Title	Optimization of Many-Body Wave Function
Author(s)	Maezono, Ryo
Citation	Journal of Computational and Theoretical Nanoscience, 6(12): 2474-2482
Issue Date	2009
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/16061
Rights	Copyright (C) 2009 American Scientific Publishers. This material is posted here with permission of American Scientific Publishers. Maezono Ryo, Journal of Computational and Theoretical Nanoscience, 6(12), 2009, 2474-2482. http://dx.doi.org/10.1166/jctn.2009.1308
Description	

Journal of
Computational and Theoretical Nanoscience
Vol. 6, 1–9, 2009

Optimization of Many-Body Wave Function

Ryo Maezono

Japan Advanced Institute of Science and Technology, Asahidai 1-1 Nomi, Ishikawa 923-1292, Japan

An overview on recent technical aspects of *ab-initio* electronic structure calculations using Quantum Monte Carlo is given in this article. In order to treat realistic, inhomogeneous, and larger systems within the practical calculational cost, several important techniques have been developed for the method. The article explains variety of many-body wavefunction form and the optimization schemes applied to the recent works mainly on extended periodic systems. Selected data are given to show how the optimization works to improve ground state energy.

Keywords: Quantum Monte Carlo, Many-Body Wave Function, Jastrow Function, Backflow Transformation, Variational Optimization.

CONTENTS

1. INTRODUCTION

Electronic properties of extended solid systems have been an interest of modern science, forming established research communities.[1] Orthodox interests would be on those such as mechanical, magnetic, optical, and transport properties etc. of non-relativistic interacting electrons of inhomogeneous systems including ionic cores. The major approach for such problem has been constructing simplified model of many-body interactions and elucidating possible mechanisms on the model on comparison with experiments. On the other hand, such approaches not employing explicit modeling but based on the *ab-initio* simulation of the fundamental equation have recently formed a growing research community.[2] The density functional theory (DFT) provides a firm theoretical foundation to assure the treatment of complicated many-body interactions in a affordable form, and stimulates the *ab-initio* research field, being helped by the progress of high performance computing techniques. The objectivity of *ab-initio* approach gets more in demand in the recent trend of materials science to seek novel mechanisms and functionalities in more complicated interacting systems.

Studies on spectrum, response and excitation by external fields would be a majority of materials research though, there still remains challenging research targets even at the ground state properties. Quantitative investigations of delicate balance betweeen the electronic correlation in many-body electron systems and the effect of inhomogeneous potentials due to ionic cores have been an important research topics in this field.[3] Besides the original interest in how the unique fundamental equation brings about the variety of materials properties, the fundamental/precise *ab-initio* research about the ground state gets contemporary meaning to support the foundation of DFT through the construction and calibration of exchange-correlation (XC) potentials.[2,4] The research could then be regarded as an infra-structure in the stream of materials science through DFT, which is now commonly used to analyze experimental researches and to justify simplified models accounting for mechanisms of excitation properties.

The method used for the calibration mission ought to possess the universal reliability excluding arbitrariness as possible. The quantum Monte Carlo (QMC) method discussed in this article is one of the most appropriate framework for this purpose.[5] The method is used to evaluate the ground state mean value of physical quantities,

$$\langle\hat{O}\rangle = \frac{\int d\vec{X}\cdot\Phi^*(\vec{X})\cdot\hat{O}\Phi(\vec{X})}{\int d\vec{X}\cdot\Phi^*(\vec{X})\Phi(\vec{X})} \tag{1}$$

where $\Phi(\vec{X}) = \Phi(\vec{x}_1,\ldots,\vec{x}_N)$ is the many-body wavefunction of N-electron system, and $\vec{x}_j = (\vec{r}_j,\sigma_j)$ denotes a set of spacial and spin coordinates of j-th particle. Using the sampling distribution $P(\vec{X}) = |\Phi(\vec{X})|^2/\int d\vec{X}\cdot|\Phi(\vec{X})|^2$

in $3N$-dimensional space the mean value can be evaluated as

$$\langle \hat{O} \rangle = \int d\vec{X} \cdot P(\vec{X}) \cdot \{\Phi^{-1}(\vec{X}) \cdot \hat{O}\Phi(\vec{X})\} \tag{2}$$

in a form of statistical sampling. In the form of sampling the many-body interaction is evaluated with minimal approximation and hence we can avoid the arguments suspecting the reliability of practical approximations of electronic correlations used in conventional SCF (self-consistent-field) treatments. This is, as a matter of course, at the expense of introducing other sources of error and bias such as a statistical error, time-step error etc.[6] Furthermore we have to consider how to prepare $\Phi(\vec{X})$ first of all. In variational Monte Carlo method (VMC) some trial guess $\Psi_T(\vec{X})$, such as a Slater determinant of Hartree-Fock (HF) orbitals, is used to sample Eq. (2). In diffusion Monte Carlo method (DMC), the trial guess is driven in imaginary time as,

$$\Phi(\vec{X}, \tau) = \exp[-\tau\hat{H}] \cdot \Psi_T(\vec{X}) \tag{3}$$

to get converged to be closer to exact solution, and hence used to sample Eq. (2).[7] Here $\hat{H}$ denotes the many-body Hamiltonian[8]

$$\hat{H} = -\frac{1}{2}\sum_j \nabla_j^2 + V(\vec{X}) \tag{4}$$

where $V(\vec{X})$ is the many-body potential energy for a set of position $\vec{X} = (\vec{x}_1, \ldots, \vec{x}_N)$, and ∇_j^2 is the Laplacian operator with respect to the j-th electron position. The methodology for the practical implementation of Eqs. (2) and (3) have formed an established research field, for which several representative review articles and textbooks are already available.[5,9] Those include important techniques such as the fixed node approximation, importance sampling, branching method, statistical estimators etc. In this article these details of well-established/common issues are omitted and only those of updated ones relevant to my recent QMC studies of extended solid state systems are reported. Though these should include recent pseudo potential schemes, finite size correction schemes etc. This paper concentrates only on the issues about the variational optimization of many-body wavefunction. The article is organized as follows: In Section 2, the form of many-body wavefunction used in my recent calculation is described. The variational parameters to describe the electronic correlations are introduced in each subsection. Section 3 treats how to optimize the variational parameters. The procedure is sometimes technically complicated for which the example of carbon atomic system is taken for explanation. Conclusion is summarized in Section 4.

