Complex Langevin equation and the many-fermion problem

Chris Adami* and Steven E. Koonin

W. K. Kellogg Radiation Laboratory, California Institute of Technology, Pasadena, California 91125
(Received 19 October 2000; published 20 February 2001)

We study the utility of a complex Langevin (CL) equation as an alternative for the Monte Carlo (MC) procedure in the evaluation of expectation values occurring in fermionic many-body problems. We find that a CL approach is natural in cases where nonpositive definite probability measures occur, and remains accurate even when the corresponding MC calculation develops a severe ''sign problem.'' While the convergence of CL averages cannot be guaranteed in principle, we show how convergent results can be obtained in two simple quantum mechanical models, as well as a nontrivial schematic shell model path integral with multiple particles and a noncommuting interaction (the Lipkin model).

I. INTRODUCTION

There has been recent significant progress in the large-scale numerical computation of nuclear properties in the shell model, using the auxiliary field path-integral (AFPI) Monte Carlo method [1]. Due to the benign scaling of the computational effort with the single-particle basis, properties of large nuclei can now be calculated that are out of the reach of conventional diagonalization methods [2,3].

Quite generally, AFPI treatments of interacting fermion systems with the Monte Carlo (MC) method are difficult for certain realistic Hamiltonians, and for odd-particle configurations.[1] This difficulty, also known as the ''sign problem,'' is the prime impediment to large scale computational efforts both in nuclear and condensed matter physics calculations. Briefly, repulsive interactions and/or odd-particle configurations can lead to probability distributions (integration measures for the auxiliary fields) that are negative, or even complex. As the MC update algorithm relies on a positive-definite measure, the sign of the distribution is made part of the observable being calculated. Under circumstances where the average sign of the distribution is small, the expectation value is the ratio of two very small numbers, that converges only asymptotically.

A number of alternatives to or extensions of the MC method have been proposed over the years, among them hybrid methods combining the MC method with replication mechanisms for importance sampling [5], random-walk branching [6], and diagonalization over optimal bases using variational techniques [7]. Here, we investigate the possibility of replacing the MC method altogether with one based on the complex Langevin (CL) equation, at least in those cases where the sign problem is prominent. The CL equation has received considerable attention in connection with lattice gauge theory calculations, where either static charges or a nonzero chemical potential give rise to complex actions [8]. It has been abandoned mostly because of the perception that complex Langevin averages ought not to be trusted, due to

the possibility that they can be nonstationary. Here, we show that the problem of nonstationarity can often be avoided, and does not play a role in a large class of Hamiltonians which give rise to Langevin equations with fixed points in the complex plane.

By way of introduction and to establish notation, we briefly review the AFPI Monte Carlo method in the next section. Section III discusses the application of the CL method to simple one-dimensional integrals abstracted from those occurring in fermionic many-body systems, while Sec. IV expands this to a simple toy Hamiltonian with characteristics reminiscent of shell models. In Sec. V we apply the method to the Lipkin model (the MC treatment of which is similar in character to the full shell model), and close with conclusions, in Sec. VI.

II. AUXILIARY FIELD PATH-INTEGRAL MONTE CARLO

In the AFPI method, the significant savings in computational effort are obtained through a Hubbard-Stratonovich (HS) transformation [9] (see below) of the imaginary-time evolution operator

$$\hat{U} = \exp(-\beta\hat{H}). \tag{2.1}$$

The thermal expectation value of an operator $\hat{O}$ is given by

$$\langle\hat{O}\rangle_\beta = Z^{-1}\,\hat{\mathrm{Tr}}[\hat{O}\exp(-\beta\hat{H})], \tag{2.2}$$

where

$$Z = \hat{\mathrm{Tr}}\exp(-\beta\hat{H}) \tag{2.3}$$

is the partition function, $\hat{\mathrm{Tr}}$ is the many-body trace, and β is the inverse temperature. Ground-state (zero-temperature) properties are obtained in the limit $\beta\to\infty$.

For a many-body operator $\hat{H}$ in the quadratic form

$$\hat{H} = \sum_\alpha \epsilon_\alpha\hat{\mathcal{O}}_\alpha + \frac{1}{2}\sum_\alpha V_\alpha\hat{\mathcal{O}}_\alpha^2 \tag{2.4}$$

---

*Corresponding author. Electronic address: adami@caltech.edu
Phone (+) 626 395-4256. Fax (+) 626 564-8708.
[1]See for example [4].

containing one-body operators $\hat{\mathcal{O}}_\alpha$ and two-body operators $\hat{\mathcal{O}}_\alpha^2$, a HS transformation of the imaginary-time evolution operator (2.1) leads to

$$\exp(-\beta\hat{H}) \approx \int_{-\infty}^{\infty} \prod_{\alpha,n} d\sigma_{\alpha,n} \left(\frac{\Delta\beta|V_\alpha|}{2\pi}\right)^{1/2}$$

$$\times G(\sigma) \prod_n \exp[-\Delta\beta\hat{h}(\vec{\sigma}_n)], \quad (2.5)$$

by splitting β into N_t time slices such that $\Delta\beta = \beta/N_t$.

Here, $\vec{\sigma}_n$ denotes a set of auxiliary fields [one for each two-body operator appearing in Eq. (2.4)] at time-slice n, and σ stands for the totality of fields at all time-slices. Furthermore, $G(\sigma)$ is the Gaussian weight factor

$$G(\sigma) = \exp\left(-\sum_{\alpha,n} \frac{1}{2}\Delta\beta|V_\alpha|\sigma_{\alpha n}^2\right), \quad (2.6)$$

and $\hat{h}(\vec{\sigma}_n)$ is the *one-body* Hamiltonian

$$\hat{h}(\vec{\sigma}_n) = \sum_\alpha (\epsilon_\alpha + s_\alpha V_\alpha \sigma_{\alpha n})\hat{\mathcal{O}}_\alpha, \quad (2.7)$$

where $s_\alpha = \pm 1(=\pm i)$ if $V_\alpha < 0(>0)$. Thus, the HS transformation has the effect of replacing the quadratic dependence on $\hat{\mathcal{O}}_\alpha$ in Eq. (2.4) with a linear one, at the expense of an integral over auxiliary fields.

