Symmetry projection schemes for Gaussian Monte Carlo methods

F. F. Assaad
Institut für theoretische Physik und Astrophysik, Universität Würzburg, Am Hubland D-97074 Würzburg, Germany

P. Werner, P. Corboz, E. Gull, and M. Troyer
Institut für theoretische Physik, ETH Hönggerberg, CH-8093 Zürich, Switzerland
(Received 11 September 2005; revised manuscript received 7 November 2005; published 30 December 2005)

A sign-free Monte Carlo method for the Hubbard model has recently been proposed by Corney and Drummond. High-precision measurements on small clusters show that ground-state correlation functions are not correctly reproduced. We argue that the origin of this mismatch lies in the fact that the low-temperature density matrix does not have the symmetries of the Hamiltonian. Here we show that supplementing the algorithm with symmetry projection schemes provides reliable and accurate estimates of ground-state properties.

DOI: 10.1103/PhysRevB.72.224518 PACS number(s): 71.27.+a, 71.10.Fd

I. INTRODUCTION

The understanding of low-temperature properties of doped Mott insulators is a central challenge in solid-state physics. From the numerical point of view, this problem has remained elusive due to the sign problem. Recently Corney and Drummond[1] have proposed a stochastic method in which the sign problem does not explicitly occur. They show that the density matrix of an arbitrary model Hamiltonian may be expressed as a positive sum over Gaussian operators. The imaginary time propagation of the density matrix boils down to a Fokker-Planck equation governing the time evolution of the probability distribution in the space of Gaussian operators. One can then solve the Fokker-Planck equation by integrating, numerically, the corresponding stochastic differential equation (SDE). For the Hubbard model on arbitrary lattice topologies and at arbitrary band fillings, the SDE has real stochastic and drift forces thereby leading to no explicit sign problem.

The aim of this paper is to test the precision of the method with respect to ground state properties. We will see—on the basis of simple examples—that the low-temperature density matrix obtained by numerically solving the SDE does not have the symmetry of the Hamiltonian, thereby producing biased ground-state properties. The problem stems from the fact that a single Gaussian operator breaks spin, lattice as well as translation symmetries. Since the weighted summation over the Gaussian operators produces the density matrix, the summation has to restore the symmetries of the problem at hand. However, the summation is carried out stochastically, and it is *a priori* not clear that the sampling is efficient enough to restore symmetries. At *high* temperatures this poses no further problems, but as the temperature is lowered symmetry restoration fails. We show that one can solve this problem by projecting the density matrix onto the symmetry sector of the ground state.

The organization of this paper is as follows. In Sec. II, we briefly review the formulation of the Gaussian quantum Monte Carlo approach (GQMC) applied to the Hubbard model and, in Appendix C, present the generalization to four-fermion terms arising from Coulomb interactions. We will highlight the implicit assumptions used to derive the Fokker-Planck equation, namely the absence of boundary terms. In Appendix B, we discuss stochastic gauges which may serve as a means of suppressing the possible occurrence of boundary terms. In Sec. III we show, in detail, how to implement the symmetry projections and demonstrate their efficiency in Sec. IV. Finally, we draw conclusions.

II. GAUSSIAN QUANTUM MONTE CARLO METHOD FOR THE HUBBARD MODEL

In this section we summarize the results of Ref. 1. Although the GQMC is general and can be generalized to arbitrary Coulomb interactions as discussed in Appendix C, we will concentrate here on the Hubbard model

$$\hat{H} = \hat{\boldsymbol{c}}^{\dagger} \boldsymbol{T} \hat{\boldsymbol{c}} - \frac{U}{2} \sum_{\vec{i}} (\hat{\boldsymbol{c}}_{\vec{i}}^{\dagger} \boldsymbol{\sigma} \hat{\boldsymbol{c}}_{\vec{i}})^2, \tag{1}$$

where $\hat{\boldsymbol{c}}^{\dagger} = (\hat{c}_1^{\dagger}, \ldots, \hat{c}_{N_s}^{\dagger})$. $\hat{c}_x^{\dagger}$ creates a fermion with quantum numbers $x = (\vec{i}, \sigma)$ where $\vec{i}$ denotes the lattice site and σ the z component of spin. Hence x runs from 1 to $N_s \equiv 2N$, N being the number of lattice sites. $\boldsymbol{T}$ is the hopping matrix. It is diagonal in spin indices and takes the value $-t(-t')$ for next neighbors (next-nearest neighbors). Finally $\boldsymbol{\sigma}$ denotes a Pauli spin matrix and $\hat{\boldsymbol{c}}_{\vec{i}}^{\dagger} = (\hat{c}_{\vec{i},\uparrow}^{\dagger}, \hat{c}_{\vec{i},\downarrow}^{\dagger})$. Setting $\boldsymbol{\sigma} = \mathbf{1}$ yields the attractive Hubbard model whereas setting $\boldsymbol{\sigma}$ to $\boldsymbol{\sigma}^x$ or $\boldsymbol{\sigma}^z$ the repulsive case.[12]

Corney and Drummond[1] propose to expand the density matrix in terms of Gaussian operators

$$\hat{\Lambda}(\mathbf{n}) = \det(\mathbf{1} - \mathbf{n}) : e^{-\hat{\mathbf{c}}^{\dagger}[\mathbf{2} + (\mathbf{n}^T - \mathbf{1})^{-1}]\hat{\mathbf{c}}} : \tag{2}$$

with $\mathbf{n}$ an $N_s \times N_s$ real matrix. The Gaussian operators are normalized, $\mathrm{Tr}[\hat{\Lambda}(\mathbf{n})] = 1$ and obey Wick's theorem such that

$$\mathrm{Tr}[\hat{\Lambda}(\mathbf{n}) \hat{c}_x^{\dagger} \hat{c}_y] = \boldsymbol{n}_{x,y},$$

$$\mathrm{Tr}[\hat{\Lambda}(\mathbf{n}) \hat{c}_x^{\dagger} \hat{c}_y \hat{c}_w^{\dagger} \hat{c}_z] = \boldsymbol{n}_{x,y} \boldsymbol{n}_{w,z} + \boldsymbol{n}_{x,z} (\mathbf{1} - \boldsymbol{n})_{w,y}. \tag{3}$$

The major result of Ref. 1 is that one can expand the density matrix in terms of a positive sum of Gaussian operators

$$\hat{\rho}(\tau) = \sum_i P_i(\tau)\hat{\Lambda}(\mathbf{n}_i), \quad P_i \geqslant 0. \tag{4}$$

Clearly $\mathrm{Tr}[\hat{\rho}(\tau)] \equiv \Sigma_i P_i(\tau)$ grows exponentially with τ. One can account for this exponential growth by attaching a weight factor to the Gaussian operators thereby obtaining

$$\hat{\rho}(\tau) = \int \mathrm{d}\underline{\boldsymbol{\lambda}} P(\underline{\boldsymbol{\lambda}}, \tau)\hat{\Lambda}(\underline{\boldsymbol{\lambda}}), \tag{5}$$

with $\underline{\boldsymbol{\lambda}} = (\Omega, \boldsymbol{n})$, $\hat{\Lambda}(\underline{\boldsymbol{\lambda}}) = \Omega\hat{\Lambda}(\boldsymbol{n})$, and $\int \mathrm{d}\underline{\boldsymbol{\lambda}} P(\underline{\boldsymbol{\lambda}}, \tau) = 1$.

The aim is now to formulate a stochastic process which samples the probability distribution $P(\underline{\boldsymbol{\lambda}}, \tau)$ in the space of Gaussian operators. To this end one considers the imaginary time evolution of the density matrix

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\hat{\rho}(\tau) = -\frac{1}{2}[\hat{H}, \hat{\rho}(\tau)]_+, \tag{6}$$

so that in conjunction with Eq. (5) we are left with the evaluation of the anticommutator $-\frac{1}{2}[\hat{H}, \hat{\Lambda}(\underline{\boldsymbol{\lambda}})]_+$. The anticommutator can be transformed into a differential form acting on $\hat{\Lambda}(\underline{\boldsymbol{\lambda}})$:

$$\begin{aligned} -\frac{1}{2}[\hat{H}, \hat{\Lambda}(\underline{\boldsymbol{\lambda}})]_+ = \Bigg(& -\Omega h(\boldsymbol{n})\frac{\partial}{\partial\Omega} - \sum_{x,y} \boldsymbol{A}_{x,y}\frac{\partial}{\partial \boldsymbol{n}_{x,y}} \\ & + \frac{1}{2}\sum_{\vec{i},x,y,w,z} \boldsymbol{B}^{(\vec{i})}_{x,y}\boldsymbol{B}^{(\vec{i})}_{w,z}\frac{\partial^2}{\partial \boldsymbol{n}_{x,y}\,\partial \boldsymbol{n}_{w,z}} \\ & + \frac{1}{2}\sum_{\vec{i},x,y,w,z} \boldsymbol{C}^{(\vec{i})}_{x,y}\boldsymbol{C}^{(\vec{i})}_{w,z}\frac{\partial^2}{\partial \boldsymbol{n}_{x,y}\,\partial \boldsymbol{n}_{w,z}} \Bigg)\hat{\Lambda}(\underline{\boldsymbol{\lambda}}), \end{aligned} \tag{7}$$

with

$$h(\boldsymbol{n}) = \mathrm{Tr}(\hat{\Lambda}(\boldsymbol{n})\hat{H}), \tag{8}$$

$$\boldsymbol{A} = \frac{1}{2}\boldsymbol{n}(\boldsymbol{T} - U\boldsymbol{M})\bar{\boldsymbol{n}} + \frac{1}{2}\bar{\boldsymbol{n}}(\boldsymbol{T} - U\boldsymbol{M})\boldsymbol{n},$$

$$\boldsymbol{B}^{(\vec{i})}_{x,y} = \sqrt{\frac{U}{2}}\sum_{\sigma,\sigma'} \boldsymbol{n}_{x,(\vec{i},\sigma)}\boldsymbol{\sigma}_{\sigma,\sigma'}\bar{\boldsymbol{n}}_{(\vec{i},\sigma'),y},$$

$$\boldsymbol{C}^{(\vec{i})}_{x,y} = \sqrt{\frac{U}{2}}\sum_{\sigma,\sigma'} \bar{\boldsymbol{n}}_{x,(\vec{i},\sigma)}\boldsymbol{\sigma}_{\sigma,\sigma'}\boldsymbol{n}_{(\vec{i},\sigma'),y}. \tag{9}$$

