DIFFUSION MONTE CARLO

LUBOS MITAS
National Center for Supercomputing Applications,
University of Illinois at Urbana-Champaign,
Urbana, IL 61801, USA

1. Introduction.

These lecture notes will describe the basic steps of the diffusion Monte Carlo (DMC) method for solving the Schrödinger equation for many-particle systems. The DMC method is one of several approaches which are known under a common name of quantum Monte Carlo (QMC)[1, 2].
The imaginary time Schrödinger equation is given by

$$-\partial_t \Phi(R,t) = (H - E_T)\Phi(R,t) \tag{1}$$

where E_T is an energy offset, H is a Hamiltonian and $R = (\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ is a $3N$-dimensional vector which denotes positions of N particles in continuous space. For a system of interacting electrons and ions the Hamiltonian is given by

$$H = T + U = T + V_{ee} + V_{eI} + E_{II} \tag{2}$$

where T, V_{ee}, V_{eI} are operators of kinetic energy, electron-electron and electron-ion interactions, respectively. For the sake of simplicity the ions will be static with a constant contribution to the energy E_{II} (the approach can be generalized to include the quantum character of ions as well). The equation can be written in an integral form

$$\Phi(R, t+\tau) = \int G(R', R, \tau)\Phi(R', t)dR' \tag{3}$$

where G is a Green function (GF) given as a solution of the next equation

$$-\partial_t G(R', R, t) = (H - E_T)G(R', R, t); \tag{4}$$

with the initial condition $G(R', R, 0) = \delta(R' - R)$. One can verify that GF can be expressed as

$$G(R', R, \tau) = \langle R'| \exp[-\tau(H - E_T)|R\rangle = \sum_\alpha \exp[-\tau(E_\alpha - E_T)]\phi^*_\alpha(R')\phi_\alpha(R) \tag{5}$$

where $\{\phi_\alpha\}$, $\{E_\alpha\}$ denote the complete set of eigenvectors and eigenvalues of H, respectively. One can show the equivalence of the Eqs. (1) and (3) by differentiating (3) according to τ and then setting $\tau = 0$.

Using the above equations it is easy to see that in the limit $\tau \to \infty$ the operator $\exp[-\tau(H - E_T)]$ projects out the lowest eigenstate φ_0 of a given symmetry from an arbitrary trial function Ψ with the same symmetry and nonzero overlap:

$$\lim_{\tau\to\infty} \exp[-\tau(H - E_T)]\Psi = \lim_{\tau\to\infty} \int G(R', R, \tau)\Psi(R')dR' =$$

$$= \lim_{\tau\to\infty} \langle\Psi|\varphi_0\rangle \exp[-\tau(E_0 - E_T)]\varphi_0(R) \tag{6}$$

where by setting $E_T = E_0$ the exponential factor will be constant while the higher states will be exponentially damped. This is the fundamental relation of the diffusion (projector) Monte Carlo method which we will describe in a more detail in the coming text.

1.1. IMAGINARY TIME SCHRÖDINGER EQUATION AND ITS RELATION TO DIFFUSION PROCESSES WITH BRANCHING; REPRESENTATION OF THE WAVEFUNCTION BY WALKERS; GREEN FUNCTION FOR PARTICLES IN A POTENTIAL

Let us for a moment neglect the potential terms in (2) so that the Hamiltonian is equal to the kinetic energy operator. The equation (1) simplifies into a *diffusion equation in a 3N-dimensional space*

$$-\partial_t\Phi(R, t) = -(1/2)\sum_{i=1}^N \nabla_i^2\Phi(R, t) \tag{7}$$

The corresponding GF is a gaussian with variance proportional to τ

$$G(R', R, \tau) = (2\pi\tau)^{-3N/2} \exp\left[\frac{-(R - R')^2}{2\tau}\right] \tag{8}$$

