A constrained path auxiliary-field quantum Monte Carlo method for the homogeneous electron gas

1995 J. Phys.: Condens. Matter 7 L371

(http://iopscience.iop.org/0953-8984/7/28/001)

J. Phys.: Condens. Matter 7 (1995) L371–L376. Printed in the UK

LETTER TO THE EDITOR

A constrained path auxiliary-field quantum Monte Carlo method for the homogeneous electron gas

M T Wilson and B L Györffy

H H Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK

Received 10 April 1995

Abstract. Using a method suggested by Zhang, Carlson and Gubernatis, we have formulated an auxiliary-field method for the electron gas (jellium) in the form of a random walk through the space of Slater determinants. A calculation on a model jellium shows that this formulation removes some of the large statistical errors that have afflicted previous calculations.

The stochastic quantum Monte Carlo (QMC) methods allow, in principle, the calculation of exact ground-state properties of systems of interacting electrons. These calculations are attractive because the numerical cost scales with size significantly better than for 'exact' approaches such as the configuration interaction or exact diagonalization methods. There are many forms of QMC which have proved successful for many problems in physics and chemistry. However, since all QMC methods are stochastic approaches, they are afflicted with statistical uncertainty. In the case of fermions, one also has to cope with the famous 'fermion sign problem', which renders an accurate calculation impossible in the limit of large inverse temperature β.

Although the most successful QMC techniques for realistic electron systems are the diffusion (DMC) and Green's function (GFMC) methods [1], recently there has been some interest in the 'auxiliary-field quantum Monte Carlo' (AFQMC) method [2, 3, 4]. This method has been much used to study the Hubbard model where it has proved very successful [5, 6]. In its simplest form, the density matrix $e^{-\beta\hat{H}}$ is decomposed into a sum of one-body operators over all possible auxiliary fields. These fields must be summed over with an importance sampling method. The formulation exhibits many features that make it attractive compared to the DMC and GFMC methods. In particular the method is, in principle, generalizable to dealing with relativistic problems described by the Dirac Hamiltonian [4]. Moreover, it allows the straightforward computation of ground-state expectation values of operators which do not commute with $\hat{H}$. However, the auxiliary-field method suffers from much larger statistical error than the DMC or GFMC methods, and this has, up to now, severely limited its use.

Recently, Zhang, Carlson and Gubernatis [7] have proposed a new importance sampling technique for the AFQMC method and illustrated its use via calculations based on the Hubbard model. By introducing some ideas from the GFMC method, they have been able to reduce significantly the statistical error associated with the random fluctuations in the auxiliary fields and the fermion sign problem. In this letter we demonstrate that their constrained path auxiliary-field QMC method (CPAFQMC) works well for the case of the long-range Coulomb interaction in the jellium model, by showing that the fluctuations are much reduced compared to those in the AFQMC calculations of Silvestrelli *et al* [3].

0953-8984/95/280371+06$19.50

For clarity, we briefly recall the AFQMC method. We start with the real-space Hamiltonian

$$\hat{H} = \sum_{ij} T_{ij} c_i^\dagger c_j + \frac{1}{2} \sum_{ij} V_{ij} c_i^\dagger c_j^\dagger c_j c_i \tag{1}$$

where $T_{ij} = \langle \delta(\boldsymbol{r}-\boldsymbol{r}_i) | -\hbar^2 \nabla_r^2 / 2m | \delta(\boldsymbol{r}-\boldsymbol{r}_j) \rangle$ and $V_{ij} = e^2/4\pi\epsilon_0 |\boldsymbol{r}_i - \boldsymbol{r}_j|$. We further define $\hat{n}_i$ as the number operator $\hat{n}_i = c_i^\dagger c_i$. We can then use the density matrix (or imaginary time propagator) $e^{-\beta\hat{H}}$ to project out the ground state from a trial wavefunction $|\psi_T\rangle$:

$$E_0 = \lim_{\beta\to\infty} \frac{\langle \phi_T | \hat{H} e^{-\beta\hat{H}} | \psi_T \rangle}{\langle \phi_T | e^{-\beta\hat{H}} | \psi_T \rangle} \tag{2}$$