2. MANY-BODY WAVE FUNCTIONS

2.1. Orbital Functions

Many-body wavefunctions for electronic systems should be anti-symmetric about the exchange amongst the coordinates of electrons.[10] The most popular way to setup such wavefunction is to anti-symmetrize the product of one-particle orbital functions $\{\psi_j(\vec{x})\}$ as

$$\Psi\left(\vec{x}_1, \ldots, \vec{x}_N\right) = \hat{A}\left\{\psi_1\left(\vec{x}_1\right), \ldots, \psi_N\left(\vec{x}_N\right)\right\} \tag{5}$$

The Slater determinant,

$$\Psi_{\mathrm{D}}\left(\vec{x}_1, \ldots, \vec{x}_N\right) = \begin{vmatrix} \psi_1\left(\vec{x}_1\right) & \cdots & \psi_N\left(\vec{x}_1\right) \\ \vdots & \ddots & \vdots \\ \psi_1\left(\vec{x}_N\right) & \cdots & \psi_N\left(\vec{x}_N\right) \end{vmatrix} \tag{6}$$

is a representative one of this type.[10] Other type of the anti-symmetric wavefunction is introduced later in Section 3. Popular choice of one-body orbital functions are Hartree-Fock orbitals by self-consistent field (SCF) calculations, Kohn-Sham (KS) orbitals by DFT, or Natural orbitals (NO) by molecular orbital (MO) methods.[11]

The orbital function is usually treated in the expanded form by some set of basis functions.[12] Plain wave (PW) basis sets are popular for periodic extended systems whilst the Gaussian basis sets for isolated molecular systems. For isolated atomic systems with spherical symmetry direct numerical representation of radial component is sometimes used, corresponding to taking δ-function as the expanding

Dr. Ryo Maezono is a Lecturer of the School of Information Science at Japan Advanced Institute of Science and Technology (JAIST). He received his Ph.D. degree from University of Tokyo in March 2000. He was a research fellow of Japan Society for the Promotion of Science from 1999–2001, a research associate (EPSRC-UK) at Cavendish laboratory, University of Cambridge from 2001–2002, a researcher at National Institute for Materials Science in Japan from 2001–2007, and moved to JAIST at March 2007. His research interests are on the many-body theory and its applications to real materials using quantum Monte Carlo calculation, as well as on High Performance Computing for the simulations.

basis sets. If the basis sets expand the Hilbert space perfectly the choices are equivalent in principle, being able to transformed into each other. In practice the cost of calculation largely depends on the choice. In QMC the dominating part of the CPU time is the repeating evaluation of the determinant at each step after electrons move. Delocalizing orbitals such as PW are quite disadvantageous in this aspect because every orbital spreads over the system and any local change is reflected to all the orbitals and hence all the matrix elements should be updated.[13] Using localized orbitals in space is therefore vital to achieve QMC for larger systems at practical cost.[14]

For extended solid systems PW basis is commonly used. QMC using such PW orbital functions is, however, proved to be quite inefficient and costly.[13] In a modern procedure we used the PW orbitals are re-expanded by local spline basis sets, and used in QMC.[14] Representative results taken in silicon crystal (diamond structure) case are summarized in Table I. For given spatial resolution of PW, $\delta l \sim 1/G_{\rm max}$ ($G_{\rm max}$ denotes the cutoff wave number of the PW basis sets), the space is divided into box elements of side length $L \sim \delta l / x_{\rm mul}$. Within each box element the orbital is expressed by local spline functions called as 'blip.' The parameter $x_{\rm mul}$ adjusts the resolution of splined expression. From the comparison of kinetic energies in Table I the curvature of wavefunction is proved to be enough preserved after the transformation from PW into blip expansion. The CPU time and error bar show the spline basis calculation achieve around 200 times faster QMC calculation than that with PW to get the ground state energy with the same amount of statistical error bar. CPU time of spline basis VMC is almost independ on $x_{\rm mul}$. What matters depending on $x_{\rm mul}$ is the vast amount of file size of wavefunction, roughly proportional to $x_{\rm mul}$, causing inability of calculation due to I/O capacity. For this reason we choose minimum possible $x_{\rm mul}$ to preserve the kinetic energy.

2.2. Correlation Described by Amplitude

VMC using a Slater determintant composed of Hartree-Fock orbitals is equivalent to the Hartree-Fock self-consistent field (HFSCF) calculation.[10] VMC with KS orbitals in the single determinant gives higher energy than that with Hartree-Fock orbitals because the latter is variationally optimum by definition. The Jastrow factor $\exp(J(\vec{X}))$ [15] multiplied by the determinant as,

$$\Psi_{\rm SJ}(\vec{X}) = e^{J(\vec{X})} \cdot \Psi_{\rm D}(\vec{X}) \tag{7}$$

is used to take into account the correlation at VMC level, allowing degrees of freedom to modify the amplitude of many-body wavefunction.[5,9] Because it only adjusts the amplitude the amount of the correlations described by the Jastrow function can be described by the fixed node DMC[16–18] even without it in principle. In practice, however, the Jastrow factor has critical importance on the feasibility of DMC rather than a mere role to accelerate the convergence: The local energy, $E_L(\vec{X}) = \Psi^{-1}(\vec{X})\hat{H}\Psi(\vec{X})$ sampled throughout QMC, can have the singularity due to the divergence of Coulomb potential term $\sim 1/\left|\vec{r}_i - \vec{r}_j\right|$ whenever particles coalesce unless Ψ satisfies the Kato's cusp condition[19,20] exactly. The singularity brings about significant instability of population control of walkers[5] as well as large variance[21] and disables stable sequence of statistical accumulations. To satisfy the cusp condition by the wavefunction expanded by finite number of one-body functions, Ψ should posess explicit dependence on the distance between particles, such as $r_{ij} = |\vec{r}_i - \vec{r}_j|$. For this purpose the Jastrow factor is formed by such functions depending on inter-particle distances, keeping the stability of DMC accumulation.

