

Quantum Monte Carlo Methods in Chemistry

Synonyms and Acronyms

Fixed-node diffusion Monte Carlo (FN-DMC); Green's function Monte Carlo (GFMC); Pure diffusion Monte Carlo (PDMC); Reptation Monte Carlo (RMC); Stochastic reconfiguration Monte Carlo (SRMC); Variational Monte Carlo (VMC)

Description of the Problem

The problem considered here is to obtain accurate solutions of the time-independent Schrödinger equation for a general molecular system described as N electrons moving within the external potential of a set of fixed nuclei. This problem can be considered as the central problem of theoretical and computational chemistry. Using the atomic units adapted to the molecular scale the Schrödinger equation to solve can be written as

$$H\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad (1)$$

where H is the Hamiltonian operator given by

$$H = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + V(\mathbf{r}_1, \dots, \mathbf{r}_N), \quad (2)$$

$\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ the spatial positions of the N electrons, $\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}$ the Laplacian operator for

electron i of coordinates $\mathbf{r}_i = (x_i, y_i, z_i)$, Ψ the wavefunction, E the total energy (a real constant), and V the potential energy function expressed as

$$V(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i < j} \frac{1}{r_{ij}} - \sum_{i, \alpha} \frac{Z_\alpha}{r_{i\alpha}} + \sum_{\alpha < \beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \quad (3)$$

In this formula $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the interelectronic distance, Z_α the charge of nucleus α (a positive integer), \mathbf{R}_α its vector position, $r_{i\alpha} = |\mathbf{r}_i - \mathbf{R}_\alpha|$, and $R_{\alpha\beta} = |\mathbf{R}_\alpha - \mathbf{R}_\beta|$. The Schrödinger equation being invariant under complex conjugation, we can restrict without loss of generality the eigensolutions to be real-valued. The boundary conditions are of Dirichlet-type: Eigenfunctions Ψ are imposed to vanish whenever one electron (or more) goes to infinity

$$\Psi \rightarrow 0 \text{ as } \sqrt{\mathbf{r}_1^2 + \dots + \mathbf{r}_N^2} \rightarrow +\infty \quad (4)$$

In addition, the mathematical constraints resulting from the Pauli principle must be considered. Within a space-only formalism as employed in QMC, two types of electron – usually referred to as the “spin-up” and “spin-down” electrons – are distinguished and the Pauli principle is expressed as follows. Among all eigenfunctions verifying (1)–(4) only those that are antisymmetric under the exchange of any pair of spin-like electrons are physically allowed. Because of the permutational invariance, the N_\uparrow spin-up electrons can be arbitrarily chosen as those having the first labels and the mathematical conditions can be written as

$$\begin{aligned} & \Psi(\dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots | \mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N) \\ &= -\Psi(\dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots | \mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N) \end{aligned} \quad (5a)$$

and

$$\begin{aligned} & \Psi(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow} | \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots) \\ &= -\Psi(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow} | \dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots) \end{aligned} \quad (5b)$$

for all pairs (i, j) of spin-like electrons. Equations 1–5b define the mathematical problem discussed here. Although such a mathematical model results from a number of physical approximations, it contains the bulk of most chemical phenomena and solving it with enough accuracy (=chemical accuracy) can be considered as the major problem of computational chemistry. The two standard approaches to deal with the electronic structure problem in chemistry are the density functional theory (DFT) (Density Functional Theory) and the post-Hartree–Fock wavefunction approaches (Post-Hartree Fock Methods and Excited States Modelling, Coupled-Cluster Methods). Quantum Monte Carlo (QMC) presented here may be viewed as an alternative approach aiming at circumventing the limitations of these two well-established methods (for a detailed presentation of QMC, see, e.g., [1]). In contrast with these deterministic approaches, QMC is based on a stochastic sampling of the electronic configuration space. In the recent years, a number of remarkable applications have been presented, thus establishing QMC as a high potential approach although a number of limitations are still present. Here, we shall present the two most popular approaches used in chemistry, namely, the variational Monte Carlo (VMC) and the fixed-node diffusion Monte Carlo (FN-DMC) methods.

