Finite-temperature auxiliary-field quantum Monte Carlo: Self-consistent constraint and systematic approach to low temperatures

Yuan-Yao He,[1,2,3] Mingpu Qin,[2] Hao Shi,[3] Zhong-Yi Lu,[1] and Shiwei Zhang[2,3]

[1]*Department of Physics, Renmin University of China, Beijing 100872, China*
[2]*Department of Physics, College of William and Mary, Williamsburg, Virginia 23187, USA*
[3]*Center for Computational Quantum Physics, Flatiron Institute, New York, New York 10010, USA*

We describe an approach for many-body calculations with a finite-temperature, grand canonical ensemble formalism using auxiliary-field quantum Monte Carlo (AFQMC) with a self-consistent constraint to control the sign problem. The usual AFQMC formalism of Blankenbecler, Scalapino, and Sugar suffers from the sign problem with most physical Hamiltonians, as is well known. Building on earlier ideas to constrain the paths in auxiliary-field space [Zhang, Phys. Rev. Lett. **83**, 2777 (1999)] and incorporating recent developments in zero-temperature, canonical-ensemble methods, we discuss how a self-consistent constraint can be introduced in the finite-temperature, grand-canonical-ensemble framework. This together with several other algorithmic improvements discussed here leads to a more accurate, more efficient, and numerically more stable approach for finite-temperature calculations. We carry out a systematic benchmark study in the two-dimensional repulsive Hubbard model at 1/8 doping. Temperatures as low as $T = 1/80$ (in units of hopping) are reached. The finite-temperature method is exact at very high temperatures, and approaches the result of the zero-temperature constrained-path AFQMC as temperature is lowered. The benchmark shows that systematically accurate results are obtained for thermodynamic properties.

I. INTRODUCTION

Quantum Monte Carlo (QMC) methods have become a key numerical technique for studying interacting quantum many-body systems in various areas, including condensed-matter [1,2], high-energy [3,4] and nuclear [5] physics as well as quantum chemistry [6]. Among them, auxiliary-field quantum Monte Carlo (AFQMC) methods [1,2,7–12] decouple the two-body interactions in the Hamiltonian by the Hubbard-Stratonovich transformation [13], and sample the resulting external auxiliary fields by Monte Carlo. AFQMC methods can produce accurate, sometimes even numerically exact, solutions in correlated fermion systems, by explicitly accessing only a small fraction of the whole auxiliary-field space, whose size grows exponentially with system size.

The determinantal quantum Monte Carlo (DQMC) algorithm, formulated by Blankenbecler, Scalapino, and Sugar [1] (BSS), is commonly applied for both finite-temperature and ground-state [2,14] calculations. Direct DQMC calculations, while formally exact, suffer from the minus sign problem [15–17] in general. As a result, the computational cost of DQMC calculations, for fixed statistical accuracy, scales exponentially with system size and inverse temperature [16,17], instead of the polynomial scaling expected for systems free of the sign problem. Though many specific models have turned out to be free of the sign problem [18–27], the vast majority of correlated electron systems are not. The sign problem hinders or prevents DQMC studies of a variety of interesting problems in correlated fermion systems, including in Hubbard-like models with doping and repulsive interactions, and almost all realistic Hamiltonians of molecules and solids.

More recently, constrained-path (CP) AFQMC methods [9,10,28] have been developed as an alternative approach. Building on the basic formalism of the DQMC algorithm, these methods introduce a computational framework of random walks in the manifold of mean-field or independent-particle solutions, which connects with concepts in electronic structure and quantum chemistry [12]. CP-AFQMC controls the sign problem by applying a constraint which restricts the Monte Carlo sampling in auxiliary-field space. The constraint identifies and removes redundant contributions to the ground-state wave function or the finite-temperature density matrix by paths in auxiliary-field space which would cancel with explicit, full path integration but appear as random noise in Monte Carlo sampling. The formalism is exact when the path identification is exact. In practice it is implemented approximately with a trial wave function or trial density matrix, which introduces a possible systematic bias in the numerical results but in turn removes the exponentially growing computational cost and recovers the algebraic complexity. Through many tests and developments in the last two decades [29–41], the zero-temperature (ZT) CP-AFQMC method has been proved to be a highly accurate, general numerical approach for studying ground-state properties of various interacting fermion models as well as molecules and realistic materials by its generalization, the phaseless AFQMC method [28,42,43].

The finite-temperature (FT) CP-AFQMC method was formulated to study thermodynamic properties in correlated fermion systems in Ref. [10]. The constraint applied in the FT-CP-AFQMC method to control the sign problem involves an input trial Hamiltonian or trial density matrix, corresponding to the trial wave function used in the ZT method. The

FT-CP-AFQMC formalism has been extended to Bose-Fermi mixtures [44], and more recently applied to molecular systems [45] by replacing the sign constraint with the phaseless constraint [28]. However, applications of the FT-CP-AFQMC have been limited, since the benchmark has not been nearly extensive as for ZT, and the many recent developments in understanding and improving the constraint have not been realized within the FT framework.

The ability to compute the temperature dependence of various physical properties in many fermion systems is of fundamental importance. Examples in correlated fermion models include pseudogap physics in cuprates, as well as BCS-BEC crossover for ultracold fermions in optical lattice. More recently, the direct measurements of the equation of state [46], Mott insulator [47], as well as short-range charge and spin correlations [48,49] of the Hubbard model are realized in optical lattice. As experiments start to access lower temperatures, there is a growing need for accurate computations across temperature ranges. It is especially timely to revisit the difficulty posed by the sign problem and extend the reach of AFQMC to finite but sufficiently low temperatures.

We address this need in the present paper. Building on the FT-CP-AFQMC ideas in Ref. [10] and recent developments in ZT methods [38–40], we introduce a self-consistent constraint in the FT framework, investigate the systematic error and its relation to spin symmetry breaking in the trial density matrix, and perform systematic benchmark studies in the two-dimensional Hubbard model. We also describe in detail several other algorithmic improvements in the FT framework which allow stable and efficient computations to low temperatures (as low as $T = 1/80$ in units of hopping in the Hubbard model).

In the benchmark, we mainly concentrate on 1/8 hole doping in the one-band Hubbard model, which has the most severe sign problem and is algorithmically very difficult. Physically it is a crucial regime to understand in the context of the phase diagram of cuprates. The results demonstrate that FT-CP-AFQMC is exact at very high temperatures and approaches ZT-CP with decreasing temperature. They show that the new FT-CP algorithm is capable of computing thermodynamic properties systematically with accuracy comparable to or even better than that of ZT-CP for ground-state properties.

In the rest of the paper, we will refer to the standard algorithm of BSS as DQMC and, unless necessary, refer to the CP-AFQMC framework as AFQMC for brevity, distinguishing the zero- and finite-temperature formalisms by ZT and FT respectively. As we will see in more detail below, the CP-AFQMC framework shares the basic formalism of DQMC; however, it reformulates the Monte Carlo process as branching random walks, using importance-sampling transformation, with an ensemble or population of walkers in the imaginary-time (inverse temperature) direction instead of the usual Metropolis sampling of the entire path in auxiliary-field space. The reason for the reformulation is twofold. One is to allow the imposition of the constraints without incurring ergodicity problems [10,29,50]. The other is that, for ground-state or ZT calculations, it provides a clear and formal connection [28,51] with standard electronic structure methods within density-functional theory, which has enabled general computations in solids and molecular systems.

The rest of the paper is organized as follows. In Sec. II, we summarize the DQMC algorithm, and establish the necessary concepts and formalisms which are also important parts of FT-AFQMC method. In Sec. III, we describe the FT-AFQMC method, focusing on the constrained path approximation, process of random walks with importance sampling, and several implementation issues for improved efficiency and stability in the numerical procedures. Then the self-consistent constraint and the forms of the constraining trial Hamiltonian or density matrix are discussed in Sec. IV, using the two-dimensional (2D) doped Hubbard model with repulsive interaction as a concrete example. In Sec. V, we present the benchmark results for this system, which demonstrates the performance and accuracy of the FT-AFQMC method. Finally, Sec. VI summarizes this work, and discusses applications of the method in other models and realistic molecules and materials.

II. DETERMINANTAL QUANTUM MONTE CARLO METHOD AND THE SIGN PROBLEM

There are various overlapping numerical details in DQMC and FT-AFQMC methods. Here we present a brief review of the DQMC method including the basic formalism and the sign problem. This will help illustrate the connection, and also introduce several concepts and set up our notation for describing AFQMC in the following sections.

A. Formalism of DQMC method

We begin with a general interacting model of spin-1/2 fermions, $\hat{H} = \hat{H}_0 + \hat{H}_I$, where $\hat{H}_0$ is the noninteracting part and $\hat{H}_I$ represents fermion-fermion interactions. $\hat{H}_0$ can be written as $\sum_{ij,\sigma\sigma'}(\mathbf{H}_0)_{i\sigma,j\sigma'}c^+_{i\sigma}c_{j\sigma'}(\sigma = \uparrow, \downarrow$ is the spin index), with $c^+_{i\sigma}$ ($c_{i\sigma}$) as the creation (annihilation) operator on site i. A chemical potential term is included in $\hat{H}_0$ implicitly. Given a finite-size lattice (or basis) with N_s sites, $\mathbf{H}_0 = \{(\mathbf{H}_0)_{i\sigma,j\sigma'}\}$ is the $2N_s \times 2N_s$ hopping matrix. If there are no spin-flip terms, $\mathbf{H}_0$ is block diagonal with respect to spin species: $\mathbf{H}_0 = \text{Diag}(\mathbf{H}^\uparrow_0, \mathbf{H}^\downarrow_0)$, where $\mathbf{H}^\sigma_0$ is the $N_s \times N_s$ matrix.

