EUROPHYSICS LETTERS 14 September 1991

Europhys. Lett., 16 (3), pp. 249-254 (1991)

Lanczós-Type Algorithm for Quantum Monte Carlo Data.

M. CAFFAREL(*)(**), F. X. GADEA(***) and D. M. CEPERLEY(**)

(*) *Laboratoire Dynamique des Interactions Moléculaires, Université Paris VI 75252 Paris Cedex 05, France*

(**) *Department of Physics and NCSA, University of Illinois at Urbana-Champaign Urbana, IL 61801*

(***) *Laboratoire de Physique Quantique, Université Paul Sabatier 31062 Toulouse Cedex, France*

(received 11 March 1991; accepted 11 July 1991)

PACS. 31.15 – General mathematical and computational developments.
PACS. 31.50 – Molecular solids.
PACS. 05.30F – Fermion systems and electron gas.

Abstract. – A method for accelerating the rate of convergence of the long-time (or small-temperature) limit of quantum Monte Carlo approaches is presented. To do that, a variation of the Lanczós algorithm suitable for QMC data is introduced. This algorithm allows one to extract more information from correlation functions at small times, thus avoiding large statistical fluctuations associated with large times. It is first applied to an exactly soluble system and then to the LiH molecule. Calculations using both the fixed-node and nodal-release approaches are discussed.

Quantum Monte Carlo (QMC) methods have proved to be powerful techniques for solving the Schrödinger equation. They have been applied to a variety of problems [1] such as the study of quantum liquids and solids, the electron gas or the electronic structure of small molecules. In each case, very accurate results for some properties of these systems have been obtained. Although there exists a number of variants of QMC methods, the common idea in the approaches we consider here consists in projecting out the ground-state component of a known trial wave function, Ψ_T, by applying a suitable projection operator to this function ($\exp[-tH]$ in diffusion Monte Carlo (DMC) or $1/(H-E)^n$ in Green's function Monte Carlo (GFMC) methods, H denoting the Hamiltonian operator) and then letting the projecting parameter (t or n) go to infinity. Within the framework of DMC methods used in this work, this projection procedure takes the form

$$\exp[-tH]\,\Psi_T \to \Psi_0 + O(\exp[-t\Delta E]), \quad \text{as } t \to \infty, \tag{1}$$

where Ψ_0 denotes the ground-state wave function and ΔE is the gap in energy between the first two eigenstates having a nonzero overlap with the trial wave function.

This long-time limit may be difficult to perform. Certainly the most well-known illustration of such a difficulty is the so-called sign problem occurring in exact simulations of

fermion systems. This problem has been discussed in detail in many places (see, *e.g.*, [2]. It can be summarized as follows. Fermion matrix elements of the operator $\exp[-$ decompose as a difference of two boson contributions corresponding to even and permutations of the particle labels. At large times t (or low temperatures), the two bo contributions nearly cancel and the resulting fermion contribution becomes ra exponentially smaller than the statistical fluctuations. Accordingly, only reasonably sm values of t may be used in eq. (1) in a fermionic simulation and convergence of the limit t not be possible in practice. Even in bosonic-type calculations the long-time limit can difficult to handle for quantities other than the energy, particularly for systems involvin large number of bosons. For example, to compute ground-state expectations of operat not commuting with H requires a similar projection at large time (see, for instance, discussion in [5]).

In this work we propose a new procedure for taking advantage of the informa contained in data at small values of the projecting time t, thus minimizing the effect of statistical fluctuations at large times. We shall present this procedure within the framewor of a variant of the DMC approach—the pure diffusion Monte Carlo method [6]—altho any other Monte Carlo scheme could be employed without essential changes. Consider projected trial wave function at time t:

$$\widetilde{\Psi}_{\mathrm{T}}(t) \equiv \exp[-tH]\,\Psi_{\mathrm{T}}.$$

With quantum Monte Carlo techniques, quantum averages with respect to $\widetilde{\Psi}_{\mathrm{T}}$ may computed. In what follows the norm $n(t)$ of $\widetilde{\Psi}_{\mathrm{T}}(t/2)$ and the average $h(t)$ of H over $\widetilde{\Psi}_{\mathrm{T}}$ will be used:

$$n(t) \equiv \langle \widetilde{\Psi}_{\mathrm{T}}(t/2) | \widetilde{\Psi}_{\mathrm{T}}(t/2) \rangle = \langle \Psi_{\mathrm{T}} | \exp[-tH] | \Psi_{\mathrm{T}} \rangle$$

and

$$h(t) \equiv \langle \widetilde{\Psi}_{\mathrm{T}}(t/2) | H | \widetilde{\Psi}_{\mathrm{T}}(t/2) \rangle = \langle \Psi_{\mathrm{T}} | H \exp[-tH] | \Psi_{\mathrm{T}} \rangle .$$

