Monte Carlo technique for finding the lowest eigenvalue of a modified Schrodinger equation

1985 J. Phys. A: Math. Gen. 18 1327

(http://iopscience.iop.org/0305-4470/18/9/016)

J. Phys. A: Math. Gen. 18 (1985) 1327-1335. Printed in Great Britain

Monte Carlo technique for finding the lowest eigenvalue of a modified Schrödinger equation

Jan Vrbik

Department of Mathematics, Brock University, St Catharines, Ontario, L2S 3A1, Canada

Received 3 July 1984

Abstract. An efficient and accurate algorithm for Monte Carlo simulation of a solution to the diffusion equation with drift and branching terms is derived by expanding the Green function operator in time.

1. Introduction

It has been shown (Ceperley and Adler 1980) that the task of finding a ground-state energy of a quantum system of particles can be reduced to solving

$$-\tfrac{1}{2}\nabla^2 f(\boldsymbol{r})+\nabla\cdot(\boldsymbol{F}(\boldsymbol{r})f(\boldsymbol{r}))+E(\boldsymbol{r})f(\boldsymbol{r})\overset{\text{def}}{=}\hat{T}f(\boldsymbol{r})=E_0 f(\boldsymbol{r}) \tag{1}$$

where $\boldsymbol{r}$ is an N-dimensional vector, $E(\boldsymbol{r})$ and $\boldsymbol{F}(\boldsymbol{r})$ are given functions of $\boldsymbol{r}$ (scalar and vector, respectively), and E_0 is an eigenvalue of the linear operator $\hat{T}$. Let the conditions of the problem imply that there is a real minimum eigenvalue (E_0 will, from now on, represent only this eigenvalue) which corresponds to a non-negative integrable eigenfunction $f(\boldsymbol{r})$. This enables us to further require that

$$\int f(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}=1 \tag{2}$$

and interpret $f(\boldsymbol{r})$ as a probability density function of an N-dimensional statistical distribution (Kalos 1962).

Finding an accurate numerical approximation to $f(r)$ and E_0 can be achieved by employing a Monte Carlo technique. To this end, one needs to modify equation (1) in the following fashion (Grimm and Storer 1969):

$$-\partial f(\boldsymbol{r},t)/\partial t=(\hat{T}-E_0)f(\boldsymbol{r},t) \tag{3}$$

where f is now considered a function of $\boldsymbol{r}$ and t (called time, for convenience). Obviously, the stationary ($t\to\infty$) behaviour of a solution to (3) will supply a solution to (1).

2. Monte Carlo simulation

A simulation of a solution to (3) can be carried out easily once the Green function $G(\boldsymbol{r}'\leftarrow\boldsymbol{r},t)$ of the equation is known (Kalos 1962, Grimm *et al* 1971). Such a simulation requires generating a set of M (usually hundreds) random vectors (called, in this

0305-4470/85/091327+09$02.25

context, configurations ${}^1\boldsymbol{r}, {}^2\boldsymbol{r}, \ldots, {}^M\boldsymbol{r}$ from an initial distribution $f(\boldsymbol{r}, 0)$ (which, in our case, can be chosen arbitrarily); then, for a specific value of t (called the time step), repeatedly advancing the M configurations according to $G(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t)$ until a target time is reached.

The Green function itself should meet

$$-\partial G(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t)/\partial t = (\hat{T}' - E_0)G(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t) \tag{4}$$

subject to

$$G(\boldsymbol{r}' \leftarrow \boldsymbol{r}, 0) = \delta(\boldsymbol{r}' - \boldsymbol{r}) \tag{4a}$$

where $\hat{T}'$ is an exact analogue of $\hat{T}$, operating on $\boldsymbol{r}'$ instead of $\boldsymbol{r}$ (in (4), E_0 should be viewed as a parameter rather than an eigenvalue). Unfortunately, in the case of our specific operator $\hat{T}$ (defined in (1)) there is no general way of solving (4) analytically. Aiming to remedy the situation, we rewrite (3) as

$$-\partial f(\boldsymbol{r}, t)/\partial t = (\hat{T}_0 + \hat{T}_1 + \hat{T}_2)f(\boldsymbol{r}, t) \tag{5}$$

where

$$\hat{T}_0 f = -\tfrac{1}{2}\boldsymbol{\nabla}^2 f \tag{5a}$$

$$\hat{T}_1 f = \boldsymbol{\nabla} \cdot (\boldsymbol{F}(\boldsymbol{r})f) = (\boldsymbol{\nabla} \cdot \boldsymbol{F})f + \boldsymbol{F} \cdot \boldsymbol{\nabla} f \tag{5b}$$

