# Quantum Monte Carlo in Chemistry

D. Bressanini, M. Mella and G. Morosi

*Dipartimento di Chimica Fisica ed Elettrochimica, Via Golgi 19, 20133 Milano, Italy*

Quantum Monte Carlo methods are rapidly gaining acceptance in the chemistry community due to their simplicity and to their capability to cope with the many body problem. While computationally demanding, QMC methods are not affected by problems common to other computational methods, such as the choice of a basis set for electronic structure or vibrational calculations.

In this overview, we include Variational Monte Carlo techniques, that use a Monte Carlo procedure to compute expectation values for the energy and other properties of a trial wavefunction, and Diffusion or Green's Function methods which try to directly simulate the Schrödinger equation with very few or no approximations at all.

We will first introduce the main ideas of QMC methods, and then describe how these techniques have been applied in the last few years to chemical problems.

## Introduction

Chemistry problems are intrinsically many-body problems. This means that simple trial wavefunction like monodeterminantal Hartee-Fock model are not able to describe correctly some part of the relevant physics of the system like potential energy curves and surface, low interaction energy between molecules and transiction matrix elements. To correct the Hartree-Fock description of a many electron system it's needed to introduce particles "correlation". This means that electron are not independent particles and that their description need to contain "implicitly or explicitly" interparticles distance.

Starting from this reality many methods was developed to overcome this difficulty: perturbation teory, multiconfiguration self consistent field, complet active space self consistent field, configuration interaction, multi reference configuration interaction and so on.

All this methods introduce in implicit way interparticles distance and their convergency speed are very low so that very long expanction are needed to recover relevant part of correlation energy of the system. If it's possible to introduce directly interelectronic distance short expanction can be used, but this method is limitated to very few electron systems becouse needed integrals are very difficult to compute in analitical way and very time consuming if computed by numerical methods.

Another way to treat correlation problem is the density functional approch: this method is much simpler than ab initio methods but the exact density functional is not known and there is not sistematic way to improve present approximation.

In recent years another very different class of methods was developed to cope with correlation problems. These methods are all based on Monte Carlo use of random numbers to compute multidimensional integral or to solve in a stocastic way integral equation. Some of these are intrisically "exact" meaning that they can give right results and that it's possible to improve statistical precision in a way dependent only on the computational time. The simplicity of the ideas and of the computer programs, coped with the ability of recover near 100% of the correlation energy and the absence of the basis set problem like in ab initio methods, are some reason for the recent spread of these methods in the chemistry comunity.

For now Monte Carlo methods are in their early childhood, so it's possible that their applications and accurancy can be greatly improved in near future and that they equalize ab-initio methods diffusion.

## Variational Monte Carlo and optimisation of a trial wavefunction

The general quantum chemistry problem is to obtain an analitical approximation of the exact wavefunction for the states of a given atomic or molecular system. By mean of this function it can be possible to compute all the relevant information needed to predict or to interpretate the result of some experiment, so that chemical steps can be guided by theoretic ideas.

Usually ab initio methods, by means of Rayleigh-Ritz variational theorem, try to optimize trial wavefunction minimizing the aspectation value of the energy, i.e. given a parameter dependent trial function search the parameters such that

$$\langle E \rangle = \min_p \left[ \frac{\int \Psi(\mathbf{R},\mathbf{p})^* \hat{H} \Psi(\mathbf{R},\mathbf{p}) d\mathbf{R}}{\int \Psi(\mathbf{R},\mathbf{p})^* \Psi(\mathbf{R},\mathbf{p}) d\mathbf{R}} \right] \tag{1}$$

where **p** is the parameters vector and **R** is a point in configuration space. The variational method needs the computation of the integrals apparing in the equation and so, generally, one can only use trial wavefunction whome analytical form is costructed using linear combination of Slater determinant.