The Variational Monte Carlo (VMC) Method

The variational Monte Carlo (VMC) method is the simpler and the most popular quantum Monte Carlo approach. From a mathematical point of view, VMC is a standard Markov chain Monte Carlo (MCMC) method. Introducing an approximate trial wavefunction $\Psi_T(\mathbf{r}_1, \dots, \mathbf{r}_N)$ known in an analytic form (a good approximation of the unknown wavefunction), the Metropolis-Hastings algorithm is used to generate sample points distributed in the 3N-dimensional configuration space according to the quantum-mechanical probability density π associated with Ψ_T

$$\pi(\mathbf{R}) = \frac{\Psi_T^2(\mathbf{R})}{\int d\mathbf{R} \Psi_T^2(\mathbf{R})} \quad (6)$$

where \mathbf{R} is a compact notation representing the positions of the N electrons, $\mathbf{R}=(\mathbf{r}_1, \dots, \mathbf{r}_N)$. Expectation values corresponding to various physical properties can be rewritten as averages over π . As an important example, the total energy defined as

$$E_{VMC}(\Psi_T) \equiv \frac{\int d\mathbf{R} \Psi_T(\mathbf{R}) H \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T^2(\mathbf{R})} \quad (7)$$

may be rewritten under the form

$$E_{VMC}(\Psi_T) = \int d\mathbf{R} \pi(\mathbf{R}) E_L(\mathbf{R}) \quad (8)$$

where $E_L(\mathbf{R})$ is the local energy defined as

$$E_L(\mathbf{R}) = \frac{H \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}. \quad (9)$$

In VMC, the total energy is thus estimated as a simple average of the local energy over a sufficiently large number K of configurations $\mathbf{R}^{(k)}$ generated with the Monte Carlo procedure

$$E_{VMC} \simeq \frac{1}{K} \sum_{k=1}^K E_L[\mathbf{R}^{(k)}], \quad (10)$$

the estimator becoming exact as K goes to infinity with a statistical error decreasing as $\sim \frac{1}{\sqrt{K}}$. Properties other than the energy can be obtained in a similar way.

In the case of the energy, it can be shown that there exists a variational principle expressed as $E_{VMC}(\Psi_T) \geq E_0$ for any Ψ_T , the equality being obtained for the exact ground-state wavefunction of energy E_0 . In addition, there also exists a zero-variance principle stating that the closer the trial wavefunction is from the exact solution, the smaller the fluctuations of the local energy are, the statistical error vanishing in the limit of an exact trial wavefunction. In practice, both principles – minimization of the energy and/or of the fluctuations of the local energy – are at the basis of the various approaches proposed for optimizing the parameters entering the trial wavefunction.

The Diffusion Monte Carlo (DMC) Method

The fundamental idea is to introduce a formal mathematical connection between the quantum and stochastic worlds by introducing a fictitious time dynamics as follows

$$\frac{\partial \Psi(\mathbf{R}, t)}{\partial t} = -[H(\mathbf{R}) - E_T] \Psi(\mathbf{R}, t) \quad (11)$$

where t plays the role of a time variable, $\Psi(\mathbf{R}, t)$, a time-dependent real wavefunction, and E_T , some constant reference energy. The solution of this equation is uniquely defined by the choice of the initial wavefunction, $\Psi(\mathbf{R}, t=0)$. Using the spectral decomposition of the self-adjoint (hermitic) Hamiltonian operator, the solution of (11) can be written as

$$\Psi(\mathbf{R}, t) = \sum_i c_i e^{-t(E_i - E_T)} \psi_i(\mathbf{R}) \quad (12)$$

where the sum is performed over the complete set of the eigensolutions of the Hamiltonian operator

$$H(\mathbf{R}) \psi_i(\mathbf{R}) = E_i \psi_i(\mathbf{R}), \quad (13)$$

and $c_i = \int d\mathbf{R} \psi_i^*(\mathbf{R}) \Psi(\mathbf{R}, 0)$.