Table I. Comparison between plain wave (PW) and spline basis (BW) calculations. VMC using a Slater determinant trial wavefunction is performed for a silicon diamond structure (3 × 3 × 3, 54 atoms, 216 electrons, Ne-core pseudo potential[38,39] calculation). The cutoff of PW basis sets is $|G_{\rm max}|^2 = 0.0341$ (a.u.). CPU time is measured on Hitachi SR11000, 32 core MPI calculation.

Method	$x_{\rm mul}$	Kinetic energy (a.u./primitive cell)	CPU time (seconds)	File size of wavefunction (GB)
PW-SCF[59]		3.0776	—	—
PW-VMC		3.0777(5)	685905	0.26
BW-VMC	1.50	3.0776(2)	21919	0.78
	1.75	3.0774(2)	—	1.27
	2.00	3.0775(2)	—	1.80
	2.25	3.0779(2)	—	2.62
	2.50	3.0779(2)	—	3.46

The symmetry as a whole many-body wavefunction should be considered carefully with the Jastrow factor. In periodic systems the Jastrow factor should be truncated within the simulation cell to maintain the translational symmetry.[22,23] This enforces to introduce the cutoff lengths of Jastrow function $J(X)$ as extra variational parameters. The spin symmetry is also affected by the Jastrow factor, which should be carefully handled especially in precise molecular calculations.[24,25] It is a trade-off issue because the Jastrow factor is used to ensure the cusp condition while it contaminates the spin symmetry, both of which are important in precise handling of calculations.

The jastrow function is usually decomposed as,

$$J(\vec{X}) = \sum_{i>j} u(r_{ij}) + \sum_{i,I} \chi_I(r_{iI}) + \sum_{I,j>j} f_I(r_{iI}, r_{jI}, r_{ij}) \tag{8}$$

where i and I are the index for electrons and ions, respectively.[26] The two-body term $u(r_{ij})$ denotes the correlation between electrons whilst the one-body term $\chi_I(r_{iI})$ plays a role to correct the over modification of charge density by u-term in inhomogeneous systems.[5,27] The three-body term $f_I(r_{iI}, r_{jI}, r_{ij})$ is essential to remove the singularity of E_L due to the three-body coalescence of two electrons and ionic core,[28] which can be omitted in pseudo potential calculations.

REVIEW

Table II. VMC and DMC energies of sodium solid,[30] $8\times 8\times 8$ simulation cell (512 electrons) with Ne-core pseudo potentials[31] at experimental equilibrium volume. All values are raw data without any corrections such as core polarization potentials, zero point vibration energies, finite size error corrections etc. Jastrow parameters are optimized by reweighted variance minimization.

Method	Energy (a.u./atom)
SCF (UHF)	−0.1896(1)
VMC (Slater-Jastrow)	−0.21964(3)
DMC (Slater-Jastrow)	−0.22138(2)

Table III. VMC and DMC energies of carbon diamond,[36] $5\times 5\times 5$ simulation cell (250 atoms, 1000 electrons) with He-core pseudo potentials[38, 39] (core polarization potential applied) at equilibrium volume. Jastrow parameters are optimized by unreweighted variance minimization. All values are raw data without any corrections such as zero point vibration energies, finite size error corrections etc.

Method	Energy (a.u./primitive cell)
SCF (HF)	−11.0194(6)
VMC (Slater-Jastrow)	−11.3392(2)
DMC (Slater-Jastrow)	−11.3662(2)

In the calculations of solid sodium[29, 30] with Ne-core pseudo potentials[31] and gaussian basis sets[32] (Table II), the following type of Jastrow functions are used,

$$u(r_{ij}) = \frac{A}{r_{ij}}[1-e^{-(r_{ij}/F_{\sigma\cdot\sigma'})}]\cdot e^{-(r_{ij}/L_0)^2} + S_1(r_{ij}) \tag{9}$$

$$\chi_I(r_{iI}) = S_2(r_{iI}) \tag{10}$$

The two-body term $u(r)$ has the primary A-term which is based on the asymptotic form of Bohm-Pines RPA,[33–35] truncated with a cutoff length L_0.[22, 23] S_1 is introduced to compensate the error due to the truncation of the asymptotic form. The ratio $A/F_{\sigma\sigma'}$ is determined by the electron–electron cusp conditions. The functions $S_{1,2}(r)$ are expanded by Chebychev polynomials multiplied by cutoff factor $(r-L_{1,2})^2$. The reason to use Chebychev polynomials is the superior property in the round-off error when the expansion order gets larger.[23] The variational parameters, in this case, are A, expansion coefficients in $S_{1,2}$, and the cutoff lengths for each terms. The optimization of A parameter is the hardest step because it is a non-linear optimization for the most dominating ingredient of correlations.