The DQMC method deals with the partition function of the system as

$$Z = \text{Tr}(e^{-\beta\hat{H}}) = \text{Tr}(\underbrace{e^{-\Delta\tau\hat{H}}\dots e^{-\Delta\tau\hat{H}}e^{-\Delta\tau\hat{H}}}_{M}), \quad (1)$$

where $\Delta\tau = \beta/M$ and M is the number of imaginary-time slices. For a small $\Delta\tau$, the Trotter-Suzuki decomposition, such as the asymmetric one $e^{-\Delta\tau\hat{H}} = e^{-\Delta\tau\hat{H}_I}e^{-\Delta\tau\hat{H}_0} + \mathcal{O}[(\Delta\tau)^2]$, and the Hubbard-Stratonovich transformation, generally expressed as $e^{-\Delta\tau\hat{H}_I} = \sum_{\mathbf{x}} p(\mathbf{x})\hat{B}_I(\mathbf{x})$, are applied to transform the many-body propagator $e^{-\Delta\tau\hat{H}}$ in Eq. (1) into single-particle operators expressed as free fermions coupled to auxiliary fields $\mathbf{x} = (x_1, x_2, \dots, x_{N_f})$ with N_f (comparable with N_s) components. The error $\mathcal{O}[(\Delta\tau)^2]$ originates from $[\hat{H}_0, \hat{H}_I] \neq 0$, and $\hat{B}_I(\mathbf{x}) = \exp\{\sum_{ij,\sigma\sigma'}[\mathbf{H}_I(\mathbf{x})]_{i\sigma,j\sigma'}c^+_{i\sigma}c_{j\sigma'}\}$ where $\mathbf{H}_I(\mathbf{x}) = \{[\mathbf{H}_I(\mathbf{x})]_{i\sigma,j\sigma'}\}$ is a $2N_s \times 2N_s$ Hermitian or anti-Hermitian matrix. Combining with the kinetic propagator $\hat{B}_K = e^{-\Delta\tau\hat{H}_0}$ which has no dependence on imaginary time,

we can rewrite the propagator $e^{-\Delta\tau\hat{H}^{(\ell)}}$ at the ℓth time slice as

$$e^{-\Delta\tau\hat{H}} = \sum_{\mathbf{x}_\ell} p(\mathbf{x}_\ell)\hat{B}_\ell + \mathcal{O}[(\Delta\tau)^2], \qquad (2)$$

where $\hat{B}_\ell = \hat{B}_I(\mathbf{x}_\ell)\hat{B}_K$. Applying this to all time slices in Eq. (1), we arrive at $Z \simeq \sum_{\mathbf{X}} P(\mathbf{X})\mathrm{Tr}(\hat{B}_M \ldots \hat{B}_2\hat{B}_1)$, where $P(\mathbf{X}) = \prod_{\ell=1}^{M} p(\mathbf{x}_\ell)$ is a probability density function and the auxiliary-field configuration $\mathbf{X} = \{\mathbf{x}_M, \ldots, \mathbf{x}_2, \mathbf{x}_1\}$ contains MN_f components. Since $\hat{B}_\ell$ is a single-particle propagator, the trace in the partition function can now be evaluated explicitly to yield

$$Z \simeq \sum_{\mathbf{X}} P(\mathbf{X})\det(\mathbf{I}_{2N_s} + \mathbf{B}_M \ldots \mathbf{B}_2\mathbf{B}_1), \qquad (3)$$

where $\mathbf{B}_\ell = \mathbf{B}_I(\mathbf{x}_\ell)\mathbf{B}_K$ with $\mathbf{B}_I(\mathbf{x}_\ell) = e^{\mathbf{H}_I(\mathbf{x}_\ell)}$ and $\mathbf{B}_K = e^{-\Delta\tau\mathbf{H}_0}$. In Eq. (3), $D(\mathbf{X}) = P(\mathbf{X})\det(\mathbf{I}_{2N_s} + \mathbf{B}_M \ldots \mathbf{B}_2\mathbf{B}_1)$ is the so-called "weight" of auxiliary-field configuration $\mathbf{X}$. With this, we have formally mapped the study of a D-dimensional quantum system for fermions with N_s lattice sites into solving a $(D+1)$ classical system with MN_f classical variables (or sites on a space-time lattice).

If the determinant satisfies $D(\mathbf{X}) \geqslant 0$, we can define a probability density function

$$W(\mathbf{X}) = \frac{D(\mathbf{X})}{\sum_{\mathbf{X}'} D(\mathbf{X}')} \qquad (4)$$

and sample it by Monte Carlo (for example with the Metropolis algorithm) to calculate physical observables

$$\langle\hat{O}\rangle = \frac{1}{Z}\mathrm{Tr}[e^{-(\beta-\tau)\hat{H}}\hat{O}e^{-\tau\hat{H}}] = \sum_{\mathbf{X}} O(\mathbf{X})W(\mathbf{X}), \qquad (5)$$

where the first and second equalities give the definition and the formula used in Monte Carlo calculations, respectively. In Eq. (5), the extra $\tau = \ell\Delta\tau$ with $0 \leqslant \ell \leqslant M$ is used to indicate that, although there is overall translational invariance in τ for $\langle\hat{O}\rangle$, the evaluation of $O(\mathbf{X})$, the measurement of $\hat{O}$ within configuration $\mathbf{X}$, can depend on imaginary time. For example, the static Green's-function matrix $\mathbf{G}(\tau, \tau) = \{G_{i\sigma, j\sigma'} = \langle c_{i\sigma}c_{j\sigma'}^+\rangle_\tau\}$ takes the simple form $\mathbf{G}(\tau, \tau) = (\mathbf{I}_{2N_s} + \mathbf{RL})^{-1}$ where $\mathbf{R} = \mathbf{B}_\ell \ldots \mathbf{B}_2\mathbf{B}_1$ and $\mathbf{L} = \mathbf{B}_M\mathbf{B}_{M-1} \ldots \mathbf{B}_{\ell+1}$. With this, the $O(\mathbf{X})$ for all single-particle and two-particle observables can be calculated directly or via Wick decomposition. Equation (5) is then used to obtain the statistical average and perform standard analysis to estimate the statistical error.

We remark on several technical aspects. (1) For the special case of spin decoupled systems, the matrices involved become block diagonal with respect to spin species, as mentioned earlier. The determinant $D(\mathbf{X})$ can be separated as $D(\mathbf{X}) = D_\uparrow(\mathbf{X})D_\downarrow(\mathbf{X})$. (2) We used the asymmetric Trotter-Suzuki decomposition in the discussion above. It is straightforward to apply the symmetric one, e.g., $e^{-\Delta\tau\hat{H}} = e^{-\Delta\tau\hat{H}_0/2}e^{-\Delta\tau\hat{H}_I}e^{-\Delta\tau\hat{H}_0/2} + \mathcal{O}[(\Delta\tau)^3]$. In any case, the systematic error from a finite $\Delta\tau$ can be removed by extrapolating several calculations with different $\Delta\tau$ values. (3) The DQMC approach is typically realized by a Markov-chain Monte Carlo algorithm. Ergodicity is required in the sampling of new auxiliary-field configurations and detailed balance

must be maintained in the update procedure. (4) The computational complexity of the DQMC algorithm is $\mathcal{O}(MN_s^3)$. (5) There can be an infinite variance problem [52] in sign-problem-free simulations with $D(\mathbf{X}) \geqslant 0$ which needs to be controlled. We will defer the discussion of numerical stabilization and efficient updating of auxiliary-field configurations to the sections on AFQMC, and discuss the sign problem in DQMC next.

B. Minus sign problem

For systems free of the sign or phase problem, special symmetries are typically present to keep $D(\mathbf{X}) \geqslant 0$ for all configurations, such as the antiunitary symmetry for the one-particle Hamiltonian after HS transformation [18,19]. Recently, the guiding principles for interacting fermion models which can be made free of the sign or phase problem have been largely extended [20–27]. However, these cases are still rare. In general, the determinant $D(\mathbf{X})$ cannot be made non-negative for all configurations, which leads to the minus sign problem. More generally, the $\mathbf{B}_\ell$ matrix can be complex, because of hopping terms in $\mathbf{H}_0$ from a twist angle or magnetic field, or spin-orbit coupling, and/or because of specific choices of the Hubbard-Stratonovich transformation in $\mathbf{H}_I$. This leads to a phase problem.

In the presence of a sign or phase problem, $W(\mathbf{X})$ in Eq. (4) can no longer be used as a probability density. It is of course straightforward to choose, for example, $W'(\mathbf{X}) = \frac{|D(\mathbf{X})|}{\sum_{\mathbf{X}'} |D(\mathbf{X}')|}$ instead, and replace the Monte Carlo average of the observable $\hat{O}$ in Eq. (5) with

$$\langle\hat{O}\rangle = \frac{\sum_{\mathbf{X}} \mathrm{sgn}(\mathbf{X})O(\mathbf{X})W'(\mathbf{X})}{\sum_{\mathbf{X}'} \mathrm{sgn}(\mathbf{X}')W'(\mathbf{X}')}, \qquad (6)$$

where $\mathrm{sgn}(\mathbf{X}) = D(\mathbf{X})/|D(\mathbf{X})|$ is the phase or sign of $D(\mathbf{X})$. The denominator in Eq. (6) gives the average of the sign (or phase):

$$\langle\mathrm{sgn}\rangle = \sum_{\mathbf{X}} \mathrm{sgn}(\mathbf{X})W'(\mathbf{X}) = \frac{\sum_{\mathbf{X}} D(\mathbf{X})}{\sum_{\mathbf{X}'} |D(\mathbf{X}')|}, \qquad (7)$$

which can be thought of as the ratio between two partition functions, one for our actual interacting fermion system and the other a fictitious system defined by the absolute value of the determinants. Note that $\langle\mathrm{sgn}\rangle$ is always a real number and below we will loosely refer to it vanishing as a sign problem, although in the case of a phase problem the imaginary part can grow to be of comparable magnitude as the real part, causing large fluctuations in the phase.