where, in principle, $|\psi_T\rangle$ and $|\phi_T\rangle$ are *any* wavefunctions not orthogonal to the ground state. We now use the well-known Hubbard–Stratonovich transformation [8] to write $e^{-\Delta\tau\hat{H}}$ as a functional integral over one-body operators:

$$\exp(-\Delta\tau\,\hat{H}) = \int d\sigma(\boldsymbol{r})\; G[\sigma] \exp\left(-\hat{h}[\sigma(\boldsymbol{r})]\Delta\tau\right) + O(\Delta\tau^2) \tag{3}$$

where $G[\sigma]$ is a gaussian measure, $G[\sigma] \propto \exp(\frac{1}{2}\Delta\tau \sum_{ij} V_{ij}\sigma_i\sigma_j)$ and $\Delta\tau$ is a small time slice. $\hat{h}[\sigma(\boldsymbol{r})]$ is the one-body Hamiltonian $\hat{T} + \sum_{ij} V_{ij}\sigma_j\hat{n}_i$. We see that since V_{ij} is a positive definite operator σ_i must be chosen to be purely imaginary for the functional integral to converge. Writing $e^{-\beta\hat{H}} = (e^{-\Delta\tau\hat{H}})^N$, we then are left with an auxiliary-field representation for $e^{-\beta\hat{H}}$ which involves a local, time-varying field $\sigma(\boldsymbol{r}, t)$. Putting together equations (2) and (3) we obtain an expression suitable for a Monte Carlo calculation:

$$E_0 = \lim_{\beta\to\infty} \frac{\int d\sigma(\boldsymbol{r},t) \left(\prod_{t=1}^N G[\sigma(\boldsymbol{r},t)]\right) \langle \phi_T | \hat{H} \prod_{t=1}^N \exp\left(-\hat{h}[\sigma(\boldsymbol{r},t)]\Delta\tau\right) | \psi_T \rangle}{\int d\sigma(\boldsymbol{r},t) \left(\prod_{t=1}^N G[\sigma(\boldsymbol{r},t)]\right) \langle \phi_T | \prod_{t=1}^N \exp\left(-\hat{h}[\sigma(\boldsymbol{r},t)]\Delta\tau\right) | \psi_T \rangle}. \tag{4}$$

We now comment on the significance of $|\psi_T\rangle$ and $|\phi_T\rangle$. They can in principle be selected as any wavefunction not orthogonal to the true ground state of $\hat{H}$, $|\psi_0\rangle$. However, it is clear that the method works best if $|\psi_T\rangle$ is chosen to be the best single Slater determinant (i.e. the Hartree–Fock wavefunction.) This is simply because under the action of $\exp(-\hat{h}[\sigma(\boldsymbol{r},t)]\,\Delta\tau)$ any Slater determinant evolves into another Slater determinant. The significance of $|\phi_T\rangle$ is in controlling the sign problem, as explained below. Zhang *et al* [7] suggest choosing $|\phi_T\rangle$ as the best available simple approximation to $|\psi_0\rangle$.

The basic approach of Silvestrelli *et al* is to sample the probability distribution $p[\sigma] = (\prod_{t=1}^N G[\sigma(\boldsymbol{r},t)])\ \langle \phi_T | \prod_{t=1}^N \exp(-\hat{h}[\sigma(\boldsymbol{r},t)]\Delta\tau) | \psi_T \rangle$, using a simple Monte Carlo method in two stages. Firstly one can generate σ-fields from the gaussian distribution G efficiently with the Box–Müller method [9]. Secondly, once the σ-fields have been chosen one can calculate the overlap $D[\sigma(\boldsymbol{r},t)] = \langle \phi_T | \prod_{t=1}^N \exp(-\hat{h}[\sigma(\boldsymbol{r},t)]\Delta\tau) | \psi_T \rangle$. Applying a Metropolis test [10] to successive $D[\sigma(\boldsymbol{r},t)]$ enables one to recover the full probability distribution $p[\sigma]$. One then evaluates the energy with the expression

