Exact Fixed-node Quantum Monte Carlo　Self-optimizing Procedure

HUANG　Hong-Xin[*]

College of Chemistry and Chemical Technology　Hunan Normal University　Changsha　Hunan 410081　China

In this paper　a novel exact fixed-node quantum Monte Carlo　EFNQMC　algorithm was proposed　which is a self-optimizing and self-improving procedure. In contrast to the previous EFN-QMC method　the importance function of this method is optimized synchronistically in the diffusion procedure　but not before beginning the EFNQMC computation. In order to optimize the importance function　the improved steepest descent technique is used　in which the step size is automatically adjustable. The procedure is quasi-Newton type and converges super linearly. The present method also uses a novel trial function　which has correct electron-electron and electron-nucleus cusp conditions. The novel EFNQMC algorithm and the novel trial function are employed to calculate the energies of $1\ ^1A_1$ state of CH_2　^1A_g state of C_8 and the ground-states of H_2　LiH　Li_2 and H_2O.

Keywords　　exact fixed-node quantum Monte Carlo method　electronic correlation energy　self-optimizing procedure

Introduction

The fixed-node quantum Monte Carlo　FNQMC method　which solves the Schrödinger equation by an appropriate diffusing random walk of electrons　has been successfully utilized in the domain of quantum chemistry.[1] A given　generally optimized　trial wave function Ψ_T is used as an" importance function" in the stochastic approach that solves the" diffusion equation". Recently　a novel method for the FNQMC has been proposed　which is called the exact FNQMC　EFNQMC　method.[2] EFNQMC method usually recovers about 80%—90% of the correlation energy of small molecules.

There are several obstacles in the EFNQMC calculation

1　Before beginning the EFNQMC computation　the parameters of the trial function　importance function　Ψ_T must have been optimized by the variational Monte Carlo VMC　method. Optimization of these parameters is both time-consuming and expensive.

2　The optimization is not efficient　because there are two different samplings required in the VMC and EFN-QMC methods.

3　Previously Umrigar　Huang　Bueckert and others proposed methods for optimizing the trial functions.[3-5]

Umrigar employed the variance minimization technique over a fixed set of configurations of the electron samples from $|\Psi_T|^2$. In practice there is a" weight ill-condition" about the node of the trial function. A steepest descent technique was employed by Huang and Bueckert. The technique converges linearly　and even for the optimization of a quadratic function it often converges very slowly and encounters the ill-condition　especially near the optimum point.

In this paper an approach for improving the EFNQMC method is proposed　which has the following advantages

1　It is a self-optimizing diffusion procedure. In the novel EFNQMC computation　the parameters of the importance function Ψ_T are optimized synchronistically. It is economical in CPU time.

2　This novel algorithm is a self-improving Monte Carlo scheme. With this method one could construct well-behaved trial function and improve the accuracy and convergence of EFNQMC computations.

3　In order to optimize the importance function　the improved steepest descent technique is used　in which the step size is automatically adjustable. The procedure is quasi-Newton type and converges super linearly.

In the present paper a novel trial function is also used　which has correct electron-electron and electron-nucleus cusp conditions. The electron-nucleus cusp condition proves the most significant.

In order to test the correctness of the novel EFNQMC algorithm　the values of the zeroth and the first approximation of the energies of $1\ ^1A_1$ state of CH_2　^1A_g state of C_8 and the ground-states of H_2　LiH　Li_2 and H_2O have been calculated using the self-optimizing procedure.

EFNQMC calculation

H represents Hamilton operator for a system　Ψ_A represents a trial function for the FNQMC method　and Ψ_B is an" exact wave function" which possesses the same node structure as that possessed by Ψ_A and has" node approximation"　then

$$\overline{H_i} = \frac{\Psi_A | H^i | \Psi_B}{\Psi_A | \Psi_B} \qquad i = 0 \quad 1 \quad 2 \quad \ldots \qquad 1$$

It is proved that the value of the energy calculated using the traditional FNQMC method is only the zeroth order approximation of the eigenvalue of the energy and the first order approximation value is[2]

$$E_1 = \overline{H_1} - \frac{\overline{H_2} - \overline{H_1}\,^2\,^2}{H_3 - 3\,\overline{H_2 H_1} + 2\,\overline{H_1}\,^3} \qquad 2$$

As can be seen from Eqs. 1 and 2 that calculation of the first order approximation of the eigenvalue of the energy for a system using EFNQMC method needs not revise the traditional FNQMC program but only needs to calculate $\overline{H_i}$ $i = 1$ 2 3 according to the original FNQMC process.