Let $\hat{U}_\sigma$ denote the *one-body* evolution operator

$$\hat{U}_\sigma = \exp[-\Delta\beta\hat{h}(\vec{\sigma}_n)] \quad (2.8)$$

and $F(\sigma)$ its trace:

$$F(\sigma) = \hat{\mathrm{Tr}}(\hat{U}_\sigma). \quad (2.9)$$

Expectation values can then be written using the above path-integral decomposition of the partition function:

$$\langle\hat{O}\rangle_\beta = \frac{\int \mathcal{D}[\sigma]G(\sigma)\hat{\mathrm{Tr}}(\hat{O}\hat{U}_\sigma)}{\int \mathcal{D}[\sigma]G(\sigma)F(\sigma)}. \quad (2.10)$$

An effective action S_σ can be defined such that Eq. (2.10) appears as a simple expectation value:

$$\langle\hat{O}\rangle_\beta = \frac{\int \mathcal{D}[\sigma]e^{-S_\sigma}\langle\hat{O}\rangle_\sigma}{\int \mathcal{D}[\sigma]e^{-S_\sigma}} \quad (2.11)$$

with

$$S_\sigma = \sum_{\alpha,n} \frac{1}{2}\Delta\beta|V_\alpha|\sigma_{\alpha,n}^2 - \ln F(\sigma), \quad (2.12)$$

where $\langle\hat{O}\rangle_\sigma = \hat{\mathrm{Tr}}(\hat{O}\hat{U}_\sigma)/F(\sigma)$.

Further, since $\exp(-S_\sigma)$ is not positive definite, we define the sign function

$$\Phi_\sigma = \frac{F(\sigma)}{|F(\sigma)|}. \quad (2.13)$$

Then, writing $e^{-S_\sigma} = \Phi(\sigma)e^{-\tilde{S}_\sigma}$ with a *real* $\tilde{S}_\sigma$, the expectation value (2.10) appears as a *ratio* of expectation values

$$\langle\hat{O}\rangle_\beta = \frac{\int \mathcal{D}[\sigma]e^{-\tilde{S}_\sigma}\Phi(\sigma)\langle\hat{O}\rangle_\sigma}{\int \mathcal{D}[\sigma]e^{-\tilde{S}_\sigma}\Phi(\sigma)} = \frac{\langle\!\langle\Phi_\sigma\langle\hat{O}\rangle_\sigma\rangle\!\rangle}{\langle\!\langle\Phi_\sigma\rangle\!\rangle}$$

$$(2.14)$$

each calculated with a positive definite probability distribution $e^{-\tilde{S}_\sigma}$. The Monte Carlo average over samples is denoted as $\langle\!\langle \ldots \rangle\!\rangle$. If the function $F(\sigma)$ is negative on a substantial part of the σ-field manifold, the expectation values $\langle\!\langle\Phi_\sigma\langle\hat{O}\rangle_\sigma\rangle\!\rangle$ and $\langle\!\langle\Phi_\sigma\rangle\!\rangle$ can each become very small, and the ratio only converges asymptotically. This is the essence of the sign problem in quantum Monte Carlo calculations. Below, we construct expectation values susceptible to the sign problem, in order to gain insight into the circumstances in which a CL equation approach can be applied successfully.

III. COMPLEX LANGEVIN EQUATION AND SIMPLE INTEGRALS

The HS transformation is nothing but the Gaussian integral identity

$$e^{-(1/2)z^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\sigma e^{-(1/2)\sigma^2} e^{\pm i\sigma z}. \quad (3.1)$$

Note that this case corresponds to a repulsive interaction [$V > 0$ in Eq. (2.7)]. As Eq. (3.1) has no imaginary part, this reduces to

$$e^{-(1/2)z^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\sigma e^{-(1/2)\sigma^2} \cos(z\sigma). \quad (3.2)$$

This is precisely the form appearing in the denominator of Eq. (2.10), but in one dimension, with $\cos(z\sigma)$ playing the role of $F(\sigma)$.

Generalizing this, we would like to examine expectation values

$$\langle\sigma^2\rangle_N = \frac{\int d\sigma \sigma^2 e^{-(1/2)\sigma^2}[\cos(\sigma z)]^N}{\int d\sigma e^{-(1/2)\sigma^2}[\cos(\sigma z)]^N} \quad (3.3)$$

as a function of the real number z and the ''particle number'' N, using MC evaluation and the complex Langevin equation approach.

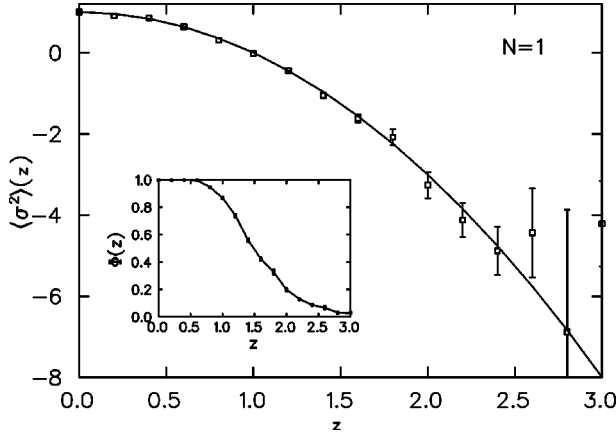Clearly, the MC procedure will suffer from the sign problem only for odd N. Figure 1 shows a straightforward

FIG. 1. Monte Carlo average of integral $\langle\sigma^2\rangle_{N=1}$ Eq. (3.3) (squares), the exact solution (solid line), and the sign of the calculation (inset).

MC evaluation of this integral for $N=1$, where the inset shows the development of the average sign $\Phi(\sigma)$ $=\cos(z\sigma)/|\cos(z\sigma)|$. As expected, the accuracy of the MC estimate deteriorates as $\Phi(\sigma)\to 0$.

As an alternative to the MC procedure, consider the Langevin equation. For systems with real actions, expectation values such as

$$\langle O\rangle=\frac{1}{Z}\int d\sigma O(\sigma)e^{-S(\sigma)} \qquad (3.4)$$

with the partition function

$$Z=\int d\sigma e^{-S(\sigma)} \qquad (3.5)$$

can be calculated by creating a stochastic process using the Langevin equation, with an equilibrium distribution $P_0(\sigma)$ $=Z^{-1}\exp[-S(\sigma)]$.

The Langevin equation is given by

$$\frac{d\sigma(t)}{dt}=-\frac{1}{2}\frac{\partial S}{\partial\sigma}+\eta(t), \qquad (3.6)$$

where t is a fictitious time and $\eta(t)$ is stochastic noise with zero mean and unit variance:

$$\langle\eta(t)\eta(t')\rangle=\delta(t-t'). \qquad (3.7)$$

To each Langevin equation corresponds a Fokker-Planck equation for the probability density $P(\sigma,t)$

$$\frac{\partial P(\sigma,t)}{\partial t}=H_{FP}P(\sigma,t) \qquad (3.8)$$

with a Hermitian Fokker-Planck Hamiltonian

$$H_{FP}=\frac{1}{2}\frac{\partial}{\partial\sigma}\left(\frac{\partial}{\partial\sigma}+\frac{\partial S}{\partial\sigma}\right). \qquad (3.9)$$