Implementing the above reweighting technique does not prevent the sign problem, of course. It can be shown both theoretically and numerically [16,17] that $\langle\mathrm{sgn}\rangle \propto e^{-\xi\beta N_s}$ for β larger than a specific value, where ξ is a positive constant dependent on the filling and interaction strength of the system. Thus, as β increases, $D(\mathbf{X})$ tends to approach an antisymmetric function and the sign average $\langle\mathrm{sgn}\rangle$ vanishes exponentially. The computational cost of DQMC simulations, for fixed statistical accuracy, scales exponentially with system size and inverse temperature for any $\langle\hat{O}\rangle$. As pointed out in Ref. [10] and further discussed below in Sec. III A, the symptoms and the origin of the sign problem can be understood with a simple

intuitive picture in terms of the path integrals in auxiliary-field space.

III. FINITE-TEMPERATURE CONSTRAINED PATH AUXILIARY-FIELD QUANTUM MONTE CARLO METHOD

The FT-AFQMC approach is an alternative method. A constraint is applied, based on exact considerations of the nature of the paths in auxiliary-field space but implemented approximately using a trial density matrix, to control the sign/phase in $D(\mathbf{X})$. The approximation can introduce a systematic error, but removes completely any decay of $\langle \text{sgn} \rangle$, and restores algebraic scaling. In this section, we discuss the basic formalism of FT-AFQMC method in the first part. Then we introduce the formulation of branching random walk with importance sampling, which is necessary for an efficient implementation. The last part contains several important technical aspects for improving numerical efficiency and stability.

A. Constrained path approximation

The CP approaches at both finite and zero temperatures are based on understanding the structure of the paths in auxiliary-field space, and devising rigorous constraints which effectively sum positive and negative contributions prior to the Monte Carlo sampling. Below we illustrate the idea at finite temperature using the formalism of DQMC discussed earlier.

Instead of starting from a full initial configuration $\mathbf{X}$ as in DQMC, we imagine that the complete path $\mathbf{X} = \{\mathbf{x}_M, \ldots, \mathbf{x}_2, \mathbf{x}_1\}$ is generated by M successive steps in imaginary time, from $\mathbf{x}_1$ to $\mathbf{x}_M$. Beginning with the partition function $Z = \text{Tr}(\hat{\mathcal{B}}\hat{\mathcal{B}} \ldots \hat{\mathcal{B}})$ where $\hat{\mathcal{B}} = e^{-\Delta\tau\hat{H}}$, we replace $\hat{\mathcal{B}}$ by Eq. (2) one at a time from the right to the left. After ℓ such steps, the partial paths $\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}$ are constructed and the partition function can be written as $Z \simeq \sum_{\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}} \mathcal{P}_\ell(\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}, \hat{\mathcal{B}})$ with $\mathcal{P}_\ell$ as

$$\mathcal{P}_\ell(\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}, \hat{\mathcal{B}})$$
$$= \prod_{n=1}^{\ell} p(\mathbf{x}_n) \cdot \text{Tr}(\underbrace{\hat{\mathcal{B}} \ldots \hat{\mathcal{B}}\hat{\mathcal{B}}}_{M-\ell} \underbrace{\hat{B}_\ell \ldots \hat{B}_2 \hat{B}_1}_{\ell}). \quad (8)$$

If we include all possible values of $\mathbf{x}_\ell$, all complete paths will be generated after M steps, and we recover the full path-integral representation of the partition function in Eq. (3).

Let us consider a case with a particular partial path $\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}$ of length ℓ, when $\mathcal{P}_\ell(\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}, \hat{\mathcal{B}}) = 0$. Mathematically this is equivalent to

$$\sum_{\{\mathbf{x}_M, \ldots, \mathbf{x}_{\ell+1}\}} \prod_{n=1}^{M} p(\mathbf{x}_n) \cdot \text{Tr}(\hat{B}_M \ldots \hat{B}_{\ell+1}\hat{B}_\ell \ldots \hat{B}_2 \hat{B}_1) = 0, \quad (9)$$

where we have replaced the remaining $(M - \ell)$ many-body operators $\hat{\mathcal{B}}$ in Eq. (8) by the corresponding $\hat{B}(\mathbf{x})$ operators with auxiliary fields, using Eq. (2). Equation (9) implies that the contributions of all complete paths whose first ℓ elements are the particular partial path $\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}$ simply cancel in the partition function Z, as the summation over all possible $\{\mathbf{x}_M, \ldots, \mathbf{x}_{\ell+2}, \mathbf{x}_{\ell+1}\}$ is equal to zero.

This observation allows us to divide all the complete paths (or the auxiliary-field space) into two categories according to

their contributions to the partition function Z: an antisymmetric part and the residual part. All the complete paths containing the partial path $\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}$ discussed above belong to the antisymmetric category. Complete paths in this category share the common characteristic that $\mathcal{P}_\ell$ defined in Eq. (8) encounters $\mathcal{P}_\ell = 0$ for at least one $\ell \in [1, M]$. Statistically, the complete paths belonging to the antisymmetric category are "noise," since the contributions of such paths cancel in Z. These noise paths increasingly dominate the auxiliary-field space with lowering temperature; they contribute to the exponential growth in the variance of the numerical results. On the other hand, complete paths belonging to the other category, the residual part, are responsible for the actual signal in the partition function; their contributions vanish exponentially with the length of the path. The overall effect gives rise to the behavior of the sign average discussed in Sec. II B, namely an exponential decay of signal-to-noise ratio that is the signature of the sign problem.

The recognition of the different categories of paths provides an understanding of the origin of the sign problem. It shows that a complete path contributes if and only if all of the following M conditions are satisfied in the generation of the path:

$$\mathcal{P}_\ell(\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}, \hat{\mathcal{B}}) > 0, \quad \ell = 1, 2, \ldots, M. \quad (10)$$

The constraints in Eq. (10) represent an absorbing boundary condition (BC) at $\mathcal{P}_\ell = 0$ with increasing ℓ. Under this BC, the probability distribution of the generated paths vanishes smoothly close to the boundary. Thus application of these constraints will eliminate all noise paths while keeping the contributing ones. Since the absorbing BC filters out the antisymmetric part in auxiliary-field space, the algorithm remains exact. This is the basic idea of the FT-AFQMC method.

To realize the FT-AFQMC method numerically, several additional issues must be addressed. First, a finite $\Delta\tau$ is always used in practical simulations, which makes it difficult to determine $\mathcal{P}_\ell = 0$ precisely. Since $\mathcal{P}$ is continuous in ℓ only at the $\Delta\tau \to 0$ limit, the random walk will typically cross the absorbing boundary "in between" two successive time steps. This problem turns out to be straightforward to handle. In the simplest approximation, the absorbing boundary can be identified by ℓ corresponding to the first passage to $\mathcal{P}_\ell < 0$. Alternatively, a higher-order approach called mirror correction [9,29,53] can be applied, which introduces a finite probability to terminate the path at $(\ell - 1)$ for the case when $\mathcal{P}_{\ell-1} > 0$ and $\mathcal{P}_\ell < 0$. These two approaches locate the absorbing boundary with controllable systematic errors of $\mathcal{O}(\Delta\tau)$ and $\mathcal{O}[(\Delta\tau)^2]$, respectively, which vanish as $\Delta\tau \to 0$ and can be extrapolated away with finite $\Delta\tau$ calculations.

A second, much more significant issue is that $\hat{\mathcal{B}}$ in Eq. (10) is a many-body operator, which makes the exact calculation of $\mathcal{P}_\ell(\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}, \hat{\mathcal{B}})$ in Eq. (8) a many-body problem. In practice, we replace it by a known trial propagator $\hat{B}_T = e^{-\Delta\tau\hat{H}_T}$, where $\hat{H}_T$ can be thought of as a trial Hamiltonian. This substitution causes a systematic error, since now the absorbing BC is only determined approximately. This is the constrained path approximation and it results in the only systematic error of the FT-AFQMC method. (All other numerical errors can be systematically removed, such as the

finite-$\Delta\tau$ error, as discussed above, or population control error.) The constrained path approximation becomes exact if $\hat{B}_T$ is exact. In practice the trial propagator $\hat{B}_T$ is typically of single-particle form. In other words the trial $\hat{H}_T$ is often a single-particle Hamiltonian: $\hat{H}_T = \sum_{ij,\sigma\sigma'} (\mathbf{H}_T)_{i\sigma,j\sigma'} c_{i\sigma}^+ c_{j\sigma'}$, where $\mathbf{H}_T = \{(\mathbf{H}_T)_{i\sigma,j\sigma'}\}$ is a $2N_s \times 2N_s$ Hermitian matrix. Using the trial $\hat{B}_T$, the $\mathcal{P}_\ell$ in Eq. (8) can be rewritten in the operator and matrix determinant forms as

$$
\begin{aligned}
\mathcal{P}_\ell^T(\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}) &= \mathcal{P}_\ell(\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}, \hat{B}_T) \\
&= \prod_{n=1}^{\ell} p(\mathbf{x}_n) \cdot \mathrm{Tr}\Big(\underbrace{\hat{B}_T \ldots \hat{B}_T \hat{B}_T}_{M-\ell} \hat{B}_\ell \ldots \hat{B}_2 \hat{B}_1 \Big) \\
&= \prod_{n=1}^{\ell} p(\mathbf{x}_n) \cdot \det\left[\mathbf{I}_{2N_s} + \left(\prod_{n=1}^{M-\ell} \mathbf{B}_T\right) \mathbf{B}_\ell \ldots \mathbf{B}_2 \mathbf{B}_1 \right],
\end{aligned}
\tag{11}
$$

where the matrix $\mathbf{B}_T = e^{-\Delta\tau \mathbf{H}_T}$. Then the constraints in Eq. (10) are replaced by the following:

$$
\mathcal{P}_\ell^T(\{\mathbf{x}_\ell, \ldots, \mathbf{x}_2, \mathbf{x}_1\}) > 0, \quad \ell = 1, 2, \ldots, M, \tag{12}
$$

which is imposed successively in imaginary time from $\ell = 1$ to M.