Optimization theory

For optimizing the importance function we minimize Δ

$$\Delta = \frac{1}{N} \sum_{i=1}^{N} E_L - E_T\,^2{}_i$$
$$= \overline{E_L - E_T\,^2} \qquad 3$$

where E_L is the" local energy" $E_L = H\Psi_T / \Psi_T$ and E_T is the" trial energy". The sum is over the N configurations in the diffusion process.[1] Suppose Ψ_T P R is the importance function in the EFNQMC computation where $P = p_1$ p_2 $\ldots$ p_J denotes the set of parameters to be optimized and R identifies the coordinate space of electrons. The first and the second derivatives of Δ with respect to the parameter P take the forms

$$g = \frac{\partial \Delta}{\partial P}$$
$$= \frac{1}{N} \sum_{i=1}^{N} \Big[2\, E_L - E_T\, \frac{\partial E_L}{\partial P} \Big]_i$$
$$= 2\, \overline{E_L - E_T\, \frac{\partial E_L}{\partial P}} \qquad 4$$

$$q = \frac{\partial^2 \Delta}{\partial P^2}$$
$$= \frac{1}{N} \sum_{i=1}^{N} 2\Big[\Big(\frac{\partial E_L}{\partial P}\Big)^2 + E_L - E_T\, \frac{\partial^2 E_L}{\partial P^2} \Big]_i$$
$$\approx \frac{1}{N} \sum_{i=1}^{N} 2\Big(\frac{\partial E_L}{\partial P}\Big)^2_i$$
$$= 2\, \overline{\Big(\frac{\partial E_L}{\partial P}\Big)^2} \geqslant 0 \qquad 5$$

The approximation in Eq. 5 is reasonable. First for most cases it has been proved to be a good approximation. Second the extent of the approximation is not important

here if Eq. 5 is positive. In the following discussion it can be seen that the variance of q might be corrected by using an adjustable step size.

The classical steepest descent method for unconstrained minimization of Δ P is of the form

$$P_{k+1} = P_k - \lambda_k g_k \qquad 6$$

The search direction g_k is chosen as the gradient of Δ P at P_k

$$g_k = g\ P_k$$
$$= \Big(\frac{\partial \Delta}{\partial P}\Big)_k \qquad 7$$

where k indicates the kth iteration. λ_k is a step-size factor obtained with a one-dimensional search which was an empirical parameter in previous work.[3-5] We start from Eq. 6 and derive the expression of the automatically adjustable λ_k. Eq. 6 is rewritten as

$$P_{k+1} = P_k - M_k g_k \qquad 8$$

where $M_k = \lambda_k I_k$ I_k is considered temp as a unit matrix. Let

$$S_{k-1} = P_k - P_{k-1} \qquad 9$$

$$y_{k-1} = g_k - g_{k-1} \qquad 10$$

Note that the quasi-Newton equation takes the form

$$y_{k-1} = H_k S_{k-1} \qquad 11$$

where H_k is Hessian matrix of Δ P . We attempt to adjust the step size λ_k so that M_k closes to H_k^{-1} which implies that the Euclidean norm of $S_{k-1} - M_k y_{k-1}$ is minimized

$$\min \| S_{k-1} - M_k y_{k-1} \|_2 \qquad 12$$

From Eq. 12 one can easily obtain the relationship

$$\lambda_k = S_{k-1}^T y_{k-1} \ / \ y_{k-1}^T y_{k-1}$$
$$= \frac{P_k - P_{k-1}\,^T\, g_k - g_{k-1}}{g_k - g_{k-1}\,^T\, g_k - g_{k-1}} \qquad 13$$

where S_{k-1}^T denotes the transposition of the column vector $S_{k-1} \ldots$

Eq. 13 suggests that λ_k could be adjusted automatically by the value of P and g from the $k-1$ th and kth iterations. This step size implies that the steepest descent behaves like a partial quasi-Newton method.

