PHYSICAL REVIEW B **105**, 045107 (2022)

Deconvolving the components of the sign problem

S. Tarat,[1,*] Bo Xiao,[2,†] R. Mondaini,[1,‡] and R. T. Scalettar[3,§]
[1]*Beijing Computational Science Research Center, Beijing 100193, China*
[2]*Center for Computational Quantum Physics, Flatiron Institute, New York, New York 10010, USA*
[3]*Department of Physics, University of California, Davis, California 95616, USA*

(Received 9 August 2021; revised 10 November 2021; accepted 21 December 2021; published 4 January 2022)

Auxiliary field quantum Monte Carlo simulations of interacting fermions require sampling over a Hubbard-Stratonovich field h introduced to decouple the interactions. The weight for a given configuration involves the products of the determinant of matrices $M_\sigma(h)$, where σ labels the species, and hence is typically not positive definite. Indeed, the average sign $\langle \mathcal{S} \rangle$ of the determinants goes to zero exponentially with increasing spatial size and decreasing temperature for most Hamiltonians of interest. This statement, however, does not explicitly separate two possible origins for the vanishing of $\langle \mathcal{S} \rangle$. Does $\langle \mathcal{S} \rangle \to 0$ because *randomly* chosen field configurations have $\det[M(h)] < 0$, or does the sign problem arise because the specific subset of configurations chosen by the weighting function have a greater preponderance of negative values? In the latter case, the process of weighting the configurations with $|\det[M(h)]|$ might steer the simulation to a region of configuration space of h where positive and negative determinants are equally likely, even though randomly chosen h would preferentially have determinants with a single dominant sign. In this paper, we address the relative importance of these two mechanisms for the vanishing of $\langle \mathcal{S} \rangle$ in quantum simulations.

DOI: 10.1103/PhysRevB.105.045107

I. INTRODUCTION

Auxiliary field quantum Monte Carlo (AFQMC) [1–8] relies on the observation that traces over products of exponentials of quadratic forms of fermionic operators can be done analytically. Thus, in QMC, for a Hamiltonian like the Holstein model, where only such quadratic forms are present, the resulting simulation samples over the remaining space and imaginary time-dependent bosonic (phonon) degrees of freedom $x(r, \tau)$, using a weight which combines a boson action $S_{\rm Bose}$ and the fermion determinants (one for each fermionic species). If, as in the Hubbard model, quartic (interaction) terms in the fermions are present, they are decoupled via an auxiliary field $h(r, \tau)$. In either case, after the fermionic trace is performed, the sampling is now over these classical fields and may be implemented by utilizing a variety of standard numerical techniques.

With the proliferation of computing resources over the last few decades, such QMC simulations have become indispensable tools for investigating difficult problems involving strong correlations in a variety of topics in condensed matter [9,10], high energy [11] and nuclear physics [12] as well as in chemistry [13,14], providing many breakthroughs in these fields. These successes notwithstanding, a pervasive problem afflicting such methods, limiting their scope of application considerably, is the sign problem (SP), which occurs when the fermion determinants become negative for certain quantum configurations, leading to a negative probability. This has given rise to a considerable body of research aimed at solving, or alleviating, the SP [8,15–30]. Nevertheless, it remains unsolved, and one of the central issues in this regard is to understand how the underlying physics of the problem in consideration affects the SP. In the case of AFQMC, this is intimately related to the individual configurations of the bosonic or auxiliary fields and how they affect the sign of the fermion determinant.

The fermion determinant itself sums over all the possible quantum mechanical world lines of fermions moving in the instantaneous value of the physical bosonic field or the artificially introduced auxiliary field. For a particular world line configuration, if the fermions wind around each other an odd number of times, the contribution to the determinant is negative. This picture provides one view of the origin of the SP: to the extent that different regions of the space-imaginary time lattice are uncorrelated, there is a constant density of winding, and one expects that, as the spatial lattice size N and inverse temperature β grow, the likelihood of positive and negative world line configurations becomes equal, and the average sign vanishes exponentially with both β and N.

The assumption that different regions are uncorrelated is nontrivial. One motivation is the absence of any intrinsic dynamics in the auxiliary field: $h(r, \tau)$ couples to the fermionic degrees of freedom, but different components of $h(r, \tau)$ are not coupled to each other. Indeed, it is known that, if $h(r, \tau)$ do interact, the SP can be mitigated [1,31,32]. As an extreme example, if h has no τ dependence, the sign of the determinant is positive. Similarly, in electron-phonon models, the phonon kinetic energy $\hat{p}^2/2m$ induces correlations in $x(r, \tau)$ on

---

[*] tarats@csrc.ac.cn

[†] bxiao@flatironinstitute.org

[‡] rmondaini@csrc.ac.cn

[§] scalettar@physics.ucdavis.edu

adjacent τ by penalizing imaginary time fluctuations, especially at low phonon frequencies ω. Although the details are complex, typically, one expects the SP to be reduced.

Symmetries might allow simulations to avoid the SP in special situations [15,33,34]. The simplest scenario is one in which there are two fermionic species (e.g., spin up and spin down) which couple to the auxiliary or bosonic field in the same way. Then if they also share a common structure in the other pieces of the Hamiltonian (the same hopping and chemical potential terms, for example), the two matrices arising when the species are traced out are identical. Although the individual determinants can (and do) change sign, their product is always a square and hence positive. Such situations are, however, not generic, and in most QMCs, the SP is significant.

This traditional argument [15] makes no explicit reference to how the fields are selected and hence suggests that the average sign of the determinants of *randomly chosen* $h(r,\tau)$ should vanish. The intent of this paper is to investigate this issue further and quantitatively. Before doing so, it is useful to make some analogies with Monte Carlo for classical degrees of freedom.

In classical statistical mechanics, the expectation value of an observable $\mathcal{A}$ takes the following form:

$$\langle\mathcal{A}\rangle = \mathcal{Z}^{-1}\,\mathrm{Tr}_{\{h\}}\,\mathcal{A}(\{h\})\,e^{-\beta E(\{h\})},$$
$$\mathcal{Z} = \mathrm{Tr}_{\{h\}}\,e^{-\beta E(\{h\})}, \tag{1}$$

where $E(\{h\})$ is the energy of the system described by some collection of degrees of freedom $\{h\}$, and β is the inverse temperature. Implicit in the structure of Eq. (1) is that $\mathcal{A}$ does not depend on β. However, certain $\mathcal{A}$ do have (trivial) β dependence. For example, in the paramagnetic phase of the classical Ising model, the magnetic susceptibility $\chi = \beta\sum_{ij}\langle S_iS_j\rangle$, where $S_i = \pm 1$ are the Ising spins. Similarly, energy-fluctuation-based measurements of the specific heat measure the observable $C = \beta^2(\langle E^2\rangle - \langle E\rangle^2)$. In such cases, β comes out of the trace over the degrees of freedom. Its presence does not affect the critical properties. In the case of the susceptibility, as one approaches T_c, the β factor merely multiplies the sum of the spin-spin correlations $\langle\sum_{ij} S_iS_j\rangle$ by β_c but does not alter the power law describing its divergence.

In such a situation, a rescaling of the weight $\mathcal{W} = e^{-\beta E} \to (e^{-\beta E})^g$ merely amounts to a shift in inverse temperature $\beta \to g\beta$ (or equivalently, in the energy scales $E \to gE$) in the calculation of $\langle\mathcal{A}\rangle$. This analogy also makes clear that $g \neq 1$ will alter the expectation values measured since it changes the temperature of the simulation. As we shall see in detail below, in QMC, the situation is more complex. There, the observable $\hat{\mathcal{A}}$ can depend in a complicated way on the inverse temperature. As a consequence, the change in $\langle\hat{\mathcal{A}}\rangle$ upon a weight rescaling $\mathcal{W} \to \mathcal{W}^g$ can be highly nontrivial.

The rescaling parameter g has the effect of tuning the configurations sampled in a simulation. Here, $g = 1$ gives the appropriate expectation values for the energy (Hamiltonian) in question. The limit $g = 0$ makes all configurations of $\{h\}$ equally likely. In this paper, we investigate the effect of tuning g in determinant QMC (DQMC) [1]. Focusing on the average sign $\langle\mathcal{S}\rangle$, our purpose here is to understand whether the fermion SP in which $\langle\mathcal{S}\rangle \to 0$ occurs because the value of the sign itself for random configurations is becoming increasingly badly behaved as β increases, as suggested by the winding argument above, or whether the sampling is preferentially guiding the system to a region in phase space where the weighted configurations yield vanishing average sign. To this end, we analyze three different realizations of the Hubbard model, viz., on the square lattice, on the honeycomb lattice, and the ionic Hubbard model with a staggered potential, investigating the average sign as well as several physical variables, over a large range of parameter values such as the chemical potential μ and the interaction U, as we vary g systematically.

Our results show that the SP depends on the weight and temperature in a nontrivial manner. In the square lattice model, we find that the SP originates mainly from the the fact that *weight guides the simulation to regions of phase space with a low average sign*. In this situation, random Hubbard-Stratonovich (HS) fields result in a larger value of the overall sign at all temperatures and densities. In the honeycomb lattice, on the other hand, the SP becomes progressively worse for random sampling as the system approaches the interaction-induced antiferromagnetic Mott insulator (AFMI), underscoring the complex relationship between the origin of the SP and the underlying physical states in DQMC. The ionic Hubbard model shows qualitatively similar results; here, the sign becomes worse again for random sampling as we navigate across the correlated metal (CM) phase and approach the AFMI state. In addition, analysis of various physical quantities as well as the detailed nature of the sign curves reveals that, in many respects, a reduced g pushes the system to weaker coupling. Nevertheless, the system retains nontrivial signatures of the interaction in certain characteristics even in the fully random $g = 0$ case.

Next, we briefly consider the complementary oversampling case where $g > 1$ for the square lattice model. In this case, signatures of a stronger effective coupling as g is increased are even more prominent, as the Mott plateau around half-filling becomes progressively stronger, leading to a swift reduction in the SP accompanied by a shift in the sign minimum as g is increased. We end with a quantitative analysis of the effective interaction with changing g, confirming the qualitative observations made earlier.

II. DQMC AND THE HUBBARD MODEL

We investigate the SP within the context of the Hubbard Hamiltonian [35,36]:

$$\hat{H} = \hat{K} + \hat{V},$$
$$\hat{K} = -t\sum_{\langle ij\rangle}(c^\dagger_{j\sigma}c_{i\sigma} + c^\dagger_{i\sigma}c_{j\sigma}) - \sum_i \mu_i(n_{i\uparrow} + n_{i\downarrow}),$$
$$\hat{V} = U\sum_j\left(n_{j\uparrow} - \frac{1}{2}\right)\left(n_{j\downarrow} - \frac{1}{2}\right). \tag{2}$$

Here, $c^\dagger_{j\sigma}(c_{j\sigma})$ are fermion creation(destruction) operators at spatial site j and with spin σ; $n_{j\sigma} = c^\dagger_{j\sigma}c_{j\sigma}$ is the number operator; t is hopping matrix element between nearest neighbor sites; μ is the chemical potential; and U is an onsite repulsion. We set $t = 1$ as our unit of energy.

As remarked earlier, we will explore several different contexts, beginning with the two-dimensional square lattice at constant $\mu_i = \mu$, which is the most famous geometry owing to its relevance to cuprate physics [37]. We will then simulate the honeycomb lattice to understand if the effect of g is linked in any way to the quantum critical point $U_c/t \simeq 3.8$ which separates antiferromagnetic from semimetallic behavior [38–43]. Finally, we will turn to the ionic Hubbard model [$\mu_i = (-1)^i \mu$] on a square lattice to further test our conjectures.