B. Random walk with importance sampling

We next discuss an algorithm for sampling the paths efficiently while imposing the constraints in Eq. (12). Our goal is to generate complete paths (configurations) $\mathbf{X}$ which both satisfy the constraints and are distributed according to $D(\mathbf{X})$. We construct such contributing paths following exactly the process outlined in the thought experiment in Sec. III A. Instead of sweeping through the entire path for updates as in the DQMC method, we construct the complete path $\mathbf{X} = \{\mathbf{x}_M, \ldots, \mathbf{x}_2, \mathbf{x}_1\}$ from $\mathbf{x}_1$ to $\mathbf{x}_M$ using a branching random walk, while imposing the constraints in Eq. (12).

To introduce importance sampling into the random-walk process to improve sampling efficiency, we note that the partition function can be rewritten as

$$
Z \simeq \sum_{\{\mathbf{x}_M, \ldots, \mathbf{x}_2, \mathbf{x}_1\}} \frac{\mathcal{P}_M^T}{\mathcal{P}_{M-1}^T} \frac{\mathcal{P}_{M-1}^T}{\mathcal{P}_{M-2}^T} \cdots \frac{\mathcal{P}_2^T}{\mathcal{P}_1^T} \frac{\mathcal{P}_1^T}{\mathcal{P}_0^T} \mathcal{P}_0^T, \tag{13}
$$

where $\mathcal{P}_\ell^T$ is shorthand for the full expression in Eq. (11), and $\mathcal{P}_M^T = D(\mathbf{X})$. Starting from $\mathcal{P}_0^T = \det(\mathbf{I}_{2N_s} + \prod_{n=1}^M \mathbf{B}_T) > 0$, we first use $\lambda(\mathbf{x}_1) = \max[\mathcal{P}_1^T/\mathcal{P}_0^T, 0]$ to construct a normalized probability density function $\eta(\mathbf{x}_1) = \lambda(\mathbf{x}_1)/\sum_{\mathbf{x}_1'} \lambda(\mathbf{x}_1')$. We then draw a sample for $\mathbf{x}_1$ from $\eta(\mathbf{x}_1)$, and assign the normalization factor $\sum_{\mathbf{x}_1'} \lambda(\mathbf{x}_1')$ as a weight of the newly sampled path. Note that the constraints in Eq. (12) for $\ell = 1$ have automatically been implemented by our choice of the probability density function. We then repeat the same procedure from $\mathbf{x}_2$ to $\mathbf{x}_M$. At the ℓth step, with the partial path $\{\mathbf{x}_{\ell-1}, \ldots, \mathbf{x}_2, \mathbf{x}_1\}$ already constructed, we use the conditional PDF $\eta(\mathbf{x}_\ell) = \lambda(\mathbf{x}_\ell)/\sum_{\mathbf{x}_\ell'} \lambda(\mathbf{x}_\ell')$ to sample $\mathbf{x}_\ell$, where $\lambda(\mathbf{x}_\ell) = \max[\mathcal{P}_\ell^T/\mathcal{P}_{\ell-1}^T, 0]$. The weight of the path is multiplied by the normalization $\sum_{\mathbf{x}_\ell'} \lambda(\mathbf{x}_\ell')$.

In the calculation, we carry an ensemble of $N_\mathbf{X}$ samples and propagate them in parallel. Each sample is called a random walker, and the random walk is carried out for M steps. During the random walk, the walker weights can fluctuate and a population control procedure is applied periodically, as discussed in the next section. At the end of the random walk, $N_\mathbf{X}$ complete paths are obtained, given by $\mathbf{X}_k$ with weight w_k for $k = 1$ to $N_\mathbf{X}$. They provide Monte Carlo samples of a modified probability density function $W^c(\mathbf{X})$ as defined in Eq. (4) but with $D(\mathbf{X})$ replaced by $D^c(\mathbf{X})$, where the superscript "c" means "under the constraint" of Eq. (12). Observables can then be computed as in Eq. (5), which is further discussed in the next section. We repeat the M-step random walk procedure as needed to reach the desired statistical accuracy.

The FT-AFQMC method, as is now evident, does not use the Markov-chain Monte Carlo employed in DQMC. The choice is driven by the difficulty in imposing the constraints in the path-integral formalism in DQMC. The constraining conditions are nonlocal in imaginary time, since the condition at the ℓth step depends on the path history from 1 to $(\ell - 1)$, or alternatively from M to $(\ell + 1)$ depending on how one views the reference point (or sweeping direction in the path sampling). This means that the sampling could get "stuck" with a configuration which violates the absorbing boundary condition. As we see from the analysis in Sec. III A, this will occur with higher and higher probability as the path becomes longer (lower temperature). The one-directional random walk with branching adopted in FT-AFQMC, which is similar to ZT-AFQMC, solves this problem.

In the discussion so far, we have implied that trial Hamiltonian $\hat{H}_T$ (or trial propagator $\hat{B}_T$) has no imaginary-time dependence. (Note that the time dependence of $\hat{B}_T$ should be distinguished from that of the constraint; because of the product form in Eq. (11), the constraint is time dependent even if $\hat{B}_T$ is not.) It is straightforward to generalize the procedure to an imaginary-time dependent $\hat{H}_T$.

C. Implementation issues for numerics

In this section we discuss several issues in a general implementation of the FT-AFQMC method for correlated fermion systems. They include numerical stabilization, growth estimator and population control, measurements, and the form of $\hat{H}_T$ and its implementation. Additional details are provided in Appendix A.

Finite-temperature AFQMC calculations, like DQMC, are more challenging to keep numerically stable than zero-temperature calculations. The instability grows more severe at lower temperatures, and is caused by numerical round-off errors [8,54,55] from the multiplication of many $\mathbf{B}$ matrices. In our calculations, we use the column-pivoted QR algorithm to stabilize the matrix products and we also implement a very stable way [54,55] to calculate the single-particle Green's-function matrix shown in Appendix A. Combining them, we can access temperatures as low as $1/80$ in units of the hopping parameter (or $1/20$ in units of Fermi energy).

In the FT-AFQMC method, the weights of walkers, w_k, can fluctuate. The procedure to keep these fluctuations under control and maintain statistical accuracy is similar to that used

in ground-state AFQMC calculations [29] and in diffusion Monte Carlo calculations [6]. Several schemes are possible whose details differ somewhat and can affect the statistical accuracy and population control bias, but the effects are negligible (much smaller than the statistical errors) in our calculations with $\mathcal{O}(10^3–10^4)$ random walkers. Here we will not make distinctions of the technical details of the population control schemes, but simply describe the simplest approaches.

In our calculations, a fixed population of N_X random walkers are kept. The weights of the walkers are carried. We often use a combing algorithm, which periodically resets all weights to unity by resampling them according to the normalized distribution $w_k / \sum_k w_k$. Alternatively, we monitor the weights for large and small values which are predefined thresholds (e.g., 5.0 and 0.2, respectively). If the overall weight $\sum_k w_k$ systematically increases or decreases with imaginary time, we can adjust all weights by a constant factor as needed, similar to the growth estimator of ground-state calculations [29]. Walkers with large weights are duplicated and those with small weights are eliminated with the appropriate probability (to maintain a statistically identical population).

The measurement of observables are quite straightforward after the contributing paths are generated with importance sampling. In the branching random-walker formulation, every single random walker generates a contributing path X_k with the final overall weight w_k as its contribution to the partition function Z. Thus the physical observable $\langle \hat{O} \rangle$ can be computed as

$$\langle \hat{O} \rangle = \frac{\sum_{k=1}^{N_X} \omega_k O_k}{\sum_{k=1}^{N_X} \omega_k}, \tag{14}$$

where O_k is the measurement result within the configuration X_k. Since the contributing path is generated from $\ell = 1$ to $\ell = M$ time slices, the simplest way to calculate O_k is to measure it at the Mth time slice (or $\tau = \beta$):

$$O_k = \frac{\text{Tr}(\hat{O}\hat{B}_M \hat{B}_{M-1} \ldots \hat{B}_2 \hat{B}_1)}{\text{Tr}(\hat{B}_M \hat{B}_{M-1} \ldots \hat{B}_2 \hat{B}_1)}. \tag{15}$$

The evaluation of O_k in Eq. (15) is the same as discussed in Sec. II A for DQMC. Although imaginary-time translational invariance is broken in FT-AFQMC, it is reasonable to expect that measurements at different times are of comparable quality and will become more equivalent with better trial $\hat{B}_T$. We generally take the average of the multiple measurements at different imaginary times, which can be easily achieved by wrapping the $\hat{B}_\ell$'s after the complete path has been sampled. The extra computational cost is often more than compensated for by the gain in statistical accuracy. Furthermore, we find that the time-averaged results, which partially recover translational invariance in imaginary time, tend to have smaller systematic error than the single-time measurement at β. The procedure for measuring imaginary-time correlation functions is straightforward [11,56].