For solutions with exponential time dependence,

$$P(\sigma,t)=e^{-Et}P_E(\sigma), \qquad (3.10)$$

Eq. (3.8) reverts to an eigenvalue equation. For $t\to\infty$, $P(\sigma,t)\to P_0(\sigma)$, the solution for the lowest eigenvalue E $=0$. Accordingly, for $t\to\infty$, if $\sigma_0(t)$ is the solution to the Langevin equation (3.6),

$$O(\sigma_0(t))\to\langle O\rangle \quad (t\to\infty) \qquad (3.11)$$

with $\langle O\rangle$ given by Eq. (3.4). Finally, ergodicity assures that $\langle O\rangle$ is also obtained by averaging over the path $\sigma_0(t)$:

$$\langle O\rangle=\langle\!\langle O(\sigma_0)\rangle\!\rangle=\lim_{T\to\infty}\frac{1}{T}\int_0^T O(\sigma_0(t))dt. \qquad (3.12)$$

In principle, nothing prevents us from using the Langevin equation in the case where the action is complex:

$$S(\sigma)=S_R(\sigma)+iS_I(\sigma). \qquad (3.13)$$

Then, we obtain two equations, for the real and the imaginary part of σ:

$$\frac{\partial\sigma_R(t)}{\partial t}=-\frac{1}{2}\mathrm{Re}\left(\frac{\partial S}{\partial\sigma}\right)+\eta(t), \qquad (3.14)$$

$$\frac{\partial\sigma_I(t)}{\partial t}=-\frac{1}{2}\mathrm{Im}\left(\frac{\partial S}{\partial\sigma}\right). \qquad (3.15)$$

However, the Fokker-Planck Hamiltonian loses its Hermiticity, and the eigenvalues can acquire imaginary parts. As a consequence, the probability distribution $P(\sigma_R,\sigma_I,t)$ need not converge anymore, nor does the expectation value $\langle O\rangle$. While the lowest eigenvalue is still $E_0=0$, the E_n with n >0 are in general complex, and the asymptotic condition $P(\sigma)\to P_0$ is violated whenever there are any E_n with Re $E_n>0$ $(n>0)$ [10]. As a rule of thumb then, expectation values obtained via the CL equation should only be trusted if the ensemble averages become time independent [11].

For the numerical solution of Eqs. (3.14),(3.15) we use the two-step algorithm of Greenside and Helfand [12]. Defining the complex gradient as

$$\nabla S(t)=\frac{\partial S}{\partial\sigma}[\sigma_R(t),\sigma_I(t)], \qquad (3.16)$$

the stochastic differential equation is discretized via

$$\sigma_R(t_{1/2})=\sigma_R(t_0)-\Delta t\,\mathrm{Re}[\nabla S(t_0)]+\sqrt{\Delta t}\,\eta(t_0), \qquad (3.17)$$

$$\sigma_R(t_1)=\sigma_R(t_0)-\frac{1}{2}\Delta t\,\mathrm{Re}[\nabla S(t_0)+\nabla S(t_{1/2})]+\sqrt{\Delta t}\,\eta(t_0) \qquad (3.18)$$

and analogously for the imaginary part.[2]

For the expectation value (3.3), the effective action is

_____

[2]For this particular choice of variables, the corresponding equation for the imaginary part does not include a noise term. For other choices, see, e.g., [13].

$$S_N = \frac{1}{2}\sigma^2 - N\ln[\cos(\sigma z)], \qquad (3.19)$$

while the associated Langevin equation reads

$$\dot{\sigma} = -\frac{1}{2}[\sigma + Nz\tan(z\sigma)] + \eta. \qquad (3.20)$$

It is immediately apparent that the fixed points of this equation ($\dot{\sigma}=0$) lie on the real axis. On the other hand it is also clear that, since for example $\langle\sigma^2\rangle_{N=1} = 1 - z^2$, the correct solution $\sigma_0(t)$ needs to spend a considerable amount of time away from the real line, at least for odd N. Analytically, this must happen because (for odd N) there is a delta-function-like drift term that has been ignored in Eq. (3.20) which is due to the imaginary part of S_N. The additional drift term is

$$\delta\nabla S_N = \pm i\pi\delta(\cos(z\sigma)) \qquad (3.21)$$

but cannot be adequately modeled numerically. As a consequence, the solution spends most of its time on the real line between the first turning points, and the resulting average is inaccurate, as has been noted previously [14]. The solution to this dilemma is also not new. Going back to Eq. (3.1) and $N=1$ we see that the action can also be written as

$$S_1 = \frac{1}{2}\sigma^2 \pm iz\sigma, \qquad (3.22)$$

in which case the fixed point is away from the real line in the complex plane: $\sigma_I = \pm z$. Since then $P(\sigma_R, \sigma_I)$ $\rightarrow \exp(-\frac{1}{2}\sigma_R^2)\delta(\sigma_I\mp z)$, the complex average reduces to

$$
\langle\sigma^2\rangle_{N=1} = \frac{\displaystyle\int_{-\infty}^{\infty}\sigma^2 P(\sigma_R,\sigma_I)d\sigma_R d\sigma_I}{\displaystyle\int_{-\infty}^{\infty}P(\sigma_R,\sigma_I)d\sigma_R d\sigma_I}
$$

$$
= \frac{\displaystyle\int_{-\infty\mp iz}^{\infty\pm iz}d\sigma\,\sigma^2\exp(-\frac{1}{2}\sigma^2)}{\displaystyle\int_{-\infty\mp iz}^{\infty\pm iz}d\sigma\,\exp(-\frac{1}{2}\sigma^2)}. \qquad (3.23)
$$

The integral over the complex path removed from the real line by $\pm iz$ equals the one over the real line if the observable has no poles in the enclosed area.

In Fig. 2 we show the result of a CL evaluation of $\langle\sigma^2\rangle_{N=1}$ using the ''extended'' action (3.22) (squares) and the original action (3.19) (triangles). The solid line represents the exact result. As expected, a complex Langevin simulation with fixed points on the real line does not converge, while the calculation with the extended action is robust even in the regime where the MC (Fig. 1) fails.

Can this procedure be extended to arbitrary N? A canonical extension of the method exists, where $\cos^N(\sigma z)$ is decomposed into single powers of cosines at multiples of σz. Subsequently, the cosines are replaced by exponentials. However, this procedure results in a shift of the fixed point away from the real line *only* for odd N. Figure 3 shows results for the cases $N=2$ and $N=3$, with exact results
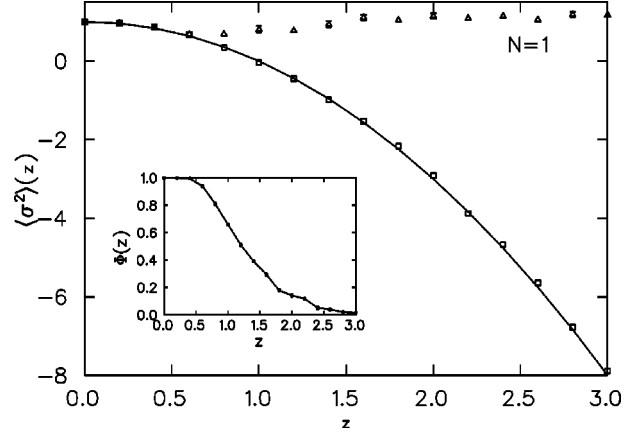

FIG. 2. Langevin average of $\langle\sigma^2\rangle_{N=1}$. Squares: extended action; triangles: original action; solid line: exact solution.

$$\langle\sigma^2\rangle_{N=2} = 1 - \frac{4z^2}{1+e^{2z^2}}, \qquad (3.24)$$

$$\langle\sigma^2\rangle_{N=3} = 1 - z^2\frac{1+3e^{-4z^2}}{1+\frac{1}{3}e^{-4z^2}}. \qquad (3.25)$$

For even N the complex Langevin equation does not converge at all z due to the problems described. We do not need to worry about this, however, as the MC procedure is very accurate there.