$$E_0 = \lim_{\beta\to\infty} \lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} \frac{\langle \phi_T | \hat{H} \prod_{t=1}^N \exp\left(-\hat{h}[\sigma_m(\boldsymbol{r},t)]\Delta\tau\right) | \psi_T \rangle}{\langle \phi_T | \prod_{t=1}^N \exp\left(-\hat{h}[\sigma_m(\boldsymbol{r},t)]\Delta\tau\right) | \psi_T \rangle} \tag{5}$$

where the $\sigma_m(\boldsymbol{r},t)$ are taken from the distribution $p[\sigma_m]$. Unfortunately, this method for sampling $p[\sigma]$ is rather inefficient, as is evident from the very small Metropolis acceptance rate in the limit of large β. This inefficiency occurs because one looks only at the value of

the overlap $D[\sigma(\boldsymbol{r}, t)]$ *after* a propagation up to imaginary time β, whereas in reality it is often clear whether a configuration is likely to be accepted at much earlier times.

To benefit from this observation we follow the idea of Zhang *et al* [7] and instead of considering the problem as one of summing over random auxiliary fields we approach it in a manner similar to that of the GFMC method: namely by studying a selection of 'walkers', undergoing a random walk through the space of Slater determinants, whose motion is governed by the random auxiliary fields $\sigma(\boldsymbol{r}, t)$.

We consider an ensemble of M 'walkers' $\{|\phi_m^{(0)}\rangle\}$, each starting initially at the point $|\psi_T\rangle$ in the space of Slater determinants. This ensemble represents a many-body wavefunction, defined by $|\psi^{(0)}\rangle = \frac{1}{M}\sum_{m=1}^{M}|\phi_m^{(0)}\rangle$. We then consider applying the operator $e^{-\beta\hat{H}}$ on $|\psi^{(0)}\rangle$ one time step at a time, that is, applying each $e^{-\Delta\tau\hat{H}}$ individually. This is achieved by propagating each walker for a time $\Delta\tau$ via a different auxiliary field $\sigma_m(\boldsymbol{r})$, thus making all the $|\phi_m^{(1)}\rangle$ different. This application generates a new ensemble of walkers, $\{|\phi_m^{(1)}\rangle\}$. In fact, if we denote the ensemble at a time slice n by $\{|\phi_m^{(n)}\rangle\}$, and the many-body wavefunction it represents by $|\psi^{(n)}\rangle = \frac{1}{M}\sum_{m=1}^{M}|\phi_m^{(n)}\rangle$, we have an iterative relation for the wavefunction:

$$|\psi^{(n+1)}\rangle = \int_{\sigma} \mathrm{d}\sigma(\boldsymbol{r}; n+1)\; G[\sigma(\boldsymbol{r}; n+1)] \exp\left(-\hat{h}[\sigma(\boldsymbol{r}; n+1)]\,\Delta\tau\right)|\psi^{(n)}\rangle. \tag{6}$$

In this way $|\psi^{(n)}\rangle \to |\psi_0\rangle$, *the ground-state wavefunction, as* $n \to \infty$. *For the individual* walkers $|\phi_m^{(n)}\rangle$ in the ensemble we define the relationship $|\phi_m^{(n+1)}\rangle = \exp(-\hat{h}[\sigma(\boldsymbol{r}; n+1)]\,\Delta\tau)|\phi_m^{(n)}\rangle$. We could generate $|\psi^{(n)}\rangle$ simply by propagating the ensemble of $|\phi_m^{(n)}\rangle$ up to the imaginary time β in which we are interested, picking an auxiliary field from the distribution $G[\sigma(\boldsymbol{r})]$ at every time step k; however, this method is inefficient. Instead, we weight each walker $|\phi_m^{(n+1)}\rangle$ with its overlap with $|\phi_T\rangle$, that is with $W(|\phi_m^{(n+1)}\rangle) = \langle\phi_T|\phi_m^{(n+1)}\rangle$. In order that one may consider the weight properly at each time step we must define the 'transition probability' for the step $n \to n+1$ as