We next comment on the operations involved in substituting the trial $\hat{B}_T$ by the interacting propagators in the sampling process. The Hamiltonian can be written as $\hat{H} = \hat{H}_0 + \hat{H}_I = \hat{H}_T + (\hat{H}_I + \hat{H}_0 - \hat{H}_T)$. The trial propagator is $\hat{B}_T = e^{-\Delta\tau \hat{H}_T}$, and we need to insert the $e^{-\Delta\tau(\hat{H}_I + \hat{H}_0 - \hat{H}_T)}$ operator. Operationally, this means replacing $\hat{B}_T$ by $\hat{B} = \hat{B}_I(\mathbf{x})\hat{B}_K$ at

every time slice. The overall weight of the generated path comes from two parts: the importance sampling of the auxiliary fields and the ratio of the determinants when changing $\mathbf{B}_T$ to $\mathbf{B}_K$. In the special case when we choose $\hat{H}_T = \hat{H}_0$, as is the case when a restricted Hartree-Fock form is used, there is an additional simplification, and we only need to insert the corresponding $\mathbf{B}_I(\mathbf{x}_\ell)$ matrix. Further details and the formulas for $\mathcal{P}_\ell^T / \mathcal{P}_{\ell-1}^T$ are provided in Appendix A.

IV. FT-AFQMC WITH SELF-CONSISTENT CONSTRAINT

In this section, we present a self-consistent constraint in FT-AFQMC, after first carrying out a systematic study of the effect of different choices of mean-field trial density matrices. For concreteness, we will use the doped two-dimensional repulsive Hubbard model as an example. However, much of the discussion can be generalized to other Hamiltonians, including realistic electronic Hamiltonians under the phaseless formalism of the constraint [28].

A. Trial density matrix: Illustration in the Hubbard model

The one-band Hubbard model is a representative model for studying correlation effects of interacting electrons. The model Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_I$ is as follows [57]:

$$\hat{H} = -t \sum_{\langle ij \rangle \sigma} (c_{i\sigma}^+ c_{j\sigma} + c_{j\sigma}^+ c_{i\sigma}) + \mu \sum_i (\hat{n}_{i\uparrow} + \hat{n}_{i\downarrow})$$
$$+ U \sum_i \left(\hat{n}_{i\uparrow}\hat{n}_{i\downarrow} - \frac{\hat{n}_{i\uparrow} + \hat{n}_{i\downarrow}}{2} \right), \tag{16}$$

where $\hat{n}_{i\sigma} = c_{i\sigma}^+ c_{i\sigma}$ is the density operator on the lattice site $i = (i_x, i_y)$. The nearest-neighbor hopping t, on-site Coulomb interaction U, and chemical potential μ are model parameters. In this work, we focus on repulsive interaction, $U > 0$. The Hamiltonian above is written such that $\mu = 0$ gives half filling, with $\mu > 0$ for hole doping and $\mu < 0$ for electron doping. The overall electron density is given as $n = (N_\uparrow + N_\downarrow)/N_s$, and the hole density, or doping, is then $(1 - n)$.

To choose a trial density matrix, the simplest $\hat{H}_T$ to consider is the restricted Hartree-Fock (RHF) type:

$$\hat{H}_T = -t \sum_{\langle ij \rangle \sigma} (c_{i\sigma}^+ c_{j\sigma} + c_{j\sigma}^+ c_{i\sigma}) + \sum_i \mu_{i,T}(\hat{n}_{i\uparrow} + \hat{n}_{i\downarrow}), \tag{17}$$

where $\mu_{i,T}$ are free parameters. In this work, we only consider the simplified case of $\mu_{i,T} = \mu_T$, namely keeping translational invariance. Taking $\mu_T = \mu$ simply gives $\hat{H}_T = \hat{H}_0$. More importantly, μ_T can be used to tune the electron filling $\langle \hat{n} \rangle_T$ for the trial Hamiltonian.

Another trial Hamiltonian for Hubbard model is the unrestricted Hartree-Fock (UHF) type:

$$\hat{H}_T = -t \sum_{\langle ij \rangle \sigma} (c_{i\sigma}^+ c_{j\sigma} + c_{j\sigma}^+ c_{i\sigma})$$
$$+ \sum_{i\sigma} \left[U_T \left(\langle \hat{n}_{i\bar{\sigma}} \rangle - \frac{1}{2} \right) + \mu_T \right] \hat{n}_{i\sigma}, \tag{18}$$

where $\bar{\sigma}$ denotes the opposite of σ, and $\{\langle \hat{n}_{i\uparrow} \rangle, \langle \hat{n}_{i\downarrow} \rangle, i = 1, 2, \ldots, N_s\}$ are from a self-consistent solution with $U = U_T$. The μ_T parameter here is similar to that in the RHF

trial Hamiltonian in Eq. (17). We have omitted a constant term which only affects the overall weights as discussed in the previous section.

Both the RHF and UHF trial Hamiltonians resemble (or are identical) to the $\hat{H}_0$ term in the many-body Hamiltonian in Eq. (16). Thus computational simplifications exist in the updating procedure as discussed in the previous section. Of course in principle any other form of single-particle $\hat{H}_T$ can be used as trial Hamiltonian.

B. FT-AFQMC calculations with self-consistent procedure

The self-consistent procedure we develop here will allow us to improve the constraint based on feedback from the AFQMC calculation. It generalizes the method discussed in Ref. [40] under the ZT-AFQMC framework to finite temperatures. We will optimize the trial Hamiltonian (or trial propagator) step by step iteratively through a sequence of FT-AFQMC calculations so as to reduce the error from the constraint.

For illustrating the method and testing our results, we will consider a more general form of the Hamiltonian of Eq. (16), by including an extra external potential term $\sum_{i\sigma} v_{i\sigma}\hat{n}_{i\sigma}$. This will allow us to add magnetic pinning fields [58–60] to break translational invariance in illustrating the method and testing our results. With pinning fields, spin and charge correlations in periodic systems can be probed by one-body spin and charge order parameters, whose measurements are straightforward from the single-particle Green's functions.

The basic idea of a self-consistent procedure is to couple the QMC calculation to an independent-particle (IP) calculation. The IP calculation deals with the following Hamiltonian, including an external pinning field:

$$
\hat{H}_{\mathrm{IP}} = -t \sum_{\langle ij \rangle \sigma} (c_{i\sigma}^+ c_{j\sigma} + c_{j\sigma}^+ c_{i\sigma}) + \sum_{i\sigma} v_{i\sigma}\hat{n}_{i\sigma}
$$
$$
+ \sum_{i\sigma} \left[U_{\mathrm{eff}}\left(\langle \hat{n}_{i\bar{\sigma}} \rangle - \frac{1}{2} \right) + \mu_{\mathrm{eff}} \right] \hat{n}_{i\sigma}, \quad (19)
$$

where U_{eff} and μ_{eff} are tuning parameters, and the input densities $\{\langle \hat{n}_{i\uparrow} \rangle, \langle \hat{n}_{i\downarrow} \rangle\}$ are from a preceding QMC calculation. The solution is used to construct the input trial Hamiltonian for the next QMC calculation in the self-consistent procedure.

We outline the method in concrete steps:

(1) Start from a FT-AFQMC calculation using any typical choice of trial Hamiltonian $\hat{H}_T$, for example RHF or UHF.

(2) $\hat{H}_{\mathrm{IP}}$ in Eq. (19) is solved using the densities obtained from the AFQMC calculation in the previous step: $\langle \hat{n}_{i\sigma} \rangle_{\mathrm{QMC}} \rightarrow \langle \hat{n}_{i\sigma} \rangle_{\mathrm{IP}}$. We vary U_{eff} to find an optimal value, with which the computed densities *at the targeted temperature* are closest to the input from QMC, by minimizing the following function:

$$
\chi = \sqrt{ \frac{1}{2N_s} \sum_{i\sigma} (\langle \hat{n}_{i\sigma} \rangle_{\mathrm{IP}} - \langle \hat{n}_{i\sigma} \rangle_{\mathrm{QMC}})^2 }. \quad (20)
$$

(3) The IP Hamiltonian $\hat{H}_{\mathrm{IP}}$ with the optimal U_{eff} and input densities from the previous AFQMC calculation is taken as trial Hamiltonian $\hat{H}_T$ to perform a new FT-AFQMC calculation.

(4) Return to step (2) or stop if convergence criteria is reached.

As mentioned, μ_{eff} in Eq. (19) is a tuning parameter similar to μ_T in the RHF and UHF trial Hamiltonians in Eqs. (17) and (18). In our calculations, we tune μ_{eff} to make $\langle \hat{n} \rangle_{\mathrm{IP}}$ equal to the desired electron filling of the many-body system.

Two somewhat different schemes can be used to obtain $T > 0$ results at each electron filling $\langle \hat{n} \rangle$. We can perform the above self-consistent procedure at T within FT-AFQMC calculations. Alternatively, we can first apply the self-consistent procedure using ZT-AFQMC method [40] to obtain the converged results of U_{eff} and densities at $T = 0$. Substituting these converged results into the IP Hamiltonian $\hat{H}_{\mathrm{IP}}$ in Eq. (19), we can then take $\hat{H}_{\mathrm{IP}}$ to generate a UHF trial Hamiltonian once for all to perform FT-AFQMC calculations. The temperature dependence of the UHF trial Hamiltonian only lies in μ_{eff} parameter, which is again tuned so that $\langle \hat{n} \rangle_{\mathrm{IP}}$ match the filling $\langle \hat{n} \rangle$. We have done a careful comparison of these two schemes in the benchmark study. Both were found to yield accurate results with only slight differences visible at high and low temperatures. We will only present the results of finite-T self-consistent calculations in the following.