In diamond calculations[36, 37] (He-core pseudo potential,[38, 39] spline basis sets. Table III, we used other form of u-term all expanded in power series rather than the compact A-term form,

$$u\left(r_{ij}\right) \sim \left(r_{ij}-L_u\right)^2 \left\{\alpha_0 + \alpha_1\cdot r_{ij} + \sum_{l=2}^{N_u} \alpha_l \cdot r_{ij}^l\right\} \tag{11}$$

χ- and f-terms are similarly expanded in power series.[26] The cusp condition enforces a relation between α_0 and α_1. In this form, though the number of parameters is increased, robust optimization schemes are available.[40] It is proved that the new form performs better than the old A-term form, earning more correlation energy.[40] The cutoff lengths are still remaining as non-linear variational parameters being hard to be optimized robustly. For periodic systems we usually fix L_u to be Wigner-Seitz radius of simulation cell,[22, 23] and L_χ to be less than L_u. Amongst several trials of L_χ we pick up the variationally best one. Such cutoff lengths less than the Wigner-Seitz radius would be difficult to describe the density structure near to the simulation cell edge reflecting the correlation. To complement this more variational terms expanded in Fourier series of simulation cell period is used.[26]

Other type of truncated Jastrow factor is available[24, 41] in the form of,

$$u(r_{ij}) = \Phi_u(r_{ij}) - \Phi_u(L_u) \tag{12}$$

$$\chi_I(r_{iI}) = \Phi_\chi(r_{iI}) - \Phi_\chi(L_\chi) \tag{13}$$

where $\Phi_{u,\chi}(r) = \Phi_{u,\chi}(R_{u,\chi}(r))$ is expressed in terms of some normalized radius $R(r)$ that maps the system dependent cutoff interval $[0, L_{u,\chi}]$ into finite constant interval. The functional form of $R(r)$ is determined so that the whole many-body wavefunction can smoothly be continuous at the cutoff length. The functional form of $\Phi_{u,\chi}(R)$ is given, for example, as

$$\Phi_{u,\chi}(R) = \frac{a_1 R}{1+a_2 R} + a_3 R^2 + a_4 R^3 + \cdots \tag{14}$$

so that it can satisfy the cusp conditions.[24] The three-body f-term is expressed as a polynomial of

$$\bar{R}_j(r) = 1 - \frac{R_j(r)}{R_j(L_j)}, \quad (j = u, \chi) \tag{15}$$

by which the truncation is realized by vanishing $\bar{R}_j$ at the cutoff length. QMC calculation of immersed atom system is treated in this form of Jastrow functions (Table IV). Other implementation[42] employs the Jastrow function expanded by the basis sets $\{b_j(r)\}$ which is designed to vanish at and ensure for many-body wavefunction to be smoothly continuous at the cutoff lengths. This type of Jastrow functions are used for carbon pseudo atom calculation given in Table V.

Table IV. VMC and DMC energies of immersed atom system where a positive point charge $Z = 2$ is immersed into a jellium sphere ($r_S = 5.25$, $N_{\mathrm{elec}} = 60$) keeping charge neutrality. The system has an infinite potential wall at surface of sphere. Jastrow functions are optimized by reweighted variance minimization.

Method	Energy (a.u.)
SCF (HF)	−111.99(9)
VMC (Slater-Jastrow)	−112.81(6)
DMC (Slater-Jastrow)	−113.367(3)

Table V. Ground state energy of silicon diamond structure, 3 × 3 × 3 simulation cell (54 atoms, 216 electrons) with Ne-core pseudo potentials (core polarization potential applied). All values are raw data without any corrections such as zero point vibration energies, finite size error corrections etc. See the text accounting for cost factors (1) and (2).

Method	Energy (a.u./atom)	Cost factor (1)	Cost factor (2)
SCF (PBE-GGA)	−3.9264		
VMC (Slater)	−3.8244(2)		
VMC (Slater-Jastrow)	−3.9502(6)		
VMC (Slater-Jastrow-Backflow)	−3.96630(9)		
DMC (Slater-Jastrow)	−3.96651(9)	1.0	1.0
DMC (Slater-Jastrow-Backflow)	−3.96897(7)	61.1	14.8

2.3. Correlation Described by Nodes

Projection operation in DMC updates many-body wavefunction to be closer to the exact solution. It is well known that updating nodal degrees of freedom is technically difficult problem, known as negative sign problem.[9] Practical implementations of DMC therefore employs the fixed node approximation,[16–18] where the nodal surface is approximated to be fixed as that of trial initial guess Ψ_T. In some cases the nodal surface by single determinant is not sufficient, and the strategies to improve the nodal structure become important issue. There has been efforts to develop methodologies for updating nodal surface by DMC[43] but they are still far practical to be applied to larger systems of practical interests in materials science. Recent promising strategies are, instead, the VMC optimization of variational degrees of freedom which modifies the nodal surface based on the variational principle about it. Representative ones include the backflow transformation of Slater determinant[44, 45] and Pfaffian many-body wavefunction[46, 47] as well as the optimization of multi determinant coefficients, orbital functions, and basis functions themselves.[48]

The backflow transformation refers the shift of arguments of orbital functions consisting Slater determinant, $\vec{r}_j \rightarrow \vec{r}_j + \vec{\xi}_j$.[5, 44, 45, 47] The name 'backflow' is coming historically from the fact that the above shift is used to describe the enhancement of effective mass by a variational method, which physical picture corresponds to the name.[49] It allows the degrees of freedom to shift the argument to a direction preferred by interactions. In inhomogeneous systems the transformation is then generalized as,

$$\vec{r}_i \rightarrow \vec{r}_i + \vec{\xi}_i^{ee}(\{r_{ij}\}) + \vec{\xi}_i^{eN}(\{r_{iI}\}) + \vec{\xi}_i^{eeN}(\{r_{ij}\}, \{r_{iI}\}) \quad (16)$$

to have different preferred directions due to possible interactions. The transformation changes the nodal structure of many-body wavefunction to be closer to exact one as variational parameters. In the practical implementation we use each $\vec{\xi}$ expanded in the power series of each argument multiplied by the truncating factor with cutoff lengths.[44] The expanding coefficients and cutoff lengths are hence to be optimized variationally. As in the case of Jastrow functions we try several possible choice of cutoff lengths and pick up the variationally best ones rather than trying numerically unstable non-linear optimization. QMC calculation with the backflow transformation is quite expensive typically taking 20 times longer than that without it.[45] As a trade-off it improves the statistical error bar instead, reducing required steps to achieve specified accuracy. Table V shows an example of silicon diamond crystal system. If we compare the CPU time taken for the same amount of steps the backflow calculation costs 61.1 times more expensive. The same amount of accumulation, however, gives reduced error bar with backflow, giving the effective factor 14.8 estimated to achieve the same amount of error bar.