We use a commonly employed DQMC formulation [1,2]. In computing the partition function, the inverse temperature is discretized $\beta = L_\tau \Delta\tau$, and the Trotter approximation is used to separate out the exponentials of the kinetic and potential energies. Thus, the partition function may be written as

$$\begin{aligned}\mathcal{Z} &= \mathrm{Tr}(e^{-\beta \hat{H}}) \\ &= \mathrm{Tr}(e^{-\Delta\tau \hat{H}})^{L_\tau} \\ &\approx \mathrm{Tr}[e^{-\Delta\tau \hat{K}(1)} e^{-\Delta\tau \hat{V}(1)} e^{-\Delta\tau \hat{K}(2)} e^{-\Delta\tau \hat{V}(2)} \cdots \\ &\quad e^{-\Delta\tau \hat{K}(L_\tau)} e^{-\Delta\tau \hat{V}(L_\tau)}]. \end{aligned} \tag{3}$$

The quartic onsite interaction term $\hat{V}$ is now decoupled by a discrete HS transformation [44]:

$$\begin{aligned}&\exp\left[-\Delta\tau U\left(n_{i\uparrow} - \frac{1}{2}\right)\left(n_{i\downarrow} - \frac{1}{2}\right)\right] \\ &\quad = \mathcal{C} \sum_{h_i = \pm 1} \exp\left[\alpha h_i (n_{i\uparrow} - n_{i\downarrow})\right], \end{aligned} \tag{4}$$

where $\cosh(\alpha) = \exp(\Delta\tau U/2)$, $\mathcal{C} = \frac{1}{2}\exp(-\Delta\tau U/4)$, and h_i are discrete classical variables that only take values ± 1. This converts a quartic fermionic term into a quadratic one, while adding a sum over the new variables h_i. Introducing this transformation for each lattice site i at each time slice l, the partition function may be rewritten as

$$\mathcal{Z} = \mathcal{C}^{NL_\tau} \mathrm{Tr}_{\{h\}} \mathrm{Tr}\big[e^{\Delta\tau \hat{K}} e^{\hat{\mathcal{V}}(1)}\, e^{\Delta\tau \hat{K}} e^{\hat{\mathcal{V}}(2)} \cdots e^{\Delta\tau \hat{K}} e^{\hat{\mathcal{V}}(L_\tau)}\big], \tag{5}$$

where N denotes the number of lattice sites, the operator $\hat{\mathcal{V}}_{ij\sigma}(l) = c^\dagger_{i\sigma}\, v^{ij}_\sigma(l)\, c_{j\sigma}$, with $v^{ij}_\sigma(l) = \alpha\, \sigma h_i(l)\, \delta_{ij}$, where $\sigma = \pm 1$, and $\hat{K}_{ij\sigma}(l) = c^\dagger_{i\sigma} k^{ij}_\sigma(l) c_{j\sigma}$, $k^{ij}_\sigma(l) = (t)\delta_{\langle ij \rangle} + \mu_i \delta_{ij}$. For a given configuration $\{h\}$, the terms in the exponentials are all quadratic, allowing us to perform the fermionic trace analytically, yielding [1,2]

$$\mathcal{Z} = \mathcal{C}^{NL_\tau} \mathrm{Tr}_{\{h\}} [\mathrm{det} M_\uparrow(h)][\mathrm{det} M_\downarrow(h)], \tag{6}$$

where the matrix $M_\sigma(h) = [I + \prod_l e^{\Delta\tau k} e^{v_\sigma(l)}]$.

All physical observables can be expressed in terms of the fermion Green's function $\mathcal{G}_{\sigma,ij} = \langle c_{i\sigma} c^\dagger_{j\sigma} \rangle = M^{-1}_{\sigma ij}$. For example, for the fermion density on site i with spin σ, $n_{i\sigma} = 1 - \mathcal{G}_{\sigma,ii}$, the kinetic energy (excluding the chemical potential) $\langle \mathcal{K} \rangle = (8t)\mathcal{G}_{i\,i+\hat{x},\sigma}$, where the factor of eight arises from the two spin species, the two directions x, y to hop, and the Hermitian conjugate pair associated with hopping $i \leftrightarrow j$. Finally, the pair correlator $P^d_{ij} = \langle \Delta^d_i \Delta^{d\dagger}_j \rangle$, where $\Delta^{d\,\dagger}_j = c^\dagger_{j\uparrow}(c^\dagger_{j+\hat{x}\downarrow} - c^\dagger_{j+\hat{y}\downarrow} + c^\dagger_{j-\hat{x}\downarrow} - c^\dagger_{j-\hat{y}\downarrow})$ for d-wave symmetry.

It is important for us to reemphasize the goal of this paper in introducing the parameter g. We are *not* seeking to find an improved importance sampling scheme which would reduce the SP and associated error bars while leaving physical observables with the same expectation values. Such work has been productively undertaken by several groups in the community, especially within the context of constrained path AFQMC [45–50]. Instead, our objective is solely focused on gaining insight into the origin of *SP itself* and isolating whether it is better or worse for *randomly selected* field configurations compared with those chosen preferentially according to the weight (fermion determinants).

III. WEAK AND STRONG COUPLING LIMITS

In the $U = 0$ limit, the matrices M_σ and their inverses, the Green's functions $\mathcal{G}_\sigma$, are independent of the HS field configuration. As a consequence, expectation values of any observable $\hat{A}$ are exact and also independent of g. The proof is straightforward:

$$\begin{aligned}\langle \hat{A} \rangle &= \mathcal{Z}^{-1} \int \mathcal{D}h \hat{A}(h)[\mathrm{det} M_\uparrow(h) \mathrm{det} M_\downarrow(h)]^g, \\ \mathcal{Z} &= \int \mathcal{D}h [\mathrm{det} M_\uparrow(h) \mathrm{det} M_\downarrow(h)]^g \\ \Rightarrow \langle \hat{A} \rangle &= \frac{\int \mathcal{D}h \hat{A}(h)}{\int \mathcal{D}h}. \end{aligned} \tag{7}$$

Here, the notation $\int \mathcal{D}h$ is used as a general symbol for summation or integration over the degrees of freedom h, including both the cases where they represent continuous variables (for a continuous HS transformation) and discrete ones (as in our calculations below).

The strong coupling (single site) limit $t = 0$ is more interesting. In this case, the sites are completely decoupled, and the solution reduces to a product of single site calculations. The partition function for a general value of g is given by

$$\mathcal{Z} = \mathcal{C}^{NL_\tau} \prod_i \sum_{\{h_i\}} \left(\prod_\sigma \left\{ 1 + \exp\left[\beta\mu_i + \alpha\sigma \sum_l h_i(l) \right] \right\} \right)^g. \tag{8}$$

This can be evaluated analytically at $g = 1$ (and trivially at $g = 0$). The general case with $g \neq 1$ can be easily evaluated numerically. In Appendix E, we show detailed results for the number density $\langle n \rangle$ and the double occupancy $\langle n_\uparrow n_\downarrow \rangle$ for several values of g, U, and β. These results show that this simple case already demonstrates aspects of the nontrivial effects of varying the nature of the high probability configurations in such problems.

IV. HUBBARD MODEL ON A SQUARE LATTICE

In this section, we consider the Hubbard model on a square lattice of size L with a constant $\mu_i = \mu$. As is well known, at half-filling, the onsite interaction U results in an AFMI. Away from half-filling, the possibility of superconducting correlations mediated by spin fluctuations makes this iconic Hamiltonian relevant as a prospective model for high-temperature superconductors [37]. Thus, it provides a rich background for investigating the effect of the rescaling

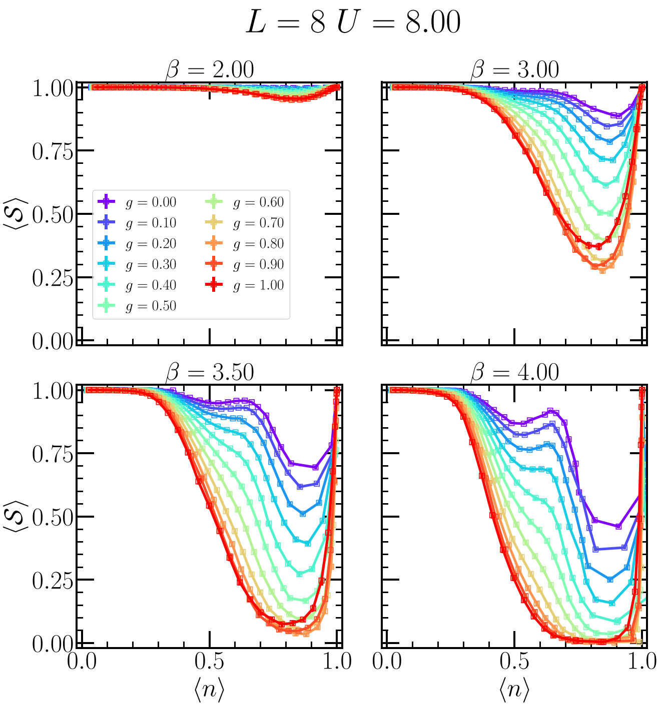

FIG. 1. The average sign $\langle \mathcal{S} \rangle$ vs $\langle n \rangle$ at $U = 8$ as a function of g on an 8×8 lattice for four values of β. As β increases, the minimum in $\langle \mathcal{S} \rangle$ becomes wider and deeper, as expected, displaying a minimum around $\langle n \rangle_c \sim 0.85$. While the sign remains close to 1 for $g \to 0$ at $\beta \lesssim 3$, it starts reducing noticeably as β is increased further, eventually forming a valley-peak-valley structure, as seen in the bottom panels. This suggests that, while the bulk of the sign problem at low temperature originates from the restricted configuration space available to the system, random configurations of the Hubbard-Stratonovich (HS) fields also result in a perceptible reduction in $\langle \mathcal{S} \rangle$ at low enough temperatures. The bottom panels, where the sign is the worst, show a sharp upturn near half-filling since the total sign is constrained to be equal to 1 at $\mu = 0$ due to particle-hole symmetry.

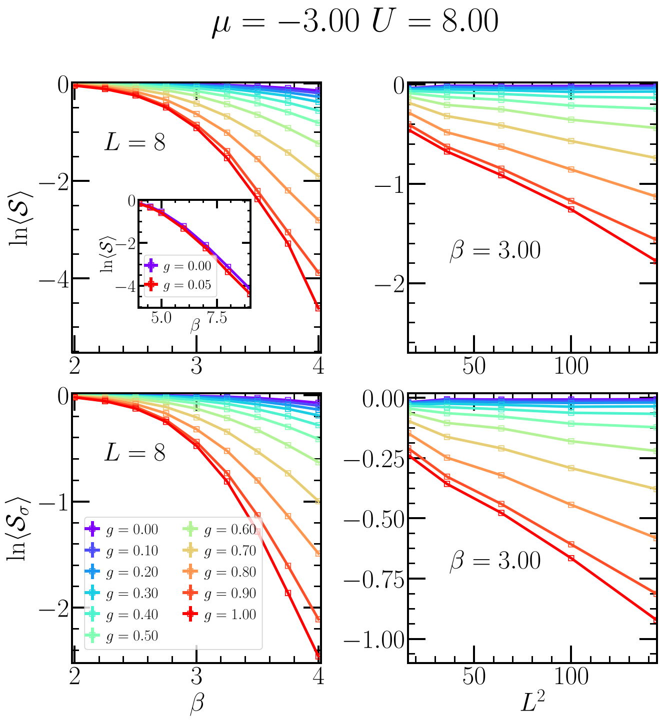

FIG. 2. Scaling of the total and spin-resolved sign with temperature and system size at $\mu = -3.0$, $U = 8$ on an 8×8 lattice. Left column shows $\ln(\langle \mathcal{S} \rangle)$ and $\ln(\langle \mathcal{S}_\sigma \rangle)$ vs the inverse temperature β for a range of values of g. We find a faster-than-exponential reduction in the sign with decreasing temperature, both for the total and spin-resolved quantities. The inset in the top panel extends this to temperatures $\beta \sim 9$ for $g = 0.0$ and 0.05, demonstrating that even the randomly sampled case shows a faster-than-exponential scaling over a large temperature range. The right column plots the same quantities as a function of the lattice size L^2. The scaling in this case is exponential for all g values.

parameter g on the evolution of the system in various parameter regimes.