Monte Carlo methods [Kal86] overcome this difficulty by means of numerical integration schemes not based over the usual integration grid methods. Rewriting the equation (1) we obtain

$$\langle E \rangle = \left[ \frac{\int \left|\Psi(\mathbf{R},\mathbf{p})\right|^2 \frac{\hat{H}\Psi(\mathbf{R},\mathbf{p})}{\Psi(\mathbf{R},\mathbf{p})} d\mathbf{R}}{\int \left|\Psi(\mathbf{R},\mathbf{p})\right|^2 d\mathbf{R}} \right] = \int P(\mathbf{R},\mathbf{p}) \frac{\hat{H}\Psi(\mathbf{R},\mathbf{p})}{\Psi(\mathbf{R},\mathbf{p})} d\mathbf{R} \qquad (2)$$

where $P$ is a normalized probability distribution and the other term is defined "local energy". This means that it is possible to compute aspectation value directly sampling $P$ and computing the mean value over the sampled point of the local energy. The simpler way in which $P$ can be sample is to generate a Markov chain and to satisfied detailed balance like in Metropolis algorithm [Met53]. This means that it is possible to use any analytical trial function to compute energy value, overcoming in such way the limitation due to the analytical computation of the integrals. The computational effort for the integration is much lesser than in normal quadrature schemes because the statistical precision is a function of $N^{-\frac{1}{2}}$, where N is the number of indipendent local energy value, and the Metropolis method sample only relevant configuration space zones.

Even if choosing the parameters value so that mean energy is minimized can be used to optimize trial function the computational effort needed to follow this way is very large because it's necessary to obtain very high precision to see the effect over the energy of a parameter change. This problem can be overcomed using "correlated sampling", i.e. computing the integral over the same set of points, because correlation reduce the statistical variance of the computed energy difference [Wel].

Better then optimize using mean energy is to use "local energy variance" [Umr88]

$$\sigma^2 = \frac{\sum_i w_i \left[ E_{loc}(\mathbf{R}_i) - \bar{E} \right]^2}{\sum_i w_i}$$

where $w$ are some integration weight and $\bar{E}$ is the mean energy value computed over the fixed sample of points. This method is much more numerically stable and has the advantage that its lower bound is known.

Using very flexible trial wave function it is possible in this way to obtain accurated energy values for atoms and molecules with two to ten electron or to treat systems like noble gas clusters[Umr88, Mos90, Nig94] . Physical sense of the wavefunction can help to understand the way by which electrons interact and to construct better *ansatz* for other systems. The optimized wavefunction can be used also to compute other aspectation value[Ale92], even if the accurancy of the results are much stronger dependent from the function form than energy and from the kind of observable in which one is interested[Bia94].

In the last years much of the research over Variational Monte Carlo was directed to solve the large atomic number problem[Les94,Sot94]: Monte Carlo losts its efficency when the nuclear charge of the system atoms blows up. Some improvements was proposed but other work is needed to solve the problem and to speed up Monte Carlo computation.

**Diffusion Monte Carlo**

Variational Monte Carlo can be seen like an improved way to treat correlated wavefunction and to solve integration problem that are in other way intractable. A complete different method to treat correlation problem can be developed using Monte Carlo tecnique starting from the theory of the integral equation and of the Green's function for partial differential equation [Kol82,Rey82].

Schrödinger equation written in immaginary time t=iτ

$$\hat{H}\Psi(\mathbf{R},t) = -\frac{\partial\Psi(\mathbf{R},t)}{\partial t} \qquad \mathbf{(4)}$$

can be trasformed in an integral equation of form

$$\Psi(\mathbf{R},t+\delta t) = \int d\mathbf{R}' G(\mathbf{R}' \to \mathbf{R},\delta t)\Psi(\mathbf{R}',t) \qquad (5)$$

where $G$ is the Green's function for the differential equation, i.e. is a solution of the equation having boundaries condition such that

$$G(\mathbf{R}' \to \mathbf{R},0) = \delta(\mathbf{R}' - \mathbf{R}) \qquad (6).$$

Its formal solution is

$$\Psi(\mathbf{R},t)=\sum_i c_i\psi_i(\mathbf{R})e^{-(E_i-E_t)t} \qquad (7)$$

where $\psi$ are eigenvector of the differential equation and E are their eigenvalue. This means that starting from a trial solution it's possible to project out the excitated states of the system, leaving present only the ground state component. Problems of this approch to wavefunction computation are the multidimensionality of the needed integral and the uknown analitical form of the Green's function. While the first of this problems can be overcomed by using Monte Carlo methods obtaining in such a way a statistical sampling of the integral, the latter forces us to use an approximated form for the Green's function projector. This approximation can be written in a way that lead to the exact form if propagation time step go to zero such that by means of an extrapolation we can recover the right results.