IV. QUANTUM MECHANICAL TOY MODEL

In this section we show how shifting the fixed points in a CL evaluation of integrals can be used in a quantum mechanical model which describes a single shell of angular momentum j with variable filling. This toy model is defined by the Hamiltonian

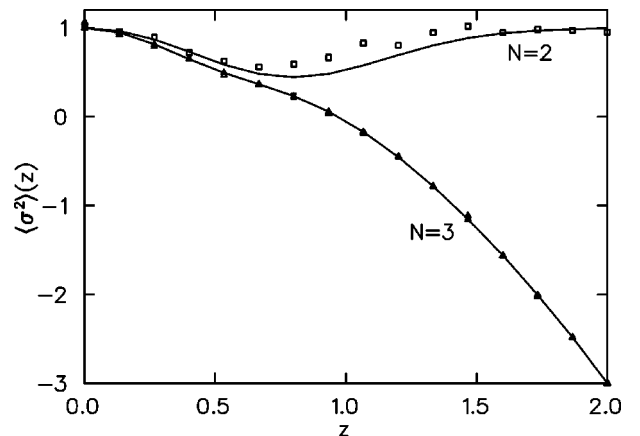$$\hat{H} = -\frac{1}{2}\epsilon\hat{N}^2 + \frac{1}{2}V\hat{J}_z^2, \qquad (4.1)$$

FIG. 3. Complex Langevin average of $\langle\sigma^2\rangle_N$ for $N=2$ and $N=3$.

where $\hat{N}$ is the number operator and $\hat{J}_z$ the third component of the angular momentum. A HS transformation on the imaginary-time evolution operator yields

$$\hat{\mathrm{Tr}}_N(e^{-\beta\hat{H}})$$
$$\propto \int d\sigma_0 d\sigma_1 e^{-(1/2)(\sigma_0^2+\sigma_1^2)} \hat{\mathrm{Tr}}_N(e^{\sqrt{\beta\epsilon}\hat{N}\sigma_0 - i\sqrt{\beta V}\hat{J}_z\sigma_1}),$$
(4.2)

where $\hat{\mathrm{Tr}}_N$ is the many-body trace for fixed particle number N and angular momentum j:

$$\hat{\mathrm{Tr}}_N(\hat{O}) = \sum_{\substack{m_1 + \ldots + m_N = -j \\ m_i \neq m_j}}^{j} \langle m_1 \ldots m_N | \hat{O} | m_1 \ldots m_N \rangle.$$
(4.3)

Since $\hat{N}$ and $\hat{J}_z$ commute, the evolution operator does not need to be decomposed into time-slices. Also, the contribution from the number operator is a constant factor

$$\hat{\mathrm{Tr}}_N(e^{-\beta\hat{H}}) = \exp(\sqrt{N\beta\epsilon}) \hat{\mathrm{Tr}}_N(U_{\sigma_1})$$
(4.4)

that drops out of the ratios, and will thus be ignored in the following. Above, we defined the one-body evolution operator

$$U_{\sigma_1} = e^{-i\sqrt{\beta V}\hat{J}_z\sigma_1}.$$
(4.5)

Defining $\phi = \sqrt{\beta V}\sigma_1$ and as before

$$F_N(\phi) = \hat{\mathrm{Tr}}_N(U_{\sigma_1}),$$
(4.6)

we can write the N-particle traces in terms of the one-particle trace:

$$F_1(\phi) = F(\phi) = \frac{\sin(j + \frac{1}{2})\phi}{\sin(\phi/2)},$$
(4.7)

$$F_2(\phi) = F^2(\phi) - F(2\phi),$$
(4.8)

$$F_3(\phi) = F^3(\phi) - 3F(2\phi)F(\phi) + 2F(3\phi),$$
(4.9)

and so on. Note that unlike in the previous section, the even-N trace is not positive-definite, while still being mostly positive.

We shall focus on the expectation value[3]

$$\langle J_z^2 \rangle_N = \frac{\int d\sigma e^{-(1/2)\sigma^2} \hat{\mathrm{Tr}}_N(\hat{J}_z^2 U_\sigma)}{\int d\sigma e^{-(1/2)\sigma^2} F_N(\sigma)}$$
(4.10)

which we rewrite in terms of an effective action as follows:

_____

[3]Note that we now write σ instead of σ_1 for simplicity.

$$\langle \hat{J}_z^2 \rangle = \frac{\int d\sigma e^{-(1/2)\sigma^2} \langle \hat{J}_z^2 \rangle_\sigma e^{-S_N(\sigma)}}{\int d\sigma e^{-(1/2)\sigma^2} e^{-S_N(\sigma)}},$$
(4.11)

where

$$\langle \hat{J}_z^2 \rangle_\sigma = \frac{\hat{\mathrm{Tr}}_N(\hat{J}_z^2 U_\beta)}{F_N(\sigma)}$$
(4.12)

and

$$S_N = \frac{1}{2}\sigma^2 - \ln(F_N(\sigma)).$$
(4.13)

For $N = 1$ and arbitrary j, we find for the observable ($\phi = \sqrt{\beta V}\sigma$)

$$\langle \hat{J}_z^2 \rangle_\sigma = j(j+1) + \cot\frac{\phi}{2}\left[\left(j + \frac{1}{2}\right)\cot\left(j + \frac{1}{2}\right)\phi - \frac{1}{2}\cot\frac{\phi}{2}\right]$$
(4.14)

while the Langevin equation is

$$\dot{\sigma} = -\frac{1}{2}\sigma + \frac{1}{2}\sqrt{\beta V}\left[\left(j + \frac{1}{2}\right)\cot\left(j + \frac{1}{2}\right)\phi - \frac{1}{2}\cot\frac{\phi}{2}\right] + \eta.$$
(4.15)

From the oscillatory nature of $F(\phi)$ we expect that the MC procedure will become imprecise at large β. In Fig. 4 (left panel) we show the result of a MC calculation of $\langle \hat{J}_z^2 \rangle_\beta$ for a $j = 5/2$ shell and particle numbers $N = 1$, $N = 2$, and $N = 3$. Note that $N = 3$ corresponds to half-filling, so that higher N's can be described in terms of ''hole'' numbers, and revert to the cases displayed. For this simple case, the sign (not shown) does not deteriorate too much before the ground state has been reached (β large), and the calculation is consequently reliable. Let us test nevertheless how a complex Langevin approach fares.