In the above, $\bar{\boldsymbol{n}} = \mathbf{1} - \boldsymbol{n}$ and

$$\begin{aligned} \boldsymbol{M}_{(\vec{i},\sigma),(\vec{j},\sigma')} = {} & \delta_{\vec{i},\vec{j}}\sum_{\eta,\eta'} \boldsymbol{n}_{(\vec{i},\eta),(\vec{i},\eta')}\boldsymbol{\sigma}_{\sigma,\sigma'}\boldsymbol{\sigma}_{\eta,\eta'} \\ & + \left(\frac{\mathbf{1}}{\mathbf{2}} - \boldsymbol{n}\right)_{(\vec{i},\eta),(\vec{i},\eta')}\boldsymbol{\sigma}_{\sigma,\eta}\boldsymbol{\sigma}_{\eta',\sigma'}. \end{aligned}$$

Partial integration, under the assumption that boundary terms vanish, yields the Fokker-Planck equation for the probability distribution $P(\underline{\boldsymbol{\lambda}}, \tau)$

$$\begin{aligned} \frac{\partial}{\partial\tau}P(\underline{\boldsymbol{\lambda}}, \tau) = \Bigg[& \frac{\partial}{\partial\Omega}\Omega h(\boldsymbol{n}) + \sum_{x,y}\frac{\partial}{\partial \boldsymbol{n}_{x,y}}\boldsymbol{A}_{x,y} \\ & + \frac{1}{2}\sum_{\vec{i},x,y,w,z}\frac{\partial^2}{\partial \boldsymbol{n}_{x,y}\,\partial \boldsymbol{n}_{w,z}}\boldsymbol{B}^{(\vec{i})}_{x,y}\boldsymbol{B}^{(\vec{i})}_{w,z} \\ & + \frac{1}{2}\sum_{\vec{i},x,y,w,z}\frac{\partial^2}{\partial \boldsymbol{n}_{x,y}\,\partial \boldsymbol{n}_{w,z}}\boldsymbol{C}^{(\vec{i})}_{x,y}\boldsymbol{C}^{(\vec{i})}_{w,z}\Bigg]P(\underline{\boldsymbol{\lambda}}, \tau). \end{aligned} \tag{10}$$

The form of the diffusion matrices, $\boldsymbol{D}^{C}_{(x,y),(w,z)} = \Sigma_{\vec{i}}\boldsymbol{C}^{(\vec{i})}_{x,y}\boldsymbol{C}^{(\vec{i})}_{w,z}$ and $\boldsymbol{D}^{B}_{(x,y),(w,z)} = \Sigma_{\vec{i}}\boldsymbol{B}^{(\vec{i})}_{x,y}\boldsymbol{B}^{(\vec{i})}_{w,z}$ is important. It depends on the manner in which we have written the Hubbard interaction term or, in other words, on the choice of the Fermi gauge. The fact that the diffusion matrices factor out as above allows us to formulate the SDE. Furthermore, the fact that $\boldsymbol{B}^{\vec{i}}$ and $\boldsymbol{C}^{\vec{i}}$ are real for real values of $\boldsymbol{n}$ will lead to positiveness of the weights. The appropriate choice of Fermi gauges for general Hamiltonians is considered in Appendix C.

The assumption of vanishing boundary terms is essential to justify the approach and boils down to the requirement that the probability $P(\underline{\boldsymbol{\lambda}}, \tau)$ has tails decaying sufficiently fast as $|\vec{\lambda}| \to \infty$. At this point, one has to recall that the Gaussian basis is overcomplete, such that different probability distributions $P(\underline{\boldsymbol{\lambda}}, \tau)$ will yield the the same density matrix. This degree of freedom is reflected in a stochastic gauge invariance, which is reviewed in Appendix B. Hence, even if boundary terms appear the hope remains of eliminating them by an appropriate stochastic gauge choice.

To proceed, let us assume that we can neglect the boundary terms. The Fokker-Planck equation can conveniently be transformed into an Ito SDE,[2]

$$\begin{aligned} \mathrm{d}\Omega &= -\Omega h(\boldsymbol{n})\mathrm{d}\tau, \\ \mathrm{d}\boldsymbol{n} &= -\boldsymbol{A}\mathrm{d}\tau + \sum_{\vec{i}}\boldsymbol{B}^{(\vec{i})}\mathrm{d}W_{\vec{i}} + \sum_{\vec{i}}\boldsymbol{C}^{(\vec{i})}\mathrm{d}W'_{\vec{i}}, \end{aligned} \tag{11}$$

with Wiener increments $\langle \mathrm{d}W_{\vec{i}}\rangle = \langle \mathrm{d}W'_{\vec{i}}\rangle = \langle \mathrm{d}W_{\vec{i}}\mathrm{d}W'_{\vec{j}}\rangle = 0$, and $\langle \mathrm{d}W'_{\vec{i}}\mathrm{d}W'_{\vec{j}}\rangle = \langle \mathrm{d}W_{\vec{i}}\mathrm{d}W_{\vec{j}}\rangle = \mathrm{d}\tau\delta_{\vec{i},\vec{j}}$. Equation (11) describes the time evolution of walkers in the space of Gaussian operators. At $\tau = 0$, $\rho(\tau = 0) \propto \mathbf{1}$ such that all the Walkers can be parametrized by $\underline{\boldsymbol{\lambda}} = (1, \frac{1}{2})$. At imaginary time τ they are distributed according to $P(\underline{\boldsymbol{\lambda}}, \tau)$ so that we have access to the density matrix. In particular, any equal time observable is given by

$$\langle \hat{O} \rangle \simeq \frac{\sum_i \mathrm{Tr}[\hat{\Lambda}(\underline{\boldsymbol{\lambda}}_i)\hat{O}]}{\sum_i \mathrm{Tr}[\hat{\Lambda}(\underline{\boldsymbol{\lambda}}_i)]}, \tag{12}$$

where the sum runs over the set of walkers generated by the SDE. Since Wick's theorem applies for a single Gaussian operator the numerator of the above equation may easily be calculated.

As apparent from Eq. (11), the weight of a Walker at imaginary time τ reads

$$\Omega(\tau) = e^{-\int_0^\tau d\tau' h[\boldsymbol{n}(\tau')]}. \tag{13}$$

Since the "equal-time Green functions" $\boldsymbol{n}$ are real, $h(\boldsymbol{n})$ is real and the weight remains positive. Hence the algorithm shows no explicit manifestation of the sign problem. However, the weights grow exponentially with imaginary time, thus yielding an exponential increase in the variance. To circumvent this problem, we have adopted the reconfiguration scheme proposed in Ref. 3. In this approach the population of walkers is kept constant. Walkers with large weights are cloned and those with small weights suppressed in such a way that in the large population limit the density matrix remains invariant. Finally, after reconfiguration the weights of all walkers is equal to their average.

We can now test the accuracy of the method on a 2×2 Hubbard model (see Fig. 1). As apparent from Fig. 1(a) at *high* temperatures the GQMC result for the energy [bullets in Fig. 1(a)] compares well with the exact result (solid line). However, at low temperatures there is a systematic deviation. We have failed to account for this mismatch by (i) enhancing the number of walkers, (ii) using different schemes for the integration of the SDE, (iii) varying the imaginary time step, (iv) setting $\boldsymbol{\sigma}$ to $\boldsymbol{\sigma}_x$ instead of $\boldsymbol{\sigma}_z$ in Eq. (1), and finally, (v) using different stochastic gauges (see Appendixes B and C). Writing the Hubbard term as $H_U = -\frac{U}{4}\Sigma_{\vec{i}}(\hat{\boldsymbol{c}}_{\vec{i}}^\dagger \boldsymbol{\sigma}^z \hat{\boldsymbol{c}}_{\vec{i}})^2 + (\hat{\boldsymbol{c}}_{\vec{i}}^\dagger \boldsymbol{\sigma}^x \hat{\boldsymbol{c}}_{\vec{i}})^2$ thereby adding additional noise terms only seemed to make things worse. Concerning point (ii) we tried implicit and explicit Euler schemes and a higher order Milstein integrator.[4] As the latter changes the order of the algorithm from $O(N^3)$ to $O(N^4)$ without reducing the systematic error, we prefer the Euler schemes.