V. BENCHMARK RESULTS

In this section we present benchmark results of FT-AFQMC calculations by comparing them with those from DQMC at $T > 0$ and ED, DMRG, and ZT-AFQMC at $T = 0$. The results are divided into two parts. In the first part, we perform a detailed test of the accuracy of "one-shot" FT-AFQMC calculations, using periodic supercells. Then in the second part we study the self-consistent procedure with FT-AFQMC; calculations are presented in the second part, with Hubbard model on a studying mostly systems with pinning fields. When applying the antiferromagnetic (AFM) pinning fields to an $L_x \times L_y$ supercells, we add a term $\sum_{i\sigma} v_{i\sigma}\hat{n}_{i\sigma}$ with $v_{i\uparrow} = -v_{i\downarrow} = (-1)^{i_y} h$ for both $i_x = 1$ (and sometimes also $i_x = L_x$ as we will specify below) in Eq. (16). The supercell remains periodic along the y direction and open along the x direction. With pinning fields, spin and charge correlations in periodic systems can be probed by one-body spin and charge order parameters, whose measurements are straightforward from the single-particle Green's functions.

We apply the symmetric decomposition $e^{-\Delta\tau\hat{H}} = e^{-\Delta\tau\hat{H}_0/2} e^{-\Delta\tau\hat{H}_U} e^{-\Delta\tau\hat{H}_0/2} + \mathcal{O}[(\Delta\tau)^3]$ and use the following discrete form [61] for Eq. (2):

$$
e^{-\Delta\tau U[n_{i\uparrow}n_{i\downarrow} - (n_{i\uparrow} + n_{i\downarrow})/2]} = \sum_{x_i = \pm 1} \frac{1}{2} e^{\lambda x_i (n_{i\uparrow} - n_{i\downarrow})} \quad (21)
$$

in both DQMC and FT-AFQMC calculations. In Eq. (21), $\lambda = \cosh^{-1}(e^{\Delta\tau U/2})$ for $U > 0$ and x_i is an auxiliary field.

Conservative values of $\Delta\tau$, typically 0.05 or 0.02, are used, and the same $\Delta\tau$ is applied in DQMC and FT-AFQMC when comparing their numerical results. The population of random walkers ranges from 10^3 to 10^4. All the results we will present are from the multiple measurements along the contributing paths as discussed in Sec. III C, unless otherwise noted.
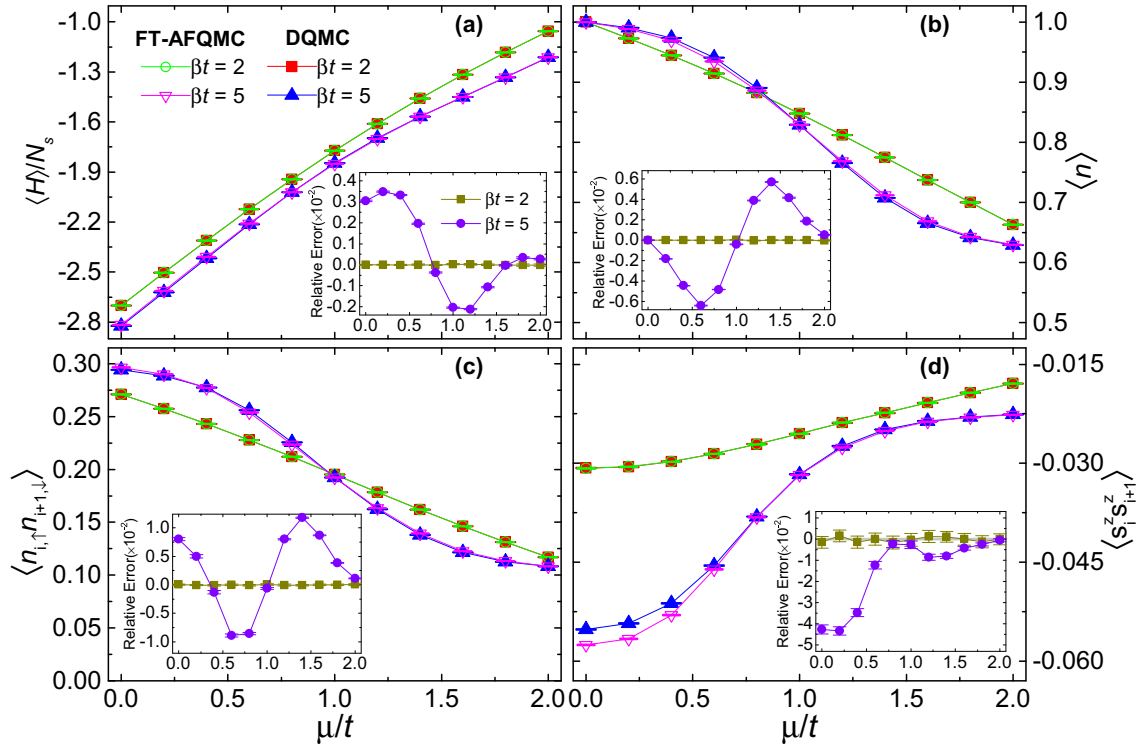
FIG. 1. Benchmark of FT-AFQMC at high temperatures. Results are shown for Hubbard model under periodic boundary conditions, on a 4×4 system with $U/t = 4$ and varying μ/t, comparing FT-AFQMC with DQMC calculations at $\beta t = 2$ and $\beta t = 5$. The four panels give (a) Total energy per site, (b) electron filling, (c) NN density-density correlation function $\langle \hat{n}_{i,\uparrow} \hat{n}_{i+1,\downarrow} \rangle$, and (d) NN spin-spin correlation function $\langle \hat{s}_i^z \hat{s}_{i+1}^z \rangle$. The relative errors of FT-AFQMC results are shown in the corresponding inset plots.

A. Numerical results from FT-AFQMC calculations with specific trial Hamiltonians

We first benchmark the FT-AFQMC at high temperatures. The Hubbard model on a 4×4 square lattice with periodic boundary conditions will be used, for availability of exact results. We study $\beta t = 2$ and $\beta t = 5$, varying the μ parameter to examine different electron fillings. The sign problem is mild in these situations, and we obtain accurate results from DQMC calculations for comparison. In AFQMC, the RHF trial Hamiltonian of Eq. (17) is applied. We determine μ_T by the condition $\langle \hat{n} \rangle_T = \langle \hat{n} \rangle$, i.e., the electron filling of $\hat{H}_T$ is equal to the targeted filling of the many-body system, as defined by DQMC. The choice of μ_T is further discussed below.

Results for the energy, electron filling, density, and spin correlations are shown in Fig. 1 for both DQMC and FT-AFQMC. It is evident that the results from both methods are perfectly consistent at $\beta t = 2$ for all μ parameters, with zero relative discrepancy within statistical error bars. This confirms that the FT-AFQMC results are exact at high temperature, even with the simplest RHF trial Hamiltonian. (For the $\beta t = 2$ results shown, we actually took $\mu_T = \mu$, i.e., $\hat{H}_T = \hat{H}_0$. Tuning it with respect to the exact density yields indistinguishable results.) For $\beta t = 5$, FT-AFQMC calculations with RHF trial Hamiltonian also generate quite accurate results, with the largest relative error <1% for total energy, electron filling, density correlation function, and about 4% for spin-correlation function. For spin-correlation function, the largest deviation appears at half filling with $\mu = 0$. This

can be improved by using a trial Hamiltonian with UHF and becomes essentially exact in ZT-AFQMC with a generalized Hartree-Fock (GHF) trial state [37,62].

We next study more closely the role of μ_T in the RHF trial Hamiltonian, with additional FT-AFQMC calculations at $\beta t = 5$. Two representative points are checked, $\mu = 0.4t$ and $\mu = 1.0t$, where relative errors shown in Fig. 1 are close to the largest. We vary μ_T and study the corresponding systematic error of FT-AFQMC, $\langle \hat{O} \rangle_{\text{FT-AFQMC}} - \langle \hat{O} \rangle_{\text{DQMC}}$. For $\mu = 0.4t$, as shown in Fig. 2(a), all the physical observables have the smallest systematic errors at the μ_T/t value determined from the $\langle \hat{n} \rangle_T = \langle \hat{n} \rangle_{\text{DQMC}}$ condition. That is not the case for $\mu = 1.0t$ as shown in Fig. 2(b), for which the smallest errors for different observables are located at different μ_T/t values. Thus no μ_T exists within the RHF trial Hamiltonian framework which "optimizes" the FT-AFQMC calculation in an absolute sense. However, at the position of μ_T/t determined from the exact electron filling condition, the systematic errors of all the quantities are very close to the minimum, validating our earlier choice. Furthermore, the errors all vary slowly (note the small scale of the errors in the plot) in a broad range of μ_T/t values, indicating that the constraint is not very sensitive to the details of the trial Hamiltonian.

Now we study a broader temperature range and examine the behavior of the FT-AFQMC algorithm for accessing sufficiently low temperatures to approach the ground state. We focus on $U = 4t$ with 1/8 hole doping, i.e., $\langle \hat{n} \rangle = 0.875$, and a temperature range from $T = t/2$ to $T = t/80$. Here we apply an AFM pinning field $h = 0.1t$ on one edge, at $i_x = 1$,
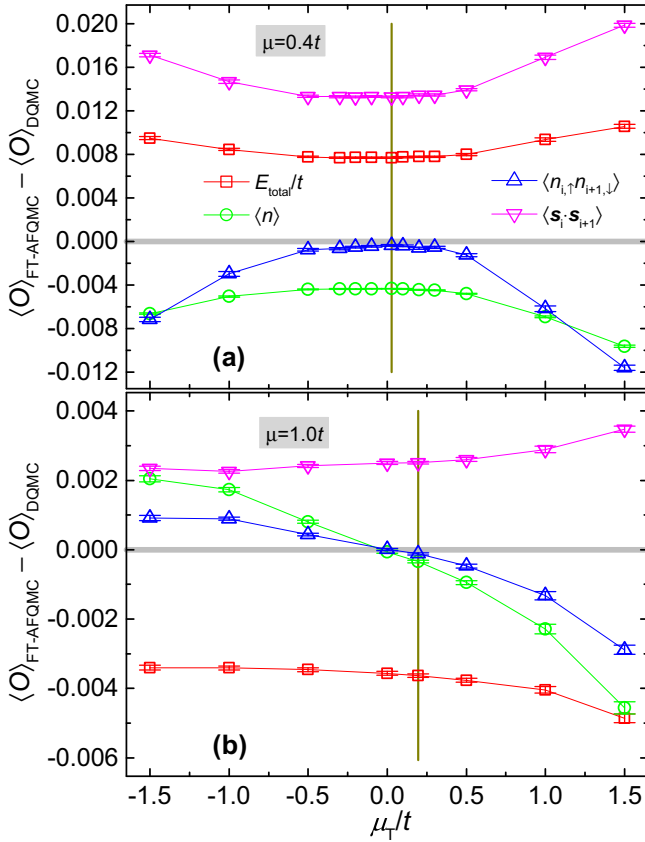
FIG. 2. Systematic errors from FT-AFQMC calculations as the μ_T/t parameter in the applied RHF trial Hamiltonian $\hat{H}_T$ is varied. The system is the same as in Fig. 1 at $\beta t = 5$. (a) $\mu = 0.4t$ and (b) $\mu = 1.0t$ are studied as representative cases. Gray, horizontal line represents zero error and the dark yellow, vertical line stands for the μ_T/t value in Fig. 1, determined from $\langle \hat{n} \rangle_T = \langle \hat{n} \rangle_{\text{DQMC}}$.