Considering this problem in more detail i.e. to make the descent procedure more similar to quasi-Newton it is necessary to make M_k approach H_k^{-1} closely. Besides

adjusting λ_k we might arrive at this result by improving the elements of I_k. In fact we can substitute the reciprocal of the second derivative of $\Delta \ P$ with respect to P for the main diagonal element 1 of the identity matrix I_k. Thus M_k becomes such a matrix that the main diagonal elements consist of the reciprocals of the components of $q = \partial^2 \Delta / \partial P^2$ and the off-diagonal elements are zero. Obviously the reconstructed M_k approaches further H_k^{-1}. In summary the process of "steepest descent" described here is governed by the following equation

$$P_{k+1} = P_k - \lambda_k \left(\frac{g}{q} \right)_k \qquad 14$$

Combining Eqs. 4 and 5 we can see that

$$\left(\frac{g}{q} \right) = \frac{\overline{E_L - E_T \ \dfrac{\partial E_L}{\partial P}}}{\overline{\left(\dfrac{\partial E_L}{\partial P} \right)^2}} \qquad 15$$

From Eqs. 4 13 14 and 15 we get

$$P_{k+1} = P_k - \frac{1}{2} \left\{ \frac{\overline{E_L - E_T \ \dfrac{\partial E_L}{\partial P}}}{\overline{\left(\dfrac{\partial E_L}{\partial P} \right)^2}} \right\}_k \times$$

$$\frac{P_k - P_{k-1} \ ^T \ \left\{ \left[\overline{E_L - E_T \ \dfrac{\partial E_L}{\partial P}} \right]_k - \left[\overline{E_L - E_T \ \dfrac{\partial E_L}{\partial P}} \right]_{k-1} \right\}}{\left\{ \left[\overline{E_L - E_T \ \dfrac{\partial E_L}{\partial P}} \right]_k - \left[\overline{E_L - E_T \ \dfrac{\partial E_L}{\partial P}} \right]_{k-1} \right\}^T \cdot \left\{ \left[\overline{E_L - E_T \ \dfrac{\partial E_L}{\partial P}} \right]_k - \left[\overline{E_L - E_T \ \dfrac{\partial E_L}{\partial P}} \right]_{k-1} \right\}} \qquad 16$$

Eq. 16 is an essential formula for optimizing wave function under the diffusion condition.

Trial wave function

The most commonly used trial wave function in the FNQMC method has the form

$$\Psi = \Phi^{\uparrow} \Phi^{\downarrow} \Phi^{C} \qquad 17$$

where $\Phi^{\uparrow}$ and $\Phi^{\downarrow}$ indicate spin-up and spin-down single Slater determinants respectively which are constructed of the molecular orbitals composed of a linear combination of Slater basis functions. Φ^{C} is a correlation function and it is taken here to be a product of electron-electron correlations as well as electron-nucleus correlations.

The particular form for the Φ^{C} that we use is one proposed by Boys and Handy [6]

$$\Phi^{C} = \exp \sum_{I} \sum_{i<j} U_{Iij} \qquad 18$$

$$U_{Iij} = \sum_{k}^{N_I} \Delta \ m_{kI} \ n_{kI} \ c_{kI} \times$$
$$\bar{r}_{iI}^{m_{kI}} \bar{r}_{jI}^{n_{kI}} + \bar{r}_{jI}^{m_{kI}} \bar{r}_{iI}^{n_{kI}} \ \bar{r}_{ij}^{o_{kI}} \qquad 19$$

where the I sum runs over nuclei the ij sum is over electron pairs and the k sum runs over the N_I terms in the correlation function used around each nucleus. The $\bar{r}$ functions are for nucleus-electron terms

$$\bar{r}_{iI} = \frac{b_I r_{iI}}{1 + b_I r_{iI}} \qquad 20$$

and for electron-electron terms

$$\bar{r}_{ij} = \frac{d_I r_{ij}}{1 + d_I r_{ij}} \qquad 21$$

The m_{kI} n_{kI} and o_{kI} are taken to be integers and the function Δ is used to maintain consistency with Boys and Handy

$$\Delta \ m \ n = \begin{cases} 1 & m \neq n \\ 1/2 & m = n \end{cases} \qquad 22$$

The parameters b and d represent the inverse of the effective range of the correlations. In our calculations only one b and d value is taken for each nucleus but clearly this restriction could be relaxed. To satisfy the electron-electron cusp condition for unlike spin electrons the only term with $o = 1$ is with $n = m = 0$ and with the coefficient $c = 1/\ 2d$.