The diffusion equation describes a range of physical phenomena such as a heat transfer, Brownian motion of tiny particles in a solution, mixing of two gases, penetration of dopant atoms from a surface of a solid into the bulk, etc. Diffusion processes involve an inherent randomness which, for the examples just mentioned, is caused by random movements of atoms or molecules at nonzero temperature. One can interpret the relation above as a *master equation* of a systems of Brownian particles undergoing *a stochastic diffusion process*[3]. This interpretation enables us to represent the function Φ by a set of "Brownian particles", *i.e.*, by a set of *discrete sampling points* also known as *random walkers*. The equation (3) then suggests that the set of random walkers which is distributed according to some initial Φ is then propagated (evolved) in time t using GF as a transition probability density. Therefore we can set the correspondence

$$\Phi(R,t) \longrightarrow \sum_k \delta(R - R_k) \tag{9}$$

and evolving the walkers for a time interval $\Delta\tau$ by substituting this into (3) we get

$$\Phi(R,t+\Delta\tau) \longrightarrow \sum_k G(R,R_k,\Delta\tau) \tag{10}$$

The result is a set of gaussians with variance equal to $3N\tau$. In order to restore the original delta function representation, the gaussians in (10) are sampled by a new set of delta functions. The whole procedure can then be repeated and after a sufficient number of iterations we reach the long time limit solution and gather enough samples to calculate averages for quantities of interest, as will be explained below.

Let us try a more realistic and complicated case when H includes also potential terms as in Eq. (2). The problem becomes nontrivial as, in general, there is no simple explicit expression for exact GF and we have to use approximations. First, we will invoke the Trotter formula, which, for some given operators A, B is given by

$$e^{-\Delta\tau(A+B)} = e^{-\Delta\tau A/2} e^{-\Delta\tau B} e^{-\Delta\tau A/2} + O[(\Delta\tau)^3] \tag{11}$$

The approximate Green function for small $\Delta\tau$ is then given by

$$G(R,R',\tau) = C_{\text{norm}} \exp\left[-\frac{(R-R')^2}{2\Delta\tau}\right] \exp\left[-\Delta\tau(U(R)+U(R')-2E_T)/2\right] \tag{12}$$

with an error proportional to $(\Delta\tau)^3$ [4]. We already know how to find the solution of the diffusion equation by simulating the diffusion process. How to deal with the more complicated case of GF we have just derived?

The term in (12) which involves potential energy

$$P = \exp\left[-\Delta\tau(U(R)+U(R')-2E_T)/2\right] \tag{13}$$

acts as a reweighting (renormalization) of GF. This can be incorporated into the stochastic process in various ways. One possibility is to assign each walker a weight and then to accumulate the product of weights as the propagation goes on. This procedure has to be further refined (see lectures of Stefano Baroni) otherwise, if the evolution goes for too long, some of the weights will grow very rapidly. Ultimately, one will end up with a single walker with exponentially dominating weight. Another possibility is to use the *branching (proliferation/decay) algorithm* in which P is interpreted as a probability to survive to the next step[4]. Therefore:

i) if $P < 1$ walker continues in the evolution with the probability P
ii) if $P > 1$ walkers continues and, in addition, with the probability $P - 1$ new walker(s) with the same coordinates are created.

Both possibilities can be conveniently coded as a single command

$$M_{\mathrm{new}} = \mathrm{INT}(P + \eta) \tag{14}$$

where M_{new} is number of walkers continuing to the next step at the given position R, INT is an integer part of a real number and η is a random number drawn from a uniform distribution on interval (0,1). From the expression for P one can infer that in the regions of high potential energy the walkers will be disappearing while in the places of low potential energy the walkers will be proliferating. Branching prevents the catastrophic weight accumulation in the favorable regions and rather changes that to a higher density of walkers at that place. The role of the energy offset E_T is now clear: E_T will be adjusted so that the overall number walkers will fluctuate around some prescribed value. Typical number of walkers used in simulations is of the order of 10^2 to 10^4.

1.2. IMPORTANCE SAMPLING AND GREEN FUNCTION WITH DRIFT; ESTIMATORS; SIMPLE DMC ALGORITHM

The outlined simple version of the algorithm is rather inefficient, because the potential in the exponential can vary significantly from step to step and can even diverge at particle collision points. This makes the process very inefficient and in case of divergences also not well defined.

One can overcome these difficulties by using importance sampling with a trial (or variational) wavefunction Ψ. The concept of importance sampling is well-known in Monte Carlo and was introduced into the QMC calculations by Malvin Kalos in sixties.