$$q^{(n+1)}[\sigma(\boldsymbol{r}; n+1)] = \frac{W(|\phi_m^{(n+1)}\rangle)}{W(|\phi_m^{(n)}\rangle)} = \frac{\langle\phi_T|\phi_m^{(n+1)}\rangle}{\langle\phi_T|\phi_m^{(n)}\rangle} \tag{7}$$

so that $W(|\phi_m^{(N)}\rangle) = \prod_{k=1}^{N} q^{(k)}\langle\phi_T|\phi_m^{(0)}\rangle$. We see that since $|\phi_m^{(0)}\rangle = |\psi_T\rangle \;\; \forall m$, $W(|\phi_m^{(N)}\rangle)$ is just equal to $D[\sigma(\boldsymbol{r}, t)]$ defined above. Thus, if we generate the fields from the same gaussian $G[\sigma]$ and weight the walkers with weight $W(|\phi_m^{(N)}\rangle)$, we are in effect sampling the same distribution $p[\sigma]$. This sampling can be carried out using a replication/deletion step instead of a Metropolis step. After every time step k we adjust the weight of a walker with $q^{(k)}$. This is most efficiently done by deleting the walker with probability $1 - q^{(k)}$ if $q^{(k)} < 1$, or replicating it with probability $q^{(k)} - 1$ if $q^{(k)} > 1$. In practice, we find that $q^{(k)}$ is always significantly less than 1, and so we adjust this criterion to keep the number of particles more or less constant.

For fermions we observe that the overlap $W(|\phi_m^{(k)}\rangle)$ at some time slice k can become negative. This manifests itself by a negative $q^{(k)}$. It is clear that if $\Delta\tau$ is sufficiently small then $q^{(l)} \to 0$ as $l \to k$. Hence the walker should be deleted if at any point $W(|\phi_m^{(k)}\rangle)$ turns negative. Whilst this deletion avoids the fermion sign problem, it is important to realize that it is, in general, an approximation. The quality of the approximation depends on the form of $|\phi_T\rangle$. Zhang *et al* suggest that if $|\phi_T\rangle$ is chosen to be identical to $|\psi_0\rangle$, the true ground-state wavefunction, then this approximation is exact. Thus if $|\phi_T\rangle$ is chosen close to $|\psi_0\rangle$ then it is reasonable for one to assume that the approximation will be good.

We have implemented this approach with calculations on a paramagnetic jellium system containing 14 electrons in a repeated box, at density $r_s = 4$ Bohr radii. We have performed

calculations for the ground-state energy using the methods of both Zhang *et al* (CPAFQMC) and Silvestrelli *et al* (AFQMC). In each case we have chosen the trial wavefunction $|\psi_T\rangle$ to be the Hartree–Fock solution (i.e. a Slater determinant of plane waves.) For simplicity we have also chosen $|\phi_T\rangle$ to be equal to the Hartree–Fock solution. (In the AFQMC calculations we followed the prescriptions of Silvestrelli *et al* as described in [3]. In particular we used a symmetric form of equations (2) and (4) with an evolution operator of time $\beta/2$ to both the left and the right of $\hat{H}$.) For each method we have plotted the correlation energy per particle against imaginary time β. In figure 1 we show the results for the method of Silvestrelli *et al*, and the results for the CPAFQMC method of Zhang *et al* are shown in figure 2. For comparison we also show the result of a GFMC calculation for the same system [3].