Figure 1 shows the average of the total sign $\langle \mathcal{S} \rangle$ vs the total density $\langle n \rangle$ for different values of the rescaling parameter g at four different values of β. For $\beta \leqslant 3$ (top panels), it is seen that, as $g \to 0$, $\langle \mathcal{S} \rangle \to 1$, with only a small dip around $\langle n \rangle \sim 0.8$. These plots suggest that, at temperatures that are not too low, simulations which sample with randomly chosen HS field configurations tend to have a considerably reduced SP. On the other hand, at lower temperatures $\beta \gtrsim 3$ (lower panels), we find that, while the sign continues improving systematically as g is reduced from 1 toward 0, the value of $\langle \mathcal{S} \rangle$ around its minimum is decreased considerably compared with larger temperatures, even near $g = 0$. In addition, we find an emerging valley-peak-valley structure as $g \to 0$, reminiscent of the shell effect in noninteracting finite-sized lattices. In Appendix D, we explore this connection more carefully by considering even lower temperatures and demonstrate that the values of $\langle n \rangle$ at the maxima of the sign correspond to the locations of the density steps on a noninteracting lattice of the same size due to the shell effect [51]. An additional feature of the data, seen at all temperature values presented here, is that there is an initial *worsening* of the SP for $0.8 \lesssim g < 1$.

As explained earlier, the traditional argument about the connection of the sign with the fermionic world lines suggests that, with lowering temperature (i.e., increasing β) and increasing system size, the fermion world lines are more likely to wind around each other and provide determinants of both sign more frequently, resulting in a progressively worse SP on the average. To see how the choice of g affects the scaling properties of the sign, we show in Fig. 2 the scaling properties of $\langle \mathcal{S} \rangle$ as well as the spin resolved sign $\langle \mathcal{S}_\sigma \rangle = (\langle \mathcal{S}_\uparrow \rangle + \langle \mathcal{S}_\downarrow \rangle)/2$, with increasing lattice size $N = L^2$ and inverse temperature β. We find that, for $g \lesssim 1$, where the HS fields are sampled according to the correct thermal weights, both show a faster-than-exponential drop with increasing β, as found in earlier work [27]. As g is reduced, the overall sign gradually improves at all β values, but the reduction with increasing β remains faster than exponential. The plots near $g = 0$ seem to show little change in this temperature range, but as the inset demonstrates, at smaller temperatures, they show the same faster-than-exponential reduction. The scaling with system size at fixed temperature, as seen in the right column, is exponential. Again, we find that, while $\langle \mathcal{S} \rangle$ is less severe

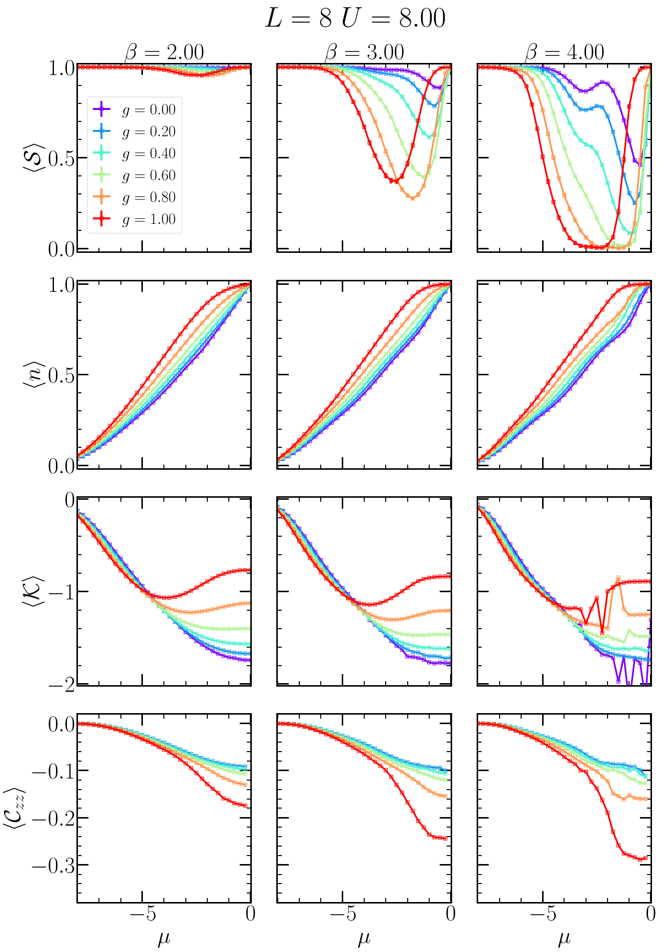

FIG. 3. Plots of the average sign $\langle \mathcal{S} \rangle$ (first row) and observables, including the density $\langle n \rangle$ (second row), kinetic energy $\langle \mathcal{K} \rangle$ (third row), and nearest neighbor spin-spin correlator C_{zz} (fourth row), vs μ as a function of g for three different values of β. These physical observables clearly demonstrate signatures of a reduced effective interaction (see text for more details).

at lower g values, it continues to reduce exponentially with increasing lattice size.

These results indicate that, while the SP (both the total value as well as the spin-resolved ones) itself is dominated by the reduced phase space at lower temperatures and larger lattice sizes, the scaling dependence on both parameters remains very similar throughout the whole range of the sampling parameter g.

Near $g = 1$, the thermodynamic sampling restricts the HS field configurations to a suitably confined region of the total phase space, reflecting the underlying physics, especially at low temperatures. As g is reduced from 1, the increasingly random sampling occupies progressively larger regions of phase space, affecting the physical properties of the system. To study this more carefully, we show a number of physical variables, including the number density $\langle n \rangle \equiv \frac{1}{N} \sum_{i,\sigma} \langle n_{i\sigma} \rangle$, the kinetic energy $\langle \mathcal{K} \rangle$, and the nearest neighbor spin-spin correlations $C_{zz} \equiv \frac{1}{4} \langle (n_{i\uparrow} - n_{i\downarrow})(n_{j,\uparrow} - n_{j,\downarrow}) \rangle$ (i, j are nearest neighbors), as a function of the chemical potential μ in Fig. 3 for three temperature values.

The second row demonstrates some of this physics for $g = 1$ by plotting $\langle n \rangle$ vs μ. As the temperature decreases, the number density saturates close to half-filling as the AFMI sets in and the compressibility goes to zero, triggering a Mott plateau. As g is reduced and the HS fields are sampled more and more randomly, we find that this physics disappears slowly; the plots near $g = 0$ do not show any signs of saturating. Instead, we find gentle oscillations in $\langle n \rangle$, reminiscent of the oscillations we saw in the sign for similar parameter values in Fig. 1. As mentioned above, we demonstrate in Appendix D that this is due to the reemergence of finite-sized effects from the noninteracting problem.

In the third row, we show the kinetic energy $\langle \mathcal{K} \rangle$ for the same parameter values. Again, we see that, while the $g = 1$ results demonstrate the expected reduction in the kinetic energy as the density increases and the electrons start to avoid each other due to the onsite coupling U, lower values of the sampling parameter nullify this effect. We emphasize, however, that even at $g = 0$, the system is very different from a noninteracting one in many respects, as seen, for instance, in the enlarged bandwidths in these results.

The fourth row, which plots C_{zz}, also demonstrates the same suppression of the physics with lowering g. At $g = 1$, as the system approaches half-filling and the Mott state begins to appear, the nearest neighbor spin correlations become negative, their magnitude increasing with lowering temperature as the antiferromagnetic order becomes stronger. The weakening Mott insulator that results as g is lowered leads to a corresponding reduction in the spin correlations, as is clearly evident from the plots.

Hence, we see that, as g is reduced, the physical variables clearly demonstrate signatures of a reduced effective interaction in some aspects while retaining nontrivial signatures of the full interaction in others. This is not unreasonable, as the HS fields act as proxies for the electron-electron interactions [44], and randomizing them progressively is expected to dilute the effect of U. On the other hand, the matrix elements in the simulations still contain exponential factors that include the full interaction strength, and hence, certain features, such as the bandwidth, essentially retain their $g = 1$ values throughout. In Sec. VIII, we will quantify this relationship between the effective interaction $U_{\text{eff}}(g)$ and U by analyzing physical observables such as the double occupancy $\langle n_\uparrow n_\downarrow \rangle$ and the kinetic energy $\langle \mathcal{K} \rangle$ in an attempt to put these observations on a firmer footing.

Now we return to the behavior of the sign again, focusing on the temperature dependence at fixed g as well as the g dependence at fixed μ. In Fig. 4, we concentrate on the temperature dependence of the sign at fixed representative values of g. As expected, we find that the sign worsens with decreasing temperature for all g values. In the lower panels, for larger values of g, the plots are smooth and rather similar. However, the top panels with $g < 0.4$ show clear evidence of finite-size-based oscillations at lower temperatures, as remarked earlier in this paper.

In Fig. 5, we replot the data with the sign $\langle \mathcal{S} \rangle$ as a function of the parameter g with temperature at different values of μ. As the top right and lower panels with $\mu \lesssim -2.0$ show, increasing g leads to a worsening of the sign at all temperatures. In contrast, the top left panel with $\mu = -0.5$ close to

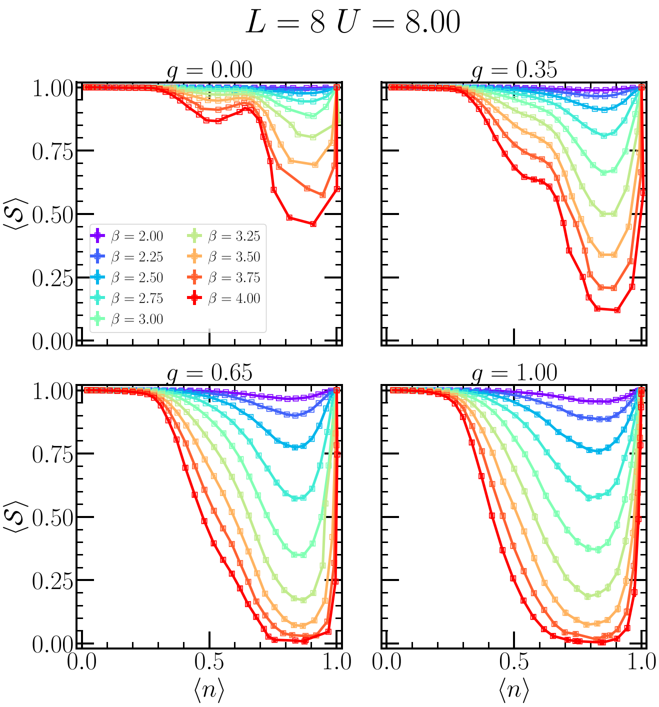

FIG. 4. The average sign $\langle \mathcal{S} \rangle$ vs $\langle n \rangle$ at $U = 8$ on an 8×8 lattice as a function of β for four different g values. Upper panels for $g < 0.4$ demonstrate finite-sized oscillations at lower temperatures.

half-filling shows a nonmonotonic behavior where the sign first reduces, reaches a minimum at a moderate value, and then increases continuously to unity as $g \to 1$, as the approaching Mott insulating state pins the density $\langle n \rangle \sim 1$, establishing particle-hole symmetry and mitigating the SP.