Some important concepts in empirical chemistry about electronic processes such as lone pairs are more fit to be described by pairwise two-body orbitals $\phi(i, j)$ called as geminal rather than one-body orbitals. The anti-symmetrized two-body orbitals to form many-body wavefunction is known as the 'geminal power.' [50] Applications of such wavefunction,

$$\Psi = \hat{A}\{\phi^{\uparrow\downarrow}(i, j), \chi^{\sigma\sigma}(i, j)\} \quad (17)$$

composed of spin singlet ($\phi^{\uparrow\downarrow}$) and triplet ($\chi^{\sigma\sigma}$; $\sigma = \uparrow, \downarrow$) pair spatial functions to QMC is becoming more practical recently.[46, 47, 51–53] Construction of such wavefunction using Pfaffian,

$$\Psi = \mathrm{Pf}\begin{bmatrix} \chi^{\uparrow\uparrow} & \phi^{\uparrow\downarrow} \\ -\phi^{\uparrow\downarrow T} & \chi^{\downarrow\downarrow} \end{bmatrix} = \mathrm{Pf}\,[\Phi_P] \quad (18)$$

satisfies the anti-symmetry where $\mathrm{Pf}[A]$ of $2n \times 2n$ matrix $A = \{a_{ij}\}$ is defined by an exterior product $\omega = \sum_{i<j} a_{ij} \cdot (e_i \wedge e_j)$ as

$$\frac{\omega^n}{n!} = \mathrm{Pf}[A] \cdot (e_1 \wedge e_2 \wedge \cdots \wedge e_{2n}) \quad (19)$$

More generalized form,

$$\Psi_{\mathrm{Pf}} = \mathrm{Pf}\begin{bmatrix} \Phi_P & \Phi_N \\ -\Phi_N^T & 0 \end{bmatrix}, \quad \Phi_N = \begin{bmatrix} 0 & B^{\uparrow} \\ B^{\downarrow} & 0 \end{bmatrix} \quad (20)$$

is implemented so that it can reduce to the conventional product of Slater determinant $D^{\sigma} = det[B^{\sigma}]$ when the amplitude of pairing contribution Ψ_P goes to zero.[42] In practical implementations in QMC the pairing orbitals are composed from natural orbitals $\{\psi_{\mu}(i)\}$ of SCF calculations as

$$\phi(i, j) = \sum_{\alpha \geq \beta} a_{\alpha\beta} \cdot [\psi_{\alpha}(i)\psi_{\beta}(j) + \psi_{\beta}(i)\psi_{\alpha}(j)] \quad (21)$$

When α and β run within the occupied orbitals of SCF the Pfaffian reduces to the conventional Slater determinant. Including virtual orbitals in Eq. (20) and optimizing expanding coefficients $\{a_{\alpha\beta}\}$ variationally we can take into

Table VI. Ground state energies of carbon atom with He pseudo potential. Jastrow (N) stands for the Jastrow factor describing upto N-body term ($N = 2, 3$). For CASSCF we used $(n, m) = (4, 8)$ for the active space, namely four electrons and eight orbitals in the active space. NO stands for the natural orbital. NMO is the matrix size of $a_{\alpha\beta}$ in Eq. (21). All the optimizations uses the scheme to minimize linear combination of energy (95%) and variance (5%).

Method	Orbitals	Energy (a.u./atom)	Note
SCF	UHF	−5.319437	
SCF	CASSCF	−5.400134	
VMC (Slater-Jastrow(2))	CASSCF-NO	−5.39848(7)	2-body Jastrow
VMC (Pfaffian-Jastrow(2))	CASSCF-NO	−5.40494(8)	Singlet only (NMO = 5)
VMC (Pfaffian-Jastrow(3))	CASSCF-NO	−5.4078(1)	Singlet only (NMO = 8)
VMC (Pfaffian-Jastrow(3))	CASSCF-NO	−5.4083(1)	Singlet only (NMO = 14)
VMC (Pfaffian-Jastrow(3))	CASSCF-NO	−5.4085(1)	Singlet/triplet (NMO = 14)
DMC (Pfaffian-Jastrow(3))	CASSCF-NO	−5.41348(5)	Singlet/triplet (NMO = 14)
DMC (Slater-Jastrow(3))	CASSCF-NO	−5.4061(3)	[47]
DMC (Pfaffian-Jastrow(3))	UHF-NO	−5.4082(1)	Singlet/triplet (NMO = 14)

Source: Reprinted with permission from [60], I. Ovcharenko et al., *J. Chem. Phys.* 114, 7790 (**2001**). © 2001.

account the correlation in a more compact and effective form than the multi determinant expansions. As in the case of Laplace expansion of determinants the Pfaffian can be evaluated recursively decomposing it into the operations on smaller matrix size, being not so expensive than determinant evaluation.[46] Generalization for multi-Pfaffian expansion is straightforward giving more compact form than multi determinant expansion.[46] The variational optimization of Pfaffian wavefunction modifies the nodal surface of many-body wavefunction going beyond the single determinant fixed nodes. Highly accurate performance for isolated atomic and molecular systems is reported recently.[46, 47]

Table VI shows how the Pfaffian works in QMC calculation of pseudo carbon atom. Compared to the case of Jastrow function the optimization of Pfaffian (or the degrees of freedom to tune the nodal surface in general, see Section 3) is easy to fall down with instability, then careful handling to choose the path to proceed and to enlarge the number of variational parameters is required. Detailed description about the procedure is given later in Section 3. The optimization is also sensitive to the choice of one-body orbitals in this case. The same optimization procedure using unrestricted HF natural orbitals (UHF-NO) cannot reach the value by complete-active-space SCF natural orbitals (CASSCF-NO). These correspond to different choice of initial guess of nodal surface since the Pfaffian can tune it during the variational optimization, showing importance of initial guess.