FIG. 3. Benchmark results over the complete range of temperatures. The system is a 4×4 square lattice with $U/t = 4$ and AFM pinning fields $h = 0.1t$ at $i_x = 1$, targeting $\langle \hat{n} \rangle = 0.875$. The computed total energy vs temperature is shown in (a), and the relative errors are shown with respect to DQMC for $T = t/15 \sim t/2$ region and then with respect to ED at $T = 0\,$K. Measure A denotes doing a single measurement at $\tau = \beta$ in FT-AFQMC, while measure B denotes averaging multiple measurements along the whole path. The computed density is shown in (b), with the numerically determined μ parameter shown in the inset together with the μ_T parameter in the RHF trial Hamiltonian $\hat{H}_T$.

with an open boundary condition the along x direction in the 4×4 system. All reference calculations, ED, ZT-AFQMC, and DQMC, are performed on the same Hamiltonian and parameter choices. The lowest temperature accessible here with DQMC is $T = t/15$, with average sign of 0.031. We use the RHF trial Hamiltonian (with the same AFM pinning fields) in FT-AFQMC calculations.

The results of this benchmark are shown in Fig. 3. In panel (a), energies computed from FT-AFQMC at high temperature, e.g., $\beta t = 2$, are indistinguishable from DQMC as seen earlier. As T decreases, the systematic error becomes visible and tends to increase, with a maximum of about 0.2% deviation at the lowest temperature available from DQMC, $T = t/15$. In the lower temperature region, we find that the energy with measure B obtained from FT-AFQMC fits accurately to an exponential, $E = E_0 + be^{-c/T}$. The extrapolated value E_0 at $T = 0$ is consistent with the result from ZT-AFQMC (measured with back propagation). This consistency unifies the FT- and ZT-AFQMC framework. Thus we expect a maximum relative error, as $T \to 0$, of $\sim 0.24\%$ as given by ZT-AFQMC, as shown in inset of Fig. 3(a). This is very accurate, especially since it is obtained with the simplest RHF form of the trial Hamiltonian. The energy results from measure A (single measurement at $\tau = \beta$) and measure B (averaging multiple
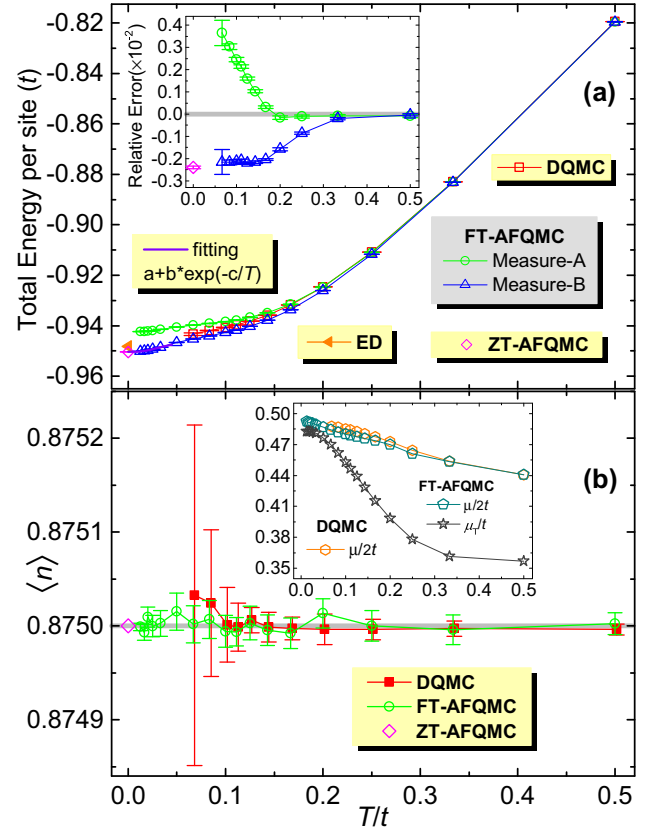
measurements along the complete path) show significant differences at low temperatures, as seen in Fig. 3(a) and the inset. The improvement from measure B at low temperatures is an effect of partially restored imaginary-time translational symmetry, as mentioned in Sec. III C. As seen in Fig. 3(b), the numerically calculated overall filling $\langle \hat{n} \rangle$ from both DQMC and FT-AFQMC methods are indeed equal to the desired value 0.875, within statistical error bars, across the entire range of temperature. The calculated electron fillings are the same for measure A and measure B, since the density operator commutes with the $\hat{B}$ used. As shown in inset of Fig. 3(b), the μ parameters determined from DQMC and FT-AFQMC calculations, to produce the desired density, are in good agreement at high temperatures, but discrepancy is visible with decreasing T. This is a consequence of the constrained path approximation in FT-AFQMC. The μ_T parameter in the RHF trial Hamiltonian for AFQMC, determined via the density condition $\langle n \rangle_T = \langle n \rangle = 0.875$, is also indicated in the plot.
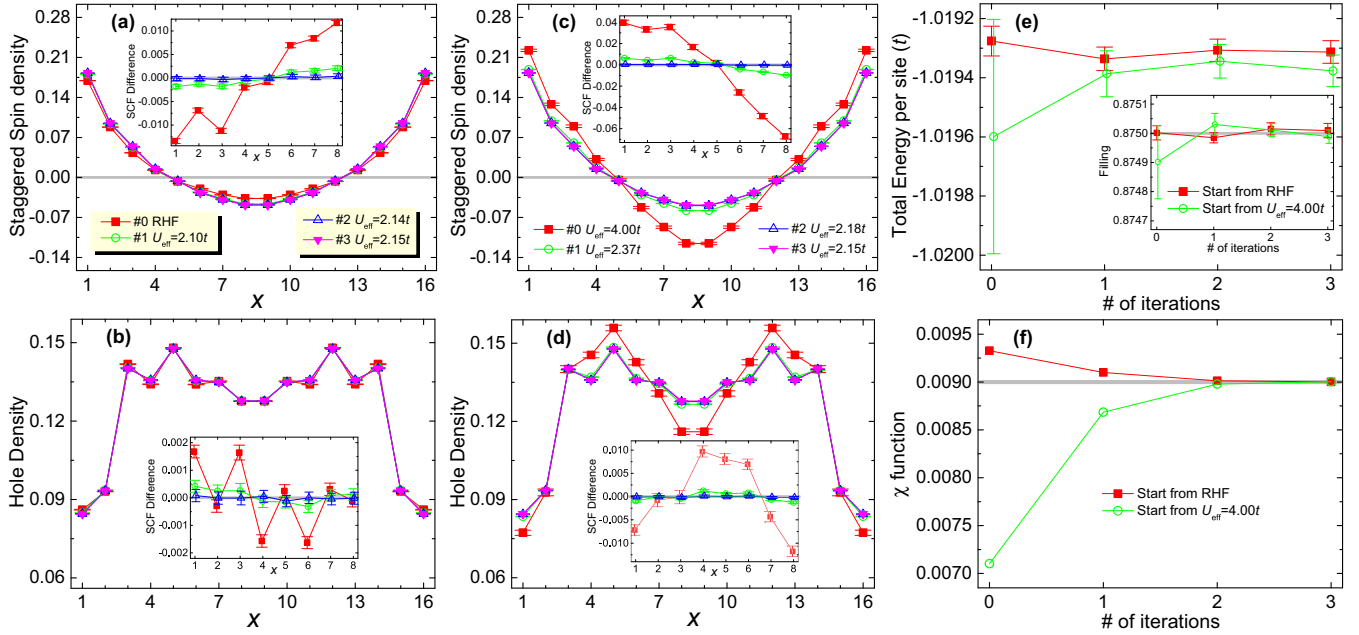
FIG. 4. Convergence of the self-consistent FT-AFQMC procedure. Results are shown for a 16×4 system with $U = 4t$, $\beta t = 40$ and AFM pinning field $h = 0.10t$ applied to both edges in x direction, targeting an overall electron filling $\langle \hat{n} \rangle = 0.875$. The left column, (a) and (b), shows results of staggered spin and hole density respectively, starting from the RHF trial Hamiltonian, while the middle column, (c) and (d), presents the corresponding results starting from the UHF trial Hamiltonian with $U_{\mathrm{eff}} = U$. The difference at each stage of the iteration with respect to the final converged results is shown in the insets. The right column shows energy (e), electron filling [inset in (e)], and χ (f) vs iteration number.