(the operation range of $\boldsymbol{\nabla}$ is indicated by parentheses) and

$$\hat{T}_2 f = (E(\boldsymbol{r}) - E_0)f. \tag{5c}$$

One can easily find the individual Green functions of the following three equations

$$-\partial f/\partial t = \hat{T}_i f \qquad i = 0, 1, 2 \tag{6}$$

to be

$$G_0(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t) = (2\pi t)^{-N/2} \exp[-(\boldsymbol{r}' - \boldsymbol{r})^2/2t] \tag{6a}$$

$$G_1(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t) = \delta(\boldsymbol{r}' - \boldsymbol{r}(t)) \tag{6b}$$

and

$$G_2(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t) = \exp[-(E(\boldsymbol{r}) - E_0)t]\delta(\boldsymbol{r}' - \boldsymbol{r}) \tag{6c}$$

respectively. $\boldsymbol{r}(t)$ in (6*b*) represents a solution to

$$\mathrm{d}\boldsymbol{r}(t)/\mathrm{d}t = \boldsymbol{F}(\boldsymbol{r}(t)) \tag{7}$$

subject to

$$\boldsymbol{r}(0) = \boldsymbol{r} \tag{7a}$$

(i.e. $\boldsymbol{r}(t)$ is a rather non-trivial function of $\boldsymbol{r}$ and t).

For a specific value of t, each of our Green functions corresponds to a linear operator defined by

$$\hat{G}(t)g(r) = \int G(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t)g(r)\,\mathrm{d}\boldsymbol{r}|_{\boldsymbol{r}'=\boldsymbol{r}} \tag{8}$$

($g(\boldsymbol{r})$ is an arbitrary function, the integration is done over the whole N-dimensional space). Thus, if $f(\boldsymbol{r}, t_0)$ is a solution to (3) at a time t_0, $\hat{G}(t)f(\boldsymbol{r}, t_0)$ will provide the

corresponding solution at t_0+t. Also, if $f(\boldsymbol{r}, \infty)$ is the stationary solution to (3),

$$\hat{G}(t)f = f. \tag{9}$$

As mentioned, the exact $G(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t)$ (and, therefore, $\hat{G}(t)$) is not known, and we have to resort to some convenient approximation of the same. One possibility is suggested by the following equation

$$\hat{G}(t) = \hat{G}_2(t) * \hat{G}_1(t) * \hat{G}_0(t) + \hat{O}(t^2) \tag{10}$$

where $*$ denotes a convolution of two operators, defined in terms of the corresponding Green function by

$$\hat{G}_1(t) * \hat{G}_0(t) \Leftrightarrow \int G_1(\boldsymbol{r}' \leftarrow \boldsymbol{r}'', t) G_0(\boldsymbol{r}'' \leftarrow \boldsymbol{r}, t)\, \mathrm{d}\boldsymbol{r}'' \tag{11}$$

etc, and $\hat{O}(t^2)$ represents an (undetermined) operator in the second and higher powers of t. Relation (10) can be derived by writing

$$\begin{aligned}\hat{G}(t) &= \exp[-t(\hat{T}-E_0)] = \exp[-t(\hat{T}_0+\hat{T}_1+\hat{T}_2)] \\ &= \hat{I} - t(\hat{T}_0+\hat{T}_1+\hat{T}_2) + t^2/2(\hat{T}_0^2+\hat{T}_1^2+\hat{T}_2^2) + t^2/2(\hat{T}_0*\hat{T}_1+\hat{T}_1*\hat{T}_0 \\ &\quad + \hat{T}_0*\hat{T}_2+\hat{T}_2*\hat{T}_0+\hat{T}_1*\hat{T}_2+\hat{T}_2*\hat{T}_1) - \ldots \end{aligned} \tag{12a}$$