Let us multiply the equation (1) by a trial function Ψ. Introducing a function $f = \Phi\Psi$ and rearranging the resulting equation, one gets

$$-\partial_t f(R,t) = Tf - \nabla \cdot [f \nabla \ln |\Psi|] + (E_L(R) - E_T)f \tag{15}$$

where the local energy is defined as $E_L(R) = H\Psi/\Psi$. The corresponding integral equation is given as

$$(R, t+\tau) = \int G^*(R', R, \tau) f(R', t) dR' \tag{16}$$

where $G^*(R', R, \tau) = \Psi(R)G(R', R, \tau)/\Psi(R')$. The short-time $G^*(R', R, \Delta\tau)$ now has the form

$$G^*(R', R, \Delta\tau) \cong (2\pi\Delta\tau)^{-3N/2} \exp\left[-\frac{[R' - R - \Delta\tau\nabla\ln|\Psi|]^2}{2\Delta\tau}\right] \times$$
$$\exp\left[-\Delta\tau(E_L(R') + E_L(R) - 2E_T)/2\right] \tag{17}$$

The key difference with the previous case is that the density of walkers in the regions where Ψ is large will be enhanced and vice versa because of the *drift* term $\Delta\tau\nabla\ln|\Psi|)$ which acts as a "quantum force" [4]. In addition, instead of fluctuations which are proportional to potentials we have the reweighting term which includes local energy $H\Psi/\Psi$ and for a good trial function we expect that the fluctuations will be much smaller. In fact, the importance sampled approach fulfills *the zero-variance property*: as Ψ approaches the exact solution, the local energy converges to the corresponding exact eigenvalue quadratically fast *for each individual walker*. This is one of the important properties which, in fact, enables simulations of large systems with acceptable computational demands. It also suggests that a reasonable human investment in improvement of the trial function is important for the feasibility of large-scale calculations. The improvement in the efficiency by importance sampling is very significant (by two to three orders of magnitude) and enables us to carry out calculations which can involve hundreds and, in a few cases, even thousands of electrons.

The result of the stochastic process described above is a set of samples which represent the solution Φ. How we can evaluate the expectation values, *e.g.*, the energy?

One possibility is to measure the energy through the offset E_T which keeps the population constant (in a statistical sense): so-called generational estimator. More often is used the mixed estimator which is given by

$$E_{DMC} = \lim_{\tau\to\infty} \frac{\langle e^{-\tau H/2}\Psi|H|e^{-\tau H/2}\Psi\rangle}{\langle e^{-\tau H/2}\Psi|e^{-\tau H/2}\Psi\rangle} = \lim_{\tau\to\infty} \frac{\langle e^{-\tau H}\Psi|H|\Psi\rangle}{\langle e^{-\tau H}\Psi|\Psi\rangle} = \frac{\langle\Phi|H|\Psi\rangle}{\langle\Phi|\Psi\rangle}$$
$$= \frac{\int f[H\Psi/\Psi]\,dR}{\int f\,dR} \cong \frac{1}{M}\sum_m \frac{H\Psi(R_m)}{\Psi(R_m)} \tag{17}$$

where $\{R_m\}$ denotes M samples of $f(R)$ resulting from the DMC run.

Expectations of quantities other than energy can be calculated using the mixed estimator and the variational estimator as follows

$$\langle\Phi|S|\Phi\rangle \approx 2\frac{\langle\Phi|S|\Psi\rangle}{\langle\Phi|\Psi\rangle} - \frac{\langle\Psi|S|\Psi\rangle}{\langle\Psi|\Psi\rangle} + O[(\Phi-\Psi)^2] \tag{19}$$

where S denotes the operator of quantity of interest. For nonnegative quantities (*e.g.* density) one can use

$$\langle\Phi|S|\Phi\rangle \approx \left[\frac{\langle\Phi|S|\Psi\rangle}{\langle\Phi|\Psi\rangle}\right]^2 \left[\frac{\langle\Psi|S|\Psi\rangle}{\langle\Psi|\Psi\rangle}\right]^{-1} + O[(\Phi-\Psi)^2] \tag{20}$$

The most basic version of the DMC algorithm with importance sampling can be briefly outlined as follows.

---

1. Generate set of delta functions drawn from some initial distribution (*e.g.* from the trial function Ψ using the variational Monte Carlo).

2. Propagate each walker by time interval τ

$$R' = R + \chi + \tau\nabla \ln|\Psi|$$

where χ is a normally distributed $3N$-dimensional vector with variance τ and zero mean. The last term describes a drift coming from the importance sampling by Ψ.