Thus, we find that, on the whole, the SP in the standard square lattice Hubbard model is mainly due to the fermion determinants steering the simulation to regions of low average sign. Random sampling typically mitigates the problem substantially over a large parameter window, even though it reappears at sufficiently low temperatures. The random fields tend to reduce the effect of the interaction, pushing the system to weaker coupling, resulting in a reemergence of certain noninteracting features such as the shell effect as well as reducing the signatures of the AFMI in observables such as the density, the kinetic energy, and the spin-spin correlations. Even so, this comparison only holds for certain characteristics, and such systems retain several nontrivial signatures of interactions even at $g = 0$.

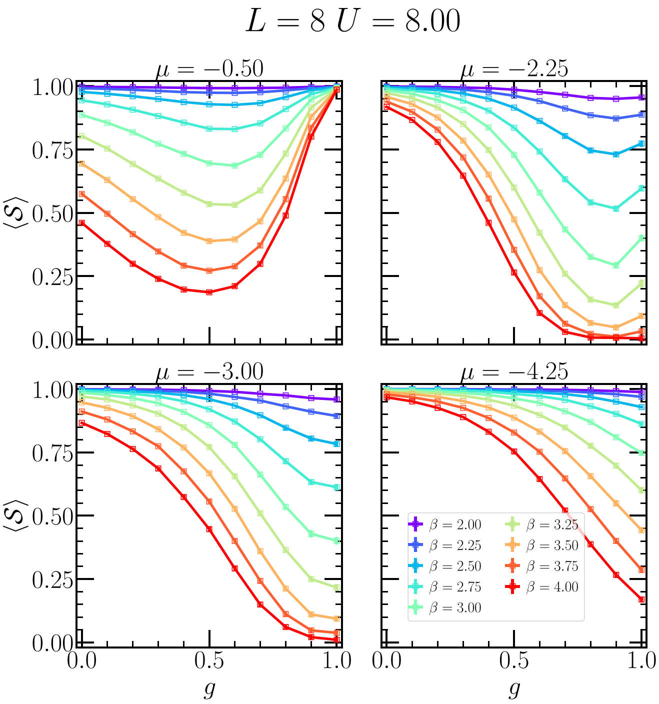

FIG. 5. The average sign $\langle \mathcal{S} \rangle$ vs g at $L = 8$, $U = 8$ as a function of β for four different μ values.

V. HUBBARD MODEL ON A HONEYCOMB LATTICE

In the previous section, we analyzed how the rescaling parameter g affected the SP and physical variables of the square lattice Hubbard Hamiltonian. While this model has a very rich phenomenology, the Hubbard model on the honeycomb lattice provides added insight into the sampling problem by presenting a sharply defined quantum phase transition.

At $U = 0$, the Hubbard model on a honeycomb lattice is a semimetal whose energy $\epsilon(k)$ disperses linearly with the momentum k close to special points on the k-space called Dirac points. This leads to a semimetallic density of states which also varies linearly with energy ω near $\omega = 0$. As U is turned on, the semimetal state persists up to a critical coupling value $U = U_c \sim 3.8$ [38–43], where it undergoes a quantum phase transition into an AFMI, unlike the square lattice version, which shows antiferromagnetic order at any nonzero value of U.

In Fig. 6, we plot the average sign $\langle \mathcal{S} \rangle$ on a lattice with $L = 9$ at a small nonzero $\mu = 0.2$ (the model is SP free at half-filling) vs the coupling U as a function of g for four different values of the temperature. The $g = 1$ plots deviate from their limiting values of $\langle \mathcal{S} \rangle \to 1$ in a broad region around U_c, tracking the transition, as we have shown in detail elsewhere [52]. As g is reduced, we find that the dip in the sign becomes more pronounced, in stark contrast to the results of the square lattice Hubbard model where this led to a reduction in the SP, and the minimum shifts to higher values of U. An intuitive explanation of the latter result follows from the observation we have already made earlier: as the sampling becomes more random, the effect of U is reduced. As a result, the semimetal-to-AFMI transition here is pushed to higher values of U. This does not lead to a reduction in the sign, however, as full inclusion of the fermion determinants ($g = 1$) in this case evidently leads the system to a phase space region (corresponding to the AFMI phase) where the SP is less severe than $g = 0$, where the determinants are ignored. Since the AFMI phase is characterized by a finite regime where $\langle n \rangle \sim 1$ is saturated at 1, reinstating particle-hole symmetry (originally broken by the small but nonzero μ), the sign may be expected to be robust in this phase.

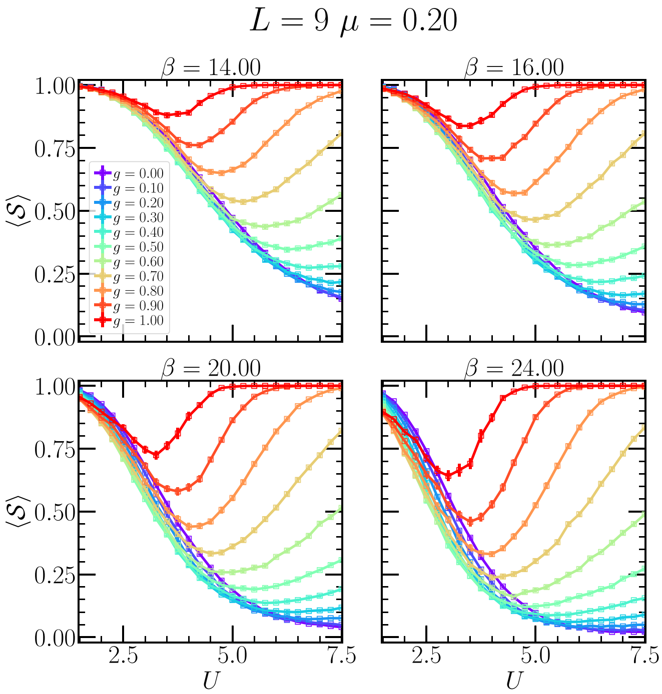

FIG. 6. The average sign $\langle \mathcal{S} \rangle$ vs U at $\mu = 0.2$, for the Hubbard model on a honeycomb lattice with $L = 9$, with varying U as a function of g, for four values of β. At $g = 1$, $\langle \mathcal{S} \rangle$ shows a minimum around U_c. Reducing g lowers the sign, shifting the minimum to the right (see text for details).

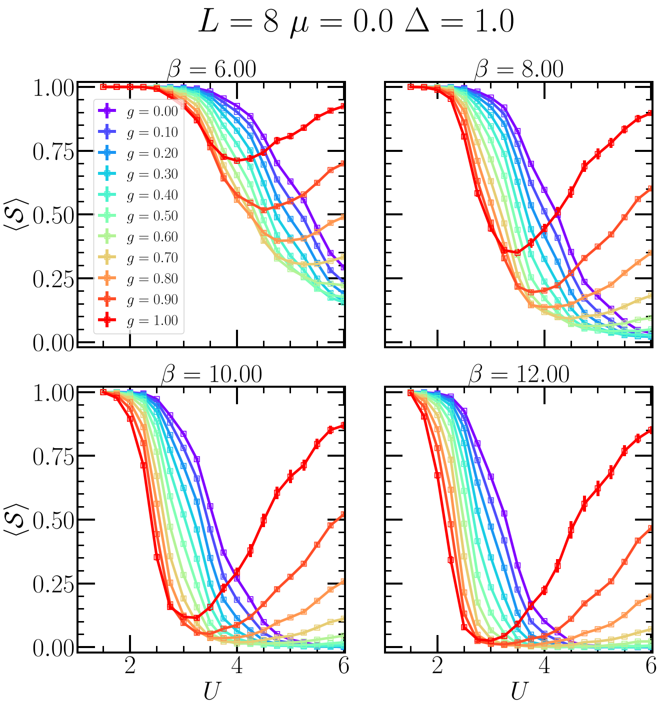

FIG. 7. The average sign $\langle \mathcal{S} \rangle$ vs U at half-filling $\mu = 0$, as a function of g on an 8×8 lattice for four values of β in the ionic Hubbard model. The staggered potential $\Delta = 1.0$. Results are qualitatively like the honeycomb model (see text).

Overall, these results again demonstrate the complex effect the underlying physics of the system can have on the SP in DQMC.

VI. IONIC HUBBARD MODEL

In this section, we shift our attention to the ionic Hubbard model at half-filling. On a square lattice, this model essentially consists of an added local staggered potential $\mu_i = (-1)^i \Delta/2$ to the square lattice Hubbard model. In the noninteracting limit $U/\Delta \rightarrow 0$, the system is a band insulator (BI), where the sites with a negative value of the potential $-\Delta/2$ have a higher occupancy than the sites with a positive value, resulting in a charge density wave order due to the breaking of the sublattice symmetry by the staggered potential. In the opposite limit $U/\Delta \gg 1$, the system is an AFMI due to the large onsite repulsion favoring single occupancy everywhere. Interestingly, as the coupling strength U is increased starting from the weak coupling limit, the BI does not undergo a direct transition to the AFMI. Instead, over a range of values of U and Δ where the two energy scales are comparable, the model displays an exotic CM phase, as past work has demonstrated [53,54].

Figure 7 plots the sign $\langle \mathcal{S} \rangle$ vs the coupling U at different g values for four values of β, as before. At $g = 1$, the sign shows a broad minimum roughly corresponding to the CM phase (the correspondence becomes sharper at lower temperatures). Like the results for the honeycomb Hubbard model, we find that increasingly random sampling worsens the SP and shifts the minimum value of the sign to the right, leading to a complete flattening of the curves (at least in the coupling range shown here) for $g \lesssim 0.5$ as the AFMI state becomes progressively weaker. The shift in the minimum suggests that the BI-CM boundary is also pushed to higher values of U, consistent with the reduction in the effective interaction.