and

$$\begin{aligned}&\hat{G}_2(t) * \hat{G}_1(t) * \hat{G}_0(t) \\ &\quad = \exp(-t\hat{T}_2) * \exp(-t\hat{T}_1) * \exp(-t\hat{T}_0) \\ &\quad = \hat{I} - t(\hat{T}_0+\hat{T}_1+\hat{T}_2) + t^2/2(\hat{T}_0^2+\hat{T}_1^2+\hat{T}_2^2) \\ &\qquad + t^2(\hat{T}_1*\hat{T}_0+\hat{T}_2*\hat{T}_0+\hat{T}_2*\hat{T}_1) - \cdots \end{aligned} \tag{12b}$$

where $\hat{I}$ is the identity operator (Grimm *et al* 1969).

Simulating the action of $\hat{G}_2(t) * \hat{G}_1(t) * \hat{G}_0(t)$ on the M configurations of the Monte Carlo procedure can be done in the following fashion (Kalos *et al* 1974, Anderson 1980):

(a) Simulate $\hat{G}_0(t)$: To each configuration ${}^m\boldsymbol{r}$, add a random vector with N components generated, independently, from the normal distribution with zero mean and the variance equal to t, in correspondence with (6*a*). It is not difficult to see that any finite statistical distribution with the above mean and variance can simulate the effect of $\hat{G}_0(t)$ with the required first-order-of-t accuracy, and can be used instead.

(b) Simulate $\hat{G}_1(t)$: Advance the new value of ${}^m\boldsymbol{r}$ (${}^m\boldsymbol{r}_0$ say) to ${}^m\boldsymbol{r}_0(t)$, in correspondence with (6*b*). This has to be done by a numerical integration of (7), preserving the t-accuracy of the algorithm. It turns out to be sufficient (as shown in § 3) to use

$${}^m\boldsymbol{r}_0(t) \approx {}^m\boldsymbol{r}_0 + t\boldsymbol{F}({}^m\boldsymbol{r}_0) \overset{\text{def}}{=} {}^m\boldsymbol{r}_1. \tag{13}$$

(c) Simulate $\hat{G}_2(t)$: Recall (6*c*) that

$$\hat{G}_2(t)g(\boldsymbol{r}, t_0) = \exp[-t(E(\boldsymbol{r})-E_0)]g(\boldsymbol{r}, t_0) \tag{14}$$

where $g(\boldsymbol{r}, t_0) = \hat{G}_1(t) * \hat{G}_0(t) f(\boldsymbol{r}, t_0)$.

The right-hand side of (14) remains a properly normalised probability density function, i.e.

$$\int \exp[-t(E(\boldsymbol{r})-E_0)]g(\boldsymbol{r}, t_0)\, \mathrm{d}\boldsymbol{r} = 1, \tag{15}$$

only for one particular value of E_0. Since the set of M configurations ${}^m\boldsymbol{r}_1$ ($m = 1, 2, \ldots, M$) represents a sample from $g(\boldsymbol{r}, t_0)$, the proper way of converting it to a sample from $\exp[-t(E(\boldsymbol{r}) - E_0)]g(\boldsymbol{r}, t_0)$ is to increase (decrease) the multiplicity of each configuration by the factor

$$B_m = \exp[-t(E({}^m\boldsymbol{r}_1) - E_0)] \qquad m = 1, 2, \ldots, M. \tag{16}$$

The only value of E_0 which leaves the total multiplicity of the M configurations unchanged is

$$E_0 = -\ln\left(\sum_{m=1}^{M} \exp[-tE({}^m\boldsymbol{r}_1)]/M\right)t^{-1} \tag{17}$$

and, as such, should be used in (16) (Grimm *et al* 1971).

Since B_m can, in general, have non-integer values, it is more convenient to use instead (Reynolds *et al* 1982)

$$M_m = \text{int}(B_m + \xi_m) \qquad m = 1, 2, \ldots, M \tag{18}$$

where the ξ_m's represent random variables, uniform over (0, 1), and int implies truncation of the argument to the nearest smaller integer. Because the expected value of each M_m equals B_m, the resulting distribution of configurations will be a random sample from (14). Simulating integer multiplicities M_m is done simply by deleting ${}^m\boldsymbol{r}_1$ from the sample when $M_m = 0$, retaining it when $M_m = 1$, duplicating it when $M_m = 2$, etc.