Results of sample calculations as shown later indicate that this type of correlation function if employed in EFNQMC runs is capable of giving much more accurate results.

Self-optimizing diffusion process

Before starting the EFNQMC computations one has to choose a suitable form of trial function the initial values of P and the trial energy E_T. The self-optimizing process can be summarized as follows

i Select N diffusion particles as a set of initial configurations according to $\Psi_T \ P_{K=0} \ R$. We call the initial configurations a configuration block 0. Use Eqs. 4 and 5 to compute g/q $_{K=0}$ over block 0. Let $\lambda_0 =$

0.05—0.1 and update the value of parameter P

$$P_{k=1} = P_{k=0} - \lambda_0 \left(\frac{g}{q} \right)_{k=0} \qquad\qquad 23$$

and thus one obtains a new $\Psi_T\ P_{K=1}\ R$.

　　ii Pick the mth configuration from the configuration block. The electrons in this configuration will move. Let them diffuse independently for a small time τ according to the Gaussian part of the Green's function G . If the current electron is the jth electron in configuration m it moves to[7]

$$r'_j{}^m = r_j{}^m + \tau \nabla \ln | \Psi_T\ R | + \chi \qquad\qquad 24$$

where $r'_j{}^m$ is the three-dimensional coordinate of the electron j χ is a three-dimensional Gaussian random variable with a mean of zero and a variance of τ . After electron j moves as Eq. 24 the acceptance probability of the movement is

$$A\ R \to R'\ \tau\ = \min\left[1\ \frac{| \Psi_T\ R'\ |^2\ G\ R' \to R\ \tau}{| \Psi_T\ R\ |^2\ G\ R \to R'\ \tau} \right]$$

$$25$$

　　When all the n electrons in the current configuration m had moved once the branching probability M_n for configuration m was calculated from the exponential prefactor of the Green's function G

$$M_m = \exp\ - \tau\ E_L\ R'\ - E_T \qquad\qquad 26$$

After the N configurations in the block were diffused according to Eqs. 24 — 26 a new configuration block block 1 was formed.

　　iii Note that the right-hand-side quantities in Eq. 16 are derived from block 0 and 1. Use Eq. 16 over block 0 and 1 to calculate $P_{k=2}$ and a newer $\Psi_T\ P_{K=2}$ R can be obtained.

　　iv Repeat steps ii and iii until the steady state is reached.

　　During the self-optimizing diffusion process the following aspects should be noted.

　　1 The " trial energy" E_T is estimated by the average of E_L from a prior block. So E_T and the block are updated synchronistically.

　　2 In the course of a circle if $\lambda \leqslant 0$ let $\lambda = 0.05$—0.1 and then restart the circle.

　　3 The criterion of convergence is chosen as $|\Delta E_T| \leqslant 10^{-5}$.

Results and discussion

　　In order to test the correctness of the novel EFNQMC method the values of the zeroth and the first approximation of the energies of $1\ {}^1A_1$ state of CH_2 ${}^1A_g$ C_{4h} acet state of C_8 and the ground-states of H_2 LiH Li_2 and H_2O

have been calculated using the self-optimizing procedure. The geometrical configurations of these states were given in the literature.[27] HF B type function is used as a trial function Ψ for the novel EFNQMC method refer to Eqs. 17 — 19 .

　　The numbers of the initial configuration taken by us are 1000 H_2 5000 CH_2 LiH and Li_2 10000 H_2O and 12000 C_8 and the time step are 0.005/h H_2 0.001/h CH_2 LiH and Li_2 and 0.0001/h C_8 and H_2O respectively. The times of calculation required on a P4 computer are 5 min H_2 51 min LiH 69 min Li_2 90 min CH_2 231 min H_2O and 3550 min C_8 respectively. The values of the zeroth and the first approximation of the energies for these states calculated using the novel EFNQMC method E_0 and E_1 are given in Table 1. For the convenience of comparison Table 1 also lists the values of the energies for these states calculated using H-F CI and the ordinary EFNQMC methods.[27] In addition the experimental data which are taken from the literatures [8] of the values of these energies are given in Table 1 where the percentage listed below each datum is a percentage of the electronic correlation energy corresponding to it.