To summarize these detailed results, we show plots of the ratio $\langle \mathcal{S} \rangle (g = 1)/\langle \mathcal{S} \rangle (g = 0)$ side by side for the three models considered here in Fig. 8. In a sense, this ratio isolates the part of the sign due to the correct sampling (at $g = 1$) from the intrinsic sign due to random sampling (at $g = 0$). While this is a somewhat crude quantity, it provides a good first impression of the above effects and how they vary from one model to another. The left plot for the square lattice Hubbard model

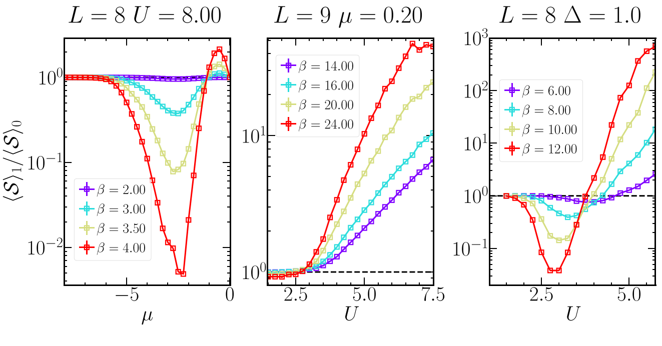

FIG. 8. The ratio $\langle \mathcal{S} \rangle (g = 1)/\langle \mathcal{S} \rangle (g = 0)$ for the square lattice Hubbard model (left), the Hubbard model on the honeycomb lattice (middle), and the ionic Hubbard model (right). The horizontal dashed black lines show $\langle \mathcal{S}_1 \rangle / \langle \mathcal{S}_0 \rangle = 1$, providing a reference for the other plots.

results clearly demonstrates how the sign gets progressively worse with increasing g in the intermediate μ regime, where the ratio becomes small, and then rises up to unity again as the system approaches half-filling. The slight increase >1 near half-filling is due to the faster increase of the $g = 1$ plots there, as the Mott plateau is the strongest in this case. The middle plot shows the corresponding results for the honeycomb Hubbard model. As seen earlier in Fig. 6, as g is lowered, the sign plots become wider and deeper systematically. This also shows up in the ratio plot, where it is almost unity for $U \lesssim 3.0$ and then increases considerably with increasing U. Crudely, the branching point marks the regime where random sampling starts to make the sign worse, leading up to the AFMI phase, even though this is somewhat smaller than the critical value U_c. In the ionic model (right column), we find that the ratio initially dips <1 and then rises, crossing unity approximately around $U \sim 4$ for the temperature values shown here. The initial dip is a consequence of the rightward shift in the $g = 0$ plot, as can be seen in Fig. 7, where the random sampling sign is larger than the $g = 1$ value. However, as the system navigates through the CM phase and approaches the AFMI, the $g = 1$ sign stabilizes again, leading to the crossing at $U < U_c$.

VII. OVERSAMPLING

In the preceding sections, we have considered $0 < g < 1$, which interpolates from random sampling of the HS fields ($g = 0$) and the exact simulation at $g = 1$. As we have discussed, this allows us to separate two sources of the SP: the inherent possibility of negative determinants for *any*, i.e., randomly selected, fermionic matrices and the preferential likelihood of such matrices induced by the sampling. Here, we explore an additional issue, namely, the behavior for $g > 1$. We are indirectly motivated by the successive overrelaxation method [55,56] to solve linear equations which suggests an iterative move from k to $k + 1$ beyond the value at $k + 1$ initially computed. The analogy here is that $g > 1$ samples the HS field beyond what the $g = 1$ determinant suggests, but admittedly, we are also motivated by plain old curiosity. We focus exclusively on the square lattice.

In Fig. 9, we show plots of $\langle \mathcal{S} \rangle$ vs μ with $1 < g < 2$ for three different lattice sizes in three panels as marked. We find that the minimum in the sign becomes shallower and drifts rapidly to the left (for $\mu < 0$, the plots are symmetric about half-filling) as g is increased. The observations in the previous sections immediately suggest an explanation: increasing g pushes the system toward stronger coupling, resulting in a larger Mott plateau which improves the sign and pushes the onset of the SP to more negative μ, beyond the Mott plateau. As expected, the SP becomes progressively worse with increasing lattice size, as the panels clearly demonstrate. The bottom right panel explores the temperature dependence of the sign at $g = 1.2$ on a 10×10 lattice. We find a broad minimum in $\langle \mathcal{S} \rangle$ which grows deeper with reducing temperature, as expected. As g is increased, the Mott plateau grows progressively bigger, and by $g = 2.0$ (not shown), the sign is practically saturated at unity throughout the range shown $-8 < \mu < 0$.

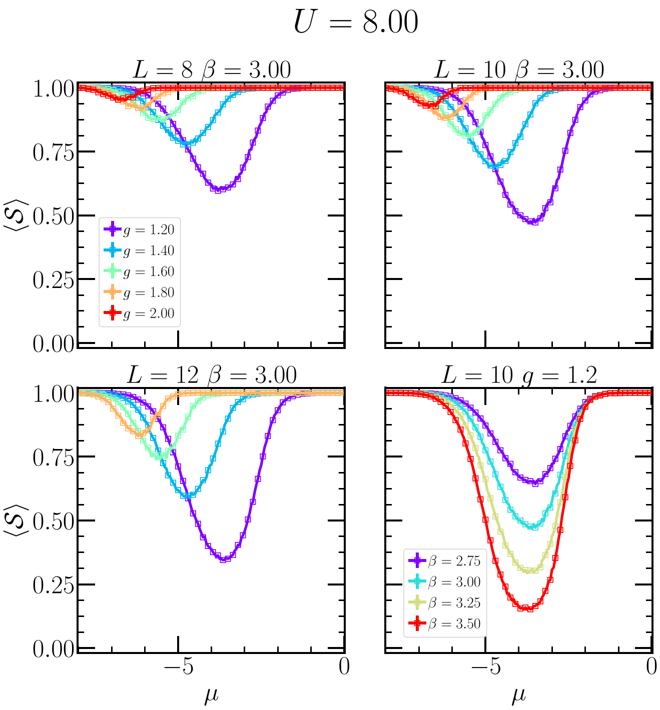

FIG. 9. The average sign $\langle \mathcal{S} \rangle$ vs chemical potential μ at fixed $\beta = 3$ as a function of reweight factor $g > 1$ at $U = 8$ on different lattice sizes $L = 8$, 10, and 12. The sign problem (SP) gets worse with increasing lattice size. Increasing g makes the sign shallower (i.e., lessens the SP) and pushes the chemical potential at which the minimum $\langle \mathcal{S} \rangle$ occurs to lower values. With fixed lattice size $L = 10$, lowering temperature results in smaller $\langle \mathcal{S} \rangle$.

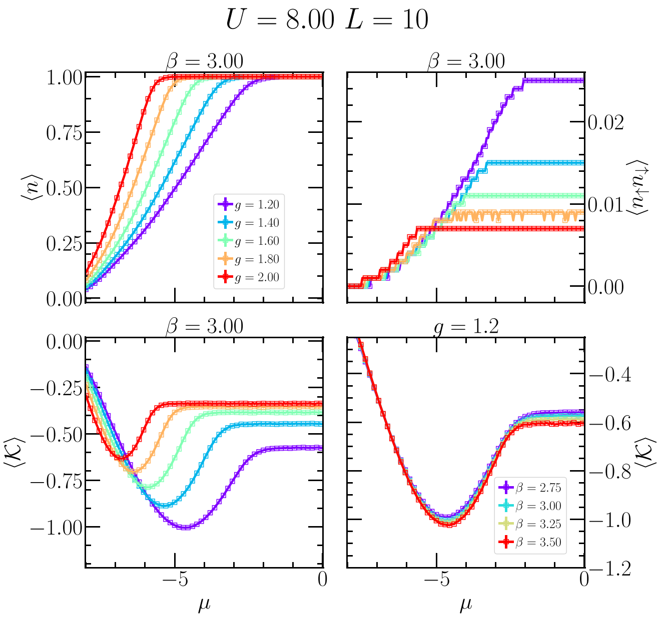

FIG. 10. Oversampling case: electron density $\langle n \rangle$, double occupancy $\langle n_\uparrow n_\downarrow \rangle$, and kinetic energy $\langle \mathcal{K} \rangle$ vs μ as a function of g at $U = 8$ on a 10×10 square lattice. Bottom right panel displays the temperature dependence of $\langle \mathcal{K} \rangle$ at $g = 1.2$ (density and double occupancy have negligible lattice and temperature dependence in this regime).

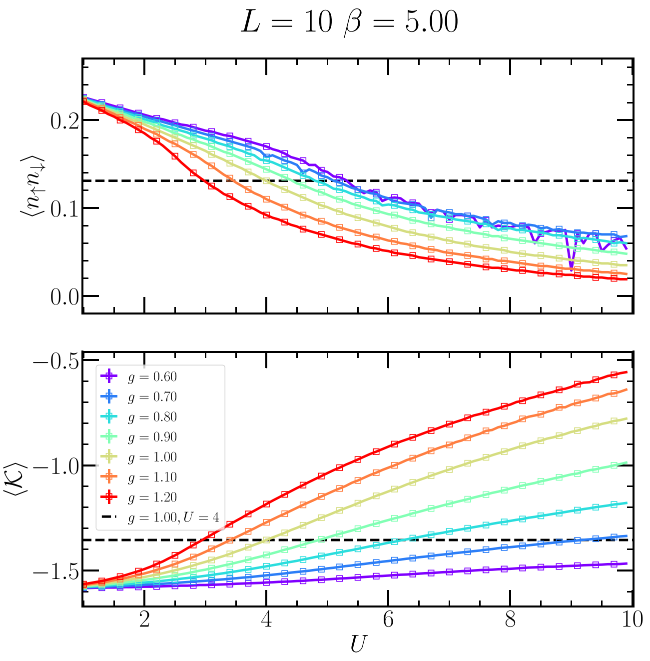

FIG. 11. Double occupancy $\langle n_\uparrow n_\downarrow \rangle$ (top) and kinetic energy $\langle \mathcal{K} \rangle$ (bottom) vs onsite interaction strength U using different reweight factors g. Here, $L = 10$ and $\beta = 5$. The dashed horizontal line intersects the $g = 1$ curve at $U = 4$. To get the same value of $\langle n_\uparrow n_\downarrow \rangle$ for $g = 0.7$ requires $U = 5.1$, and for $g = 1.2$ requires $U = 3$. This implies $U_{\text{eff}} = 4$ for $g = 1.2, U = 3.0$ and for $g = 0.7, U = 5.1$. A similar analysis of $\langle \mathcal{K} \rangle$ implies that $U_{\text{eff}} = 4$ for $g = 1.2$, $U = 2.8$ and for $g = 0.7, U = 9.2$.

In Fig. 10, we show the behavior of local physical variables such as the number density $\langle n \rangle$, the double occupancy $\langle n_\uparrow n_\downarrow \rangle$, and the kinetic energy $\langle \mathcal{K} \rangle$ vs μ in this $g > 1$ regime. As the Mott insulator becomes more extensive with increasing g, we find that $\langle n \rangle$ saturates at progressively more negative values of μ. The double occupancy and the kinetic energy both reduce in magnitude as the insulating system increasingly favors single occupancy. The bottom right panel plots $\langle \mathcal{K} \rangle$ for varying temperatures at $g = 1.2$, demonstrating a rather weak dependence on temperature. Similarly, we find negligible dependence on temperature and lattice size for all these variables in this regime (data not shown here). It is quite striking that, while the sign itself varies considerably with temperature, the physical variables show little change in comparison; a peek at Fig. 3 reveals that the same observation is largely true in the undersampling case ($g < 1$) as well.