To acquire more insight into the origin of the mismatch, we compute the charge and spin susceptibilities

$$\chi_c = \frac{\beta}{N}(\langle \hat{N}^2\rangle - \langle \hat{N}\rangle^2)$$

$$\chi_s^z = \frac{\beta}{N}(\langle \hat{S}_z^2\rangle - \langle \hat{S}_z\rangle^2)$$

$$\chi_s^{xy} = \frac{1}{2}\frac{\beta}{N}(\langle \hat{S}_x^2\rangle - \langle \hat{S}_x\rangle^2) + \frac{1}{2}\frac{\beta}{N}(\langle \hat{S}_y^2\rangle - \langle \hat{S}_y\rangle^2), \tag{14}$$

where $\hat{N} = \Sigma_{\vec{i}} \hat{\boldsymbol{c}}_{\vec{i}}^\dagger \hat{\boldsymbol{c}}_{\vec{i}}$ and $\hat{S}_\alpha = \Sigma_{\vec{i}} \hat{\boldsymbol{c}}_{\vec{i}}^\dagger \boldsymbol{\sigma}^\alpha \hat{\boldsymbol{c}}_{\vec{i}}$. Those quantities are plotted in Figs. 1(b) and 1(c). As apparent, the charge susceptibility as well as χ_s^z follow rather precisely the exact result. On the other hand, χ_s^{xy} diverges as $1/T$ thereby signaling that the low-temperature density matrix has nonvanishing overlaps with $s>0$ spin sectors. To solve this problem so as to produce accurate ground-state results, we propose to implement symmetry projection schemes.

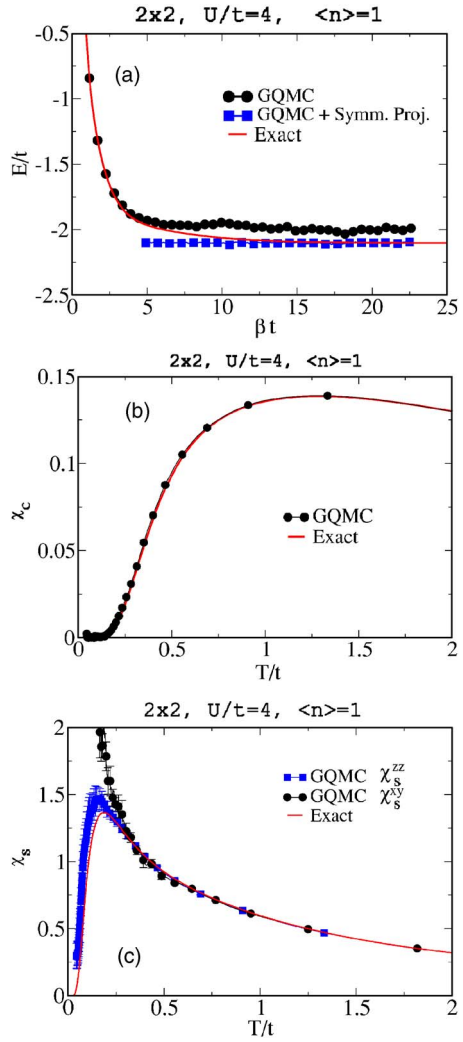

FIG. 1. (Color online) (a) Energy as a function of inverse temperature as obtained from exact diagonalization (solid line), from the GQMC (bullets) and from the GQMC supplemented by the symmetry projection (squares). Here we have projected onto the total spin $s=0$ state and d-wave lattice symmetry. (b) The charge susceptibility and (c) the longitudinal and transverse spin susceptibilities. In the above, we use 60 000 walkers, an imaginary time step of $\Delta\tau t = 0.0001$, an explicit Euler scheme with adaptive time step, and $\boldsymbol{\sigma} = \boldsymbol{\sigma}^z$ in Eq. (1)

III. SYMMETRY PROJECTIONS

Here we will assume that the low-temperature density matrix has a large *overlap* with the ground-state density matrix and a *small* admixture of excited states. If this assumption is correct, then projection onto the symmetry sector of the ground state will filter out the excited states and produce an accurate estimate of low-temperature properties. Let us note that symmetry projection schemes have been used successfully in the framework of the path-integral renormalization group approach where the ground-state wave function is approximated by a sum of Slater determinants.[5] Here, we first review the mathematics of symmetry projections and then show how to implement them in the context of the GQMC.

Let us first consider finite groups with elements R and irreducible representations $\mathcal{D}^\alpha(R)$. Group theory then tells us that

$$\sum_R \mathcal{D}^{\alpha}_{i,j}(R)^{\dagger}\mathcal{D}^{\beta}_{i',j'}(R) = \frac{\sum_R}{l_\alpha}\delta_{\alpha,\beta}\delta_{i,i'}\delta_{j,j'}, \tag{15}$$

where l_α corresponds to the dimension of the representation. For continuous groups, the sum has to be replaced by the invariant integral: $\Sigma_R \rightarrow \int dR$.[6]

To show how symmetry projections rely on the above identity let us first consider the group of translations by lattice vectors $\vec{R}$.

$$\hat{T}(\vec{R})\hat{c}^{\dagger}_{\vec{i},\sigma}\hat{T}(\vec{R})^{-1} = \hat{c}^{\dagger}_{\vec{i}+\vec{R},\sigma}, \tag{16}$$

with $\hat{T}(\vec{R}) = e^{i\vec{R}\cdot\sum_{\vec{p},\sigma}\vec{p}\hat{c}^{\dagger}_{\vec{p},\sigma}\hat{c}_{\vec{p},\sigma}}$ and $\hat{c}^{\dagger}_{\vec{p},\sigma} = (1/\sqrt{N})\Sigma_{\vec{i}}e^{i\vec{p}\cdot\vec{i}}\hat{c}^{\dagger}_{\vec{i},\sigma}$. In the above, N denotes the number of lattice sites. Since the group of translations is an Abelian group, the irreducible representations are one-dimensional and labeled by the total momentum $\vec{K}$. Classifying states in Fock space according to their total momentum $\vec{K}$ yields $\mathcal{D}^{\vec{K}}(\vec{R}) = \langle \vec{K}, \alpha^{\vec{}}_{\vec{K}}|\hat{T}(\vec{R})|\vec{K}, \alpha^{\vec{}}_{\vec{K}}\rangle \equiv e^{i\vec{R}\cdot\vec{K}}$. Here, $1 = \Sigma_{\vec{K},\alpha^{\vec{}}_{\vec{K}}}|\vec{K}, \alpha^{\vec{}}_{\vec{K}}\rangle\langle \vec{K}, \alpha^{\vec{}}_{\vec{K}}|$, where $\alpha^{\vec{}}_{\vec{K}}$ labels all the states in Fock space with total momentum $\vec{K}$. The projection operator onto the Hilbert space with total momentum $\vec{K}_0$ reads

$$\hat{P}_{\vec{K}_0} = \frac{1}{N}\sum_{\vec{R}} \langle \vec{K}_0|\hat{T}(\vec{R})|\vec{K}_0\rangle^{\dagger}\hat{T}(\vec{R}). \tag{17}$$

This expression may readily be verified:

$$P_{\vec{K}_0}|\Psi\rangle = \sum_{\vec{K},\alpha^{\vec{}}_{\vec{K}}} \overbrace{\frac{1}{N}\sum_{\vec{R}} \langle \vec{K}_0|\hat{T}(\vec{R})|\vec{K}_0\rangle^{\dagger}\langle \vec{K}, \alpha^{\vec{}}_{\vec{K}}|\hat{T}(\vec{R})|\vec{K}, \alpha^{\vec{}}_{\vec{K}}\rangle}^{\delta_{\vec{K},\vec{K}_0}} \times \langle \vec{K}, \alpha^{\vec{}}_{\vec{K}}|\Psi\rangle|\vec{K}, \alpha^{\vec{}}_{\vec{K}}\rangle = \sum_{\alpha^{\vec{}}_{\vec{K}_0}} \langle \vec{K}_0, \alpha^{\vec{}}_{\vec{K}_0}|\Psi\rangle|\vec{K}_0, \alpha^{\vec{}}_{\vec{K}_0}\rangle. \tag{18}$$

Within the very same framework, we can define the projection on the Hilbert space with total spin s. We first parametrize the rotations in terms of the Euler angles, $\boldsymbol{\omega} = (\alpha, \beta, \gamma)$, such that with

$$\hat{T}(\boldsymbol{\omega}) = e^{i\alpha\hat{S}^z}e^{i\beta\hat{S}^y}e^{i\gamma\hat{S}^z}, \tag{19}$$

a spinor transforms as

$$\hat{T}(\boldsymbol{\omega})\boldsymbol{c}^{\dagger}_{\vec{i}}\hat{T}^{-1}(\boldsymbol{\omega}) = \boldsymbol{c}^{\dagger}_{\vec{i}}e^{i(\alpha/2)\boldsymbol{\sigma}^z}e^{i(\beta/2)\boldsymbol{\sigma}^y}e^{i(\gamma/2)\boldsymbol{\sigma}^z}. \tag{20}$$

Here, $\hat{S}^z$ corresponds to the total z component of spin, $\Sigma_{\vec{i}}\frac{1}{2}\boldsymbol{c}^{\dagger}_{\vec{i}}\boldsymbol{\sigma}^z\boldsymbol{c}_{\vec{i}}$, and a similar definition holds for $\hat{S}^y$. Using Eq. (15) and noting that $\mathcal{D}^{s}_{m,m'}(\boldsymbol{\omega}) = \langle s, m|\hat{T}(\boldsymbol{\omega})|s, m'\rangle$ where the quantum numbers m, m' denote the z component of spin, the projection onto the Hilbert space with definite spin s, and vanishing z component of spin reads

$$\hat{P}_s = \frac{2s+1}{\int d\boldsymbol{\omega}}\int d\boldsymbol{\omega}\langle s,0|\hat{T}(\boldsymbol{\omega})|s,0\rangle^{\dagger}\hat{T}(\boldsymbol{\omega}). \tag{21}$$

Since we have chosen to parametrize rotations in terms of Euler angles the invariant integral reads $\int d\boldsymbol{\omega} = \int_0^{2\pi} d\alpha \int_0^{\pi} d\beta \sin(\beta) \int_0^{2\pi} d\gamma$ and

$$\langle s,0|\hat{T}(\boldsymbol{\omega})|s,0\rangle = P_s[\cos(\beta)], \tag{22}$$

where P_s denotes the s^{th} Legendre polynomial.