These matrix elements of $\exp[-tH]$ may be computed as stochastic averages over a set drifting random walks generated by using a Langevin equation. Denoting $\langle \ldots \rangle_{\mathrm{DRW}}$ stochastic average, $n(t)$ and $h(t)$ may be written in the following form:

$$n(t) = \langle \Psi_{\mathrm{G}} | \Psi_{\mathrm{G}} \rangle \left\langle w(R(0))\, w(R(t)) \exp\left[- \int_0^t E_{\mathrm{L}}^{\mathrm{G}}(R(s))\, \mathrm{d}s \right] \right\rangle_{\mathrm{DRW}}$$

and

$$h(t) = \langle \Psi_{\mathrm{G}} | \Psi_{\mathrm{G}} \rangle \cdot$$

$$\cdot \left\langle w(R(0))\, w(R(t)) \frac{1}{2} (E_{\mathrm{L}}^{\mathrm{T}}(R(0)) + E_{\mathrm{L}}^{\mathrm{T}}(R(t))) \exp\left[- \int_0^t E_{\mathrm{L}}^{\mathrm{G}}(R(s))\, \mathrm{d}s \right] \right\rangle_{\mathrm{DRW}} .$$

Here $w = \Psi_{\mathrm{T}}/\Psi_{\mathrm{G}}$ is a weight factor involving the trial wave function, Ψ_{T}, and the guidi function, Ψ_{G}, a strictly positive function responsible for importance sampling; $E_{\mathrm{L}}^{\mathrm{T}} = H\Psi_{\mathrm{T}}/\Psi_{\mathrm{T}}$ is the local energy associated with Ψ_{T}, $E_{\mathrm{L}}^{\mathrm{G}} = H\Psi_{\mathrm{G}}/\Psi_{\mathrm{G}}$ is the local energy associated with and $R(s)$ stands for the drifting random walk in the $3N$-dimensional configuration space. normalization factor appearing on the right-hand side of each expression will be immater in what follows. Equations (4) are a generalization of the well-known Feynman-Kac for

For a detailed presentation and derivation of these formulae, the reader is referred to previous works [6]. At this point, we would like to emphasize two situations that will be encountered in what follows. When Ψ_G is chosen to be $|\Psi_T|$, random walks generated by the Langevin equation are trapped in subdomains of the configuration space delimited by the $(3N-1)$-dimensional nodes of the trial wave function Ψ_T and no change of sign for the weight factor occurs. This stable approach is called fixed-node approximation, since the nodes are in general approximate. On the other hand, when Ψ_G is chosen to be strictly positive everywhere, no approximation is made but weights have no longer a definite sign for fermions. This exact but unstable method will be referred to as the nodal-release approach. More details about both approaches may be found elsewhere [3, 6].

The standard way of extracting the exact energy from a set of QMC data $\{n(t_i), h(t_i)\}_{i=1,N}$ consists in looking at the ratio

$$\frac{h(t)}{n(t)} \to E_0 \quad \text{as } t \to \infty . \tag{5}$$

To do that, matrix elements are computed up to values of t necessary to reach the convergence. The main point of this work is to use information contained in h, n at smaller t. This is important due to the increase of statistical fluctuations as t goes to infinity. This idea, which in fact takes its origin in the somewhat different context of effective Hamiltonian theory [7], is implemented here in a quite simple way.