VIII. RENORMALIZED COUPLING

It is useful to quantify these statements concerning the effect of g on the physics. To do this, we determine an effective (renormalized) interaction strength U_{eff} due to $g \neq 1$. We proceed as follows: for a given g and U, we find the value of the repulsion U_{eff} which at $g = 1$ yields the same value for local observables including the double occupancy $\mathcal{D} = \langle n_\uparrow n_\downarrow \rangle$ or the kinetic energy $\langle \mathcal{K} \rangle$. In Fig. 11, we show $\mathcal{D}$ and $\mathcal{K}$ as functions of U at half-filling ($\langle n \rangle = 1$) on a $N = 10 \times 10$ lattice. For a given value of U at $g = 1$, a horizontal cut gives the value of U_{eff} from the intersections with the curves at $g \neq 1$. Comparison of $\mathcal{D}$ and $\mathcal{K}$ reveals that the results are not universal—the inferred U_{eff} depends on the observable, especially for $g < 1$.

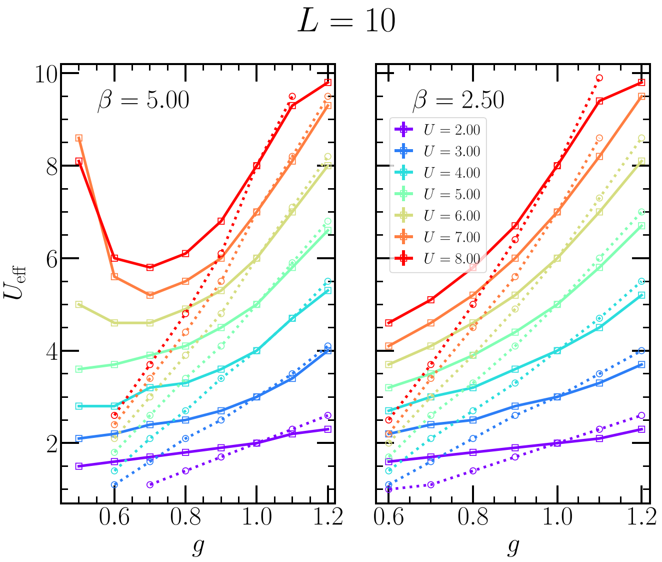

FIG. 12. Effective onsite interaction U_{eff} determined from double occupancy $\langle n_\uparrow n_\downarrow \rangle$ (solid lines) and kinetic energy $\mathcal{K}$ (dotted lines) vs reweight factor g. Inverse temperature is fixed at $\beta = 5$ (left panel) and $\beta = 2.5$ (right panel), and data were extracted using an $L = 10$ lattice. For $g > 1$, the inferred values are roughly equal, while for small g, there is a marked disagreement. U_{eff} seems to be roughly temperature independent, except for small g and large U.

In Fig. 12, we compare the values of $U_{\text{eff}}(g)$ determined in the manner described from Fig. 11. In general, the double occupancy $\mathcal{D}$ tends to predict larger U_{eff} than the kinetic energy $\mathcal{K}$ when $g < 1$, but the values are similar for $g > 1$.

IX. CONCLUSIONS

In this paper, we have argued that a simple view of the SP as originating in the independent winding of fermionic world lines across the space-time lattice is incomplete. Instead, by studying the Hubbard model in three different contexts, namely, on a square lattice, on a honeycomb lattice, and with an external staggered potential, the SP has been shown to depend nontrivially on the manner in which the fermion determinants select the HS fields. While the square lattice results show that the SP predominantly originates from determinants steering the simulation to low sign regions of phase space, the honeycomb and ionic results demonstrate that random sampling makes the sign worse, especially leading up to the AFMI phase. Given the few models considered here, it is difficult to infer a general rule that determines the extent of the different contributions to the SP under a given set of circumstances. However, as seen above, when the determinants guide the system toward a protected particle-hole symmetric point, such as the AFMI state, the SP is typically reduced. Extensions of this work to other models could shed more light on this issue.

As we have seen, shifting g away from $g = 1$ explicitly changes the values of physical observables. Indeed, it has the effect of pushing the system to weaker coupling in the sense that the effects of correlation such as the Mott gap and the suppression of double occupancy and hopping are reduced. Similarly, in the oversampling case, with $g > 1$, the effective coupling is larger. The resulting reinforcement of the Mott plateau in the square lattice case rapidly mitigates the SP. An explicit calculation of the effective coupling puts these observations on a more quantitative footing. However, unlike an analogous analysis for the renormalization of the electron-phonon coupling in a model which includes onsite interactions, where the effective coupling was the same for different observables [57], here, only whether U_{eff} is reduced or enhanced relative to U is universal. The quantitative value of U_{eff} can vary markedly depending on which observable is analyzed.

Our procedure allows us to deconvolve the ways in which the SP arises, and the alteration of the underlying physics provides us with qualitative insight into the observed trends. It is, of course, possible to formulate a protocol in which the alteration of the weight $\mathcal{W}$ by $g \neq 1$ is compensated by including an appropriate factor $\mathcal{W}^{1-g}$ in the measurement of physical observables. In this case, expectation values would be unchanged, but the error bars would be altered. Intuitively, setting $g \neq 1$ seems likely to lead to *less efficient* simulations, as it violates the spirit of proper importance sampling.

A final comment concerns connections between what we have explored here and the rational hybrid Monte Carlo (RHMC) algorithm [58] widely used in lattice gauge theory. In RHMC, the fermion determinant is split into many pieces $\det M = [\det(M^{1/n})]^n$. The purpose there has nothing to do with the SP but rather to make the computation of a pseudofermion approximant to the determinant more well conditioned, allowing larger step sizes in the integration of the equations of motion. Nevertheless, it is interesting that the $1/n$ factor plays a very similar role to the g considered here. The difference, of course, is that our approach no longer simulates the original model when $g < 1$ because we do not introduce an increased number $n = 1/g$ of copies of the determinant.

ACKNOWLEDGMENTS

R.T.S. was supported by Grant No. DE-SC0014671 funded by the U.S. Department of Energy, Office of Science. R.M. acknowledges support from the National Natural Science Foundation of China (NSFC) Grants No. U1930402, No. 11974039, No. 12050410263 and No. 12111530010. B.X. was supported by the Flatiron Institute. The Flatiron Institute is a division of the Simons Foundation. Computations were performed on the Tianhe-2JK at the Beijing Computational Science Research Center.

APPENDIX A: SIGN OF INDIVIDUAL SPIN SPECIES

Figure 13 shows the behavior of the sign of the individual weight matrices $\langle \mathcal{S}_\sigma \rangle$ with the rescaling parameter g at four values of the temperature. We find that the spin-resolved sign shows very similar qualitative behavior to the total one, except near half-filling, where the constraint of particle-hole symmetry that led to the sign approaching unity near half-filling is no longer present for the spin-resolved quantity. Interestingly, the plots with $g \gtrsim 0.8$ still show an upturn near half-filling, even though they do not reach the maximum value. The nonmonotonicity of the sign noted in Fig. 1 with reducing g, where it becomes worse initially before increasing again, is also visible here for $\langle \mathcal{S}_\sigma \rangle$.

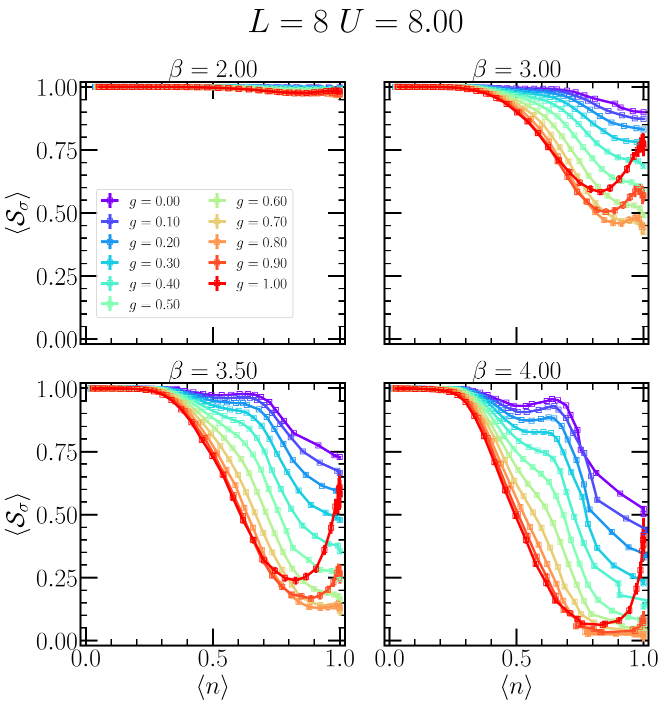

FIG. 13. The average sign of the matrices for individual spins, $\langle \mathcal{S}_\sigma \rangle$ vs $\langle n \rangle$ at $U = 8$ as a function of g on an 8×8 lattice for four values of β. As seen for the total sign in the main text, the minimum in $\langle \mathcal{S}_\sigma \rangle$ becomes wider and deeper with increasing β. The nonmonotonic behavior of the sign with reducing g, where $\mathcal{S}_\sigma$ becomes worse initially and then gets progressively better, is also evident here. In contrast to the total sign, however, the spin-resolved sign is not constrained to be equal to 1 at half-filling, and thus, the sharp upturn seen in Fig. 1 near $\mu = 0$ is absent here for most g values.

APPENDIX B: NUMBER DENSITY AT DIFFERENT LATTICE SIZES

In Fig. 14, we show the number density vs chemical potential for different lattice sizes L. Apart from the usual features of the Mott saturation near half-filling at high values of g and its absence at low values of the rescaling parameter that we have already noted in the main text, we find that the finite-sized effects, evidenced in the oscillations of $\langle n \rangle$ for $g \sim 0$, are considerably enhanced at lower lattice sizes, as expected.

APPENDIX C: SIGN VS LATTICE SIZE

In Fig. 15, we show the sign $\langle \mathcal{S} \rangle$ plotted against the density $\langle n \rangle$ as a function of the lattice size L for four different values of g. Aside from the usual observations, we clearly see the finite-sized oscillations for the $L = 4$ plots, with the sign reaching a maximum at the magic density $\langle n \rangle \sim 0.625$, corresponding to a closed-shell filling for this system size.

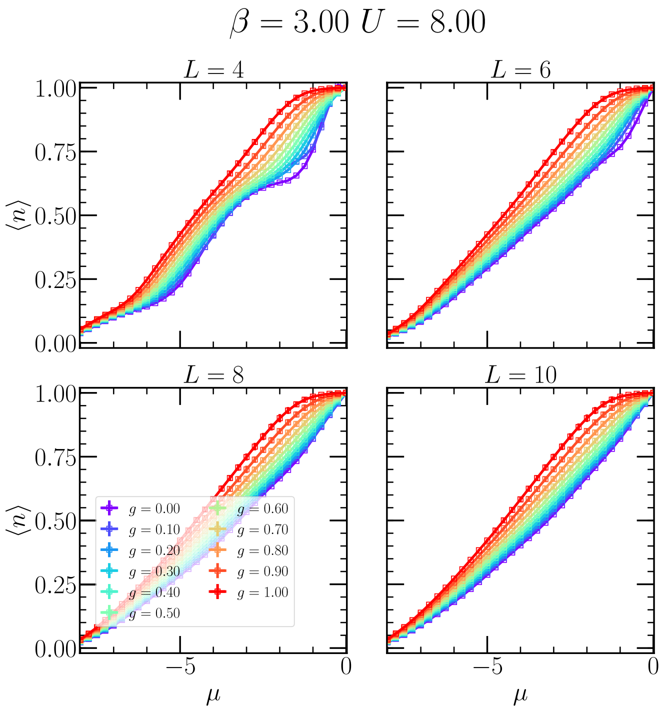

FIG. 14. Density $\langle n \rangle$ vs μ at $U = 8$, $\beta = 3.0$ on an 8×8 lattice as a function of g for four different values of the lattice size L. The oscillatory behavior at low g in the 4×4 lattice is due to the reemergence of the shell effect in small lattices. As explained in detail in the main text, a completely random sampling of the Hubbard-Stratonovich (HS) fields in quantum Monte Carlo (QMC) is more involved than a standard rescaling of T; in this case, pushing the system to weaker coupling by reducing the effect of U, resulting in a reappearance of finite-sized effects originally suppressed by the interaction.