Since the GQMC method is a grand canonical approach, we have equally implemented projection onto fixed particle number. To this purpose, we define the gauge transformation

$$\hat{T}(\phi) = e^{i\phi\sum_{\vec{i}}\boldsymbol{c}^{\dagger}_{\vec{i}}\boldsymbol{c}^{\vec{}}_{\vec{i}}} \tag{23}$$

such that $\hat{T}(\phi)\boldsymbol{c}^{\dagger}_{\vec{i}}\hat{T}^{-1}(\phi) = e^{i\phi}\boldsymbol{c}^{\dagger}_{\vec{i}}$. Projection onto a given particle number sector then reads:

$$\hat{P}_N = \frac{1}{2\pi}\int_0^{2\pi} \langle N|\hat{T}(\phi)|N\rangle^{\dagger}\hat{T}(\phi). \tag{24}$$

Finally, we have implemented the C_4 lattice symmetries to classify states according to (i) s-wave, even under parity and $\pi/2$ rotations, (ii) d-wave, even under parity and odd under $\pi/2$ rotations, and (iii) $p_x + ip_y$, odd under parity and acquires a phase factor $e^{i\pi/2}$ under $\pi/2$ rotations. We denote this projection by $\hat{P}_{\text{latt}}$.

Since the Hubbard Hamiltonian is invariant under lattice vector translations, spin rotations, gauge transformations, $\pi/2$ rotations, the ground-state density matrix will have definite momentum, spin, particle number, and lattice symmetry. Our aim is now to project the density matrix produced by the GQMC onto a given symmetry sector and then use the projected density matrix

$$\hat{P}\hat{\rho}\hat{P}^{\dagger} \tag{25}$$

to compute observables. Here, $\hat{P}$ is a product of all or only some of the above symmetry projectors with general form

$$\hat{P} = \int d\boldsymbol{x} g(\boldsymbol{x})\hat{T}(\boldsymbol{x}), \tag{26}$$

where $\hat{T}$ is unitary and $\hat{P}^{\dagger} = \hat{P}$.

To simplify the calculation, we will assume that the observable $\hat{O}$ commutes with $\hat{P}$

$$[\hat{P}, \hat{O}]_- = 0, \tag{27}$$

such that

$$\langle \hat{O}\rangle_P = \frac{\text{Tr}[\hat{P}\hat{\rho}\hat{P}O]}{\text{Tr}[\hat{P}\hat{\rho}\hat{P}]} = \frac{\text{Tr}[\hat{P}\hat{\rho}\hat{O}]}{\text{Tr}[\hat{P}\hat{\rho}]}, \tag{28}$$

since $\hat{P}^2 = \hat{P}$. Estimating the right-hand side of Eq. (28) boils down to the calculation of $\hat{P}\hat{\rho}$ where $\hat{\rho} \approx \Sigma_{\underline{\boldsymbol{\lambda}}}\hat{\Lambda}(\underline{\boldsymbol{\lambda}})$, and the sum runs over the walkers produced by integrating the SDE. Hence, using the result of Appendix A,

TABLE I. GQMC with symmetry projection for the 2×2 half-filled Hubbard model at $U/t=4$. Here we have projected onto the d-wave and spin-singlet Hilbert spaces. To impose the spin projection, we have to integrate over the three Euler angles. This integration is done numerically by replacing the three-dimensional integral by a Riemann sum over 5^3 points. The thus-produced systematic error is not included in the error bars. The results and error bars stem from averaging the data over imaginary time (squares in Fig. 1).

$2\times2, U/t=4$ $\langle n\rangle=1$	GQMC+Symm. Proj. $s=0$, d-wave	Exact
Energy/t	-2.1021 ± 0.0007	-2.1026
$S(\pi,\pi)$	2.1933 ± 0.0010	2.1947
$N(\pi,\pi)$	0.2667 ± 0.0004	0.2664

$$\langle\hat{O}\rangle_P=\frac{\sum_{\underline{\lambda}}\int d\boldsymbol{x}g(\boldsymbol{x})\mathrm{Tr}[\hat{T}(\boldsymbol{x})\hat{\Lambda}(\underline{\boldsymbol{\lambda}})\hat{O}]}{\sum_{\underline{\lambda}}\int d\boldsymbol{x}g(\boldsymbol{x})\mathrm{Tr}[\hat{T}(\boldsymbol{x})\hat{\Lambda}(\underline{\boldsymbol{\lambda}})]}=\frac{\sum_{\underline{\lambda}}\int d\boldsymbol{x}g(\boldsymbol{x})\mathrm{Tr}[\hat{\Lambda}[\underline{\boldsymbol{\lambda}}(\boldsymbol{x})]\hat{O}]}{\sum_{\underline{\lambda}}\int d\boldsymbol{x}g(\boldsymbol{x})\Omega(\boldsymbol{x})},\tag{29}$$

where $\hat{T}(\boldsymbol{x})\hat{\Lambda}(\underline{\boldsymbol{\lambda}})=\hat{\Lambda}[\underline{\boldsymbol{\lambda}}(\boldsymbol{x})]$.

We test the above procedure on the 2×2 lattice of Fig. 1. As seen in Fig. 1(a) (solid squares), by projecting onto the spin-singlet and d-wave state we obtain a very accurate estimate of the ground-state energy already at $\beta t=5$. Averaging over subsequent imaginary times yields the results presented in Table I. It is important to note that not only the ground-state energy is very well reproduced but that reliable estimates for the spin and charge structure factors,

$$S(\vec{q})=\frac{4}{3N}\sum_{\vec{i},\vec{j}}e^{\vec{q}\cdot(\vec{i}-\vec{j})}\langle\hat{\boldsymbol{S}}_{\vec{i}}\cdot\hat{\boldsymbol{S}}_{\vec{j}}\rangle,$$

$$N(\vec{q})=\frac{1}{N}\sum_{\vec{i},\vec{j}}e^{\vec{q}\cdot(\vec{i}-\vec{j})}\langle\hat{n}_{\vec{i}}\cdot\hat{n}_{\vec{j}}\rangle,\tag{30}$$

are also obtained.

IV. ACCURACY TESTS

Here we provide further tests triggered at assessing the accuracy of the method. Unless mentioned explicitly, all simulations we carry out in this section suffer from the same problem discussed in detail for the 2×2 lattice in Sec. II; that is, ground-state energy values, which are systematically too high, and spin-spin correlation functions, which take different values depending on if we measure in the transverse or longitudinal directions. We first consider systems where the sign problem is absent in auxiliary field QMC methods; that is, the particle-hole symmetric Hubbard model. Table II presents results at half filling for both 4×4 and 6×6 lattices. In both cases, one sees that the agreement with benchmark results [exact diagonalization for the 4×4 lattice and auxiliary field projector QMC (PQMC) for the 6×6 lattice] is excellent. Furthermore, the real-space spin-spin correlations agree very well with the benchmark results (see Fig. 2).

TABLE II. Comparison between GQMC and benchmark results for the 4×4 and 6×6 Hubbard model. For both parameter sets, we project onto total spin $s=0$ and total momentum $\vec{P}=0$. The $L=4$($L=6$) simulations were carried out with 12 000 (6000) walkers, an explicit Euler scheme, and an imaginary time step $\Delta\tau=0.0005$($\Delta\tau=0.001$). The exact diagonalization results for the $L=4$ lattice stem from Ref. 7. For the $L=6$ lattice, we compare to the auxiliary field projector QMC (PQMC) algorithm.

	$U/t=4, t'/t=0, \langle n\rangle=1$	
	GQMC+Symm. Proj.	
$L=4$	$s=0$, $\vec{P}=0$	Exact
Energy/t	-13.630 ± 0.016	-13.6224
$S(\pi,\pi)$	3.66 ± 0.013	3.64
$N(\pi,\pi)$	0.386 ± 0.001	0.385
	GQMC+Symm. Proj.	
$L=6$	$s=0$, $\vec{P}=0$	PQMC
Energy/t	-30.87 ± 0.04	-30.87 ± 0.02
$S(\pi,\pi)$	5.86 ± 0.05	5.82 ± 0.03
$N(\pi,\pi)$	0.400 ± 0.004	0.418 ± 0.025

The crucial point is to show that in situations where the sign problem plagues the auxiliary field QMC, the Gaussian

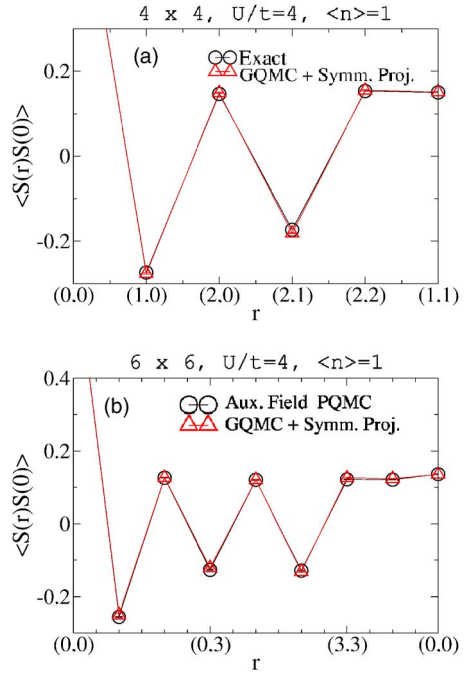

FIG. 2. (Color online) Real-space spin-spin correlations as obtained from the GQMC and comparison to benchmark results. See caption of Table II for details of the GQMC simulations.