The CL approach based on the Langevin equation (4.15) with the observable (4.14) suffers from the same problems that we noted with the simple integral: the fixed points are real and the results are consequently unreliable. Again, the remedy is to shift the fixed points such that the effective path of integration lies in the complex plane. However, here we encounter another difficulty (which is also common in more refined shell-model calculations): the expectation value (4.14) has poles [at the zeros of $F(\phi)$]. Consequently, the expectation value calculated for a shifted path will equal the real-path result plus the sum over the poles on the real-line, of which there are infinitely many. For this simple toy Hamiltonian, this can be shown to hold true exactly by calculating the residues. In more realistic models, however, the sum over the poles is not readily available. Fortunately, for most Hamiltonians the observable $\hat{\mathrm{Tr}}_N(OU_\sigma)$ can be obtained by calculating moments of $\hat{\mathrm{Tr}}_N U_\sigma$. Here, for example,
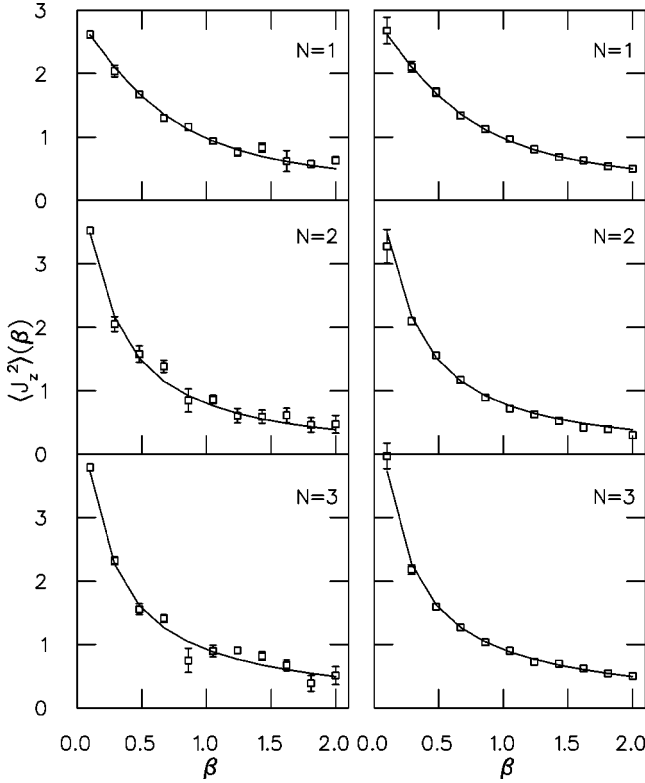
FIG. 4. Average $\langle J_z^2 \rangle$ for a $j = 5/2$ shell with $N = 1,2,3$ obtained with a Monte Carlo approach (left panel) and the complex Langevin equation (right panel). For both cases we took ten samples of 10 000 updates each.

$$\hat{\mathrm{Tr}}_N(\hat{J}_z^2 U_\sigma) = -\frac{1}{\beta V} \frac{\partial^2}{\partial \sigma^2} \hat{\mathrm{Tr}}_N U_\sigma \qquad (4.16)$$

and integrating by parts twice yields

$$\langle \hat{J}_z^2 \rangle = \frac{\int d\sigma e^{-(1/2)\sigma^2} \langle \hat{J}_z^2 \rangle_\sigma e^{-S_N(\sigma)}}{\int d\sigma e^{-(1/2)\sigma^2} e^{-S_N(\sigma)}}$$

$$= \frac{1}{\beta V} \left(1 - \frac{\int d\sigma e^{-(1/2)\sigma^2} \sigma^2 e^{-S_N(\sigma)}}{\int d\sigma e^{-(1/2)\sigma^2} e^{-S_N(\sigma)}} \right)$$

$$= \frac{1}{\beta V} (1 - \langle \sigma^2 \rangle). \qquad (4.17)$$

With this observable, the action can now be extended into the complex plane. This is achieved as in the simple integrals treated previously, by writing the trace in terms of cosines, and replacing $\cos(\phi) \rightarrow \exp(i\phi)$. Quantum mechanically, this amounts to retaining only those terms in the trace (4.3) for which the sum of the magnetic quantum numbers $m_1 + m_2 + \cdots m_N$ (with $m_i \neq m_j$) is either non-negative or non-positive. In this manner, we break time-reversal invariance by hand, since we know that it will be taken care of by the
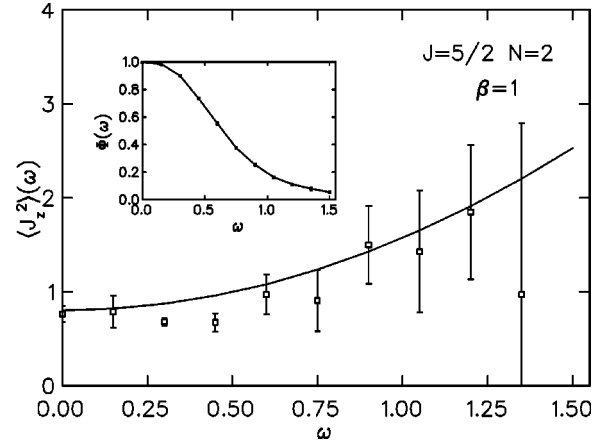

FIG. 5. Straightforward Monte Carlo average of $\langle J_z^2 \rangle$ for two particles in a cranked $j = 5/2$ shell at $\beta = 1$, and the sign (inset).

symmetry of the integral. Results of the CL calculation are shown in the right panel of Fig. 4. The CL averages are stable and accurate at large β while accuracy deteriorates only for small β where large cancellations must occur. Of course, a partial integration can also be performed for the Monte Carlo integral, which improves performance markedly because of better sampling. In that case, the Monte Carlo results are comparable to those obtained with the Langevin equation.