Let us define the following basis set of size n consisting of the projected trial wave function evaluated at n different times:

$$\{\tilde{\Psi}_T(t_1), \tilde{\Psi}_T(t_2), \ldots, \tilde{\Psi}_T(t_n)\} \equiv \{t_1, t_2, \ldots, t_n\} .$$

For finite n, such a basis set is in general linearly independent and may be used to diagonalize H. To perform the diagonalization, the matrix elements H_{ij} of H and N_{ij} of the unity operator between two arbitrary functions of the basis set are needed. It is easy to check that such matrix elements may be in fact trivially expressed in terms of the matrix elements (3) as follows:

$$H_{ij} \equiv \langle \tilde{\Psi}_T(t_1) | H | \tilde{\Psi}_T(t_2) \rangle = h(t_1 + t_2) \tag{6a}$$

and

$$N_{ij} \equiv \langle \tilde{\Psi}_T(t_1) | \tilde{\Psi}_T(t_2) \rangle = n(t_1 + t_2) . \tag{6b}$$

This is important since it means that no extra quantities beyond the usual matrix elements (eq. (3)) are required. Then once H and N are estimated with QMC, the generalized eigenvalue problem is solved by standard numerical methods. At this point, it is important to emphasize that the algorithm proposed here is nothing but a variation of the well-known Lanczós algorithm with Ψ_T playing the role of the initial vector and $\exp[-tH]$ playing the role of H. Using the terminology of Krylov spaces [8], this can be rephrased by saying that H is diagonalized within the Krylov subspace $\{\Psi_T, \exp[-t_1 H]\Psi_T, \ldots, \exp[-t_n H]\Psi_T\}$ instead of the Krylov subspace $\{\Psi_T, H\Psi_T, \ldots, H^{n-1}\Psi_T\}$ as in the Lanczós algorithm. Note that the standard method described by (5) may be viewed as a rather trivial case for which H is diagonalized within the one-dimensional subspace defined by $\tilde{\Psi}_T(t/2)$.

Let us first present the application of this approach to an exactly solvable problem, namely the harmonic oscillator described by the Hamiltonian $H = -(1/2)(d^2/dx^2) + (1/2)Kx^2$. The trial wave function is chosen to be Gaussian (different from the exact solution) and since the kernel of $\exp[-tH]$ is also Gaussian, exact expressions for matrix elements (3) may be

TABLE I. – *Comparison between the Lanczós-type algorithm and the standard method fo harmonic oscillator. Hamiltonian corresponding to* $K = 3.0$, *trial wave function*(a) *with* $k =$

This work		Standard method(b)		Exact
Basis set(c)	Eigenvalues	Basis set(c)	Eigenvalues	
{0.0}	$\lambda_0 = 1.0$	{0.0}	$\lambda_0 = 1.0$	
{0.0, 0.02}	$\lambda_0 = 0.88$ $\lambda_1 = 5.1$	{0.02}	$\lambda_0 = 0.98$	
{0.0, 0.02, 0.04}	$\lambda_0 = 0.8668$ $\lambda_1 = 4.4$ $\lambda_2 = 9.4$	{0.04}	$\lambda_0 = 0.96$	
{0.0, 0.02, 0.04, 0.06}	$\lambda_0 = 0.8661$ $\lambda_1 = 4.34$ $\lambda_2 = 7.86$ $\lambda_3 = 13.9$	{0.06}	$\lambda_0 = 0.95$	
{0.0, 0.02, 0.04, 0.06, 0.08}	$\lambda_0 = 0.86603$ $\lambda_1 = 4.331$ $\lambda_2 = 7.80$ $\lambda_3 = 12.0$ $\lambda_4 = 18.4$	{0.08}	$\lambda_0 = 0.94$	$\lambda_0 = 0.86602$ $\lambda_1 = 4.3301$ $\lambda_2 = 7.794..$ $\lambda_3 = 11.258..$ $\lambda_4 = 14.72..$

(a) $\Psi_T = (k/\pi)^{1/4} \exp[-(\sqrt{k}/2)x^2]$.
(b) Equation (5).
(c) Basis set defined as $\{t_1, t_2, ..., t_n\} \equiv \{\exp[-t_1 H]\Psi_T, \exp[-t_2 H]\Psi_T, ..., \exp[-t_n H]\Psi_T\}$, see text.

obtained. Table I presents results obtained when using a basis set of increasing size compared to those resulting from (5) using only the last component of the set. For the case the Lanczós-type algorithm all the eigenvalues are given. A few remarks are in order. First it is clear that the lowest eigenvalue in the Lanczós approach converges quite rapidly to the exact energy. This is in sharp contrast with the standard method which would require much larger times to achieve the convergence. A way of understanding this may be put as follows. Diagonalizing H within the subspace $\{t_1, t_2, ..., t_n\}$ may be viewed as constructing the best wave function written in the form of a linear combination of the projected trial wave function defined at different times, $\sum_k c_k \Psi_T(t_k)$. This combination has much more variational freedom than the one-state approach using only $\{t_n\}$ and therefore the resulting improvement in energy may be important.