Note that simulating non-integer multiplicities B_m by carrying a statistical weight with each configuration would result in a statistically unstable process. A similar instability is also observed when using the M_m values, due to the fact that the exact number of new configurations can still vary from M. This needs to be corrected (after each time step) by adjusting the number of configurations to M, using a random (discrete, uniform) deletion or duplication of configurations (Reynolds *et al* 1982).

The above procedure can be repeated arbitrarily often, thus simulating an approximate solution to (3) (with a changing value of E_0, to meet condition (15)). The stationary solution will be reached in sufficiently many time steps; from then on one will be simulating a time-independent approximate solution to the original equation (1). Collecting individual estimates of E_0 (17) over many (usually thousands of) time steps will reduce the statistical error of the overall estimate (grand mean) to any desired level.

To investigate the nature of the error introduced by the $\hat{O}(t^2)$ term of (10), one must first realise that the described technique solves, in effect, the following eigenvalue problem

$$\hat{G}_2(t) * \hat{G}_1(t) * \hat{G}_0(t) f(\boldsymbol{r}, \infty) = f(\boldsymbol{r}, \infty) \tag{19}$$

instead of the original (1). However (19) can be rewritten (see (12*b*)) as

$$f - t(\hat{T}_0 + \hat{T}_1 + \hat{T}_2)f + \hat{O}(t^2)f = f \tag{20}$$

which is equivalent to

$$(\hat{T}_0 + \hat{T}_1 + \hat{T}_2)f - \hat{O}'(t)f = 0 \tag{21}$$

where $\hat{O}'(t) = \hat{O}(t^2)/t$ is an operator in the first and higher powers of t. Using a simple

perturbation-theory argument, one can show that the relationship between the eigenvalues of (19) and (1) is

$$E_0(\text{equation (19)}) = E_0(\text{equation (1)}) + tE_0^{(1)} + t^2E_0^{(2)} + \ldots \tag{22}$$

where $E_0^{(1)}$ and $E_0^{(2)}$ are coefficients of the perturbation expansion. Thus, to obtain an unbiased estimate of E_0 (equation (1)), one needs to get a set of estimates of E_0 (equation (19)) for several (small) values of t, then find the intercept of a polynomial least-square fit to (22) (Reynolds *et al* 1982). The interesting statistical issue of optimising the selection of these t values will not be pursued here.

3. t^2-accurate Green function

The main goal of this section is modifying the simulation algorithm so as to dispose of the $tE_0^{(1)}$ term in (22), thus making the least-square fit easier and more accurate (it can be shown on statistical grounds that the resulting improvement in efficiency is dramatic.

As a starting point we use the results of Grimm *et al* (1969) and write

$$\hat{G}(t) = \hat{G}_2(t/2) * \hat{G}_1(t/2) * \hat{G}_0(t) * \hat{G}_1(t/2) * \hat{G}_2(t/2) + \hat{O}(t^3) \tag{23}$$

where $\hat{O}(t^3)$ is an operator in the third and higher powers of t (correctness of (23) can be demonstrated easily by the appropriate modification of (12*b*)).

However, the simulation of the convolution of operators on the right-hand side of (23) to the required t^2 accuracy would require three evaluations of $\boldsymbol{F}$ per time step. Since evaluating $\boldsymbol{F}$ is the costliest part of the actual computation, it is necessary to modify (23) further (note that the algorithm of § 2 required only one evaluation of $\boldsymbol{F}$ per time step).