This changes when the action is forced to acquire substantial complex pieces by ''cranking'' the Hamiltonian (4.1), which corresponds to a toy nucleus with a single shell undergoing a collective rotation. Ignoring the term involving the number operator and again setting $V_2 = 1$, the Hamiltonian becomes

$$\hat{H}_c = \hat{H} - \omega \hat{J}_z, \qquad (4.18)$$

which breaks time-reversal invariance explicitly. Cranking is notoriously difficult for quantum MC calculations because the sign-problem is exacerbated in these cases. In fact, a straightforward MC is hopeless because the sign drops very quickly with increasing ω. In Fig. 5 we show $\langle J_z^2 \rangle$ as well as
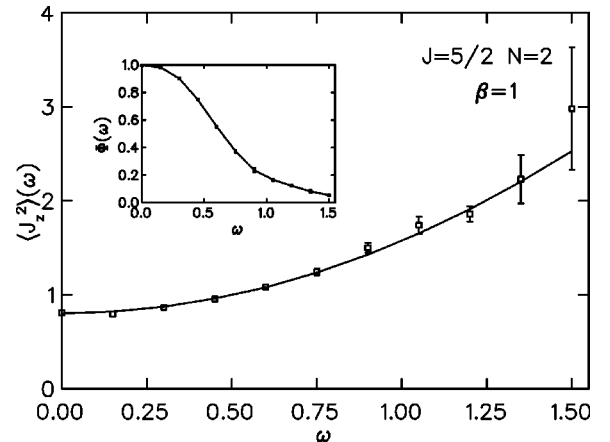

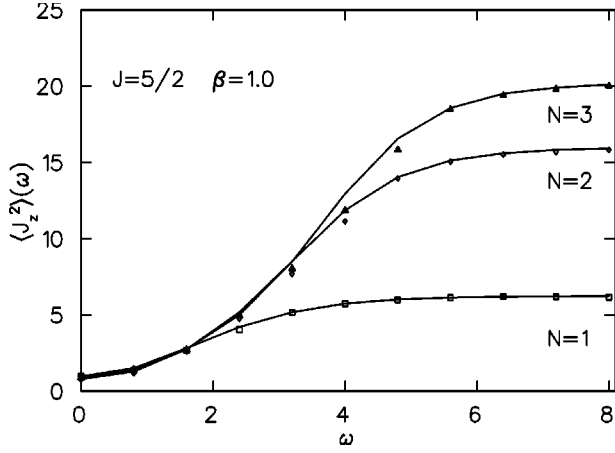FIG. 6. Monte Carlo with $\langle J_z^2 \rangle$ from partial integration, as in Fig. 5.

FIG. 7. Complex Langevin average of $\langle J_z^2 \rangle$ for a cranked $j = 5/2$ shell, at $\beta = 1.0$.

the sign for two particles in a $j = 5/2$ shell for $\omega < 1.5$, beyond which the MC calculation becomes useless. Convergence improves markedly if the observable is simplified by partial integration (Fig. 6), but of course the sign is still the same and the MC approach fails.

Let us calculate this observable with the CL approach. In Fig. 7 we plot the expectation value (4.10) as a function of cranking frequency for a $j = 5/2$ shell with $N = 1,2,3$ and $\beta = 1.0$. While the sign in the MC calculation essentially disappears for $\omega > 3$, the accuracy of the CL calculation is maintained even as the average sign is small.

V. LIPKIN MODEL

To test the CL equation approach in a more realistic situation for a system that can be exactly diagonalized, we use the Lipkin model [15]. The Lipkin model is a nontrivial schematic shell model capable of describing collective effects in nuclei. It describes N distinguishable particles labeled $1, 2, \ldots, N$, each of which can occupy one of two orbitals with energies ($\pm 1/2$) (up or down). The total number of states therefore is 2^N.

The Hamiltonian has a one-body term and two two-body terms, and suffers from the sign-problem as we shall see below. If written in terms of quasispin operators $\hat{J}$, it becomes

$$\hat{H} = \hat{J}_z - \tfrac{1}{2} V(\hat{J}_+^2 + \hat{J}_-^2) = \hat{J}_z - V(\hat{J}_x^2 - \hat{J}_y^2). \quad (5.1)$$

The eigenstates can be labeled by the total quasispin j, and classified into nondegenerate multiplets of $2j+1$ states each (see, e.g., [16]), from spin $j = N/2$ down to 0 or $\tfrac{1}{2}$, depending on whether N is even or odd.

The two-body interaction term does not commute with the free Hamiltonian, which necessitates the introduction of time-slices in the HS transformation. Writing the imaginary-time evolution operator as

$$\hat{U} = [\exp(-\Delta\beta\hat{H})]^{N_t}, \quad (5.2)$$

where $\beta = N_t \Delta\beta$, we can apply the Hubbard-Stratonovich transformation to obtain

$$e^{-\Delta\beta\hat{H}} \propto \int \mathcal{D}\sigma_x \mathcal{D}\sigma_y \exp\left(-\tfrac{1}{2}\Delta\beta V \right.$$
$$\left. \times \sum_n (\sigma_x^{(n)2} + \sigma_y^{(n)2}) \right) \Pi_n \exp(-\Delta\beta\hat{h}_\sigma^{(n)}), \quad (5.3)$$

where

$$\hat{h}_\sigma^{(n)} = \hat{J}_z + \sqrt{2}V(\hat{J}_x \sigma_x^{(n)} + i\hat{J}_y \sigma_y^{(n)}). \quad (5.4)$$

With this decomposition, $\exp(-\Delta\beta\hat{h}_\sigma)$ is accurate to order $(\Delta\beta)^2$ and $\hat{U}$ to order $\Delta\beta$. Exact results are obtained in the limit $\Delta\beta \to 0$.

To obtain averages at finite temperature, we need to take traces of such operators over the many-body basis. For small N, the model can easily be diagonalized, which we use to our advantage to compare Monte Carlo and Langevin calculations of expectation values with the known exact results. We start with the expressions for the static-path approximation (SPA), i.e., for a single time-slice. With the representation

$$h_\sigma = \begin{pmatrix} \dfrac{1}{2} & \dfrac{V}{\sqrt{2}}(\sigma_x - i\sigma_y) \\ \dfrac{V}{\sqrt{2}}(\sigma_x + i\sigma_y) & -\dfrac{1}{2} \end{pmatrix} \quad (5.5)$$
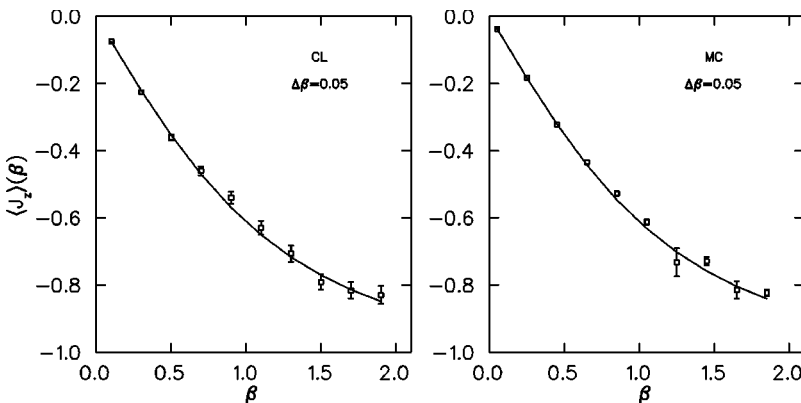

FIG. 8. Monte Carlo (right panel) and complex Langevin (left panel) average $\langle J_z \rangle$ in a three particle Lipkin model as a function of inverse temperature, with $\Delta\beta = 0.05$.