3. VARIATIONAL OPTIMIZATION

The variational optimization is given in the form of energy minimization[10] though it is difficult to perform it robustly by QMC because of the finite number of samplings available. Instead the scheme to minimize the variance of sampled local energies has developed.[28] Denoting the trial wavefunction as $\Psi(\vec{X};\vec{\alpha})$ with a set of variational parameters $\vec{\alpha}$, the variance can be rewritten as

$$\sigma_E^2(\vec{\alpha}) = \frac{\int d\vec{X}\cdot\Psi^2(\vec{X};\vec{\alpha})\cdot[E_L(\vec{X};\vec{\alpha})-\bar{E}(\vec{\alpha})]^2}{\int d\vec{X}\cdot\Psi^2(\vec{X};\vec{\alpha})} = \frac{\int d\vec{X}\cdot\Psi^2(\vec{X};\vec{\alpha}_0)\cdot w(\vec{\alpha})\cdot[E_L(\vec{X};\vec{\alpha})-\bar{E}(\vec{\alpha})]^2}{\int d\vec{X}\cdot\Psi^2(\vec{X};\vec{\alpha}_0)\cdot w(\vec{\alpha})} \quad (22)$$

where $\bar{E}$ is the average of E_L and the weight $w(\vec{\alpha}) = \Psi^2(\vec{\alpha})/\Psi^2(\vec{\alpha}_0)$ is introduced. The dependence $\sigma_E(\vec{\alpha})$ is therefore able to be evaluated by a single series of sampling generated by a trial wavefunction with a guess of variational parameter set $\vec{\alpha}_0$. The method, termed as reweighted variance minimization, is used for optimizing Jastrow functions in solid sodium calculations.[30] In this method when the fluctuation of $w(\vec{\alpha})$ gets larger the scheme becomes unstable, and this is serious when larger systems are treated. For such variational parmeter $\vec{\alpha}$ that modifies the nodal surface the scheme gets worse because of larger fluctuations of wavefunction value at the vicinity of nodal surfaces. Here we notice that the main mission using the dependence $\sigma_E(\vec{\alpha})$ is just to seek its global minimum, not to evaluate its value. Since the minimum position is not affected by the dependence of $w(\vec{\alpha})$ [54] we can modify the weight for stable behavior of optimization. A simple way is to filter out such contribution with large weight[24] from the evaluation of Eq. (22), which is used in the backflow optimization for silicon crystal calculation shown in Table V. Another way is the scheme not using the weight is used to update $\vec{\alpha}$, corresponding to setting all weights to be unity.[23, 55] With updated parameters the configuration to evaluate Eq. (22) (with $w = 0$) is generated and the $\vec{\alpha}$ is updated again repeatedly until it reaches to self-consistent minimum. In the scheme, unreweighted self-consistent optimization, the landscape of parameter space gets simpler, and for linear parameters especially it is proved that the landscape has unique minimum.[40] The scheme is used in the optimizations of our recent cases on carbon diamond[36] and silicon solids shown in Table V.

For VMC used as initial guess for DMC the developments of the variational optimization would have impacts in two-fold. The optimization of Jastrow functions, as the first issue, is vital for the feasibility of stable DMC run, otherwise worse wavefunction in its cusp quality gives much fluctuation of local energies and hence the run is easily suffered from the population explosion. Achievement of stable variance minimization in 90's establishes the feasibility of Jastrow optimization and hence the application of DMC to materials of practical interest.[5] The second is about the optimization of nodal surfaces. As mentioned above the nodal optimization is much more difficult because updating nodal positions during the optimization involves diverging fluctuations of local energy near to the nodes. The optimization techniques successful for Jastrow function are not working well for this optimization in general. This fact is one of the reason that stimulates the development of new optimization schemes. An outstanding achievement is the new energy minimization scheme which is proved to be stable even with smaller number of sampled configurations.[56] The key of the scheme is to modify the gradient and Hessian estimators from conventional ones,

$$\frac{\partial \bar{E}}{\partial c_i} = \left\langle \frac{\Psi_i}{\Psi} E_L + \frac{H\Psi_i}{\Psi} - 2\bar{E}\frac{\Psi_i}{\Psi} \right\rangle \tag{23}$$

$$\begin{aligned} \frac{\partial^2 \bar{E}}{\partial c_i \partial c_j} = 2\Bigg[& \left\langle \left(\frac{\Psi_{ij}}{\Psi} + \frac{\Psi_i \Psi_j}{\Psi^2} \right) \left(E_L - \bar{E} \right) \right\rangle \\ & - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \bar{E}_j - \left\langle \frac{\Psi_j}{\Psi} \right\rangle \bar{E}_i + \left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle \Bigg] \end{aligned} \tag{24}$$

into

$$\frac{\partial \bar{E}}{\partial c_i} = 2\left\langle \frac{\Psi_i}{\Psi} \left(E_L - \bar{E} \right) \right\rangle$$

$$\begin{aligned} \frac{\partial^2 \bar{E}}{\partial c_i \partial c_j} = 2\Bigg[& \left\langle \left(\frac{\Psi_{ij}}{\Psi} + \frac{\Psi_i \Psi_j}{\Psi^2} \right) \left(E_L - \bar{E} \right) \right\rangle \\ & - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \bar{E}_j - \left\langle \frac{\Psi_j}{\Psi} \right\rangle \bar{E}_i \Bigg] \\ & + \left[\left[\left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \left\langle E_{L,j} \right\rangle \right] + [j \leftrightarrow i] \right] \end{aligned}$$

where the simplified notation $X_j := \partial X / \partial c_j$ is used. The idea is to add such terms contributing zero average but non-zero variance such as $\langle \Psi_i/\Psi \rangle \langle E_{L,j} \rangle$ so that it can cancel out the variance coming from other terms into smaller total amount. The new scheme is proved to be robust in optimizing variational parameters those modify the nodal surface such as the backflow, Pfaffian, expanding coefficients of multi determinant, orbital functions, and basis sets.[48]