APPENDIX D: FINITE-SIZED OSCILLATIONS IN THE $g = 0$ LIMIT

In Fig. 16, we reinforce the statements made in several parts of the main text that, as we approach $g = 0$, the effect of U is reduced on certain physical attributes of the system, and noninteracting features, such as finite size shell effects, usually washed out by the strong interaction, are unmasked. The left panel shows the sign at $g = 0$ plotted for a large number of temperature values going down to $\beta = 9$. Strong oscillations in $\langle \mathcal{S} \rangle$ are observed as lower temperatures are reached. The peak positions in this limit are connected to the closed-shell densities for this system size, as confirmed by the right panel, which plots the positions of the densities corresponding to the maxima of the sign at low temperatures. This is more evident when comparing it with the dotted horizontal lines, which mark the values of the closed-shell densities for 8×8 lattices.

APPENDIX E: THE $t = 0$ LIMIT

Here, we discuss the strong coupling ($t = 0$) limit in detail. As mentioned briefly in the main text, this limit can be solved analytically at $g = 1$ ($g = 0$ is trivial, of course), whereas for a general value of g, the solution cannot be written down in a closed form.

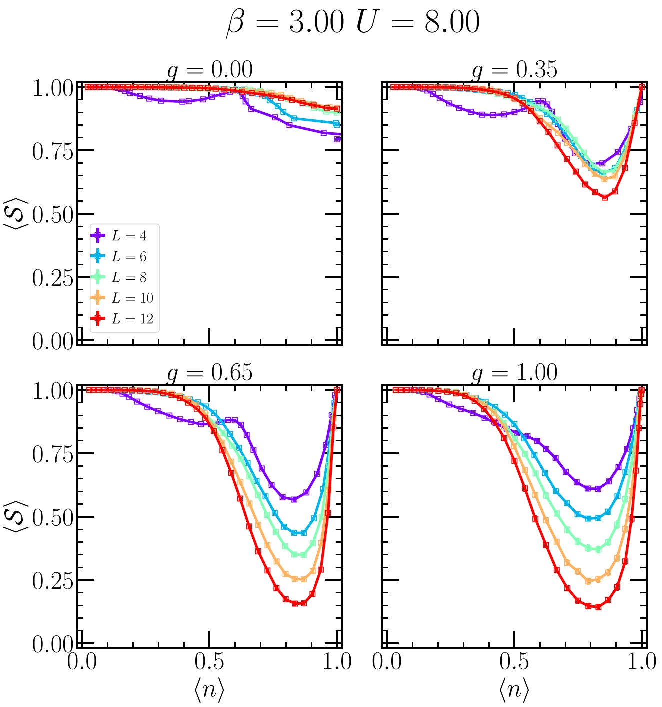

FIG. 15. $\langle \mathcal{S} \rangle$ vs $\langle n \rangle$ at $\beta = 3$, $U = 8$ as a function of L for different g values. The main observation is the finite-sized oscillations for $L = 4$, with the sign reaching a maximum at the magic density $\langle n \rangle \sim 0.625$ associated with a closed-shell filling for this system size.

The partition function for a general value of g is given by

$$\mathcal{Z} = \mathcal{C}^{NL_\tau} \prod_i \sum_{\{h_i\}} \left(\prod_\sigma \left\{ 1 + \exp\left[\beta\mu + \alpha\sigma \sum_l h_i(l) \right] \right\} \right)^g . \tag{E1}$$

For $g = 1$, this is trivially easy to calculate:

$$\mathcal{Z} = \mathcal{C}^{NL_\tau} \prod_i \sum_{\{h_i\}} \Bigg(\exp(\beta\mu) \Bigg\{ \exp\left[\alpha \sum_l h_i(l) \right]$$

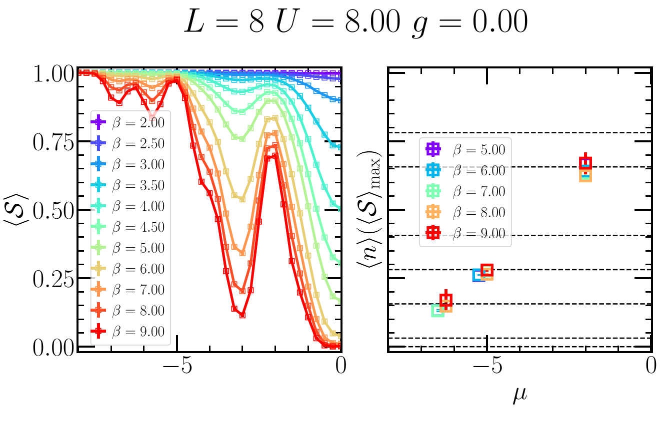

FIG. 16. Left: $\langle \mathcal{S} \rangle$ vs μ at $g = 0$, $U = 8$, $L = 8$ for different β values. As β is increased >4, $\langle \mathcal{S} \rangle$ shows strong oscillatory behavior reminiscent of shell effects. Right: values of $\langle n \rangle$ at the maxima of the sign. Dotted black lines mark the magic values of the density at this lattice size (see text).

$$
\begin{aligned}
&+\exp\left[-\alpha\sum_l h_i(l)\right]\Bigg\}+1+\exp(2\beta\mu)\Bigg)\\
&=\mathcal{C}^{NL_\tau}\prod_i\{2\exp(\beta\mu)[2\cosh(\alpha)]^{L_\tau}\\
&\quad+2^{L_\tau}[1+\exp(2\beta\mu)]\}\\
&=\prod_i\exp\left(-\frac{\beta U}{4}\right)\{1+\exp(2\beta\mu)\\
&\quad+2\exp\left[\beta\left(\mu+\frac{U}{2}\right)\right]\Big\}. \qquad \text{(E2)}
\end{aligned}
$$

On the other hand, the general case with $g\neq 1$ cannot be solved to yield an analytical expression as above and must be computed numerically. However, the calculations may be simplified by noting that the summand is a function of $\sum_l h_i(l)$ for any given site i. Since each site is independent, we will just choose one site and drop the label i in what follows below. Each of the HS fields $h(l)$ takes values ± 1. Hence, for a general situation where n of them are -1 and the rest $+1$, the sum $\sum_l h(l)=L_\tau-2n$. The degeneracy of such a situation is given by ${}^{L_\tau}C_n=\frac{L_\tau!}{n!\,(L_\tau-n)!}$, the number of ways of choosing n variables out of L_τ. Thus, the single site partition function $\mathcal{Z}_s$ may be rewritten as

$$
\mathcal{Z}_s=\mathcal{C}^{L_\tau}\sum_{n=0}^{L_\tau}{}^{L_\tau}C_n\left(\prod_\sigma\{1+\exp\left[\beta\mu+\alpha\sigma(L_\tau-2n)\right]\}\right)^g. \qquad \text{(E3)}
$$

The equal time Green's function $\mathcal{G}_\sigma(\tau,\tau)$ is independent of the imaginary time coordinate τ and is given by $\mathcal{G}_\sigma(\tau,\tau)=1/\{1+\exp[\beta\mu+\alpha\sigma\sum_l h(l)]\}$. For any given configuration $\{h\}$, this is simply the equivalent of the noninteracting expression $\{1-1/[1+\exp(\beta\epsilon_\sigma)]\}$, with $\epsilon_\sigma=-\alpha\sigma\sum_l h(l)-\mu$. Calculations of expectation values of variables $\mathcal{A}(\mathcal{G})$, expressed in terms of the equal time Green's function, may be performed in a similar manner to $\mathcal{Z}_s$:

$$
\begin{aligned}
\langle\mathcal{A}\rangle&=\mathcal{Z}^{-1}\,\mathcal{C}^{L_\tau}\sum_{n=0}^{L_\tau}{}^{L_\tau}C_n\,\mathcal{A}(\mathcal{G})\\
&\quad\times\left(\prod_\sigma\{1+\exp\left[\beta\mu+\alpha\sigma(L_\tau-2n)\right]\}\right)^g. \qquad \text{(E4)}
\end{aligned}
$$

In Fig. 17, we show the number density $\langle n\rangle$ and the double occupation $\langle n_\uparrow n_\downarrow\rangle$ as a function of μ for various values of g, U, and β. As seen in the general cases with finite coupling in the main text, we find that the Mott insulator at $g=1$ gradually disappears as g is dialed down. Similarly, the double occupancy in the bottom right panel also shows that the effect of U becomes weaker as g is reduced. However, as the $U=0$ black line indicates, even $g=0$ is very different from the noninteracting case. The top right and bottom left panels show the $g=0$ plots for $\langle n\rangle$ for different values of β (fixed U) and different U (fixed β), respectively. While these can be nominally fitted to noninteracting Fermi functions, the effective β (top right inset) is finite (the classical $\beta\to g\beta$ scaling implies infinite temperature) but is smaller than the actual value. Already at the single site level, this demonstrates the nontrivial effect that tuning g has on the physics of the interacting system.

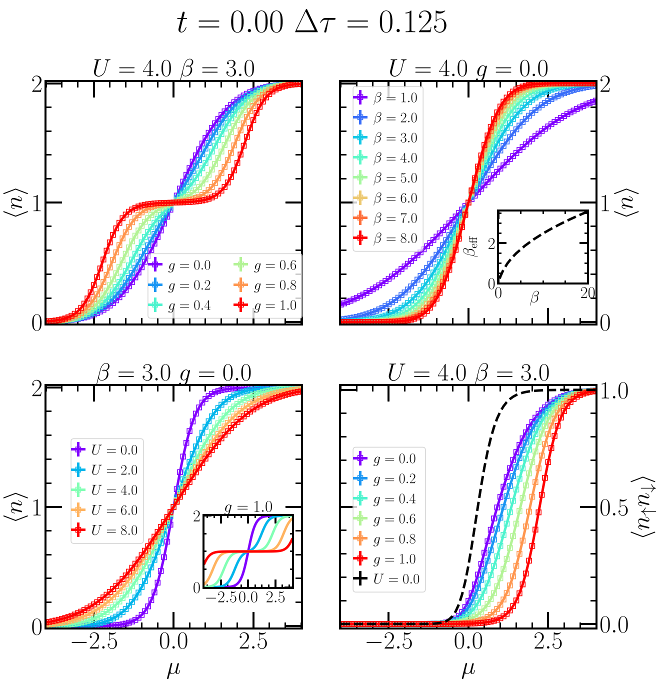

FIG. 17. The $t=0$ limit. Top left panel shows the number density $\langle n\rangle$ vs chemical potential μ at different g values, showing the Mott plateau near $g=1$ disappearing as $g\to 0$. Top right panel plots $\langle n\rangle$ vs μ for $U=4.0$, $g=0$ at different β. The inset fits these curves to a Fermi function with β_{eff}. Bottom left shows the same for fixed $\beta=3.0$, $g=0$ at different coupling U, with inset plotting the same at $g=1$. The final plot shows the double occupancy $\langle n_\uparrow n_\downarrow\rangle$ for different g values along with the $U=0$ plot for comparison.