The simplest analitical form for the Green's function is

$$G(\mathbf{R}'\to\mathbf{R},\delta t)=Ne^{-\frac{(\mathbf{R}'-\mathbf{R})^2}{2\delta t}}e^{-t(V(\mathbf{R})-E_t)} \qquad (8)$$

where $N$ is the normalization factor of the gaussian part, $V$ is the potential energy computed at arriving point. Details for the implementation of the method can be found in many papers so that we do not describe here. This form can be used only for simple system because is not very efficient and has the problem that when potential energy go to minus infinity like in electronic system it diverges. These matters can be overcomed taking advantage of prior knowledge of the system wavefunction in such a way that it is possible to reduce fluctuations of the potential part by means of 'importance sampling' method.

Multipling (4) by a trial wavefunction and defining the distribution

$$f(\mathbf{R},t)=\Phi_t(\mathbf{R})\Psi(\mathbf{R},t) \qquad (9)$$

after considerable rearrangement it's possible to obtain the 'importance sampling' version of the Schrödinger equation

$$\frac{\partial f}{\partial t}=\frac{1}{2}\nabla^2 f-\nabla\cdot\left(f\nabla\ln\Phi_t\right)-E_{loc}f \qquad (10)$$

In this equation potential energy is substituted by local energy of the trial wavefunction and there is a new drift term driving the $f$ scalar field in configuration space. The trial wavefunction can be choosen in many way, but better results are obtained by use of high accurate and optimized form like those used in Variational

Monte Carlo. From $f$ we can compute simply the energy of the system using 'mixed extimator', i.e.

$$\langle E\rangle=\frac{\int d\mathbf{R}f(\mathbf{R},t)E_{loc}(\mathbf{R})}{\int d\mathbf{R}f(\mathbf{R},t)} \qquad (11)$$

or its Monte Carlo rappresentation. In similar way it is possible to compute mean value for all the dynamical observables whome operators commute with Hamiltonian operator. Importance sampling is able to reduce statistical error of a factor of 20 and any systematic error due to the finite time step it's aspected to be similarly reduce. Detailed description of the time step-error in Diffusion Monte Carlo can be found in recent papers where many solution are addressed to reduce the statistical and sistematic error.

Starting like a toy, Diffusion Monte Carlo is now an useful methods to cope with correlation problems. Many accurate results are now published for small and medium atomic and molecular system and often these results are better than full CI calculation.

In the forcoming paragraphes some selected results are presented, varing from very little systems (2-4 electrons) to medium systems (5-10 electrons).

**Few electron systems**

In this paragraph will be given some Diffusion Monte Carlo results about very simple systems and will be introduce some of the problems that Monte Carlo methods must overcome to compute aspectation value for these.

Because Monte Carlo is a statistical method, it can sample only positive definite distribution and it has some problem to cope with system whose wavefunction change sign in configuration space. For ground state of two electron systems there is no matter because space part of the wavefuction can be always treat like positive distribution and so is possible to simulate it in a simple way. The first calculation done by means of Diffusion Monte Carlo without importance sampling was those about He atom, Hydrogen molecule and H3+ [And75,]. These systems was used to understand the method but, even if of low statistical precision, their mean energy values are of good accurancy. More precise computation was made some years later using the importance sampling Diffusion Monte Carlo obtaining 1.3439 +- .0002 hartree for H3+ equilibrium configuration. This value was at that time best Born-Openheimer estimation for the ground state energy of H3+ and is only by the use of explicitly correlated gaussian configuration that better mean energy was computed[Cle93].

To treat atomic or molecular state whose wavefunction change sign like in H2 Sigma(3,u) some trick must be done. If it is possible to define in some way the configurational space surface where the function value is zero the simulation can be done in this restricted space simply killing all the sampling moves in the negative region. This method preserves the positive sign of the distribution but is variational in nature because only if the nodal surface are exact the computed mean value are also exact, i.e. it is only possible to solve exactly the Schrodinger equation for "fixed node" and this gives variational extimation of the system energy. First simple computation was made for H 2P, H2 sigma(3,u), H4 and Be 1S state[And76]. Energy dependence from nodal location was discussed using different trial wave function to define nodal surface. By means of this way it was found that mean energy has very low dependence by nodal surface such that it is possible to obtain very accurate results even if simple wavefunction was used. The unique particularity is the fact that for Be ground state the simple nodal surface given by all Hartree-Fock wavefunction are not similar of the exact nodal structure, giving high mean energy value for that simulation.