As it is $\hat{G}_1(t)$ which calls for evaluating $\boldsymbol{F}$, we have to concentrate on the operator $\hat{G}_1(t/2) * \hat{G}_0(t) * \hat{G}_1(t/2) \overset{\text{def}}{=} \hat{G}_s(t)$. Rewriting $\hat{G}_s(t)$ in terms of the corresponding Green function gives us

$$G_s(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t) = (2\pi t)^{-N/2} \exp\left(\frac{-[\boldsymbol{r}'(-t/2) - \boldsymbol{r}(t/2)]^2}{2t}\right) J(\boldsymbol{r}' \rightarrow \boldsymbol{r}'(-t/2)). \tag{24}$$

This ensues from two simple integrations and the fact that

$$\delta(\boldsymbol{r}''(t/2) - \boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}'' = \delta(\boldsymbol{r}'' - \boldsymbol{r}'(-t/2)) J(\boldsymbol{r}' \rightarrow \boldsymbol{r}'(-t/2))\,\mathrm{d}\boldsymbol{r}'',$$

where $J(\boldsymbol{r}' \rightarrow \boldsymbol{r}'(-t/2))$ is the Jacobian of the transformation in parentheses, i.e., more explicitly

$$J(\boldsymbol{r}' \rightarrow \boldsymbol{r}'(-t/2)) = \det(\boldsymbol{\nabla}' \otimes \boldsymbol{r}'(-t/2)). \tag{25}$$

By $\boldsymbol{\nabla}' \otimes \boldsymbol{r}'(-t/2)$ we mean an N by N matrix of all spacial derivatives (with respect to $\boldsymbol{r}'$) of $\boldsymbol{r}'(-t/2)$, det stands for the determinant of the argument. Since

$$\boldsymbol{r}(t/2) = \boldsymbol{r} + t/2\boldsymbol{F}(\boldsymbol{r}) + \tfrac{1}{8}t^2\boldsymbol{F}(\boldsymbol{r}) \cdot \boldsymbol{\nabla}\boldsymbol{F}(\boldsymbol{r}) + \ldots \tag{26a}$$

(see (7)) and, similarly,

$$\boldsymbol{r}'(-t/2) = \boldsymbol{r}' - t/2\boldsymbol{F}(\boldsymbol{r}') + \tfrac{1}{8}t^2\boldsymbol{F}(\boldsymbol{r}') \cdot \boldsymbol{\nabla}'\boldsymbol{F}(\boldsymbol{r}') + \ldots \tag{26b}$$

equation (24) can be rewritten (using the shorthand notation $\boldsymbol{F} \equiv \boldsymbol{F}(\boldsymbol{r})$ and $\boldsymbol{F}' \equiv \boldsymbol{F}(\boldsymbol{r}')$) as

$$G_s(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t) = (2\pi t)^{-N/2} \exp\left(-\frac{[\boldsymbol{r}' - \boldsymbol{r} - t/2(\boldsymbol{F} + \boldsymbol{F}') + t^2/8(\boldsymbol{F}' \cdot \boldsymbol{\nabla}' \boldsymbol{F}' - \boldsymbol{F} \cdot \boldsymbol{\nabla} \boldsymbol{F}) + \ldots]^2}{2t} \right)$$
$$\times \det(I - t/2 \boldsymbol{\nabla}' \otimes \boldsymbol{F}' + t^2/8 \boldsymbol{\nabla}' \otimes (\boldsymbol{F}' \cdot \boldsymbol{\nabla}' \boldsymbol{F}') + \ldots) \tag{27}$$

where I is the unit matrix.

Here, we have neglected terms contributing to the third and higher powers of t in the $\hat{G}_s(t)$ expansion. To justify this statement, let us investigate the exact relationship between an expansion of a Green function $G_s(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t)$, and the t-expansion of the corresponding operator $\hat{G}_s(t)$. To this end, we define $G_R(\boldsymbol{r}, t; \boldsymbol{r}')$ by

$$G_s(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t) \overset{\text{def}}{=} G_0(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t) G_R(\boldsymbol{r}, t; \boldsymbol{r}') \tag{28}$$

with $G_0(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t)$ of (6*a*). When $G_R(\boldsymbol{r}, t; \boldsymbol{r}')$ is expanded as a generalised power series in t (with respect to zero) and $\boldsymbol{r}$ (with respect to $\boldsymbol{r}'$), one can easily determine the contribution of each individual term of this expansion to $\hat{G}_s(t)$. Only nine of such terms (summarised in table 1) contribute within the t^2 accuracy of the $\hat{G}_s(t)$ operator. The contribution of each of them can be determined from