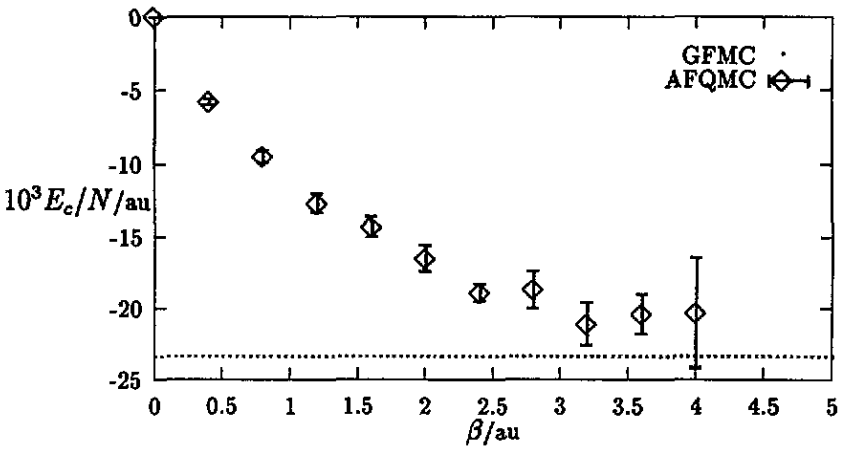

Figure 1. Correlation energy per electron against β for method of Silvestrelli *et al*, using 14 electrons at a density corresponding to $r_s = 4$ Bohr radii. $\Delta\tau = 0.2$ au. Energy is measured in atomic units (au), with 1 au = 2 rydbergs. The point $\beta = 4$ au corresponds to a total of 4500 σ-fields.

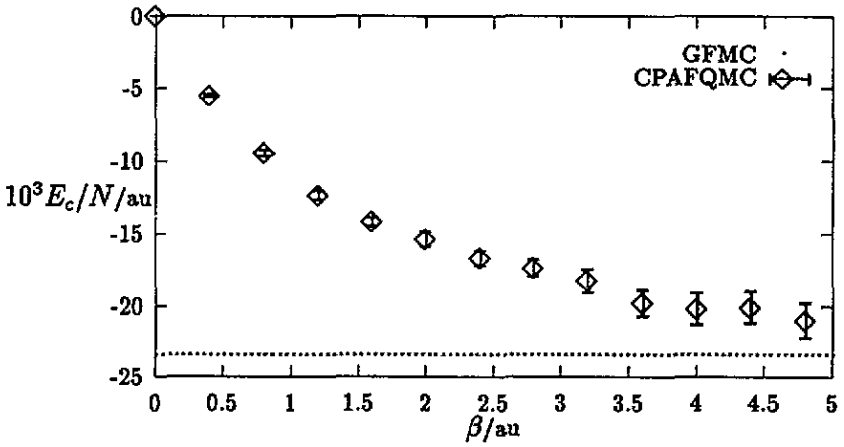

Figure 2. Correlation energy per electron against β for method of Zhang *et al*, using 14 electrons at a density corresponding to $r_s = 4$ Bohr radii. $\Delta\tau = 0.2$ au. All points correspond to a total of 400 walkers.

For the calculations we have chosen an imaginary time step of $\Delta\tau = 0.2$ au and defined the one-particle wavefunctions in the Slater determinants by specifying their lowest 343 ($=7^3$) Fourier components. The energy is given in hartrees and the imaginary times in inverse hartrees. The calculations have been performed on a CRAY Y-MP8 supercomputer. The AFQMC calculation using the method of Silvestrelli *et al* required approximately 24 hours of CPU time, whereas the CPAFQMC method of Zhang *et al* took just 4 hours.

It is clear from the figures that the new CPAFQMC method of Zhang *et al* is considerably more efficient than the AFQMC method of Silvestrelli *et al*. We note that at $\beta = 4$ au the new method gives only the order of one third of the statistical error of the method of Silvestrelli *et al*, using only one tenth of the number of auxiliary fields (or walkers.) Most notable is that the size of the statistical error no longer grows exponentially as β increases, making a much higher value of β obtainable.

We have also performed a calculation on a similar system containing 38 electrons. In this case we have been limited to just 100 walkers. We found that the overall correlation energy per particle is very similar to the 14-electron case. More significantly, the statistical error for 38 electrons is no higher than that for 14 electrons, when we take into account the number of walkers. The results are shown in figure 3. The calculation took about 16 hours, which corresponds to a scaling of the order of the number of electrons to the third power.