we have

$$\hat{T}r_1(e^{-\Delta\beta h_\sigma}) = 2\cosh\left(\frac{\Delta\beta}{2}W\right), \qquad (5.6)$$

where

$$W = \sqrt{1 + 2V^2(\sigma_x^2 - \sigma_y^2)}, \qquad (5.7)$$

and indeed quite generally

$$\hat{T}r_N\exp(-\Delta\beta h_\sigma) = \left[2\cosh\left(\frac{\Delta\beta}{2}W\right)\right]^N. \qquad (5.8)$$

Similarly, we obtain, for example,

$$\langle\hat{J}_z\rangle_N = -\frac{N}{2W}\tanh\left(\frac{\Delta\beta}{2}W\right). \qquad (5.9)$$

The many-time-slice expressions are straightforward extensions. For time-slice n, define

$$\exp(-\Delta\beta h_\sigma^{(n)}) \equiv A^{(n)} = a_0^{(n)} + \vec{a}^{(n)}\cdot\vec{\tau} \qquad (5.10)$$

and

$$U_\sigma = \exp(-\Delta\beta h_\sigma^{(1)}), \ldots, \exp(-\Delta\beta h_\sigma^{(N_t)}) \equiv u_0 + \vec{u}\cdot\vec{\tau}, \qquad (5.11)$$

where $\vec{\tau}$ are the usual Pauli matrices and

$$a_0^{(n)} = \cosh\left(\frac{\Delta\beta}{2}W_n\right), \qquad (5.12)$$

$$\vec{a}^{(n)} = \frac{V}{W^n}\sinh\left(\frac{\Delta\beta}{2}W_n\right)\begin{pmatrix} -\sqrt{2}\sigma_x^{(n)} \\ -i\sqrt{2}\sigma_y^{(n)} \\ -\frac{1}{V} \end{pmatrix}. \qquad (5.13)$$

Then

$$\hat{T}r_N(U_\sigma) = (2u_0)^N, \qquad (5.14)$$

and, for example,

$$J_z(\sigma) \equiv \hat{T}r_N(\hat{J}_z U_\sigma)/\hat{T}r_N(U_\sigma) = N\frac{u_3}{2u_0}. \qquad (5.15)$$

The action can be written as

$$S_\sigma = \frac{\Delta\beta V}{2}\left(\sum_n^{N_t}\sigma_x^{(n)2} + \sigma_y^{(n)2}\right) - \ln(\hat{T}r_N(U_\sigma)), \qquad (5.16)$$

and observables are obtained as usual. Here, we choose to examine

$$\langle\hat{J}_z\rangle_\beta = \frac{\displaystyle\int \mathcal{D}\sigma_x\mathcal{D}\sigma_y e^{-S_\sigma}J_z(\sigma)}{\displaystyle\int \mathcal{D}\sigma_x\mathcal{D}\sigma_y e^{-S_\sigma}}. \qquad (5.17)$$

The complex Langevin equation requires the gradients of $\hat{T}r(U_\sigma)$ with respect to σ_x and σ_y for each time-slice:

$$\frac{\partial\hat{T}r(U_\sigma)}{\partial\sigma_x^{(n)}} = \hat{T}r(A^{(1)}\cdots A_x'^{(n)}\cdots A^{(N_t)}), \qquad (5.18)$$

where $A_x'^{(n)}$ is the derivative of the nth slice matrix $A^{(n)}$. Thus,

$$\dot{\sigma}_x^{(n)}(t) = -\frac{1}{2}\left(\sigma_x^{(n)} - N\frac{\partial\hat{T}r(U_\sigma)}{\partial\sigma_x^{(n)}}\middle/\hat{T}r(U_\sigma)\right) + \eta \qquad (5.19)$$
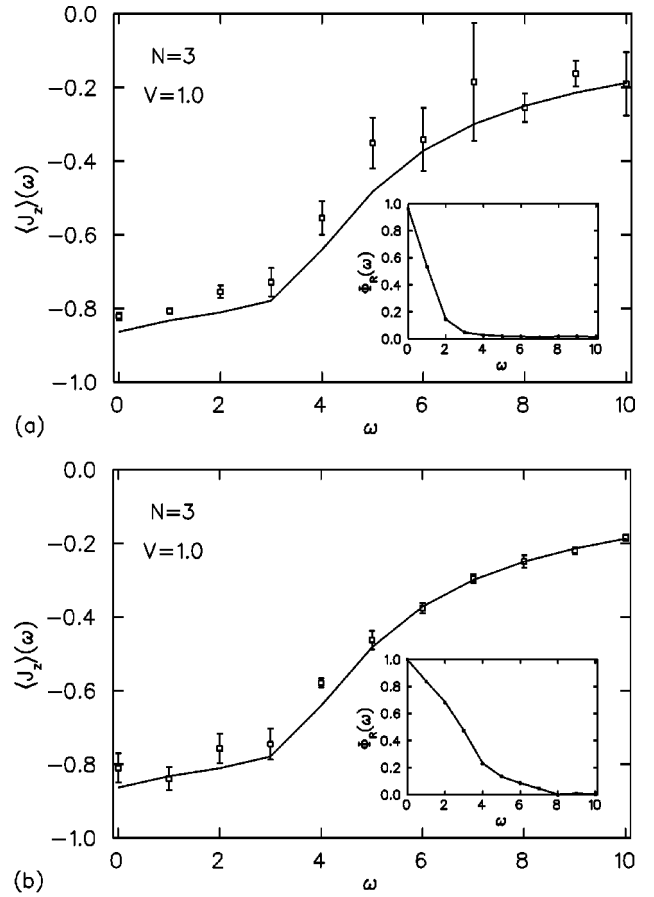

(a)

(b)

FIG. 9. (a) Monte Carlo calculation of $\text{Re}(\langle J_z\rangle)$ as in Fig. 8 (ten samples of 10 000 points), at fixed inverse temperate $\beta = 2.0$ (with $\Delta\beta = 0.1$) and as a function of cranking frequency ω. The inset shows the real part of the sign function Φ_R. (b) Complex Langevin calculation of $\langle J_z\rangle$ with same parameters as (a).

since

$$\frac{\partial}{\partial \sigma_x^{(n)}} \hat{\mathrm{Tr}}(\hat{U}_\sigma)^N = N\ \hat{\mathrm{Tr}}(A^{(1)}\cdots A'^{(n)}\cdots A^{(N_t)})\cdot \hat{\mathrm{Tr}}(U_\sigma)^{N-1} \quad (5.20)$$

if we choose a basis of multiparticle product states, and similarly for σ_y. In this manner, all remaining traces are over the two-dimensional single-particle space only. Naturally, this equation has to be separated into its real and imaginary parts as earlier, since $\hat{\mathrm{Tr}}_N(U_\sigma)$ can become negative if N is odd. This is most easily seen examining the single particle trace (5.6), which becomes oscillatory if the argument $W = \sqrt{1 + 2V^2(\sigma_x^2 - \sigma_y^2)}$ becomes complex.