$$\hat{G}_s(t) g(\boldsymbol{r}) = \int G_0(\boldsymbol{r}' \leftarrow \boldsymbol{r}, t) G_R(\boldsymbol{r}, t; \boldsymbol{r}') g(\boldsymbol{r}) \, \mathrm{d}\boldsymbol{r}|_{\boldsymbol{r}'=\boldsymbol{r}}$$
$$= \hat{G}_0(t) G_R(\boldsymbol{r}, t; \boldsymbol{r}') g(\boldsymbol{r})|_{\boldsymbol{r}'=\boldsymbol{r}}$$
$$= (\hat{I} + t/2\boldsymbol{\nabla}^2 + t^2/8(\boldsymbol{\nabla}^2)^2 + \ldots) G_R(\boldsymbol{r}, t; \boldsymbol{r}') g(\boldsymbol{r})|_{\boldsymbol{r}'=\boldsymbol{r}} \tag{29}$$

and the following standard formulae for differentiating a product of two functions:

$$\boldsymbol{\nabla}^2(g(\boldsymbol{r})h(\boldsymbol{r})) = (\boldsymbol{\nabla}^2 g)h + 2(\boldsymbol{\nabla} g) \cdot (\boldsymbol{\nabla} h) + g(\boldsymbol{\nabla}^2 h) \tag{30a}$$

$$\nabla_i \boldsymbol{\nabla}^2(gh) = (\nabla_i \boldsymbol{\nabla}^2 g)h + (\boldsymbol{\nabla}^2 g)(\nabla_i h) + 2(\nabla_j g)(\nabla_i \nabla_j h) + 2(\nabla_i \nabla_j g)(\nabla_j h)$$
$$+ (\nabla_i g)(\boldsymbol{\nabla}^2 h) + g(\nabla_i \boldsymbol{\nabla}^2 h) \qquad \text{for } i = 1, 2, \ldots, N \tag{30b}$$

and

$$\boldsymbol{\nabla}^2 \boldsymbol{\nabla}^2(gh) = (\boldsymbol{\nabla}^2 \boldsymbol{\nabla}^2 g)h + 4(\nabla_i \boldsymbol{\nabla}^2 g)(\nabla_i h) + 2(\boldsymbol{\nabla}^2 g)(\boldsymbol{\nabla}^2 h)$$
$$+ 4(\nabla_i \nabla_j g)(\nabla_i \nabla_j h) + 4(\nabla_i g)(\nabla_i \boldsymbol{\nabla}^2 h) + g(\boldsymbol{\nabla}^2 \boldsymbol{\nabla}^2 h) \tag{30c}$$

where summation (from 1 to N) over any duplicate index is understood.

Table 1 summarises the results.

Now, going back to (27), it is easy to see that the last term in the square brackets of the exponent will appear, in the $G_R(\boldsymbol{r}, t; \boldsymbol{r}')$ expansion, as

$$-t/8(\boldsymbol{r}' - \boldsymbol{r})(\boldsymbol{F}' \cdot \boldsymbol{\nabla}' \boldsymbol{F}' - \boldsymbol{F} \cdot \boldsymbol{\nabla} \boldsymbol{F}) = -t/8 A_{ij}(\boldsymbol{r}')(r' - r)_i (r' - r)_j + \ldots \tag{31}$$

with $A_{ij}(\boldsymbol{r}') = \nabla'_i (F'_k \nabla'_k F'_j)$. Deleting this term from the Green function requires an automatic deletion of

$$t^2/8 \boldsymbol{\nabla}' \otimes (\boldsymbol{F}' \cdot \boldsymbol{\nabla}' \boldsymbol{F}') \tag{32}$$

from the determinant of (27) (as Jacobian of the corresponding transformation, it will always consist of all spatial derivatives $\boldsymbol{\nabla}'$ of the square-bracket vector). Note that the

Table 1.