---

[1] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Monte Carlo calculations of coupled boson-fermion systems. I, Phys. Rev. D **24**, 2278 (1981).

[2] S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar, Numerical study of the two-dimensional Hubbard model, Phys. Rev. B **40**, 506 (1989).

[3] L. Chen and A.-M. Tremblay, Determinant Monte Carlo for the Hubbard model with arbitrarily gauged auxiliary fields, Int. J. Mod. Phys. B **06**, 547 (1992).

[4] F. F. Assaad, Quantum Monte Carlo methods on lattices: the determinantal approach, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, Lecture Notes NIC Series, Vol. 10 (NIC-Directors, Jülich, 2002), pp. 99–156.

[5] J. Gubernatis, N. Kawashima, and P. Werner, *Quantum Monte Carlo Methods: Algorithms for Lattice Models* (Cambridge University Press, Cambridge, 2016).

[6] Y. Alhassid, Auxiliary-field quantum Monte Carlo methods in nuclei, in *Emergent Phenomena in Atomic Nuclei from Large-*

Scale Modeling (World Scientific, Singapore, 2017), Chap. 9, pp. 267–298.

[7] H. Hao, B. M. Rubenstein, and H. Shi, Auxiliary field quantum Monte Carlo for multiband Hubbard models: Controlling the sign and phase problems to capture Hund's physics, Phys. Rev. B **99**, 235142 (2019).

[8] Y.-Y. He, M. Qin, H. Shi, Z.-Y. Lu, and S. Zhang, Finite-temperature auxiliary-field quantum Monte Carlo: Self-consistent constraint and systematic approach to low temperatures, Phys. Rev. B **99**, 045108 (2019).

[9] D. Ceperley, Path integrals in the theory of condensed helium, Rev. Mod. Phys. **67**, 279 (1995).

[10] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, Quantum Monte Carlo simulations of solids, Rev. Mod. Phys. **73**, 33 (2001).

[11] T. Degrand and C. DeTar, *Lattice Methods for Quantum Chromodynamics* (World Scientific, Singapore, 2006).

[12] J. Carlson, S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. E. Schmidt, and R. B. Wiringa, Quantum Monte Carlo methods for nuclear physics, Rev. Mod. Phys. **87**, 1067 (2015).

[13] B. L. Hammond, W. A. Lester, and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, Singapore, 1994).

[14] R. Needs, M. Towler, N. Drummond, P. López Ríos, and J. Trail, Variational and diffusion quantum Monte Carlo calculations with the CASINO code, J. Chem. Phys. **152**, 154106 (2020).

[15] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Sign problem in the numerical simulation of many-electron systems, Phys. Rev. B **41**, 9301 (1990).

[16] G. Ortiz, D. M. Ceperley, and R. M. Martin, New Stochastic Method for Systems with Broken Time-Reversal Symmetry: 2D Fermions in a Magnetic Field, Phys. Rev. Lett. **71**, 2777 (1993).

[17] S. Zhang, J. Carlson, and J. E. Gubernatis, Constrained path Monte Carlo method for fermion ground states, Phys. Rev. B **55**, 7464 (1997).

[18] S. Chandrasekharan and U.-J. Wiese, Meron-Cluster Solution of Fermion Sign Problems, Phys. Rev. Lett. **83**, 3116 (1999).

[19] P. Henelius and A. W. Sandvik, Sign problem in Monte Carlo simulations of frustrated quantum spin systems, Phys. Rev. B **62**, 1102 (2000).

[20] S. Bergkvist, P. Henelius, and A. Rosengren, Reduction of the sign problem using the meron-cluster approach, Phys. Rev. E **68**, 016122 (2003).

[21] M. Troyer and U.-J. Wiese, Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations, Phys. Rev. Lett. **94**, 170201 (2005).

[22] M. Nyfeler, F.-J. Jiang, F. Kämpfer, and U.-J. Wiese, Nested Cluster Algorithm for Frustrated Quantum Antiferromagnets, Phys. Rev. Lett. **100**, 247206 (2008).

[23] Y. Nomura, S. Sakai, and R. Arita, Multiorbital cluster dynamical mean-field theory with an improved continuous-time quantum Monte Carlo algorithm, Phys. Rev. B **89**, 195146 (2014).

[24] A. Mukherjee and M. Cristoforetti, Lefschetz thimble Monte Carlo for many-body theories: A Hubbard model study, Phys. Rev. B **90**, 035134 (2014).

[25] H. Shinaoka, Y. Nomura, S. Biermann, M. Troyer, and P. Werner, Negative sign problem in continuous-time quantum Monte Carlo: Optimal choice of single-particle basis for impurity problems, Phys. Rev. B **92**, 195126 (2015).

[26] R. K. Kaul, Marshall-positive SU(N) quantum spin systems and classical loop models: A practical strategy to design sign-problem-free spin Hamiltonians, Phys. Rev. B **91**, 054413 (2015).

[27] V. I. Iglovikov, E. Khatami, and R. T. Scalettar, Geometry dependence of the sign problem in quantum Monte Carlo simulations, Phys. Rev. B **92**, 045110 (2015).

[28] M. Fukuma, N. Matsumoto, and N. Umeda, Applying the tempered Lefschetz thimble method to the Hubbard model away from half filling, Phys. Rev. D **100**, 114510 (2019).

[29] M. Ulybyshev, C. Winterowd, and S. Zafeiropoulos, Lefschetz thimbles decomposition for the Hubbard model on the hexagonal lattice, Phys. Rev. D **101**, 014508 (2020).

[30] A. J. Kim, P. Werner, and R. Valentí, Alleviating the sign problem in quantum Monte Carlo simulations of spin-orbit-coupled multiorbital Hubbard models, Phys. Rev. B **101**, 045108 (2020).

[31] S. Johnston, E. A. Nowadnick, Y. F. Kung, B. Moritz, R. T. Scalettar, and T. P. Devereaux, Determinant quantum Monte Carlo study of the two-dimensional single-band Hubbard-Holstein model, Phys. Rev. B **87**, 235133 (2013).

[32] M. Bercx, F. Goth, J. S. Hofmann, and F. F. Assaad, The ALF (Algorithms for Lattice Fermions) project release 1.0. documentation for the auxiliary field quantum Monte Carlo code, SciPost Phys. **3**, 013 (2017).

[33] C. Wu and S.-C. Zhang, Sufficient condition for absence of the sign problem in the fermionic quantum Monte Carlo algorithm, Phys. Rev. B **71**, 155115 (2005).

[34] Z.-X. Li, Y.-F. Jiang, and H. Yao, Majorana-Time-Reversal Symmetries: A Fundamental Principle for Sign-Problem-Free Quantum Monte Carlo Simulations, Phys. Rev. Lett. **117**, 267002 (2016).

[35] P. Fazekas, *Lecture Notes on Electron Correlation and Magnetism* (World Scientific, Singapore, 1999).

[36] J. Hubbard and B. H. Flowers, Electron correlations in narrow energy bands, Proc. R. Soc. London A **276**, 238 (1963).

[37] D. Scalapino, Does the Hubbard model have the right stuff? in *Proceedings of the International School of Physics*, edited by R. Broglia and J. Schrieffer (North-Holland, Amsterdam, 1994).

[38] T. Paiva, R. T. Scalettar, W. Zheng, R. R. P. Singh, and J. Oitmaa, Ground-state and finite-temperature signatures of quantum phase transitions in the half-filled Hubbard model on a honeycomb lattice, Phys. Rev. B **72**, 085123 (2005).

[39] Z. Meng, S. Wessel, A. Muramatsu, T. Lang, and F. Assaad, Quantum spin liquid emerging in two-dimensional correlated Dirac fermions, Nature (London) **464**, 847 (2010).

[40] S. Sorella, Y. Otsuka, and S. Yunoki, Absence of a spin liquid phase in the Hubbard model on the honeycomb lattice, Sci. Rep. **2**, 992 (2012).

[41] F. F. Assaad and I. F. Herbut, Pinning the Order: The Nature of Quantum Criticality in the Hubbard Model on Honeycomb Lattice, Phys. Rev. X **3**, 031010 (2013).

[42] F. Parisen Toldin, M. Hohenadler, F. F. Assaad, and I. F. Herbut, Fermionic quantum criticality in honeycomb and π-flux Hubbard models: Finite-size scaling of renormalization-group-invariant observables from quantum Monte Carlo, Phys. Rev. B **91**, 165108 (2015).

[43] Y. Otsuka, S. Yunoki, and S. Sorella, Universal Quantum Criticality in the Metal-Insulator Transition of Two-Dimensional

Interacting Dirac Electrons, Phys. Rev. X **6**, 011029 (2016).

[44] J. E. Hirsch, Discrete Hubbard-Stratonovich transformation for fermion lattice models, Phys. Rev. B **28**, 4059 (1983).

[45] J. Carlson, J. E. Gubernatis, G. Ortiz, and S. Zhang, Issues and observations on applications of the constrained-path Monte Carlo method to many-fermion systems, Phys. Rev. B **59**, 12788 (1999).

[46] S. Zhang, J. Carlson, and J. E. Gubernatis, Constrained Path Quantum Monte Carlo Method for Fermion Ground States, Phys. Rev. Lett. **74**, 3652 (1995).

[47] S. Zhang and H. Krakauer, Quantum Monte Carlo Method using Phase-Free Random Walks with Slater Determinants, Phys. Rev. Lett. **90**, 136401 (2003).

[48] W. Purwanto and S. Zhang, Quantum Monte Carlo method for the ground state of many-boson systems, Phys. Rev. E **70**, 056702 (2004).

[49] M. Motta, D. E. Galli, S. Moroni, and E. Vitali, Imaginary time correlations and the phaseless auxiliary field quantum Monte Carlo, J. Chem. Phys. **140**, 024107 (2014).

[50] M. Motta and S. Zhang, Communication: Calculation of interatomic forces and optimization of molecular geometry with auxiliary-field quantum Monte Carlo, J. Chem. Phys. **148**, 181101 (2018).

[51] R. Mondaini, K. Bouadim, T. Paiva, and R. R. dos Santos, Finite-size effects in transport data from quantum Monte Carlo simulations, Phys. Rev. B **85**, 125127 (2012).

[52] R. Mondaini, S. Tarat, and R. T. Scalettar, Quantum critical points and the sign problem, arXiv:2108.08974.

[53] N. Paris, K. Bouadim, F. Hebert, G. G. Batrouni, and R. T. Scalettar, Quantum Monte Carlo Study of an Interaction-Driven Band-Insulator–to–Metal Transition, Phys. Rev. Lett. **98**, 046403 (2007).

[54] A. Chattopadhyay, S. Bag, H. R. Krishnamurthy, and A. Garg, Phase diagram of the half-filled ionic Hubbard model in the limit of strong correlations, Phys. Rev. B **99**, 155127 (2019).

[55] D. Young, Iterative methods for solving partial difference equations of elliptic type, Trans. Am. Math. Soc. **76**, 92 (1954).

[56] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations* (Springer, Cham, 1994), Vol. 95.

[57] C. Feng, B. Xing, D. Poletti, R. Scalettar, and G. Batrouni, Phase diagram of the Su-Schrieffer-Heeger-Hubbard model on a square lattice, arXiv:2109.09206.

[58] M. A. Clark and A. D. Kennedy, Accelerating Dynamical-Fermion Computations using the Rational Hybrid Monte Carlo Algorithm with Multiple Pseudofermion Fields, Phys. Rev. Lett. **98**, 051601 (2007).