3. For each walker calculate the number of copies which will continue from

$$M_{\text{new}} = \text{INT}\left\{\exp[-\Delta\tau(E_L(R) + E_L(R') - 2E)/2] + \eta\right\}$$

where η is a random number from uniform distribution on interval (0,1)

4. Accumulate samples of quantities of interest for calculations of averages, *e.g.*, energy is calculated as an average of E_L over the walkers

5. The steps 2-4 are repeated until the error bars as sufficiently small with occasional adjustment of E_T to keep on average the population constant.

---

2. DMC for fermions

2.1. ANTI-SYMMETRY AND FIXED-NODE APPROXIMATION AND FERMION NODES; FIXED-NODE BIAS AND CORRELATION ENERGY

One of the fundamental concepts not considered so far is antisymmetry which implies that any fermionic wavefunction for more than two electrons will have both positive and negative values. However, in presenting the probabilistic method of solution above we tacitly assumed that the distribution (wavefunction) we are sampling is nonnegative everywhere. How to deal with this contradiction?

Naive generalization of the DMC algorithm by assignment of signs to walkers does not work as it is explained in the lecture of Malvin Kalos devoted to the fermion sign problem. Therefore we have to find some other way. An approximate solution, which is employed in much of the current large-scale QMC calculations, is the fixed-node approximation suggested in 1975 by James Anderson [5]. In order to introduce it we will define the *fermion node* for a given wavefunction Ψ:

The fermion node of Ψ *is a subset of the* $3N$*-dimensional space for which* $\Psi(R) = 0$.

It is obvious from the definition that, in general, the fermion node is a $(3N-1)$-dimensional hypersurface. *The fixed-node approximation* can be now stated as follows: we enforce the fermion nodes of our solution $\Phi(R,t)$ (Eq.15) to be identical to the fermion nodes of a given, usually the best available, trial wavefunction Ψ. This immediately solves the nonnegativity problem as the product of $\Phi\Psi$ can then be chosen to be nonnegative everywhere and the whole machinery of DMC described above can be used without difficulties.

The fixed-node approximation involves, to a certain extent a leap in faith, especially for QMC non-practitioners: it changes the problem of (anti)symmetry into a problem of boundary conditions. Any symmetry condition is difficult to fulfill in the DMC algorithm because, *in general, symmetry is a nonlocal property: distant points in the configuration space are required to fulfill certain condition(s).* On the other hand boundary conditions are perfectly local and are very easy to incorporate into the simulation.

An obvious question to ask is what are the general properties of this approximation, how does it work in practice and how much is known about the fermion nodes? One can prove the following:

i) the fixed-node energy is an upper bound to the exact energy, *i.e.*, the method is variational

ii) the fixed-node error decreases quadratically in the error of the nodal surface of the trial function

The obvious consequence is that for the Ψ with exact fermion nodes the energy will be also exact.

Therefore in the DMC method the role of fermion nodes becomes very prominent as the fixed-node condition is the only fundamental approximation (the rest of approximations can be called 'technical' as they just involve how clever the programmer or scientist is in coding a well-defined numerical problem). The nodes are affected by the choice of the trial function and significant effort is being expended to keep the fixed-node error small (see,*e.g.*, lectures of Cyrus Umrigar).

An interesting question is what are the general properties of fermion nodes? Not that much is known about this, for some, fascinating subject. Some general properties are analyzed in paper by Ceperley, still others are given in the paper by Caffarel and coworkers [6].

How can one enforce the fixed-node condition, *i.e.*, prevent the walker to cross the node of the trial function?

First, the importance sampling will, to a large extent, take care of this task. One can see this from the fact that the quantum drift is divergent at the nodal surfaces, meaning that the quantum force is pushing the walkers away as they get close to the nodal region. However, it occasionally can and does happen that the walker crosses the node. One can see this from the fact that the drift is proportional to the $\Delta\tau$ while the diffusion vector components are proportional to $\sqrt{\Delta\tau}$ therefore for small time step the diffusion move can be sufficiently large for the walker to cross the node. This will be manifested by the trial function sign change for positions before and after the move. The simplest remedy is to reject the move and keep the walker into the next iteration at the same position. The bias caused by this and its elimination with a more elaborate algorithm can be found in the paper of Umrigar, Nightingale and Runge [7].