Term of the $G_R(r, t; r')$ expansion	Contribution to $\hat{G}_s(t)$
$A(\boldsymbol{r}')$	$A(\boldsymbol{r})[\hat{I}+t/2\nabla^2+t^2/8(\nabla^2)^2+\ldots]$
$A_i(\boldsymbol{r}')(r-r')_i$	$A_i(\boldsymbol{r})(t\nabla_i+t^2/2\nabla_i\nabla^2+\ldots)$
$A_{ij}(\boldsymbol{r}')(r-r')_i(r-r')_j$	$A_{ij}(\boldsymbol{r})[t\delta_{ij}+t^2/2(\delta_{ij}\nabla^2+2\nabla_i\nabla_j)+\ldots]$
$A_{ijk}(\boldsymbol{r}')(r-r')_i(r-r')_j(r-r')_k$	$A_{ijk}(\boldsymbol{r})t^2(\delta_{ij}\nabla_k+\delta_{ik}\nabla_j+\delta_{jk}\nabla_i)+\ldots$
$A_{ijkl}(\boldsymbol{r}')(r-r')_i(r-r')_j(r-r')_k(r-r')_l$	$A_{ijkl}(\boldsymbol{r})t^2(\delta_{ij}\delta_{kl}+\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})+\ldots$
$tA(\boldsymbol{r}')$	$tA(\boldsymbol{r})(\hat{I}+t/2\nabla^2+\ldots)$
$tA_i(\boldsymbol{r}')(r-r')_i$	$tA_i(\boldsymbol{r})\nabla_i$
$tA_{ij}(\boldsymbol{r}')(r-r')_i(r-r)_j$	$tA_{ij}(\boldsymbol{r})$
$t^2A(\boldsymbol{r}')$	$t^2A(\boldsymbol{r})$

contribution of the matrix (32) to $\hat{G}_s(t)$ effectively equals to its trace

$$t^2/8\nabla'\cdot(\boldsymbol{F}'\cdot\nabla'\boldsymbol{F}')+\ldots. \tag{33}$$

A quick look at table 1 reveals that (31) cancels against (33) within the t^2 accuracy of $\hat{G}_s(t)$. The new t^2 accurate version of $G_s(\boldsymbol{r}'\leftarrow\boldsymbol{r}, t)$ is therefore

$$(2\pi t)^{-N/2}\exp\left(-\frac{[\boldsymbol{r}'-\boldsymbol{r}-t/2(\boldsymbol{F}+\boldsymbol{F}')]^2}{2t}\right)\det(I-t/2\nabla'\otimes F'). \tag{34}$$

Expression (34) represents a probability density function of the random variable $\boldsymbol{r}'$, defined by

$$\boldsymbol{n}=\boldsymbol{r}'-\boldsymbol{r}-\tfrac{1}{2}t(\boldsymbol{F}(\boldsymbol{r})+\boldsymbol{F}(\boldsymbol{r}')) \tag{35}$$

where $\boldsymbol{n}$ corresponds to N-dimensional normal distribution with mean zero and the variance-covariance matrix equal to $t\cdot\boldsymbol{I}$ (and as such can be easily simulated).

Solving (35) for $\boldsymbol{r}'$ (discarding contributions beyond table 1) results in

$$\boldsymbol{r}'=\boldsymbol{r}+\boldsymbol{n}-\tfrac{1}{2}t[\boldsymbol{F}(\boldsymbol{r})+F(\boldsymbol{r}+\boldsymbol{n}+t\boldsymbol{F}(\boldsymbol{r}))]+\ldots \tag{36}$$

which clearly suggests the way of simulating $\boldsymbol{r}'$. Since $\boldsymbol{F}(\boldsymbol{r}+\boldsymbol{n}+t\boldsymbol{F}(\boldsymbol{r}))$ equals, within the required accuracy, $\boldsymbol{F}(\boldsymbol{r}')$, it can be used as $\boldsymbol{F}(\boldsymbol{r})$ of the next time step.