As seen from (12) the knowledge of the time-dependent solution of the Schrödinger equation allows to have direct access to information about the time-independent eigensolutions, $\psi_i(\mathbf{R})$. As an important example, the exact ground-state wavefunction (corresponding to the smaller eigenvalue E_0) can be obtained by considering the large-time limit of the time-dependent wavefunction

$$\lim_{t \rightarrow +\infty} \Psi(\mathbf{R}, t) = \psi_0(\mathbf{R}) \quad (14)$$

up to an unessential multiplicative factor.

In practice, to have an efficient Monte Carlo simulation of the original time-dependent equation, we need to employ some sort of importance sampling, that is, a practical scheme for sampling only the regions of the very high-dimensional configuration space where the quantities to be averaged have a non-vanishing contribution. Here, it is realized by

introducing a trial wavefunction Ψ_T (usually optimized in a preliminary VMC step) and by defining a new time-dependent density as follows

$$\pi(\mathbf{R}, t) \equiv \Psi_T(\mathbf{R})\Psi(\mathbf{R}, t). \quad (15)$$

The equation that π obeys can be derived without difficulty from (11) and (15), we get

$$\frac{\partial \pi(\mathbf{R}, t)}{\partial t} = L\pi(\mathbf{R}, t) - [E_L(\mathbf{R}) - E_T]\pi(\mathbf{R}, t), \quad (16)$$

where L is a forward Fokker-Planck operator defined as (see, e.g., [2])

$$L\pi = \frac{1}{2} \nabla^2 \pi - \nabla[\mathbf{b}(\mathbf{R})\pi] \quad (17)$$

and $\mathbf{b}(\mathbf{R})$ the drift vector given by

$$\mathbf{b}(\mathbf{R}) = \frac{\nabla \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}. \quad (18)$$

In order to define a step-by-step Monte Carlo algorithm, the fundamental equation (16) is rewritten under the following equivalent integral form describing the evolution of the density during a time interval τ

$$\pi(\mathbf{R}, t + \tau) = \int d\mathbf{R}' K(\mathbf{R}, \mathbf{R}', \tau) \pi(\mathbf{R}', t) \quad (19)$$

where K is the following integral kernel (or imaginary-time propagator)

$$K(\mathbf{R}, \mathbf{R}', \tau) = \langle \mathbf{R}, e^{\tau L - \tau(E_L - E_T)} \mathbf{R}' \rangle. \quad (20)$$

For an arbitrary value of τ , the kernel is not known. However, for small enough time-step accurate approximations of K can be obtained and sampled. To see this, let us first split the exponential operator into a product of exponentials by using the Baker-Campbell-Hausdorff formulas [3]

$$e^{\tau L - \tau(E_L - E_T)} = e^{-\frac{\tau}{2}(E_L - E_T)} e^{\tau L} e^{-\frac{\tau}{2}(E_L - E_T)} + O(\tau^3) \quad (21)$$

and then introduce a short-time gaussian approximation of the Fokker-Planck kernel [2],

$$\langle \mathbf{R}, e^{\tau L} \mathbf{R}' \rangle \simeq \left(\frac{1}{\sqrt{2\pi\tau}} \right)^{3N} e^{-\frac{(\mathbf{R}' - \mathbf{R} - \tau\mathbf{b}(\mathbf{R}))^2}{2\tau}} \quad (22)$$

Finally, a working short-time approximation of the DMC kernel can be written as

$$K_{DMC}(\mathbf{R}, \mathbf{R}', \tau) \simeq \left(\frac{1}{\sqrt{2\pi\tau}} \right)^{3N} e^{-\frac{(\mathbf{R}' - \mathbf{R} - \tau\mathbf{b}(\mathbf{R}))^2}{2\tau}} e^{-\frac{\tau}{2}[(E_L(\mathbf{R}') - E_T) + (E_L(\mathbf{R}) - E_T)]} \quad (23)$$