TABLE III. Comparison between GQMC and exact diagonalization results. Here we have used 12 000 walkers and a time step of $\Delta\tau t=0.0005$. The GQMC is a grand canonical simulation. Hence in cases where charge fluctuations are not negligible we project onto fixed particle number Hilbert spaces so as to allow comparison with exact diagonalization results.

$U/t=4, t'/t=0$	GQMC+Symm. Proj.	
$\langle n\rangle=0.625$	$s=0$, s-wave, $N=10$	Exact
Energy/t	-19.576 ± 0.012585	-19.584
$S(\pi,\pi)$	0.737 ± 0.002	0.73
$N(\pi,\pi)$	0.5075 ± 0.001	
$U/t=8, t'/t=-0.3$	GQMC+Symm. Proj.	Exact
$\langle n\rangle=1$	$s=0$, $\vec{P}=0$, s-wave	
Energy/t	-8.498 ± 0.012	-8.4884
$S(\pi,\pi)$	5.09 ± 0.07	4.985
$N(\pi,\pi)$	0.191 ± 0.004	0.1920
$U/t=8, t'/t=-0.3$	GQMC+Symm. Proj.	Exact
$\langle n\rangle=0.875$	$s=0$, $\vec{P}=0$, $N=14$	
Energy/t	-12.01 ± 0.40	-12.50293
$S(\pi,\pi)$	0.941 ± 0.17	0.964776
$N(\pi,\pi)$	0.266 ± 0.01	0.27962

approach remains accurate. Table III presents three data sets where the sign problem in the auxiliary field approach varies from mild to very severe.

(i) Let us start with the 4×4 Hubbard model with nearest-neighbor hopping t and $\langle n\rangle=10/16$. The agreement between the GQMC and exact diagonalization is excellent. It is worth pointing out that in this specific case, the GQMC results with and without symmetry projections are identical meaning that the GQMC automatically produces the ground-state density matrix with correct symmetries. We believe that this is due to the fact that at this large doping away from half filling the ground state is very well described by a paramagnetic mean-field solution. Such a mean-field solution is exactly reproduced by the GQMC approach.

(ii) Our second example is the half-filled 4×4 frustrated Hubbard model at $U/t=8$. Here frustration stems from a next-nearest-neighbor hopping $t'/t=-0.3$. Both Table III and Fig. 3(a) show that we obtain excellent agreement with the exact result. Note that for those model parameters, the finite temperature auxiliary field approach has an average sign of $\langle\text{sign}\rangle\approx0.2$ at $\beta t=10$ and of $\langle\text{sign}\rangle\approx0.1$ at $\beta t=15$.

(iii) We now consider a parameter set that is out of reach for the auxiliary field approach, $U/t=8$, $\langle n\rangle=0.875$ and $t'/t=-0.3$. Table III shows that we are capable of reproducing the exact results. However, the fluctuations and hence the error bars in the MC data are large in comparison to the half-filled case. Those large fluctuations stem from the symmetry projection. In particular, the denominator in Eq. (28), $\mathrm{Tr}[\hat{P}\hat{\rho}]$, is *small* and has large relative fluctuations. In other words, the low-temperature density matrix (here we have propagated the walkers up to $\beta t=40$) produced by the GQMC still includes many excited states, and it is hard to filter out ground-state properties by imposing symmetries.[13] Nevertheless, and as seen in Fig. 3(b), comparisons to exact diagonalization results show that we are capable of accurately reproducing the details of real-space spin-spin correlation function.

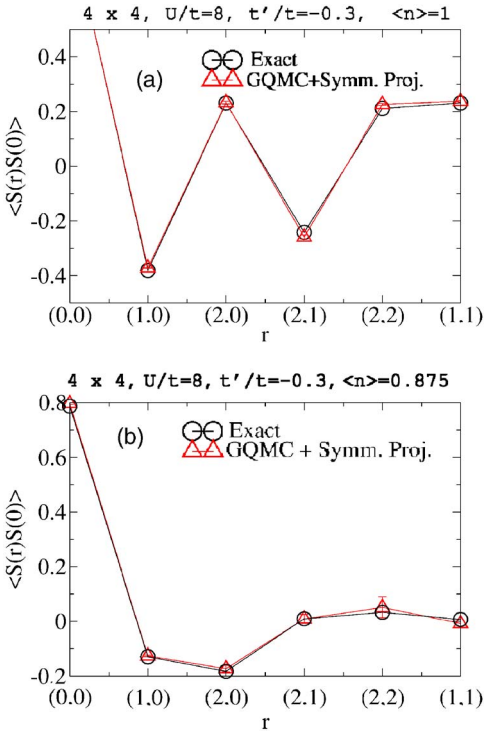

FIG. 3. (Color online) Real-space spin-spin correlations as obtained from the GQMC and comparison with exact-diagonalization results. See caption of Table III for details of the GQMC simulations.

V. CONCLUSIONS

We have shown that the GQMC method produces inaccurate ground-state properties since the numerical solution of the SDE fails to produce a low temperature density matrix with the symmetry properties of the Hamiltonian. To repair this sampling problem, we propose to *a posteriori* project the density matrix onto the symmetry sector of the ground state. We have shown ample nontrivial tests, including situations where auxiliary field methods fail due to the sign problem, where this approach yields accurate and reliable results. Those results confirm the point of view that the low-temperature density matrix produced by the GQMC has a good overlap with the exact zero-temperature density matrix, but that the GQMC density matrix contains excited states because the symmetries are not correctly reproduced. Those excited states are filtered out by the projection.

There are many open questions that deserve further work. In particular, is it possible to improve the sampling by incorporating aspects of the symmetry projections directly into the SDE? Also, we have not yet fully exploited the flexibility of the stochastic gauges. It is, at present, not clear if, with a suitable choice of stochastic gauge, the here-mentioned symmetry problems may be solved.

ACKNOWLEDGMENTS

The calculations presented here were carried out on the IBM p690 cluster of the NIC in Jülich. We would like to thank this institution for allocation of CPU time. We have greatly profited from discussions with G.G. Batrouni, C. Brünger, J. Corney, P. Drummond, W.P. Petersen, and D. Talay. Many thanks to S. Capponi who provided part of the exact diagonalization benchmark results. We acknowledge support by the Swiss National Science Foundation and DFG.

APPENDIX A: UNITARY TRANSFORMATION OF A GAUSSIAN OPERATOR

In this appendix, we show that

$$e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h} \hat{\boldsymbol{c}}} \hat{\Lambda}(\underline{\boldsymbol{\lambda}}) = \hat{\Lambda}(\underline{\tilde{\boldsymbol{\lambda}}}), \tag{A1}$$

with

$$(\tilde{\boldsymbol{n}}^T - 1)^{-1} = [(e^{i\boldsymbol{h}} - 1)\boldsymbol{n}^T + \mathbf{1}](\boldsymbol{n}^T - 1)^{-1},$$

$$\tilde{\Omega} = \Omega \det\left[(e^{i\boldsymbol{h}} - 1)\boldsymbol{n}^T + \mathbf{1}\right].$$

Here, $\boldsymbol{h}^\dagger = \boldsymbol{h}$, $\underline{\tilde{\boldsymbol{\lambda}}} = (\tilde{\Omega}, \tilde{\boldsymbol{n}})$, and $\underline{\boldsymbol{\lambda}} = (\Omega, \boldsymbol{n})$.

Before showing the above, let us fist recall some identities of the Grassmann algebra[8]

$$\langle \boldsymbol{\xi} | \boldsymbol{\xi}' \rangle = e^{\sum_x \xi_x^\dagger \xi_x'} \equiv e^{\boldsymbol{\xi}^\dagger \boldsymbol{\xi}'},$$

$$\langle \boldsymbol{\xi} | :A(\boldsymbol{c}^\dagger, \boldsymbol{c}): | \boldsymbol{\xi}' \rangle = A(\boldsymbol{\xi}^\dagger, \boldsymbol{\xi}') e^{\boldsymbol{\xi}^\dagger \boldsymbol{\xi}'},$$

$$\mathbf{1} = \int \underbrace{\prod_x \mathrm{d}\xi_x^\dagger \mathrm{d}\xi_x}_{\equiv \mathcal{D}\boldsymbol{\xi}} e^{-\boldsymbol{\xi}^\dagger \boldsymbol{\xi}} |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|. \tag{A2}$$

Here ξ_x are Grassmann variables and $|\boldsymbol{\xi}\rangle$ fermion coherent states.