A chemical interesting application of the method is the calculation of the energy of the saddle point for collinear approch of the Hydrogen exchange in Hydrogen-Hydrogen molecule reaction [Les..]. This calculation was used to check the experimental computed value of the activation energy and the potential energy surface computed by means of CI expansion. Also for H4 system in square geometry, computation was made to obtain the activation energy for the Hydrogen exchange; this is useful to chech if the mechanism passes through a square disposition of the Hydrogen atoms or if it passes through some radical species[And79, J.Comp. Phys. 31,425(1979)].

Other application of Diffusion Monte Carlo to simple systems was the computation of the non Born-Openheimer energy of the Hydrogen molecule[And..]. Because the space part of the wavefunction is positive if protons are taken of opposite spin, the simulation needs only to scale proton move by their masses.

For all this little electron systems in the last years was developed some "exact" or near "exact" methods able to treat the "sign" problem of the wavefunction. Exact calculations was made for the potential energy surface for H3[and..], for the interaction potential between two He atoms [And..] and Hydrogen-Helium atoms[And..] using a

different algorithm based over exact Green' function. Near exact calculation was dued for He 3S state, H2 in sigma(3,u) and pi(3,g) state using fluctuating nodal surface based over gaussian approximation of the wavefunction form.
For now this later approches are computational possible only for very few electrons but simple modification can be found such that computational effort can be reduced.

**Medium electrons systems**

Medium electron systems are best treated by means of "fixed node" importance sampling methods because the growing number of dimension of configurational space reduce the probability of a move to sample a given point. Furthermore the importance sampling version of the Schrodinger equation has smaller fluctuation of the potential part and so that fluctuation of the sampling are reduced.
First applications was over LiH[Kal82,Rey82], Li2 and H2O[Rey82] where time step error of the methods was analised and and extrapolation procedure was proposed to eliminate this. Also dependance from the trial function used in the calculation was discussed determinig the importance to use a compact and accurate function to obtain higly accurated energy value.
Similar calculations was made for Be and LiH systems [Han85] using trial function obtained by means of transcorrelated hamiltonian method. A good review of calculations over LiH system can be found in recent Anderson's paper over fixed node methods[]
Using simple monodeterminantal trial function to construct the approximated nodal surface, multiplied by Jastrow factor in interelectronic distance to reduce local energy variance, time step extrapolated value for the energy of F and F- systems was computed [Les86]. These computations gived electron affinity value in good accord with experimental value and was useful to recognize the importance to satisfy electron-nucleus cusp conditions in order to reduce time step dependence of the mean energy value.
For Be, using his improved algorithm and a fairly accurated four determinantal trial wave function, Umrigar computed best extrapolated non relativistic energy value only .00019+-.000003 hartree above corrected experimental value. This error is dued to uncorrected nodal position such that it ca be reduced by means of a better trial wave

function. To obtain similar accurancy more than 650,000 determinantal configuration interaction computation was needed.
For methane[Gar..] double-zeta monodeterminantal trial wave function gave energy 30 kcal/mole below the best analitical calculation, recovering 97% of the correlation energy for this system.
Similar results can be obtained for F-H2 system. This is of chemical interest because it can give "chemical laser" to use in infrared range of frequencies.

**Properties**

Having a way to obtain or to sample a wavefunction should be possible to compute some other mean value other than energy aspectation value. If the wavefunction is analitic the simpler way is to use Variational Monte Carlo to compute esactly the integrals needed. For the Diffusion Monte Carlo rappresentation things are more complicated because wavefunction is only sampled by means of the Green's function integral equation. If the dynamical observable operator commutes with system Hamiltonian "exact" mean value can be computed by means of the mixed estimator because the hamiltonian ground state wavefunction is also an eigenfunction of the observable. If it is not so, mixed extimator can be used but it has some systematic error linearly dependent by the error of the trial wavefunction used. Simple modification of the mixed estimator by means of the mean value computed by exact variational procedures gives error with quadratic dependence from the trial wave function error

$$\langle \hat{\vartheta} \rangle = 2\langle \hat{\vartheta} \rangle_{mixed} - \langle \hat{\vartheta} \rangle_{\mathrm{var}\,iational} + O(\delta)$$

Another way to compute in simple way an approximation of the mean value for the exact wave function was based over squared branching weights[Vrb..] of the sampled distribution. This method was used to compute also approximation of the vibrational frequencies for diatomics molecules.
Bookeeping the branching weights "exact" methods can be developed[Kal66]. In this way efficient algorithms can be programmed and easily used to compute coordinate dependent mean value[Rey..,spagnolo].
Recently two other methods was developed to cope with the problem of computing mean value of derivatives dependent operators. One of this uses Diffusion Monte Carlo and linear response teory to obtain the needed values, but it can compute only second

derivatives operators. Also many simulations are needed to obtain the dependence of the energy value from the perturbations.