　　It can be seen from the data given in Table 1 that for $1\ {}^1A_1$ state of CH_2 ${}^1A_g$ C_{4h} acet state of C_8 and the ground-states of H_2 LiH Li_2 H_2O the calculation only needs to be in progress until the first approximation when using our novel EFNQMC method. All the percentages of the electronic correlation energy are over 96% which are much better than the values calculated using CI and the ordinary EFNQMC methods ε_1 is the first approximation data calculated using the ordinary EFNQMC method . For example for C_8 molecule the energy values obtained using the CI method the ordinary EFNQMC methods ε_1 and the novel EFNQMC algorithm E_1 are − 303.436 − 304.258 and − 304.3342 a.u. respectively. The percentages of the correlation energy recovered by these different methods are 52.34% 92.77% and 98.58% respectively. The exact nonrelativistic energy is − 304.361 a.u. and Hartree-Fock energy is − 302.47567 a.u..[8,9] It can also be seen that for C_8 the novel EFNQMC algorithm in the present paper recovers the correlation energies of 1.7331 and 1.8577 a.u. for E_0 and E_1 the percentages of the correlation energy recovered by E_0 and E_1 are 91.97% and 98.58% and there are very small statistical errors of 0.0004 and 0.0006 a.u. respectively. While the previous EFNQMC method[2] gives the correlation energies of 1.691 1.748 and 1.847 a.u. for ε_0 ε_1 and ε_2 with the errors of 0.004 0.006 and 0.008 a.u. the percentages of the correlation energy recovered by ε_0 ε_1 and ε_2 are 89.75% 92.77% and 97.99% respectively. For the data of other molecules shown in Table 1 a similar analysis can be performed.

　　In a word the novel EFNQMC algorithm compared with the previous EFNQMC method has a higher accuracy and a smaller statistical error showing that the algorithm described in this paper is very successful.

Table 1   Energies hartrees of $1\ ^1A_1$ state of CH_2 ^1A_g C_{4h} acet   state of C_8 and the ground states of H_2 LiH Li_2 and H_2O calculated from several methods

	$1\ ^1A_1$ CH_2	^1A_g C_{4h} acet	H_2	LiH	Li_2	H_2O
Experimental	− 39.133	− 304.361	− 1.17447	− 8.07021 5	− 14.9954	− 76.4376
H-F limit	− 38.8944	− 302.477	− 1.1336	− 7.987	− 14.872	− 76.0675
Best CI	− 39.0272	− 303.436	− 1.1737	− 8.0647	− 14.903	− 76.3683
	55.66%	52.34%	98.12%	93.38%	25.12%	81.28%
Ordinary EFNQMC						
ε_0	− 39.092 3	− 304.168 4	− 1.1744 3	− 8.0468 2	− 14.9818 4	− 76.231 2
	82.79%	89.75%	99.83%	71.86%	88.98%	44.17%
ε_1	− 39.112 3	− 304.258 6		− 8.0691 3	− 14.9901 6	− 76.375 3
	91.07%	92.77%		98.67%	95.71%	83.09%
Novel EFNQMC this work						
E_0	− 39.1072 5	− 304.2097 4	− 1.1745 2	− 8.0610 2	− 14.9844 1	− 76.3308 2
	89.07%	91.97%	99.91%	88.99%	91.08%	71.15%
E_1	− 39.1261 4	− 304.3342 6		− 8.0685 2	− 14.9933 2	− 76.4264 5
	97.11%	98.58%		97.95%	98.31%	96.98%

References

1   Anderson J. B. *Int. Rev. Phys. Chem.* **1995** *14* 85.

2   Huang H.-X. Zhao Z.-H. *Chin. J. Chem.* **2001** *19* 917.

3   Umrigar C. J. Wilson K. G. Wilkins J. W. *Phys. Rev. Lett.* **1988** *60* 1719.

4   Huang S. Sun Z William A. L. *J. Chem. Phys.* **1990** *92* 597.

5   Bueckert H. Rothstein S. M. Vrbik J. *Can. J. Chem.* **1992** *70* 366.

6   Boys S. F. Handy N. C. *Proc. R. Soc. London Ser. A* **1969** *311* 309.

7   Reynolds P. J. Ceperley D. M. Alder B. J. Lester W. A. *J. Chem. Phys.* **1982** *77* 5593.

8   Raghavachari K. Binkley J. S. *J. Chem. Phys.* **1987** *87* 2191.

9   Orden A. V. Saykally R. J. *Chem. Rev.* **1998** *98* 2313.