A second point worth mentioning is that excited-state energies may also be obtained in principle. Results presented in table I show a good convergence of excited-state eigenvalues toward their respective limit, at least for the first two. Note that, according to the MacDonald variational theorem applying for linear variational calculations [9], all the eigenvalues $\lambda_i(t)$ are always greater than the corresponding exact eigenvalue of the Hamiltonian, the equality would be obtained by letting t go to infinity. How far $\lambda_i(t)$ is from $E_i = \lambda_i(\infty)$ for a given time t depends essentially on the overlap between the exact excited-state and the trial wave function. The problem of evaluating excited-state energies will not be discussed further, since the obtained results are not representative of the typical case where matrix elements have statistical errors. However, note that this approach may be readily generalized to the multiple-state method for computing excited-state properties of Ceperley and Bernu [10].

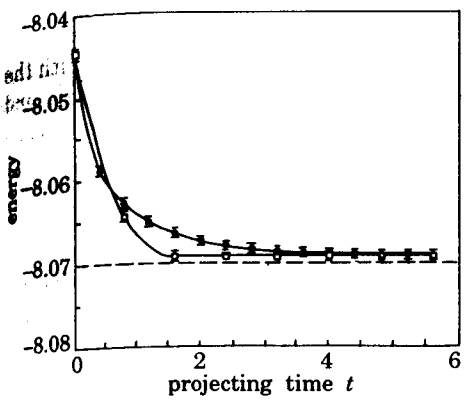

Fig. 1.

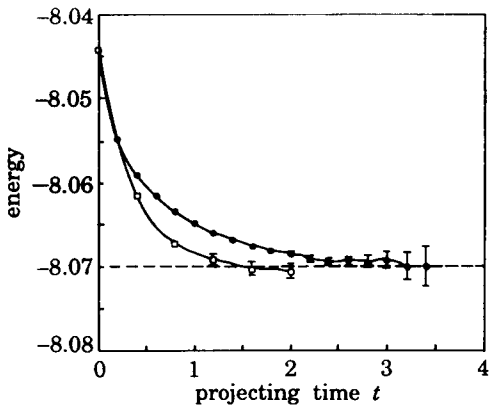

Fig. 2.

Fig. 1. – Fixed-node energy as a function of the projecting time t for the LiH molecule for both the standard method (●) and the proposed Lanczós-type method (□). The dashed line indicates the exact energy. The small difference between the energy obtained for large times and the exact energy is due to the fixed-node error. Energy and time in atomic units. The curves are only a guide to the eye.

Fig. 2. – Nodal-release energy as a function of the projecting time t for the LiH molecule for both the standard method (●) and the proposed Lanczós-type method (□). The dashed line indicates the exact energy. Large fluctuations at large times in the standard method result from the fermion sign problem. Energy and time in atomic units. The curves are only a guide to the eye.

Let us present a realistic application to the LiH molecule involving quantum Monte Carlo evaluation of matrix elements (3). In order to deal with the fermionic constraints, we have used both the approximate fixed-node and exact nodal-release approaches. Figure 1 presents the convergence of the fixed-node energy as a function of the projecting time t for both the standard method (5) (upper curve) and the proposed method (lower curve). With the Lanczós-type approach convergence is reached at times ~ 1.6 a.u., while the standard method requires times greater than 3 a.u. Statistical errors for both curves have been obtained by computing the dispersion of results over a set of independent calculations.

There is a serious numerical problem in applying this scheme to Monte Carlo results. When t goes to infinity the projected trial wave function $\tilde{\Psi}_T(t)$ converges exponentially fast to Φ_0, eq. (1). Accordingly, projected trial wave functions at large times become almost identical. Hence the matrices become nearly singular and, because of the finite precision on machine, it is not possible to use basis sets of arbitrary size if there is any statistical error on the matrix elements. We circumvented this problem by employing basis sets small enough to lead to well-conditioned matrices. For the case presented in fig. 1, the successive basis sets employed are: $\{0.0\}$, $\{0.0, 0.4\}$, $\{0.0, 0.4, 0.8\}$, $\{0.0, 0.4, 1.2\}$, ..., and $\{0.0, 0.4, 2.8\}$ with $\Delta t = 0.005$ a.u. as time step. The energy obtained in both calculations is $-8.0691(6)$ ($(99 \pm 0.7)\%$ of the correlation energy is recovered).