With several variational degrees of freedom, such as Jastrow functions with expanding coefficients and cutoff lengths, backflow transformation, pairing coefficients in Pfaffian etc., there are several possible 'paths' of optimizing procedure. Several sets are preferred to be optimized at once ideally: for example the change of nodal surface tuned by backflow should be accompanied with the change of amplitude of many-body wavefunction, which is affected by the Jastrow functions. In this sense it is preferable to optimize the backflow and Jastrow at once, though it is not always possible because of the limitation of computational resources. It is, in general, difficult to optimize at once the linear parameters such as expanding coefficients and non-linear ones such as cutoff lengths. Table VI shows the fact that we had to follow several optimization steps to arrive at the lowest possible energy. With practically available resources, the final 'quality' (variationally better or not) depends on the choice of the path of the optimizing procedure. Then the search for optimal and effective procedure is one of the most important missions for the practitioners. We begin the Jastrow function optimization because for the stable optimization of nodal surfaces the Jastrow function optimized in advance is indispensable. The optimization is better to begin with smaller numbers of variational parameters, otherwise the calculation falls down with instability. In the case shown in Table VI we gradually and carefully enlarge the size of optimization space from 2-body to 3-body for Jastrow, and NMO = 5, 8 and 14 for Pfaffian.

4. CONCLUSIONS

Helped by the inherent advantage for parallelized computing, the quantum Monte Carlo methods, including diffusion Monte Carlo, applied to *ab-initio* electronic structure calculations are rapidly expanding its applicable targets to realistic materials. The optimization of many-body wavefunction is a key and vital issue for practitioner, for which it is convenient to make arguments clear to distinguish the optimization of amplitude and of nodal surface of many-body wavefunction. For the amplitude optimization variety of Jastrow function forms are designed for inhomogeneous realistic systems and have been successful since the robust variance minimization schemes were well developed. Consideration of the Kato's cusp condition is essential for robust diffusion Monte Carlo calculation so that sampled local energies may not fluctuate largely. The optimization of Jastrow function plays an important role for such sense. There exists the variational principle for nodal surface of many-body wavefunction. We can improve the fixed node initially given by any trial form such as a Slater determinant so that the total energy gets decreased. Robust energy minimization scheme brings about practical feasibility for the optimization of such degrees of freedom to tune the nodal surface. Along with the above technical improvements to handle many-body wavefunctions we can expect further developments in research fields for deeper understanding of electronic correlation issues in realistic materials.

Acknowledgments: The results shown here are obtained in research projects financially supported by a Grant-in-Aid for Scientific Research in Priority Areas Development of New Quantum Simulators and Quantum Design (No. 17064016) of The Japanese Ministry of Education, Culture, Sports, Science, and Technology (KAKENHI-MEXT), by Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency (PREST-JST), and by the Engineering and Physical Sciences Research Council (EPSRC) of the United Kingdom. I also thank for collaborators during the researches, Michal Bajdich, Yoonseok Lee, Andrea Ma, Lubos Mitas, Richard J. Needs, Masayoshi Shimomoto, Yasutami Takada, Mike D. Towler, Cyrus J. Umrigar, and Lucas Wagner. Results shown in Table VI are obtained by Kenta Hongo in my group. The calculations were performed using facilities at the Computer Services for Academic Research (CSAR, University of Manchester), the Cambridge-Cranfield High Performance Computing Facilities (HPCF), the Center for Information Science of JAIST (Japan Advanced Institute of Science and Technology), Naka Fusion Research Establishment and Kansai Research Establishment of the Japan Atomic Energy Research Institute (JAERI), the Supercomputer Center (ISSP; Institute for Solid State Physics, University of Tokyo), the Numerical Materials Simulator (Hitachi SR11000) at the National Institute for Materials Science (NIMS, Tsukuba Japan), the High Performance Computing System of Hokkaido University, and at the Information Technology Center of the University of Tokyo. The author would like to thank Professor Teruo Matsuzawa (JAIST) for the generous provision of computing facilities during the researches.