In a first step it is convenient to transform $e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h} \hat{\boldsymbol{c}}}$ into a normal ordered form. Since h is hermitian, $\boldsymbol{h} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\dagger$ with $\boldsymbol{D}$ a diagonal and $\boldsymbol{U}$ unitary. With the canonical transformation $\hat{\boldsymbol{\gamma}}^\dagger = \hat{\boldsymbol{c}}^\dagger \boldsymbol{U}$, we obtain

$$\begin{aligned} e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h} \hat{\boldsymbol{c}}} &= \prod_x e^{i\hat{\gamma}_x^\dagger \hat{\gamma}_x D_x} = \prod_x [1 + (e^{iD_x} - 1)\hat{\gamma}_x^\dagger \hat{\gamma}_x] \\ &= \prod_x :e^{(e^{iD_x} - 1)\hat{\gamma}_x^\dagger \hat{\gamma}_x}: = :e^{\sum_x \hat{\gamma}_x^\dagger (e^{iD_x} - 1)\hat{\gamma}_x}: = :e^{\hat{\boldsymbol{c}}^\dagger (e^{i\boldsymbol{h}} - \mathbf{1})\hat{\boldsymbol{c}}}:. \end{aligned} \tag{A3}$$

We can now compute the quantity $e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h} \hat{\boldsymbol{c}}} : e^{\hat{\boldsymbol{c}}^\dagger \boldsymbol{B} \hat{\boldsymbol{c}}}:$, where $\boldsymbol{B}$ is an arbitrary matrix

$$\begin{aligned} e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h} \hat{\boldsymbol{c}}} : e^{\hat{\boldsymbol{c}}^\dagger \boldsymbol{B} \hat{\boldsymbol{c}}}: &= :e^{\hat{\boldsymbol{c}}^\dagger (e^{i\boldsymbol{h}} - 1)\hat{\boldsymbol{c}}} :: e^{\hat{\boldsymbol{c}}^\dagger \boldsymbol{B} \hat{\boldsymbol{c}}}: \\ &= \int \mathcal{D}\boldsymbol{\xi} \mathcal{D}\boldsymbol{\eta} \mathcal{D}\boldsymbol{\gamma} e^{-\boldsymbol{\xi}^\dagger \boldsymbol{\xi} - \boldsymbol{\eta}^\dagger \boldsymbol{\eta} - \boldsymbol{\gamma}^\dagger \boldsymbol{\gamma}} |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}| :e^{\hat{\boldsymbol{c}}^\dagger (e^{i\boldsymbol{h}} - 1)\hat{\boldsymbol{c}}}: \\ &\quad \times |\boldsymbol{\eta}\rangle\langle\boldsymbol{\eta}| :e^{\hat{\boldsymbol{c}}^\dagger \boldsymbol{B} \hat{\boldsymbol{c}}}: |\boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}| \\ &= \int \mathcal{D}\boldsymbol{\xi} \mathcal{D}\boldsymbol{\eta} \mathcal{D}\boldsymbol{\gamma} e^{-\boldsymbol{\xi}^\dagger \boldsymbol{\xi} - \boldsymbol{\eta}^\dagger \boldsymbol{\eta} - \boldsymbol{\gamma}^\dagger \boldsymbol{\gamma}} |\boldsymbol{\xi}\rangle e^{\boldsymbol{\xi}^\dagger e^{i\boldsymbol{h}} \boldsymbol{\eta}} e^{\boldsymbol{\eta}^\dagger (\boldsymbol{B} + \mathbf{1})\boldsymbol{\gamma}} \langle\boldsymbol{\gamma}| \\ &= \int \mathcal{D}\boldsymbol{\xi} \mathcal{D}\tilde{\boldsymbol{\eta}} \mathcal{D}\boldsymbol{\gamma} e^{-\boldsymbol{\xi}^\dagger \boldsymbol{\xi} - \tilde{\boldsymbol{\eta}}^\dagger \tilde{\boldsymbol{\eta}} - \boldsymbol{\gamma}^\dagger \boldsymbol{\gamma}} |\boldsymbol{\xi}\rangle e^{\boldsymbol{\xi}^\dagger \tilde{\boldsymbol{\eta}}} e^{\tilde{\boldsymbol{\eta}}^\dagger e^{i\boldsymbol{h}} (\boldsymbol{B} + \mathbf{1})\boldsymbol{\gamma}} \langle\boldsymbol{\gamma}| \\ &= \int \mathcal{D}\boldsymbol{\xi} \mathcal{D}\tilde{\boldsymbol{\eta}} \mathcal{D}\boldsymbol{\gamma} e^{-\boldsymbol{\xi}^\dagger \boldsymbol{\xi} - \tilde{\boldsymbol{\eta}}^\dagger \tilde{\boldsymbol{\eta}} - \boldsymbol{\gamma}^\dagger \boldsymbol{\gamma}} |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}| \tilde{\boldsymbol{\eta}}\rangle \\ &\quad \times \langle\tilde{\boldsymbol{\eta}}| :e^{\boldsymbol{c}^\dagger [e^{i\boldsymbol{h}} (\boldsymbol{B} + \mathbf{1}) - \mathbf{1}]\boldsymbol{c}}: |\boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}| = :e^{\boldsymbol{c}^\dagger [e^{i\boldsymbol{h}} (\boldsymbol{B} + \mathbf{1}) - \mathbf{1}]\boldsymbol{c}}:. \end{aligned} \tag{A4}$$

Here, we have carried out the substitution $\tilde{\boldsymbol{\eta}} = e^{i\boldsymbol{h}} \boldsymbol{\eta}$, baring in mind that $e^{i\boldsymbol{h}}$ is unitary matrix.

The result of Eq. (A1) follows from

$$\begin{aligned} e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h} \hat{\boldsymbol{c}}} \Lambda\{\underline{\boldsymbol{\lambda}} &= \Omega \det(1 - \boldsymbol{n}) e^{i\hat{\boldsymbol{c}}^\dagger \boldsymbol{h} \hat{\boldsymbol{c}}} :e^{-\hat{\boldsymbol{c}}^\dagger [\mathbf{2} + (\boldsymbol{n}^T - \mathbf{1})^{-1}]\hat{\boldsymbol{c}}}: \\ &= \Omega \det(1 - \boldsymbol{n}) :e^{-\hat{\boldsymbol{c}}^\dagger [\boldsymbol{I} + e^{i\boldsymbol{h}} (1 + (\boldsymbol{n}^T - \mathbf{1})^{-1})]\hat{\boldsymbol{c}}}: \\ &\equiv \underbrace{\Omega \frac{\det(1 - \boldsymbol{n})}{\det(1 - \tilde{\boldsymbol{n}})}}_{\tilde{\Omega}} \underbrace{\det(1 - \tilde{\boldsymbol{n}}) :e^{-\hat{\boldsymbol{c}}^\dagger [\mathbf{2} + (\tilde{\boldsymbol{n}}^T - \mathbf{1})^{-1}]\hat{\boldsymbol{c}}}:}_{\hat{\Lambda}(\tilde{\boldsymbol{n}})} \end{aligned} \tag{A5}$$

APPENDIX B: DRIFT GAUGES

Since the Gaussian operator basis is overcomplete, there are many probability distributions $P(\underline{\boldsymbol{\lambda}}, \tau)$ which will result in the same density matrix. This degree of freedom on $P(\underline{\boldsymbol{\lambda}}, \tau)$ is reflected in the choice of stochastic gauges. Clearly, the aim is to find a gauge that will suppress boundary terms that could potentially show up in the partial integration step required to obtain the Fokker-Planck equation. Here, we introduce drift gauges and then propose some ideas on how to choose the appropriate gauge.

To formulate stochastic gauge invariance, it is useful to introduce the index $\mu: 0, \ldots, N_s^2$ such that for example $\underline{\boldsymbol{\lambda}}_{\mu=0} = \Omega$ and $\underline{\boldsymbol{\lambda}}_\mu = \boldsymbol{n}_{x_\mu, y_\mu}$ for $\mu: 1, \ldots, N_s^2$. Then Eq. (7) may conveniently be written as

$$\begin{aligned} -\frac{1}{2}[\hat{H}, \hat{\Lambda}(\underline{\boldsymbol{\lambda}})]_+ = \Bigg(&- \sum_\mu \underline{\boldsymbol{A}}_\mu \frac{\partial}{\partial \underline{\boldsymbol{\lambda}}_\mu} + \frac{1}{2} \sum_{\vec{i}, \mu, \nu} \underline{\boldsymbol{B}}_\mu^{(\vec{i})} \underline{\boldsymbol{B}}_\nu^{(\vec{i})} \frac{\partial^2}{\partial \underline{\boldsymbol{\lambda}}_\mu \partial \underline{\boldsymbol{\lambda}}_\nu} \\ &+ \frac{1}{2} \sum_{\vec{i}, \mu, \nu} \underline{\boldsymbol{C}}_\mu^{(\vec{i})} \underline{\boldsymbol{C}}_\nu^{(\vec{i})} \frac{\partial^2}{\partial \underline{\boldsymbol{\lambda}}_\mu \partial \underline{\boldsymbol{\lambda}}_\nu} \Bigg) \hat{\Lambda}(\underline{\boldsymbol{\lambda}}), \end{aligned} \tag{B1}$$

with $\underline{\boldsymbol{A}} = [\Omega h(\boldsymbol{n}), \boldsymbol{A}]$, $\underline{\boldsymbol{B}}^{(\vec{i})} = (0, \boldsymbol{B}^{(\vec{i})})$ and with $\underline{\boldsymbol{C}}^{(\vec{i})} = (0, \boldsymbol{C}^{(\vec{i})})$. Eq. (B1) remains invariant under the transformation

$$\underline{\boldsymbol{B}}^{(\vec{i})} = (0, \boldsymbol{B}^{(\vec{i})}) \to (\Omega g^{(\vec{i})}, \boldsymbol{B}^{(\vec{i})}),$$

$$\underline{\boldsymbol{C}}^{(\vec{i})} = (0, \boldsymbol{C}^{(\vec{i})}) \to (\Omega f^{(\vec{i})}, \boldsymbol{C}^{(\vec{i})}),$$

$$\underline{\boldsymbol{A}} = [\Omega h(\boldsymbol{n}), \boldsymbol{A}] \to [\Omega h(\boldsymbol{n}), \boldsymbol{A} + \sum_{\vec{i}} g^{(\vec{i})} \boldsymbol{B}^{(\vec{i})} + f^{(\vec{i})} \boldsymbol{C}^{(\vec{i})}], \tag{B2}$$

where $g^{(\vec{i})}$ and $f^{(\vec{i})}$ are arbitrary functions of $\boldsymbol{n}$. This invariance stems from the fact that

$$\Omega \frac{\partial}{\partial \Omega} \hat{\Lambda}(\underline{\boldsymbol{\lambda}}) = \hat{\Lambda}(\underline{\boldsymbol{\lambda}}). \tag{B3}$$