In Fig. 8, we show Monte Carlo (right) and complex Langevin (left) calculations of the observable $\langle J_z \rangle_\beta$ as a function of the inverse temperature β, for the first non-trivial case $N=3$. Even though the sign is not strictly positive, the Monte Carlo simulation is very accurate in this case. The Langevin averages converge well also, despite the fact that

none of the tricks used in the previous examples (such as partial integration and extending the action into the complex plane) can be used for this model. In order to force a sign problem, we can revert to cranking as before.

The cranked Hamiltonian

$$H = H_0 - \omega \hat{J}_y \quad (5.21)$$

develops a sign-problem because the single particle (and single time-slice) trace (5.6) becomes complex as the argument (5.7) becomes

$$W = \sqrt{1 + \omega^2 + 2V^2(\sigma_x^2 - \sigma_y^2) - 2\sqrt{2}\mathrm{i}V\omega\sigma_y}. \quad (5.22)$$

Figure 9 compares Monte Carlo and complex Langevin calculations of the same observable as in Fig. 8, at fixed inverse temperature $\beta = 2.0$ and as a function of the cranking frequency ω. Because $\hat{\mathrm{Tr}}(U_\sigma)$ becomes complex (rather than just nonpositive), care must be given to the real and imaginary parts of the sign-function Φ. Indeed, for $\Phi = \Phi_R + \mathrm{i}\Phi_I$ and the observable $O_R(\sigma) + \mathrm{i}O_I(\sigma)$, the Monte Carlo average is

$$\langle O \rangle = \frac{\int \mathcal{D}\sigma e^{-S_\sigma}(\Phi_R O_R - \Phi_I O_I) + \mathrm{i}\int \mathcal{D}\sigma e^{-S_\sigma}(\Phi_I O_R + \Phi_R O_I)}{\int \mathcal{D}\sigma e^{-S_\sigma}\Phi_R + \mathrm{i}\int \mathcal{D}\sigma e^{-S_\sigma}\Phi_I}. \quad (5.23)$$

Figure 9(a) shows the real part of Eq. (5.23) as well as the real part of the sign Φ_R in the inset. The sign disappears quickly, both in the MC as well as the Langevin calculation in Fig. 9(b), but in the Langevin case the accuracy of the average actually increases with increasing cranking frequency. Instead, for the MC calculation, this translates into a deteriorating signal-to-noise ratio. Also, the complex Langevin calculation does not necessitate the calculation of four separate integrals such as in Eq. (5.23). However, for small cranking frequencies ω the complex Langevin averages are noticeably nonstationary, which results in larger error bars.

VI. CONCLUSIONS

The complex Langevin equation offers a new perspective on the pervasiveness of the sign-problem in fermionic quantum many-body calculations. It is not without its own problems, however, most notably the absence of a convergence proof of the Langevin averages. The root of non-stationarity for some complex Langevin averages lies in the structure of fixed points (attractors) and turning points (repellers) in the complex plane. If both the attractors and repellers lie on the real line, it is just a matter of time until the trajectory hits a pole in the gradient, and the trajectory is thrown far into the complex plane. The first two examples we have treated show how this can be avoided by modifying the action such that

the fixed points move into the complex plane, without changing the value of the average or the pole structure. The third example, the Lipkin model, showed that such a procedure is not necessary if the fixed points are naturally in the complex plane (such as is the case at finite cranking frequencies) even though the averages may become non-stationary. In these cases, the Langevin equation continues to deliver reliable averages even when the Monte Carlo averages have become meaningless.

While this study certainly suggests that the sign-problem can be overcome in particular cases, it is by no means certain that the procedure will be as successful in so-called real-life applications, with realistic interactions. However, as the payoff is potentially large, we believe that there is now enough evidence to try this approach. Another area where this approach deserves to be tested is lattice gauge calculations of matter at finite chemical potential μ, which suffer from a sign-problem because the action becomes complex. As the attractors would naturally move into the complex plane at about $\sim \mathrm{i}\mu$, the complex Langevin approach seems particularly natural in this case.

[1] G. H. Lang, C. W. Johnson, S. E. Koonin, and W. E. Ormand, Phys. Rev. C **48**, 1518 (1993).

[2] S. E. Koonin, D. J. Dean, and K. Langanke, Phys. Rep. **278**, 2 (1997).

[3] S. E. Koonin, D. J. Dean, and K. Langanke, Annu. Rev. Nucl. Part. Sci. **47**, 463 (1997).

[4] *Quantum Monte Carlo Methods in Condensed Matter Physics*, edited by M. Suzuki (World Scientific, Singapore, 1993).

[5] J. M. Deutsch, Phys. Rev. E **50**, 2411 (1994).

[6] S. W. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. B **55**, 7464 (1995).

[7] M. Honma, T. Mizusaki, and T. Otsuka, Phys. Rev. Lett. **75**, 1284 (1995).

[8] G. Parisi, Phys. Lett. **131B**, 393 (1983); J. Flower, W. Otto, and S. Callahan, Phys. Rev. D **34**, 598 (1986); J. Ambjørn, M. Flensburg, and C. Peterson, Nucl. Phys. **B275**, 375 (1986); R. Haymaker and J. Wosiek, Phys. Rev. D **37**, 969 (1988); B. Söderberg, Nucl. Phys. **B295**, 396 (1988); K. Okano, L. Schülke, and B. Zheng, Prog. Theor. Phys. Suppl. **111**, 313 (1993).

[9] J. Hubbard, Phys. Lett. **3**, 77 (1959); R. D. Stratonovich, Sov. Phys. Dokl. **2**, 416 (1958).

[10] J. R. Klauder, Phys. Rev. A **29**, 2036 (1984).

[11] S. Lee, Nucl. Phys. **B413**, 827 (1994).

[12] H. S. Greenside and E. Helfand, Bell Syst. Tech. J. **60**, 1927 (1981).

[13] M. Okamoto, K. Okano, L. Schülke, and S. Tanaka, Nucl. Phys. **B324**, 684 (1989); K. Okano, L. Schülke, and B. Zheng, Phys. Lett. B **258**, 421 (1991).

[14] J. Flower, W. Otto, and S. Callahan, Phys. Rev. D **34**, 598 (1986); J. Ambjørn, M. Flensburg, and C. Peterson, Nucl. Phys. **B275**, 375 (1986); R. Haymaker and J. Wosiek, Phys. Rev. D **37**, 969 (1988); K. Fujimura *et al.*, Nucl. Phys. **B424**, 675 (1994).

[15] H. J. Lipkin, N. Meshkov, and A. J. Glick, Nucl. Phys. **62**, 188 (1965).

[16] S. E. Koonin and D. Meredith, *Computational Physics* (Springer-Verlag, New York, 1990), Chap. 5.