Figure 2 presents our calculations using the exact nodal-release procedure. The positive guiding function used here is of the form $\Psi_G = \sqrt{\Psi_T^2 + \theta \prod_i \rho(r_i)}$, where $\prod \rho$ denotes the Hartree density corresponding to the trial wave function Ψ_T. The switching parameter θ has been chosen to have a value of 0.48 so as to minimize both statistical fluctuations on the local energy $E_L = H\Psi_G/\Psi_G$ and fluctuations arising from crossings and recrossings of nodes [3]. The upper curve of fig. 2 represents the variation of the energy *vs.* the projecting time as

obtained with the standard method (5). The fermion sign problem is evident as ... greater than 2.0 a.u. The lower curve has been obtained by applying our algorith... successive basis sets: {0.0}, {0.0, 0.2}, {0.0, 0.2, 0.4}, {0.0, 0.2, 0.6}, {0.0, 0.2, ... {0.0, 0.2, 1.0} with $\Delta t = 0.005$ a.u. Within statistical fluctuations the convergence i... for $t \sim 2.0$ u.a., that is before statistical fluctuations arising from the sign problem ... too pathological. The resulting energy is $E_0 = -8.070(1)$, compared with ... nonrelativistic energy of -8.0699 [11].

These good results should be taken with caution. The main point to emphasize is ... using a linear variational calculation the energy is expressed as a lowest eigenvalue ... a nonlinear function of the matrix elements (3). The stability of the eigenvalue with ... to statistical errors has been obtained here at the expense of a high-quality evaluatio... matrix elements. It is not clear whether such a quality can be obtained for fermionic s... involving a large number of particles. However, the results presented here are imp... since they demonstrate that QMC data at small times, that is before the sign catas... occurs, may eventually contain enough information for computing exact fermionic gr... state properties. In a forthcoming work, a more stable and general method for t... advantage of this information will be presented [12].

* * *

One of us (MC) would like to thank W. KRAUTH for a number of helpful discussion... also wishes to thank the CNRS, the NATO and the Department of Physics at ... Champaign for their financial support during his visit at the University of Illinois... computations were done on the CRAY Y-MP of the National Center for Supercomp... Applications at Urbana-Champaign. DC is supported by the NSF under NSF-D... 8808126.

REFERENCES

[1] See, *e.g.*, contributions in *J. Stat. Phys.*, **43** (1986). See also *Monte Carlo Methods in Sta... Physics*, edited by K. BINDER (Springer-Verlag, Berlin) 1977 and *ibid.*, Vol. II (1984). ... applications to the electronic structure of atoms and molecules, see, *e.g.*, LESTER jr. W. A. ... HAMMOND B. L., *Annu. Rev. Phys. Chem.*, **41** (1990) 283 and references therein.
[2] KALOS M. H., in *Monte Carlo Methods in Quantum Problems, NATO ASI Series C* (R... Dordrecht) 1982.
[3] CEPERLEY D. M. and ALDER B. J., *J. Chem. Phys.*, **64** (1984) 5833.
[4] SCHMIDT K. E. and KALOS M. H., in *Monte Carlo Methods in Statistical Physics*, edited b... BINDER (Springer-Verlag, Berlin) 1984.
[5] RUNGE K. J. and RUNGE R. J., preprint (1990).
[6] CAFFAREL M. and CLAVERIE P., *J. Stat. Phys.*, 43 (1986) 797; *J. Chem. Phys.*, 88 (1988) 1088, 1100.
[7] GADEA F. X., Thèse d'Etat, Toulouse, 1987; *Phys. Rev. A*, **36** (1987) 2557.
[8] PARLETT B. N., *The Symmetric Eigenvalue Problem, Prentice-Hall Series in Computati... Mathematics*, 1980.
[9] MACDONALD J. K. L., *Phys. Rev.*, **43** (1933) 830.
[10] CEPERLEY D. M. and BERNU B., *J. Chem. Phys.*, **89** (1988) 6316.
[11] LIE G. C. and CLEMENTI E., *J. Chem. Phys.*, **60** (1974) 1275.
[12] CAFFAREL M. and CEPERLEY D. M., to be published.