By considering small enough τ , the residual error (called the short-time error in the context of QMC) can be made arbitrarily small. In practice, the DMC simulation is performed as follows. A population of walkers [or configuration

$\mathbf{R}^{(k)}$] propagated stochastically from generation to generation according to the DMC kernel is introduced. At each

step, the walkers are moved according to the gaussian transition probability, (22). Next, each walker is killed, kept unchanged, or duplicated a certain number of times proportionally to the remaining part of the K_{DMC} kernel, namely,

$w = e^{-\frac{1}{2}[(E_L(\mathbf{R}') - E_T) + (E_L(\mathbf{R}) - E_T)]}$. In practice, an unbiased integer estimator M defining the number of copies ($M = 0, 1, \dots$) is used, $M = E[w + u]$, where E is the integer part and u is a uniform random number in $(0, 1)$ (unbiased $\Rightarrow \int_0^1 du M = w$). In contrast with the Fokker-Planck part, this branching (or birth-death) process causes

fluctuations in the number of walkers. Because of that, some sort of population control step is needed [1]. The stationary distribution resulting from these stochastic rules can be obtained as the time-independent solution of (16). After some simple algebra we get

$$\pi(\mathbf{R}) = \frac{\Psi_T(\mathbf{R})\Psi_0(\mathbf{R})}{\int d\mathbf{R}\Psi_T(\mathbf{R})\Psi_0(\mathbf{R})} \quad (24)$$

provided the reference energy E_T is adjusted to the exact value, $E_T = E_0$. From this mixed DMC distribution density, a simple and unbiased estimator of the total energy is obtained

$$E_0 = \langle E_L(\mathbf{R}) \rangle_{\pi}. \quad (25)$$

For properties other than the energy, the exact distribution density, Ψ_0^2 , must be sampled. This can be realized in different ways, for example, by using a forward walking scheme Ref.[4] or a reptation Monte Carlo algorithm, Ref.[5].

The Fixed-Node Approximation

In the preceding section, the DMC approach has been presented without taking care of the specific mathematical constraints resulting from the Pauli principle, (5b). As it is, this algorithm can be directly employed for quantum systems not subject to such constraints (bosonic systems, quantum oscillators, ensemble of distinguishable particles, etc.). An important remark is that the algorithm converges to the stationary density, (24), associated with the lowest eigenfunction

$\psi_0(\mathbf{R})$ which, in the case of a Hamiltonian of the form $H = -\frac{1}{2}\nabla^2 + V$, is known to have a constant sign (say,

positive). This property is the generalization to continuous operators of the Perron-Frobenius theorem valid for matrices with off-diagonal elements of the same sign.

For electronic systems, the additional fermionic constraints are to be taken into account and we must now force the DMC algorithm to converge to the lowest eigenfunction obeying the Pauli principle (the “physical” or fermionic ground-state) and not to the “mathematical” (or bosonic) ground-state having a constant sign. Unfortunately, up to now it has not been possible to define a computationally tractable (polynomial) algorithm implementing exactly such a property for a general fermionic system (known as the “sign problem”). However, at the price of introducing a fixed-node approximation, a stable method can be defined. This approach called fixed-node DMC (FN-DMC) just consists in choosing a trial wavefunction fulfilling the fermionic constraints, (5b). In contrast with the bosonic-type simulations where the trial wavefunction does not vanish at finite distances, the walkers are now no longer free to move within the entire configurational space. This property results directly from the fact that the nodes of the trial wavefunction [defined as the $(3N - 1)$ -dimensional hypersurface where $\Psi_T(\mathbf{R})=0$] act as infinitely repulsive barriers for the walkers [divergence of the drift vector, (18)]. Each walker is thus trapped forever within the nodal pocket cut by the nodes of Ψ_T where it starts from and the Schrödinger equation is now solved with the additional fixed-node boundary conditions defined as