For a given stochastic gauge, the Ito SDE reads

$$\mathrm{d}\Omega = -\Omega \left[h(\boldsymbol{n})\mathrm{d}\tau - \sum_{\vec{i}} g^{(\vec{i})} \mathrm{d}W_{\vec{i}} - \sum_{\vec{i}} f^{(\vec{i})} \mathrm{d}W'_{\vec{i}} \right],$$

$$\mathrm{d}\boldsymbol{n} = -\underbrace{\left[\boldsymbol{A} + \sum_{\vec{i}} g^{(\vec{i})} \boldsymbol{B}^{(\vec{i})} + f^{(\vec{i})} \boldsymbol{C}^{(\vec{i})} \right]}_{=\widetilde{\boldsymbol{A}}} \mathrm{d}\tau + \sum_{\vec{i}} \boldsymbol{B}^{(\vec{i})} \mathrm{d}W_{\vec{i}} + \sum_{\vec{i}} \boldsymbol{C}^{(\vec{i})} \mathrm{d}W'_{\vec{i}}. \tag{B4}$$

As apparent, one can modify the drift force from $\boldsymbol{A}$ to $\widetilde{\boldsymbol{A}}$ at the expense of adding noise in the weights Ω. Since our aim is to suppress the potential occurrence of boundary terms, one can follow the idea of choosing the gauge such that $\widetilde{\boldsymbol{A}}$ prohibits the walkers, $\boldsymbol{n}$, of drifting to infinity. In other words, $\widetilde{\boldsymbol{A}}_{x,y}$ should have the same sign as $\boldsymbol{n}_{x,y}$. Fulfilling this requirement for each pair of indices x,y leads to N_s^2 equalities. But since we only have have $2N$ ($g^{\vec{i}}$ and $f^{\vec{i}}$ for $\vec{i}:1,\cdots,N$) degrees of freedom, we can only fulfill the above condition on average. Defining a scalar product

$$(\boldsymbol{n},\widetilde{\boldsymbol{A}}) \equiv \sum_{x,y} \boldsymbol{n}_{x,y} \widetilde{\boldsymbol{A}}_{x,y} \tag{B5}$$

we require that

$$(\boldsymbol{n},\widetilde{\boldsymbol{A}}) > 0. \tag{B6}$$

APPENDIX C: FERMI GAUGES—POSITIVE-*P* REPRESENTATION FOR THE GENERAL ELECTRONIC STRUCTURE PROBLEM

One can also take advantage of gauge degrees of freedom by adding terms to the Hamiltonian that cancel each other (or are identically zero) and hence do not affect the observables. If these terms are of fourth and second order in the fermionic operators, they add a contribution to the diffusion part of the SDE, which is compensated in the drift part, while the equation for the weight Ω remains unaffected. In this appendix, we will show how such a Fermi gauge allows one to generalize the positive-P representation of the fermionic Hubbard model to the general electronic structure problem, including arbitrary hybridization, Coulomb, and exchange terms.

Since electronic structure calculations have important applications in quantum chemistry and the corresponding Hamiltonian

$$\hat{H} = -\sum_{i \neq j \sigma} t_{ij} \hat{c}^{\dagger}_{i\sigma} \hat{c}_{j\sigma} - \mu \sum_{i\sigma} \hat{c}^{\dagger}_{i\sigma} \hat{c}_{i\sigma} + \frac{1}{2} \sum_{ijkl\sigma\sigma'} V_{ijkl} \hat{c}^{\dagger}_{i\sigma} \hat{c}^{\dagger}_{k\sigma'} \hat{c}_{l\sigma'} \hat{c}_{j\sigma} \tag{C1}$$

has even been dubbed the "theory of everything,"[9] a simulation approach without uncontrolled approximations or systematic errors is highly desired. In Eq. (C1), $\hat{c}_{i\sigma}$ denotes the destruction operator for an electron with spin σ in the orbital i, t_{ij} the hopping amplitude and μ the chemical potential. The four fermion terms $\hat{c}^{\dagger}_{i\sigma} \hat{c}^{\dagger}_{k\sigma'} \hat{c}_{l\sigma'} \hat{c}_{j\sigma}$ arise from Coulomb interactions.

The many-body problem (C1) can be mapped to a system of stochastic differential equations with positive weights, by an appropriate choice of Fermi gauge, which generalizes the gauge proposed by Corney and Drummond for the Hubbard model.[1,10] Similar to Sec. II, we choose a basis of Gaussian operators parametrized by $\underline{\boldsymbol{\lambda}}$, which can be represented in this case by block-diagonal matrices with N_o^2 elements for spin up and spin down, respectively (N_o is the number of orbitals). The Fokker-Planck equation (10) for the probability distribution $P(\underline{\boldsymbol{\lambda}}, \tau)$ reads

$$\frac{d}{d\tau} P(\underline{\boldsymbol{\lambda}}, \tau) = L[P(\underline{\boldsymbol{\lambda}}, \tau)]. \tag{C2}$$

For a suitably chosen Fermi gauge, the operator L in Eq. (C2) takes the form

$$L = -\sum_{\alpha} \frac{\partial}{\partial \lambda_\alpha} A_\alpha + \frac{1}{2} \sum_{\alpha\beta} \frac{\partial}{\partial \lambda_\alpha} B_\alpha \frac{\partial}{\partial \lambda_\beta} B_\beta + \frac{1}{2} \sum_{\alpha\beta} \frac{\partial}{\partial \lambda_\alpha} C_\alpha \frac{\partial}{\partial \lambda_\beta} C_\beta, \tag{C3}$$

with B_α and C_α real coefficients (we present here the formulation that leads to the smallest number of noise terms). In this case, the Monte Carlo sampling can be done by integrating the Stratonovich SDE[11]

$$d\lambda_\alpha(\tau) = A_\alpha(\underline{\boldsymbol{\lambda}}) d\tau + B_\alpha(\underline{\boldsymbol{\lambda}}) dW(\tau) + C_\alpha(\underline{\boldsymbol{\lambda}}) dW'(\tau), \tag{C4}$$

where $\langle dW(\tau) dW'(\tau) \rangle = 0$ and $\langle dW(\tau)^2 \rangle = \langle dW'(\tau)^2 \rangle = d\tau$.

Gauge degrees of freedom can be used to modify the form of the operator L in (C2). In Ref. 1, the identity

$$\hat{n}^2_{ii\sigma} - \hat{n}_{ii\sigma} = 0 \tag{C5}$$

was used to map the Hubbard model to a system of real SDE. Here we will show how the identity

$$\hat{n}^2_{ij\sigma} - \delta_{ij} \hat{n}_{ij\sigma} = 0 \tag{C6}$$

can be used to obtain real SDE and positive weights Ω for the more general Hamiltonian (C1). First, we note that the latter can be written as

$$\hat{H} = -\sum_{ij\sigma} t_{ij} \hat{n}_{ij\sigma} + \sum_{ijkl\sigma\sigma'} W_{ijkl} \hat{n}_{ij\sigma} \hat{n}_{kl\sigma'}, \tag{C7}$$

with $[\hat{n}_{ij\sigma}, \hat{n}_{kl\sigma'}] = 0$. In particular, for $\sigma = \sigma'$ only terms with $i \neq k$, l and $j \neq k$, l appear. The $t_{ii\sigma}$ correspond to the chemical potential. We define

$$\hat{H}_{ij\sigma} = -t_{ij\sigma} \hat{n}_{ij\sigma}, \tag{C8}$$

and using Eq. (C6) express the $ijkl\sigma\sigma'$ contribution in Eq. (C1) in the form ($s_{W^{ij\sigma}_{kl\sigma'}}$ denotes the sign of $W^{ij\sigma}_{kl\sigma'}$)

$$\hat{H}_{ijkl\sigma\sigma'} \equiv W^{ij\sigma}_{kl\sigma'}\hat{n}_{ij\sigma}\hat{n}_{kl\sigma'} = -\frac{|W^{ij\sigma}_{kl\sigma'}|}{2}(\hat{n}_{ij\sigma} - s_{W^{ij\sigma}_{kl\sigma'}}\hat{n}_{kl\sigma'})^2 + \frac{|W^{ij\sigma}_{kl\sigma'}|}{2}(\delta_{ij}\hat{n}_{ij\sigma} + \delta_{kl}\hat{n}_{kl\sigma'}). \quad \text{(C9)}$$

Each term $\hat{H}_m$ ($m=ij\sigma$ or $ijkl\sigma\sigma'$) gives a contribution $A_m^{\Omega(n_{xy\rho})}$ to the drift term and the contributions $B_m^{n_{xy\rho}}$, $C_m^{n_{xy\rho}}$ to the diffusion term of the (stochastic) differential equations (C4) for Ω and $n_{xy\rho}$. No diffusion term appears in the equation of motion for Ω, and we can write

$$\frac{d\Omega}{d\tau} = \sum_m A_m^{\Omega}\Omega, \quad \text{(C10)}$$

$$\frac{n_{xy\rho}}{d\tau} = \sum_m (A_m^{n_{xy\rho}} + B_m^{n_{xy\rho}}\xi_m + C_m^{n_{xy\rho}}\xi'_m), \quad \text{(C11)}$$

where the ξ_m, ξ'_m are independent Gaussian random variables with variance $1/d\tau$. The hopping term (C8) yields only the drift contributions

$$A^{\Omega}_{ij\sigma} = t_{ij\sigma}n_{ij\sigma}, \quad \text{(C12)}$$

$$A^{n_{xy\rho}}_{ij\sigma} = \frac{t_{ij\sigma}}{2}[n_{xj\sigma}(\delta_{iy} - n_{iy\sigma}) + (\delta_{xj} - n_{xj\sigma})n_{iy\sigma}]\delta_{\rho\sigma}. \quad \text{(C13)}$$