In practice, the typical fixed-node error is of the order of 5% of the correlation energy if we define the correlation energy as the difference between the exact and HF energy:

$$E_{\rm corr} = E_{\rm exact} - E_{\rm Hartree-Fock} \tag{21}$$

(Correlation energy is often used as a guiding measure of the quality and accuracy of the many-body electronic structure methods.) However, the actual percentage depends on the accuracy and character of the trial function (whether one or more determinants are used, type of orbitals, *etc.* This will be further explained in other lectures). The typical amount of recovered correlation energy (95%) is much more than one gets by other methods for systems with more than 10-20 valence electrons. In most cases this error enables predictions of energy differences (binding, cohesion, excitations) with accuracy 1-4 % when compared with experiments [8, 9, 10]. However, this is not universally the case and one would like to have the fixed-node error under control. It is an active area of research and is discussed further in other lectures of this ASI.

2.2. CALCULATION OF REALISTIC SYSTEMS AND CORE ELECTRONS; DMC AND PSEUDOPOTENTIAL TECHNIQUES; DMC AND

NON-LOCAL OPERATORS

The DMC calculations of heavy atom systems are rather difficult because of large energies and corresponding energy fluctuations caused by the core electrons. Hammond and coworkers and also Ceperley derived that the DMC computer time demands grow like Z^6 where Z is an atomic number. Therefore calculations of heavy atoms systems with explicit inclusion of core electrons are very expensive. However, if we are interested in valence properties only, one can eliminate the core states from calculations while keeping accuracy .05 eV or better for the valence related energy differences. To some extent the core problem is present even in one-particle approaches and the so-called effective core potentials (ECP), sometimes called also pseudopotentials, have been developed to deal with this difficulty. In the ECP approach the core electrons are replaced by effective potentials which mimic the impact of core degrees of freedom on valence states. The most commonly used form is given by:

$$V_{Ie}^{\mathrm{ECP}} = \sum_{i,I} [V_{\mathrm{loc}}(r_{iI}) + W_{iI}] \tag{22}$$

where the first term denotes the local part while the nonlocal term is given by

$$W_{iI} = \sum_{\ell,m} v_\ell(r_{iI}) Y_{\ell m}(\Omega_{iI}) \int_{4\pi} d\Omega'_{iI} Y^*_{\ell m}(\Omega'_{iI}) \tag{23}$$

where $v_\ell(r)$ is a radial pseudopotential for the symmetry channel ℓ, $Y_{\ell m}$ is the spherical harmonic and r_{ij}, r_{iI} denote distance of i-th electron from j-th electron and I-th ion, respectively. Note that the presence of the projection operators in (22) makes W a non-local operator, *i.e.*, different symmetry channels have different potentials $v_\ell(r)$. Unfortunately, for DMC this operator is problematic because the nonlocality of W implies that the matrix element $\langle R \mid \exp(-\tau W) \mid R' \rangle$ can be negative for some R, R', τ (and, in fact, it is negative for realistic cases of ECPs). This creates a new "sign problem" with difficulties similar to the fermion sign problem mentioned before.

How to deal with this in DMC? Two possibilities have been explored so far:

a) Define a new *local operator* with the same or similar properties as W. This route had been explored by Bachelet, Ceperley and coworkers in their pseudo-Hamiltonian method.
b) Try to eliminate the nonlocality of W by projecting it onto a sufficiently accurate trial function Ψ. This will be the method which we will describe [11, 12, 13].

The technique of evaluating the nonlocal part of the Hamiltonian using a trial function (see [13]) is based on the following *localization approximation.*

Replacing the electron-ion interaction in (2) by the ECP operator (22) and by defining

$$H_{\rm loc} = T + V_{ee} + \sum_{i,I} V_{\rm loc}(r_{iI}) + E_{II} \tag{24a}$$

$$W = \sum_{Ii} W_{Ii} \tag{24b}$$

we can write Eq. (15) as

$$\partial_t f = \frac{1}{2}\sum_i \nabla_i^2 f - \sum_i \nabla_i \cdot (f \nabla_i \ln \Psi) - \left[\frac{(H_{\rm loc} - E_T)\Psi}{\Psi}\right] f - \left[\frac{W\Phi}{\Phi}\right] f \tag{25}$$

where the last term represents a "nonlocal branching". To remove the action of the nonlocal operator on the unknown wavefunction the last term in this equation is replaced as follows

$$\partial_t f = \frac{1}{2}\sum_i \nabla_i^2 f - \sum_i \nabla_i \cdot (f \nabla_i \ln \Psi) - \left[\frac{(H_{\rm loc} - E_T)\Psi}{\Psi}\right] f - \left[\frac{W\Psi}{\Psi}\right] f \tag{26}$$