References

1. N. W. Ashcroft and N. D. Mermin, Solid State Physics, Thomson Leaning (**1976**).
2. R. M. Martin Electronic Structure, Cambridge (**2004**).
3. D. G. Pettifor, Bonding and Structure of Molecules and Solids, Clarendon (**1995**).
4. G. F. Giuliani and G. Vignale, Quantum Theory of the Electron Liquid, Cambridge (**2005**).
5. W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* 73, 33 (**2001**).
6. C. J. Umrigar, M. P. Nightingale, and K. J. Runge, *J. Chem. Phys.* 99, 2865 (**1993**).
7. The methods based on this idea are called as Green function Monte Carlo (GFMC) method in broader sence. The DMC is one of the implementation of GFMC exploiting the analogy to the Green function of classical particle diffusion.
8. Hartree atomic units ($\hbar = e^2 = m_e = 4\pi\epsilon_0 = 1$) are used for all equations.
9. B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds, Monte Carlo Methods in *Ab Initio* Quantum Chemistry, World Scientific (**1994**).
10. L. I. Schiff, Quantum Mechanics, 3rd edn., McGraw-Hill (**1968**).
11. P. Fulde, Electron Correlations in Molecules and Solids, 3rd enlarged edn., Springer (**1995**).
12. T. Helgaker, P. Jorgensen, and J. Olsen, Molecular Electronic-Structure Theory, Wiley (**2000**).
13. A. J. Williamson, R. Q. Hood, and J. C. Grossman, *Phys. Rev. Lett.* 87, 246406 (**2001**).
14. D. Alfé and M. J. Gillan, *Phys. Rev. B* 70, 161101 (**2004**).
15. R. J. Jastrow, *Phys. Rev.* 98, 1479 (**1955**).
16. J. B. Anderson, *J. Chem. Phys.* 63, 1499 (**1975**).
17. J. B. Anderson, *J. Chem. Phys.* 65, 4121 (**1976**).
18. P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, Jr., *J. Chem. Phys.* 63, 5593 (**1982**).
19. T. Kato, *Commun. Pure Appl. Math.* 10, 151 (**1957**).
20. R. T. Pack and W. B. Brown, *J. Chem. Phys.* 45, 556 (**1966**).
21. Recent report points out the singularity of E_L would invade the central limiting theory of estimators of several quantities.[57,58]
22. G. Ortiz and P. Ballone, *Phys. Rev. B* 50, 1391 (**1994**).
23. A. J. Williamson, S. D. Kenny, G. Rajagopal, A. J. James, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, and P. Maccallum, *Phys. Rev. B* 53, 9640 (**1996**).
24. C. Filippi and C. J. Umrigar, *J. Chem. Phys.* 105, 213 (**1996**).
25. C.-J. Huang, C. Filippi, and C. J. Umrigar, *J. Chem. Phys.* 108, 8838 (**1998**).
26. N. D. Drummond, M. D. Towler, and R. J. Needs, *Phys. Rev. B* 70, 235119 (**2004**).
27. S. Fahy, X. W. Wang, and S. G. Louie, *Phys. Rev. Lett.* 61, 1631.(**1988**).
28. C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, *Phys. Rev. Lett.* 60, 1719 (**1988**).
29. R. J. Needs, M. D. Towler, N. D. Drummond, P. R. C. Kent, G. Rajagopal, and A. J. Williamson, CASINO Version 1.5 User's Manual, University of Cambridge, Cambridge (**2002**).
30. R. Maezono, M. D. Towler, Y. Lee, and R. J. Needs, *Phys. Rev. B* 68, 165103 (**2003**).
31. Y. Lee and R. J. Needs, *Phys. Rev. B* 67, 035121 (**2003**).
32. R. Saunders, R. Dovesi, C. Roetti, M. Causá, N. M. Harrison, R. Orlando, and C. M. Zicovich-Wilson, CRYSTAL98 User's Manual University of Torino, Torino (**1998**).
33. D. Bohm and D. Pines, *Phys. Rev.* 92, 609 (**1953**).
34. D. M. Ceperley, *Phys. Rev. B* 18, 3126 (**1978**).
35. D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* 45, 566 (**1980**).
36. R. Maezono, A. Ma, M. D. Towler, and R. J. Needs, *Phys. Rev. Lett.* 98, 025701 (**2007**).
37. R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Rios, CASINO Version 2.0 User Manual, University of Cambridge, Cambridge (**2006**).
38. J. R. Trail and R. J. Needs, *J. Chem. Phys.* 122, 174109 (**2005**).
39. J. R. Trail and R. J. Needs, *J. Chem. Phys.* 122, 014112 (**2005**).
40. N. D. Drummond and R. J. Needs, *Phys. Rev. B* 72, 085124 (**2005**).
41. CHAMP, A quantum Monte Carlo program written by C. J. Umrigar and C. Filippi with contributions by co-workers, URL: http://www.tc.cornell.edu/cyrus/champ.html.
42. L. K. Wagner, M. Bajdich, and L. Mitas, Qwalk: A quantum Monte Carlo program for electronic structure To be submitted, URL: http://www.qwalk.org/.
43. M. Kalos and F. Pederiva, Quantum Monte Carlo Methods in Physics and Chemistry of NATO Advanced-Study Institute, Series C: Mathematical and Physical Sciences, edited by M. P. Nightingale and C. J. Umrigar.
44. N. D. Drummond, P. L. Rios, A. Ma, J. R. Trail, G. G. Spink, M. D. Towler, and R. J. Needs, *J. Phys. Chem.* 124, 224104 (**2006**).
45. P. López Rios, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, *Phys. Rev. E* 74, 066701 (**2006**).
46. M. Bajdich, L. Mitas, G. Drobn', L. K. Wagner, and K. E. Schmidt, *Phys. Rev. Lett.* 96, 130201 (**2006**).
47. M. Bajdich, L. Mitas, L. K. Wagner, and K. E. Schmidt, *Phys. Rev. B* 77, 115112 (**2008**).
48. C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, *Phys. Rev. Lett.* 98, 110201 (**2007**).
49. R. P. Feynman and M. Cohen, *Phys. Rev. Lett.* 102, 1189 (**1956**).
50. H. Shull, *J. Chem. Phys.* 30, 1405 (**1959**).
51. J. Carlson, S.-Y. Chang, V. R. Pandharipande, and K. E. Schmidt, *Phys. Rev. Lett.* 91, 050401 (**2003**).
52. M. Casula and S. Sorella, *J. Chem. Phys.* 119, 6500 (**2003**).
53. M. Casula, C. Attaccalite, and S. Sorella, *J. Chem. Phys.* 121, 7110 (**2004**).

54. P. R. C. Kent, R. J. Needs, and G. Rajagopal, *Phys. Rev. B* 59, 12344 (**1999**).

55. K. E. Schmidt and J. W. Moskowitz, *J. Chem. Phys.* 93, 4172 (**1990**).

56. C. J. Umrigar and C. Filippi, *Phys. Rev. Lett.* 94, 150201 (**2005**).

57. J. R. Trail, *Phys. Rev. E* 77, 016703 (**2008**).

58. J. R. Trail, *Phys. Rev. E* 77, 016704 (**2008**).

59. S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. C. Payne, *Zeitschrift für Kristallographie* 220, 567 (**2005**); Kluwer Academic, Dordrecht, 525, 263 (**1999**).

60. I. Ovcharenko, A. Aspuru-Guzik, and W. A. Lester, Jr., *J. Chem. Phys.* 114, 7790 (**2001**).

Received: 21 November 2008. Accepted: 1 December 2008.