M. Ferrario, G. Ciccotti, K. Binder, Lecture Notes in Physics 703, 641-683, Springer Berlin Heidelberg ; physics/0510254

Reynolds PJ, Ceperley DM, Alder BJ, Lester Jr., WA (1982) Fixed-node Quantum Monte Carlo for Molecules. J Chem Phys 77:5593

Umrigar CJ, Toulouse J, Filippi C, Sorella S, Hennig RG (2007) Alleviation of the Fermion-Sign Problem by Optimization of Many-Body Wave Functions. Phys Rev Lett 98, 110201

Williamson AJ, Hood RQ, Grossman JC (2001) Linear-Scaling Quantum Monte Carlo Calculations. Phys Rev Lett 87: 246406

Zhang S, Krakauer H (2003) Quantum Monte Carlo method using phase-free walks with Slater determinants. Phys Rev Lett 90:136401