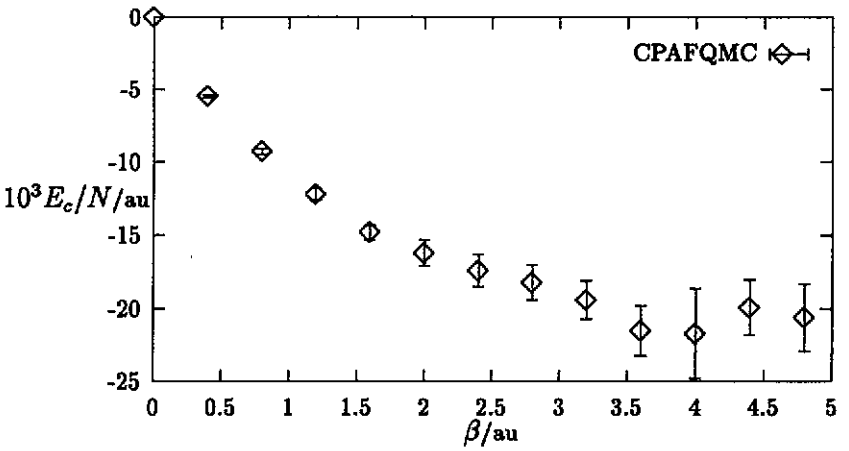

Figure 3. Correlation energy per electron against β for method of Zhang *et al*, using 38 electrons at a density corresponding to $r_s = 4$ Bohr radii. $\Delta\tau = 0.2$ au. All points correspond to a total of 100 walkers.

However, encouraging as the above results are, there is a difficulty which can be anticipated. Namely the amount of computer memory required to store the information for each walker may become prohibitive. Evidently, implementing this method requires that we store information relating to each walker for all the walkers simultaneously, whereas in the method of Silvestrelli *et al* we only need to consider one walker (auxiliary field) at a time. For example, in the simple 14-electron calculations presented in this letter, we are practically limited to about 600 walkers. Although this is sufficient for the model jellium at hand, for a more realistic system with more electrons we might require a more efficient algorithm for handling the storage of the wavefunctions.

Finally, we note some further advantages of this method. Having kept the AFQMC procedure, one is free to exploit all the advantages that using this method provides. For instance, one could straightforwardly combine the above use of walkers in the space of Slater

determinants with some of the efficient sampling algorithms for AFQMC calculations [3]. Also, since this method is not dependent on the physical system we study, given sufficient computer memory it will have application to a wide range of many-fermion problems.

To summarize, we have used the method of Zhang *et al* which combines many of the positive features of the GFMC and AFQMC algorithms to calculate the ground-state energy of a jellium system and have shown it to be superior to the previously used AFQMC method for long-range interactions.

We would like to thank Dr Matthew Foulkes for bringing the preprint of Zhang *et al* to our attention. M T Wilson would like to thank the Engineering and Physical Sciences Research Council and Daresbury Laboratory for a CASE research studentship.

References

[1] See, for example,
Binder K (ed) 1986 *Monte Carlo Methods in Statistical Physics* 2nd edn (Berlin: Springer)
[2] Sugiyama G and Koonin S 1986 *Ann. Phys., NY* **168** 1
[3] Silvestrelli P-L, Baroni S and Car R 1993 *Phys. Rev. Lett.* **71** 1148
[4] Wilson M and Gyorffy B 1995 *J. Phys. C: Condens. Matter* **7** 1565
[5] White S, Sugar R and Scalettar R 1989 *Phys. Rev.* B **38** 11 665
White S, Scalapino D, Sugar R, Gubernatis J and Scalettar R 1989 *Phys. Rev.* B **40** 506
[6] Sorella S, Baroni S, Car R and Parrinello M 1989 *Europhys. Lett.* **8** 663
[7] Zhang Shiwei, Carlson J and Gubernatis J 1995 *Phys. Rev. Lett.* **74** 3652
[8] Hubbard J 1959 *Phys. Rev. Lett.* **3** 77
Stratonovich R 1957 *Dokl. Akad. Nauk SSSR* **1115** 1097
[9] See, for example,
Negele J and Orland H 1988 *Quantum Many-Particle Systems* (New York: Addison-Wesley) ch 8
[10] Metropolis N, Rosenbluth A, Rosenbluth M, Teller A and Teller E 1953 *J. Chem. Phys.* **21** 1087