Other way[Bia95] is to compute superposition integrals between exact sampled wavefunction and a given set of esplicitly correlated basis function. Solving simple linear system it is possible to obtain the coefficient for the recostructed wavefunction and in this way compute all the needed values. Drawback of this approch is the necessity of computation of the superposition integrals between basis functions.

**Other application**

The ability of Variational and Diffusion Quantum Monte Carlo to cope with any analitical form of the potential energy means that in principle it can be used to treat any given ground state vibrational problem. In general terms DMC can compute exactly wave function for very anarmonic potential, for highly delocalized or for non Born-Oppeheimer system.

First QMC application to molecular system was the solution of ground state for dimer and trimer cluster of water molecules by Coker and Watts (Cok84, Cok87), given an intermolecular and intramolecular potential surface. Computed projections over intramolecular "local mode" coordinates of the wave function was fitted with analitical Morse function and this gives a mean field rappresentation of the interaction between intramolecular and intermolecular degree of freedom. This was subsequently used to compute analiticaly absorption frequences to assign infrared predissociation spectra of the water cluster.

The same method was used also to compute mean field infrared absorption, adiabatical dissociation energy and tunnel splitting for hydrogen fluoride isotopomers (Sun90, Qua91, Suh91).

To compute "exactly" vibrational frequences for molecules Bernu & al.(Ber90) developed an "multiple state" sampling methods based on diagonalization of QMC computed hamiltonian and superposition matrices between analitical basis functions. This was chosen to describe the vibrational molecular system. By this method they obtained converged results for some low lying states of H2CO.

Similar calculation for the same molecule was made by Watts(Lew94) using "fixed node" approximation for exitated state and a new biasing methods for intramolecular vibration.

For molecular clusters, which are composed of weakly interacting molecular subsystem, the relevant part of the total energy is dued to intramolecular vibrations. This means that to compute intermolecular interaction long simulations must be needed in order to reduce mean energy variance. To overcome the problem Buch (Buc92) developed "rigid body" DMC treating rotational degrees of freedom in a diffusive way.The method was used to simulate interaction between hydrogen and water. Subsequently Clary & al. (Gre94) compared conventional and "rigid body" diffusion Monte Carlo obtaining similar results for intermolecular degrees of freedom.

Using DMC Vegiri & al. (Veg94) studied B(H2)n cluster for n=2-8 comparing classical minimum of potential energy surface with projected distribution of the sampled wave function. Using anisotropic B - H2 interaction they found much more localization of hydrogen molecules respects results of Whaley & al.

All the reported applications of QMC to vibrational problems are much dependent from the potential energy surface used. Highly accurated surfaces can be obtained only with high computational costs and often they are biased by the fitting procedures used to generate analitical rappresentations. To overcome this matter two methods was developed: using directly computed point and interpolation formula like weighted distance (Suh93) or to adapte an analitical model surface to some experimental mean value like rotational costant(Fra94, Dyk94).


(1) Cok84 J. Chem. Phys. 82, 3554 (1984)

(2) Cok87 J. Phys. Chem. 91, 2513 (1987)

(3) Sun90 J. Chem. Phys. 92, 603 (1990)

(4) Ber90 J. Chem. Phys. 93, 552 (1990)

(5) Qua91 J. Chem. Phys. 95, 28 (1991)

(6) Suh91 Phys. Rep. 204, 293 (1991)

(7) Buc92 J. Chem. Phys. 97, 726 (1992)

(8) Suh93 Chem.Phys. Lett. 214, 373 (1993)

(9) Lew94 Mol. Phys. 81, 1075 (1994)

(10)Fra94 J. Chem. Phys. 100, 2865 (1994)

(11) Dyk94 Chem. Phys. Lett. 220, 161 (1994)

(12) Gre94 Chem. Phys. Lett. 228, 547 (1994)

(13) Veg94 J. Chem. Phys. 101, 2577 (1994)

(14) San94 J. Chem. Phys. 101, 6353 (1994)