$$\psi(\mathbf{R})=0 \text{ whenever } \Psi_T(\mathbf{R})=0. \quad (26)$$

When the nodes of ψ_T coincide with the exact nodes, the algorithm is exact. If not, a fixed-node error is introduced. Hopefully, all the nodal pockets do not need to be sampled – which would be an unrealistic task for large systems – due

to the existence of a “tiling” theorem stating that all the nodal pockets of the fermionic ground-state are essentially equivalent and related by permutational invariance [6]. For a mathematical presentation of the fixed-node approximation, see Ref.[7]. Finally, remark that in principle defining an exact fermionic DMC scheme avoiding the fixed-node approximation is not difficult. For example, by letting the walkers go through the nodes and by keeping track of the various changes of signs of the trial wavefunction. However, in practice all the schemes proposed up to now are faced with the existence of an exponentially vanishing signal-to-noise problem related to the uncontrolled fluctuations of the trial wavefunction sign. For details, the reader is referred to the work by Ceperley and Alder [8].

The Trial Wavefunction

A standard form for the trial wavefunction is

$$\Psi_T(\mathbf{R}) = e^{J(\mathbf{R})} \sum_k c_k \text{Det}_k^\uparrow(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow}) \text{Det}_k^\downarrow(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N). \quad (27)$$

where the term $e^{J(\mathbf{R})}$ is usually referred to as the Jastrow factor describing explicitly the electron-electron interactions at different level of approximations. A quite general form employed for $J(\mathbf{R})$ is

$$J(\mathbf{R}) = \sum_\alpha U^{(e-n)}(r_{i\alpha}) + \sum_{i<j} U^{(e-e)}(r_{ij}) + \sum_{\alpha i < j} U^{(e-e-n)}(r_{ij}, r_{i\alpha}, r_{j\alpha}) + \dots \quad (28)$$

where U’s are simple functions (Many different expressions have been employed). The second part of the wavefunction is quite standard in chemistry and describes the shell-structure of molecules via a linear combination of a product of two Slater determinants built from one-electron molecular orbitals. Note that several other forms for the trial wavefunction have been introduced in the literature but so far they have remained of marginal use. Finally, let us emphasize that the magnitude of the statistical error and the importance of the fixed-node bias being directly related to the quality of the trial wavefunction (both errors vanish in the limit of an exact wavefunction), it is in general quite profitable to optimize the parameters of the trial wavefunction. Several approaches have been proposed, we just mention here the recently proposed method of Umrigar and collaborators [9].

Applications

In computational chemistry, the vast majority of the VMC and FN-DMC applications have been concerned with the calculation of total energies and differences of total energies: atomization energies, electronic affinities, ionization potentials, reaction barriers, excited-state energies, etc. To get a brief view of what can be achieved with QMC, let us mention the existence of several benchmark studies comparing FN-DMC with the standard DFT and post-HF methods [10–12]. In such studies, FN-DMC appears to be as accurate as the most accurate post-HF methods and advanced DFT approaches. In addition, like DFT – but in sharp contrast with the post-HF methods – the scaling of the computational cost as a function of the system size is favorable, typically in $O(N^3)$. However, QMC simulations are much more CPU-intensive than DFT ones. To date the largest systems studied involve about 2,000 active electrons, see, e.g., [13]. Finally, note that in principle, all chemical properties can be evaluated using QMC. Unfortunately, to reach the desired accuracy is often difficult in practice. More progress is needed to improve the QMC estimators of such properties.

QMC and High-Performance Computing (HPC)

Let us end by emphasizing on one of the most important practical aspect of QMC methods, namely, their remarkable adaptation to high performance computing (HPC) and, particularly, to massive parallel computations. As most Monte Carlo algorithms, the computational effort is almost exclusively concentrated on pure CPU (“number crunching method”).