With the gauge choice (C9), the interaction terms yield a Fokker-Planck equation of the form of Eqs. (C2) and (C3) with drift terms

$$A^{\Omega}_{ijkl\sigma\sigma'} = -W^{ij\sigma}_{kl\sigma'}n_{ij\sigma}n_{kl\sigma'} + W^{ij\sigma}_{kl\sigma'}n_{il\sigma}n_{kj\sigma'}\delta_{\sigma\sigma'}, \quad \text{(C14)}$$

$$A^{n_{xy\rho}}_{ijkl\sigma\sigma'} = \frac{|W^{ij\sigma}_{kl\sigma'}|}{2}\{[n_{xj\sigma}(\delta_{yi} - n_{iy\sigma}) + (\delta_{xj} - n_{xj\sigma})n_{iy\sigma}] \times (n_{ij\sigma} - s_{W^{ij\sigma}_{kl\sigma'}}n_{kl\sigma'} - \delta_{ij}/2)\delta_{\rho\sigma} + [n_{xl\sigma'}(\delta_{ky} - n_{ky\sigma'}) + (\delta_{xl} - n_{xl\sigma'})n_{ky\sigma'}] \times (n_{kl\sigma'} - s_{W^{ij\sigma}_{kl\sigma'}}n_{ij\sigma} - \delta_{kl}/2)\delta_{\rho\sigma'}\}, \quad \text{(C15)}$$

and diffusion terms

$$B^{n_{xy\rho}}_{ijkl\sigma\sigma'} = \sqrt{|W^{ij\sigma}_{kl\sigma'}|/2}[n_{xj\sigma}(\delta_{iy} - n_{iy\sigma})\delta_{\rho\sigma} - s_{W^{ij\sigma}_{kl\sigma'}}n_{xl\sigma'} \times (\delta_{ky} - n_{ky\sigma'})\delta_{\rho\sigma'}], \quad \text{(C16)}$$

$$C^{n_{xy\rho}}_{ijkl\sigma\sigma'} = \sqrt{|W^{ij\sigma}_{kl\sigma'}|/2}[(\delta_{xj} - n_{xj\sigma})n_{iy\sigma}\delta_{\rho\sigma} - s_{W^{ij\sigma}_{kl\sigma'}} \times (\delta_{xl} - n_{xl\sigma'})n_{ky\sigma'}\delta_{\rho\sigma'}]. \quad \text{(C17)}$$

Note that the right-hand side of Eq. (C10) is $-h_m(\boldsymbol{n})\Omega$, with $h_m(\boldsymbol{n}) = \mathrm{Tr}[\hat{\Lambda}(\boldsymbol{n})\hat{H}_m]$. Furthermore, since the $n_{ij\sigma}$ are real variables, which remain real during the integration of Eq. (C11), it follows from Eqs. (C12), (C14), and (C10) that the weight Ω will always stay positive.

For the actual implementation, it is simpler to use the Ito SDE, which is also numerically more stable. In this case, the drift terms have to be modified as[11]

$$A^{\mathrm{Ito}}_{\alpha} = A^{\mathrm{Stratonovich}}_{\alpha} + \frac{1}{2}\sum_{\beta} B_{\beta}\frac{\partial}{\partial\lambda_{\beta}}B^{(n)}_{\alpha} + \frac{1}{2}\sum_{\beta} C_{\beta}\frac{\partial}{\partial\lambda_{\beta}}C^{(n)}_{\alpha}, \quad \text{(C18)}$$

while the diffusion terms remain unaffected. The four-fermion term (C9) thus yields a contribution

$$A^{n_{xy\rho},\mathrm{Ito}}_{ijkl\sigma\sigma'} = -\frac{W^{ij\sigma}_{kl\sigma'}}{2}\{[n_{xj\sigma}(\delta_{iy} - n_{iy\sigma}) + (\delta_{xj} - n_{xj\sigma})n_{iy\sigma}]n_{kl\sigma'}\delta_{\rho\sigma} + [n_{xl\sigma'}(\delta_{ky} - n_{ky\sigma'}) + (\delta_{xl} - n_{xl\sigma'})n_{ky\sigma'}]n_{ij\sigma}\delta_{\rho\sigma'} - [(n_{xl\sigma}(\delta_{iy} - n_{iy\sigma}) + (\delta_{xl} - n_{xl\sigma})n_{iy\sigma})n_{kj\sigma} + (n_{xj\sigma}(\delta_{ky} - n_{ky\sigma}) + (\delta_{xj} - n_{xj\sigma})n_{ky\sigma})n_{il\sigma}]\delta_{\sigma\sigma'}\delta_{\rho\sigma}\} \quad \text{(C19)}$$

to the right-hand side of (C11).

We have tested the Ito-SDE for the two-site model

$$H = H_t + H_{\mu} + H_u + H_c + H_x + H_s + H_{h_1} + H_{h_2}, \quad \text{(C20)}$$

where the individual terms are

$$H_t = -t\sum_{\sigma}(\hat{n}_{12\sigma} + \hat{n}_{21\sigma}), \quad \text{(C21)}$$

$$H_{\mu} = \mu\sum_{\sigma}(\hat{n}_{11\sigma} + \hat{n}_{22\sigma}), \quad \text{(C22)}$$

$$H_u = \frac{1}{2}\sum_{\sigma}(u_1\hat{n}_{11\sigma}\hat{n}_{11-\sigma} + u_2\hat{n}_{22\sigma}\hat{n}_{22-\sigma}), \quad \text{(C23)}$$

$$H_c = v_c\sum_{\sigma}(\hat{n}_{11\sigma}\hat{n}_{22\sigma} + \hat{n}_{11\sigma}\hat{n}_{22-\sigma}), \quad \text{(C24)}$$

$$H_x = v_x\sum_{\sigma}\hat{n}_{12\sigma}\hat{n}_{21-\sigma}, \quad \text{(C25)}$$

$$H_s = v_s\sum_{\sigma}(\hat{n}_{12\sigma}\hat{n}_{12-\sigma} + \hat{n}_{21\sigma}\hat{n}_{21-\sigma}), \quad \text{(C26)}$$

$$H_{h_1} = h_1\sum_{\sigma}(\hat{n}_{11\sigma}\hat{n}_{12-\sigma} + \hat{n}_{11\sigma}\hat{n}_{21-\sigma}), \quad \text{(C27)}$$

$$H_{h_2} = h_2\sum_{\sigma}(\hat{n}_{22\sigma}\hat{n}_{12-\sigma} + \hat{n}_{22\sigma}\hat{n}_{21-\sigma}). \quad \text{(C28)}$$

The (stochastic) differential equations (C10) and (C11) were integrated using an implicit Euler scheme and adaptive time steps. Every n_{reconf} steps, the family of random walkers was reconfigured according to the method of Ref. 3. Although the high-temperature behavior is correctly reproduced, similar and even more severe problems as those observed in the simulations of the Hubbard model—most notably systematic errors in the energy at low temperature—also plague the

simulation of the two-site model with Coulomb interactions. For example, if either of the amplitudes v_c, v_x, v_s, h_1, or h_2 is of the order of the hopping t, deviations in energy from the exact solution appear around $\beta t \gtrsim 1$. We have not yet checked whether symmetry projections can be successfully applied in this case.

For realistic electronic structure calculations, we propose to start with a Hartree Fock calculation of the electronic structure problem and to use both the occupied and unoccupied Hartree Fock orbitals in a subsequent quantum Monte Carlo simulation. This is the same Hamiltonian and basis set used in full-CI (configuration interaction) or coupled cluster methods (CCM) in quantum chemistry, but the algorithm described here would enable to study a larger number of basis functions than in a full-CI calculation. Using the Hartree Fock density matrix instead of a multiple of the unit matrix as initial density matrix $\hat{\rho}(0)$ has the advantage that the initial energy is already that of the Hartree-Fock solution, which will then be lowered further by the projection in imaginary time. Already a projection over a short imaginary time β (which can now not be interpreted as inverse temperature) will give an energy lower than the Hartree-Fock solution.

[1] J. F. Corney and P. D. Drummond, Phys. Rev. Lett. **93**, 260401 (2004).

[2] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena: The International Series of Mono-Graphs on physics* (Clarendon Press, Oxford, 1996).

[3] M. Calandra Buonaura and S. Sorella, Phys. Rev. B **57**, 11446 (1998).

[4] P. Kloeden, E. Platen, and H. Schurz, *Numerical Solution of SDE Through Computer Experiments* (Springer-Verlag, Berlin, 1994), with 1 IBM-PC floppy disk (3.5 in.; HD).

[5] T. Mizusaki and M. Imada, Phys. Rev. B **69**, 125110 (2004).

[6] E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra: Pure and Applied Physics* (Academic Press, New York, 1962).

[7] A. Parola, S. Sorella, M. Parrinello, and E. Tosatti, Phys. Rev. B **43**, 6190 (1991).

[8] J. W. Negele and H. Orland, *Quantum Many Body Systems: Frontiers in Physics* (Addison-Wesley, Reading, MA, 1988).

[9] R. B. Laughlin and D. Pines, Proc. Natl. Acad. Sci. U.S.A. **97**, 28 (2000).

[10] J. F. Corney and P. D. Drummond, cond-mat/0411712 (unpublished).

[11] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, New York, 1985).

[12] In principle, we could also set $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{y}$. However since $\boldsymbol{\sigma}^{y}$ is purely imaginary this leads to complex drift and stochastic forces and, in turn, to negative weights.

[13] It is known that weakly doped Mott insulators have many competing states or, in other words, a dense spectrum of low-lying states. Hence, at finite but low temperatures, the density matrix will contain excited states.