If, as we suppose, $\Psi \approx \Phi$ then the error introduced by this approximation is small — indeed we have shown that the error is proportional to the square of the error of the trial function [13]. This also means that for an exact Ψ equation (26) is exact. The important points related to this approximation can be summarized as follows:

a) the new operator $W\Psi/\Psi$ is local and therefore can be incorporated into DMC without fundamental difficulties
b) the energy converges to the exact one quadratically fast in the error of the trial function but the energy is not necessarily an upper bound
c) the method has the zero-variance property

What does it mean in practice? First we have to learn how to evaluate $H\Psi/\Psi$ in (26) therefore we need to learn how to compute the action of W_{iI} on Ψ. This is given by

$$\frac{W_{iI}\Psi}{\Psi} = \sum_\ell \frac{(2\ell+1)}{4\pi} v_\ell(r_{iI}) \int_{4\pi} P_\ell[\cos(\theta'_{iI})] \frac{\Psi(\mathbf{r}_1, ..., \mathbf{r}'_i, ..., \mathbf{r}_N)}{\Psi(\mathbf{r}_1, ..., \mathbf{r}_i, ..., \mathbf{r}_N)} \, d\Omega'_{iI} \tag{13}$$

where P_ℓ denotes a Legendre polynomial and the axes are rotated is such a way that the i-th electron is on the z-axis. The angular integral over the spherical surface with center on the I-th ion can be carried out with a special type of gaussian integration (see [13] and references therein). Typical number of integration points used in practice is 6 to 12 or more, depending on the on the highest ℓ in the ECP operator. For *sp* symmetry channels, 6 to 12 points is in most cases enough.

Because the localization approximation depends on the accuracy of the trial function, some effort has to be spent on getting sufficiently good wavefunctions so that the calculation appears to be in the "low variance/quadratic convergence regime". This is feasible, and in many cases the use of ECPs within the DMC method has opened possibilities for calculations of real systems with 100-200 valence electrons (recent record is more than 1000 valence electrons for Si systems). Although there is no upper bound property, with an accurate wavefunction method behaves in most cases like an upper bound method because the localization error is smaller that the fixed-node error as comparisons with other calculations and experiments show. (In DMC calculations both errors are combined together and it is difficult to distinguish between them solely on the basis of DMC data).

The speed-up resulting from the use of pseudopotentials can be demonstrated by the example of the iron atom [14, 9]. Table 1 gives total energies, typical values for the dispersion of the local energy, decorrelation time κ for obtaining an independent sample of energy normalized to the all electron case, and finally, the efficiency is proportional to $1/(\kappa\sigma^2)$ where σ is the variational fluctuation of the local energy defined as $\sigma^2 = \langle\Psi|[E_L(R) - E_T]^2|\Psi\rangle/\langle\Psi|\Psi\rangle$. The comparison is for the variational Monte Carlo method but qualitatively would be similar also for DMC. It is evident that with increasing size of the core the efficiency but also the systematic errors introduced by the pseudopotentials are increasing. For the given case of iron the best compromise, if we accept the accuracy level 0.1 eV, is the Ne core. This comparison gives a qualitative picture for various choices of valence space and should not be taken as a definitive one: to some extent one can always change some of these factors through improvement of the trial function, more efficient sampling and so forth.

It is interesting that very recently a new algorithm was suggested which keeps the upper bound property by combining the localization approximation and nonlocal sampling (by ten Haaf and coworkers [15]). Unfortunately, there is a price to be paid: the algorithm has no zero variance property (see [9, 8]). This might be an issue for continuous space calculations of realistic systems which are being done with very accurate wavefunctions, *i.e.*, one usually tends to be very close to the low variance regime otherwise the calculations might be very expensive. However, it is very useful in lattice models as demonstrated in the lecture of Giovanni Bachelet.