Thus, the complete algorithm to simulate

$$\hat{G}_2(\tfrac{1}{2}t)*\hat{G}_s(t)*\hat{G}_2(\tfrac{1}{2}t) \tag{37}$$

requires, in each time step:

(a) Advancing each configuration ${}^m\boldsymbol{r}$ by

(i) $t\cdot{}^m\boldsymbol{F}$, with ${}^m\boldsymbol{F}$ as evaluated in the *previous* time step

(ii) ${}^m\boldsymbol{n}(0, t)$, a random vector with independent components, normally distributed with zero mean and standard deviation equal to $t^{1/2}$ (any symmetric distribution with the same first four moments can be used instead, as it correctly simulates the operator $\hat{I}+\frac{1}{2}t\nabla^2+t^2/8(\nabla^2)^2+\ldots$)

(iii) $\frac{1}{2}t({}^m\boldsymbol{F}'-{}^m\boldsymbol{F})$, where ${}^m\boldsymbol{F}'$ is, for each configuration, evaluated at the new location (end of step (ii)).

This completes the simulation of $\hat{G}_s(t)$.

(b) Deleting, retaining, or duplicating each new configuration ${}^m\boldsymbol{r}'$ (end of step (iii)) according to

$$\text{int}\{\exp\{-t[(E({}^m\boldsymbol{r})+E({}^m\boldsymbol{r}'))/2-E_0]\}+{}^mu\} \tag{38}$$

where

$$E_0=-t^{-1}\ln\left\{\sum_{m=1}^{M}\exp\{-\tfrac{1}{2}t(E({}^m\boldsymbol{r})+E({}^m\boldsymbol{r}'))\}/M\right\} \tag{39}$$

and each ${}^mu$ is a random variable uniform over (0, 1).

As discussed in the previous section, it is necessary to statistically stabilise the simulation process by a subsequent random adjustment of the total number of configurations to M.

Repeating the time steps until the stationary solution is reached, and then collecting sufficiently many individual estimates of E_0 (equation (39)) will provide an accurate grand-mean estimate of the eigenvalue of

$$\hat{G}_2(t/2)*\hat{G}_s(t)*\hat{G}_2(t/2)\cdot f(\boldsymbol{r},\infty)=f(\boldsymbol{r},\infty). \tag{40}$$

By means of a simple perturbation-theory argument one can prove that this eingenvalue has the following time-step dependence:

$$E_0\,(\text{equation (40)})=E_0\,(\text{equation (1)})+t^2E_0^{(2)}+\ldots. \tag{41}$$

Reapplying the complete procedure with several different values of t enables us to find E_0 (equation (1)) as an intercept of a least-square fit to (41).

One should also point out that any discontinuity in $\boldsymbol{F}(\boldsymbol{r})$ will introduce an additional error (in terms of equation (41)) of the first order in t. This can be corrected by slightly modifying $\boldsymbol{F}(\boldsymbol{r})$ into a continuous function. Care should be taken to ensure that such modification represents only a t^2 perturbation of equation (1).

4. Concluding remarks

As declared in the introduction, the aim of the paper was a search for an efficient Monte Carlo procedure to solve (1). This has been achieved by developing an algorithm which, in effect, solves (3) instead. As such, the resulting technique can be applied to a variety of physical problems, not necessarily related to finding an eigenvalue of the Schrödinger equation. To this end, the simulation algorithm can be modified to allow $\hat{T}_0$ to have the more general form of $-\frac{1}{2}\nabla\cdot D(\boldsymbol{r})\nabla$ ($D(\boldsymbol{r})$ being a given function of $\boldsymbol{r}$). Also, proper treatment of various boundary conditions may be easily incorporated into the procedure.

Acknowledgment

The author would like to thank S M Rothstein for initiating this work, and helping to improve the manuscript.

References

Anderson J B 1980 *J. Chem. Phys.* **73** 3897
Ceperley D 1978 *Phys. Rev.* B **18** 3126
—— 1983 *J. Comput. Phys.* **51** 404
Ceperley D M and Adler B J 1980 *Phys. Rev. Lett.* **45** 566
Grimm R and Storer R G 1969 *J. Comput. Phys.* **4** 230
—— 1971 *J. Comput. Phys.* **7** 134
Kalos M H 1962 *Phys. Rev.* **128** 1791
Kalos M H, Levesque D and Verlet L 1974 *Phys. Rev.* A **9** 2178
Reynolds P J, Ceperley D M, Adler B J and Lester W A Jr 1982 *J. Chem. Phys.* **77** 5593