In addition, – and this is the key aspect for massive parallelism – calculations of averages can be decomposed at will: n Monte Carlo steps over a single processor being equivalent to n/p steps over p processors with no communication between the processors (apart from the initial/final data transfers). Very recently, it has been demonstrated that an almost perfect parallel efficiency up to about 100,000 compute cores is achievable in practice [14, 15]. In view of the formidable development of computational platforms: Presently up to a few hundreds of thousands compute cores (petascale platforms) and many more soon (exascale in the near future) this property could be critical in assuring the success of QMC in the years to come.

References

1. Foulkes, W.M.C., Mitas, L., Needs, R.J., Rajagopal, G.: Quantum Monte Carlo simulations of Solids. *Rev. Mod. Phys.* 73, 33–83 (2001)
2. Risken, H.: *The Fokker-Planck Equation: Methods of Solutions and Applications*. Springer Series in Synergetics, 3rd edn. Springer, Berlin (1996)
3. Gilmore, R.: Baker-Campbell-Hausdorff formulas. *J. Math. Phys.* 15, 2090–2092 (1974)
4. Caffarel, M., Claverie, P.: Development of a pure diffusion quantum Monte Carlo method using a full generalized Feynman-Kac formula. I. Formalism. *J. Chem. Phys.* 88, 1088–1099 (1988)
5. Baroni, S., Moroni, S.: Reptation quantum Monte Carlo: a method for unbiased ground-state averages and imaginary-time correlations. *Phys. Rev. Lett.* 82, 4745–4748 (1999)
6. Ceperley, D.M.: Fermion nodes. *J. Stat. Phys.* 63, 1237–1267 (1991)
7. Cancès, E., Jourdain, B., Lelièvre, T.: Quantum Monte Carlo simulation of fermions. A mathematical analysis of the fixed node approximation. *Math. Model Method App. Sci.* 16, 1403–1440 (2006)
8. Ceperley, D.M., Alder, B.J.: Quantum Monte Carlo for molecules: Green's function and nodal release. *J. Chem. Phys.* 81, 5833–5844 (1984)
9. Umrigar, C.J., Toulouse, J., Filippi, C., Sorella, S., Hennig, R.G.: Alleviation of the Fermion-sign problem by optimization of many-body wave functions. *Phys. Rev. Lett.* 98, 110201 (2007)
10. Manten, S., Lüchow, A.: On the accuracy of the fixed-node diffusion quantum Monte Carlo methods. *J. Chem. Phys.* 115, 5362–5366 (2001)
11. Grossman, J.C.: Benchmark QMCarlo calculations. *J. Chem. Phys.* 117, 1434–1440 (2002)
12. Nemeč, N., Towler, M.D., Needs, R.J.: Benchmark all-electron ab initio quantum Monte Carlo calculations for small molecules. *J. Chem. Phys.* 132, 034111-7 (2010)
13. Sola, E., Brodholt, J.P., Alfè, D.: Equation of state of hexagonal closed packed iron under Earth's core conditions from quantum Monte Carlo calculations. *Phys. Rev. B* 79: 024107-6 (2009)
14. Esler, K.P., Kim, J., Ceperley, D.M., Purwanto, W., Walter, E.J., Krakauer, H., Zhang, S.: Quantum Monte Carlo algorithms for electronic structure at the petascale; the endstation project. *J. Phys. Conf. Ser.* 125 012057 (2008)
15. Gillan, M.J., Towler, M.D., Alfè, D.: Petascale computing opens new vistas for quantum Monte Carlo Psi-k Highlight of the Month (February, 2011) (2011)

Quantum Monte Carlo Methods in Chemistry

Michel Caffarel Laboratoire de Chimie et Physique Quantiques, IRSAMC, Université de Toulouse, Toulouse, France

DOI: 10.1007/SpringerReference_333776

URL: <http://www.springerreference.com/index/chapterdbid/333776>

Part of: Encyclopedia of Applied and Computational Mathematics

Editors: -

PDF created on: July, 16, 2012 13:27

© Springer-Verlag Berlin Heidelberg 2012