TABLE 1. VMC of the Fe atom with various number of valence electrons

	all-electron	Ne core	Ar core
E_{HF}(a.u.)	-1262.444	-123.114	-21.387
E_{VMC}(a.u.)	-1263.20(2)	-123.708(2)	-21.660(1)
σ^2	≈ 50	1.54	0.16
$\kappa/\kappa_{\rm all}$	1	≈ 0.3	≈ 0.05
efficiency	0.02	2.1	125.
valence errors (eV)	0.	≈ 0.1	≈ 0.5

3. DMC: only a clever algorithm or a new mainstream method?

The DMC method has proved to be remarkably powerful. A number of applications such as calculations of carbon and silicon clusters, band gaps for insulating solids or accurate treatment of molecular reactions show the potential of the method (see, for example, our recent projects [16]). To put the the DMC method into a broader perspective we provide a comparison of methods in the Table 2 below [8]. The table compares variational and diffusion Monte Carlo (VMC and DMC) Hartree-Fock (HF), Local Density Approximation (LDA) and Coupled Cluster (CC) approaches. This is by no means an exhaustive comparison as we have used results scattered over the literature and our own tests mostly for *sp* systems with a few elements such as C,N,H and Si. Therefore the compared characteristics should be understood as *typical* values. Compared quantities are listed as follows:

i) Percentage of valence correlation energy ($E_{\rm corr}= E_{\rm exact}- E_{\rm HF}$) recovered, estimated by comparisons with high accuracy calculations for small systems and by combining HF, correlated atomic and experimental data for larger systems.
ii) Errors in binding or cohesive energies.
iii) Scaling of computational demands for total energies as a function of the number of electrons.
iv) Actual computational time for C_{10} cluster. We used the GAMESS code for the HF calculations and GAUSSIAN94 for CCSD(T); the target QMC error bar was 0.01 eV/atom and the absolute unit of time was 90 seconds of Cray C90 processor.

The last quantity is a quite ambiguous one, as it depends on implementation, type of computer, *etc.* and can vary easily by a factor of 3 or more. Therefore, for the computer time only the orders of magnitude are meaningful.

TABLE 2. Comparison of methods

Method	E_{corr}	E_{coh}; E_{bind} errors	scaling in # of corr. e^-	total time for C_{10}
HF	0	≈50%	N^3	14
LDA	N/A	15-25%	N^3 [(a)]	1
VMC	≈ 85%	2-10%	$N^3 + \varepsilon N^4$ [(b)]	16
DMC	≈ 95%	1-4%	$N^3 + \varepsilon N^4$ [(b)]	300
CCSD(T)	≈ 75% [(c)]	10-15% [(c)]	N^7	1500 [(c)]

[(a)] recently, LDA algorithms with scaling proportional to N have been proposed
[(b)] $\varepsilon \approx 10^{-3} - 10^{-4}$
[(c)] with 6-311G* basis set; in the limit of infinite basis set and order of excitations the CC method is formally exact and therefore it can provide essentially exact results for small, few-atom systems.

The results are quoted for 6-311G* basis sets and SD(T) level of excitations which was the largest run I could afford for the C_{10} systems. Obviously, the QMC method provides a new alternative for the electronic structure calculations. Although QMC is much less developed than the traditional approaches its potential to "solve" the many-body Schrödinger problem is promising.

What are the main practical bottlenecks of the DMC method? In Table 3 we outline the main tasks which are being carried out during a typical DMC calculation together with their human and computer time demands.

The table shows that from the point of view of the human time the most demanding activities involve generating the orbitals and optimization of the wavefunctions. Obviously this points out the importance of integration of the quantum chemistry and condensed matter electronic structure software with QMC methods and also the importance of better optimization methods and more efficient functional forms for variational wavefunctions.

On the other hand, savings in computer time should appear in the DMC method — we believe that a speed-up of the order of one magnitude can be obtained by several improvements in the DMC algorithms.

TABLE 3. Human and computer time costs of a typical DMC calculation

Human time	**Task**	**Computer time**
10%	Choice of basis set ECPs, etc	0 %
40%	Hartree-Fock and/or post-HF orbitals	10%
40%	corr. parameters optimization, VMC	20%
10%	DMC	70%

Acknowledgements

This work was supported by NSF, ARPA, NCSA and State of Illinois funds.

References

1. D.M. Ceperley and M.H. Kalos (1979), Quantum Monte Carlo, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder, Springer, Berlin; K.E. Schmidt and M.H. Kalos (1984), in *Monte Carlo Methods in Statistical Physics II*, edited by K. Binder, Springer, Berlin: *Introductory reviews.*
2. B.L. Hammond, W.A. Lester, Jr., and P.J. Reynolds (1994), *Monte Carlo Methods in ab initio quantum chemistry*, World Scientific, Singapore: *Very good book on quantum Monte Carlo, includes developments up to 1992.*
3. Karlin S., Taylor H.M. (1981), A Second Course in Stochastic Processes, Academic Press, New York: *Excellent book on theory of stochastic processes.*
4. P.J. Reynolds, D.M. Ceperley, B.J. Alder, and W.A. Lester (1982) J. Chem. Phys. **77**, 5593; J.W. Moskowitz, K.E. Schmidt, M.A. Lee, and M.H. Kalos (1982) J. Chem. Phys. **77**, 349; *ibid.* (1984) **81**, 5833: *The first complete description of the fixed-node DMC algorithm with branching.*
5. J.B. Anderson (1975), J. Chem. Phys. **63**, 1499; *ibid.* (1975), **65**, 4121: *Early James Anderson's papers, the second one simulating wavefunction with node.*
6. D.M. Ceperley (1991), Fermion nodes, J. Stat. Phys. **63**, 1237; M. Caffarel et al (1992), Europhys. Lett. **20**, 581: *Properties of fermion nodes.*
7. C.J. Umrigar, M.P. Nightingale, and K.J. Runge (1993), J. Chem. Phys. **99**, 2865: *New improved version of the DMC algorithm with a detailed analysis.*
8. L. Mitas (1996), Electronic structure by Quantum Monte Carlo: atoms, molecules and solids, Computer Physics Communications, **97**, 107; *Recent review and comparisons with other methods.*
9. D.M. Ceperley and L. Mitas (1996), Monte Carlo methods in quantum chemistry, Advances in Chemical Physics, Vol. XCIII, Ed. by I. Prigogine and S.A. Rice, Wiley, New York, pp. 1-38.; *Recent review.*

10. J.B. Anderson (1995), Int. Rev. Phys. Chem. **14**, 85: *Recent review.*
11. M.M. Hurley and P.A. Christiansen (1987), J. Chem. Phys. **86**, 1069; B. L. Hammond, P.J. Reynolds, and W.A.Lester, Jr. (1987), *ibid.* **87**, 1130 *The first trials to use ECPs within DMC for two electron systems.*
12. S. Fahy, X.W. Wang, and S.G. Louie (1988), Phys. Rev. Lett. **61**, 1631: *The first VMC calculations of solids using ECPs.*
13. L. Mitas, E.L. Shirley, and D.M. Ceperley (1991), Diffusion Monte Carlo and nonlocal pseudopotentials, J. Chem. Phys. **95**, 3467: *Paper demonstrating the use of ECP within DMC, theory, multi-electron applications, gaussian quadratures for sphere.*
14. L. Mitas (1994), Diffusion Monte Carlo calculation of the Fe atom, Phys. Rev. A **49**, 4411: *DMC calculations of the Fe atom.*
15. D.F.B. ten Haaf, H.J.M. van Bemmel, J.M.J. van Leeuwen, W. van Saarlos and D. M. Ceperley (1994), Phys. Rev. B, **51**, 13039 : *The fixed-node DMC method for lattice models.*
16. J.C. Grossman and L. Mitas (1997), High accuracy molecular heats of formation and reaction barriers: Essential role of electron correlation, Phys. Rev. Lett. **79**, 4353; J.C. Grossman, L. Mitas (1995), Quantum Monte Carlo Determination of Electronic and Structural Properties of Si_n Clusters ($n \leq 20$), Phys. Rev. Lett. **74**, 1323; J.C. Grossman, L. Mitas, and K. Raghavachari (1995), Structure and Stability of Molecular Carbon: Importance of Electron Correlation, Phys. Rev. Lett. **75**, 3870; L. Mitas and R.M. Martin (1994), Quantum Monte Carlo of the nitrogen: atom, dimer, atomic and molecular solids, Phys. Rev. Lett. **72**, 2438: *Some of our recent projects: N systems, Si and C clusters, molecular reactions.*