B. Numerical results from FT-AFQMC calculations with self-consistent procedure

The FT-AFQMC results presented in Sec. V A are from one-shot calculations with a specific trial Hamiltonian (namely RHF). In such calculations, there is no internal mechanism to indicate whether the results are optimal or not. The self-consistent procedure developed for $T > 0$ in this work serves as such a mechanism. In this part, we illustrate and benchmark the self-consistent approach. As mentioned, we will use a 16×4 system at $\langle \hat{n} \rangle = 0.875$ with AFM pinning fields. The local spin and hole densities, $s_i^z = \langle \hat{n}_{i\uparrow} - \hat{n}_{i\downarrow} \rangle / 2$ and $n_i^h = 1 - \langle \hat{n}_{i\uparrow} + \hat{n}_{i\downarrow} \rangle$, will serve as probes for spin and charge order. To plot the spin order, we will often use the staggered spin density $(-1)^{i_x + i_y} s_i^z$. Translational symmetry in y and the mirror symmetry in x direction are applied to reduce statistical errors.

We start by testing the general stability of the self-consistent FT-AFQMC procedure, in Fig. 4. We carry out two separate self-consistent processes, starting the calculation with RHF and UHF trial Hamiltonians, respectively. The UHF trial Hamiltonian is prepared from the UHF mean-field solution of the many-body system at the desired temperature ($\beta t = 40$), using original parameters in the many-body Hamiltonian ($U_{\mathrm{eff}} = 4t$). The results are shown in panels (a) and (b) and (c) and (d), respectively. We observe that FT-AFQMC calculations starting from both RHF and UHF trial Hamiltonians converge within four iterations. While it is possible for the self-consistent iteration to get stuck in a local minimum, the tests we have carried out at finite temperatures resulted in the same converged results for all studied

observables, independent of which initial trial Hamiltonian is used. (Interestingly, the χ function monotonically increases with the $U_{\mathrm{eff}} = 4t$ UHF starting trial Hamiltonian, which is different from the behavior in ZT-AFQMC calculations [40]. This might be from how finite temperature is accounted for in the trial density matrix.)

We now study the temperature dependence of the self-consistent FT-AFQMC procedure. We will carry out calculations for two systems, with $U = 4t$ and $U = 6t$, at temperatures as low as $\beta t = 50$. The lowest temperature accessible with standard DQMC is $\beta t \sim 8$ and 5, with average signs of 0.043 and 0.016, respectively. As we shall see, the behaviors of the system at such temperatures are very different from those at low temperatures and in the ground-state limit, highlighting the fundamental need for FT methods which control the sign problem. The ZT-AFQMC results from self-consistent calculations for these two systems are presented and discussed in Appendix B.

The computed total energy, staggered spin density, and hole density are shown in Fig. 5 for the $U = 4t$ system, after convergence of the self-consistent procedure. The results from DQMC at $\beta t = 8$ and DMRG at $T = 0$ are also presented as a benchmark. The final converged value of the U_{eff} parameter in the UHF trial Hamiltonian is also indicated for each temperature. As illustrated in Fig. 5(a), the results of total energy from AFQMC calculations are highly accurate for this system, with relative systematic error 0.07% at $\beta t = 8$ and 0.09% for the ground state. Results of a polynomial fit are shown at low temperatures, which gives an extrapolated answer at $T = 0$ fully consistent with ZT-AFQMC. At $\beta t = 8$, the FT-AFQMC
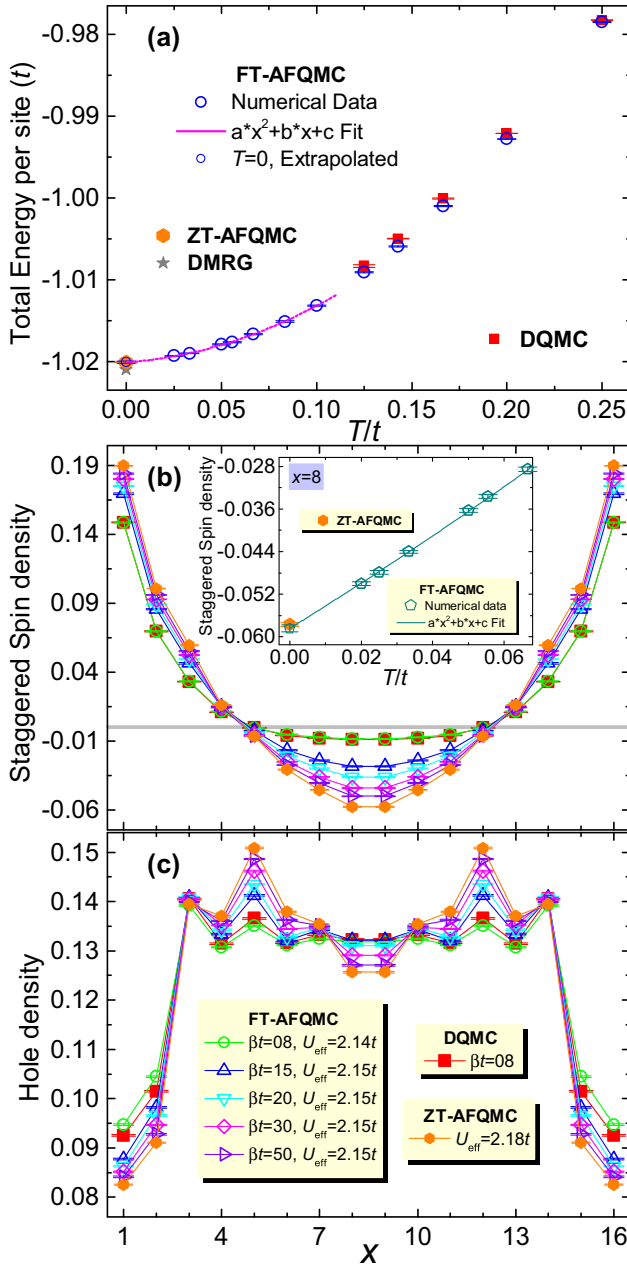
FIG. 5. Computed results of the (a) energy, (b) staggered spin density, and (c) hole density at a variety of temperatures, after convergence of self-consistency procedure. The system is 16×4 with $U = 4t$ and 1/8 hole doping, with AFM pinning fields of $h = 0.10t$ applied on both edges along the y direction. The converged U_{eff} parameter for the trial Hamiltonian is also shown at each T. FT-AFQMC results systematically approach those from ZT-AFQMC as temperature is lowered. Panels (b) and (c) share the same symbols, legends, and scales of x axis. The inset in (b) presents the staggered spin density at $i_x = 8$ and a quadratic fit at low T.

calculation yields staggered spin density which is statistically indistinguishable from that of DQMC as shown in Fig. 5(b). The hole density in Fig. 5(c) is also in good agreement but discrepancy from the constraint is visible, especially near the edges; this behavior is consistent with ZT-AFQMC calculations [40]. Both results converge monotonically toward

the ZT-AFQMC results as temperature is lowered. As shown more quantitatively in the inset of Fig. 5(b), the staggered spin density can be fitted very well by a polynomial, and the extrapolated result at $T = 0$ is in excellent agreement with the self-consistent ZT-AFQMC result. The self-consistent U_{eff} parameters shown in Figs. 5(b) and 5(c) only vary slightly as $U_{\text{eff}}/t = 2.14$–2.15 within a large range of temperature.

We next turn to a physically more relevant and challenging system with $U = 6t$. Many of the characteristics seen in the benchmark in the $U = 4t$ systems are also seen here. In Fig. 6, results are shown for the total energy, the spin order, and hole density for a temperature range $\beta = 4$–50. The sign problem prevents DQMC calculations from reaching much below $\beta = 5$, indicated by the large error bars of the results in all three panels in Fig. 6. This is too high a temperature for the antiferromagnetic correlations to develop, as seen from the results at lower temperatures. At $\beta t = 5$, the pinning fields near the edge introduce short-range antiferromagnetic correlations. The hole density in the system also responds but remains flat outside of the edge region. As temperature is lowered, the AFM order systematically increases in the self-consistent FT-AFQMC results. The hole density starts to exhibit inhomogeneities, and develops peaks near the nodal lines where the AFM order shows a phase change. The spin and hole densities approach the ZT-AFQMC results but with a visible difference even at $\beta t = 20$. The results from FT-AFQMC at $\beta t = 50$ and ZT-AFQMC are virtually indistinguishable. Compared to exact ground-state DMRG results, the spin density from AFQMC is highly accurate, while the hole density shows noticeable discrepancy. This is consistent with the observation from ground-state studies [40] where more systems were tested. As shown in Fig. 6(a), the total energy per site is also exact at high temperatures, compared with DQMC results. At lower temperatures, the computed energies extrapolate to a value of $-0.8765(3)$, consistent with the ZT-AFQMC results of $-0.8768(2)$, which has a relative error of about 0.36% compared to the exact ground-state energy of -0.8735 from DMRG.

Besides obtaining the results of thermodynamic properties, the self-consistent procedure in FT-AFQMC calculations also determines an effective interaction strength U_{eff}. The computed spin and charge densities from the IP Hamiltonian in Eq. (19) actually provide a reasonable approximation to the many-body results. As expected, the optimum U_{eff} increases with interaction strength U in the many-body Hamiltonian. Similar to the $U = 4t$ case, the U_{eff} at $T > 0$ is found to be smaller than that at $T = 0$ at $U = 6t$, although U_{eff} only varies slightly in a wide range of low temperatures.

VI. DISCUSSION AND SUMMARY

The self-consistent FT-AFQMC method developed in this work allows finite-temperature QMC computations to systematically improve the accuracy while removing the exponential computational scaling from the sign/phase problem. We have studied the form of the trial Hamiltonian or trial density matrix. Compared to ground-state calculations, for which the constraint is defined in terms of the ground-state trial wave function, the FT-AFQMC formalism involves accounting for the temperature. The self-consistent constraint we have
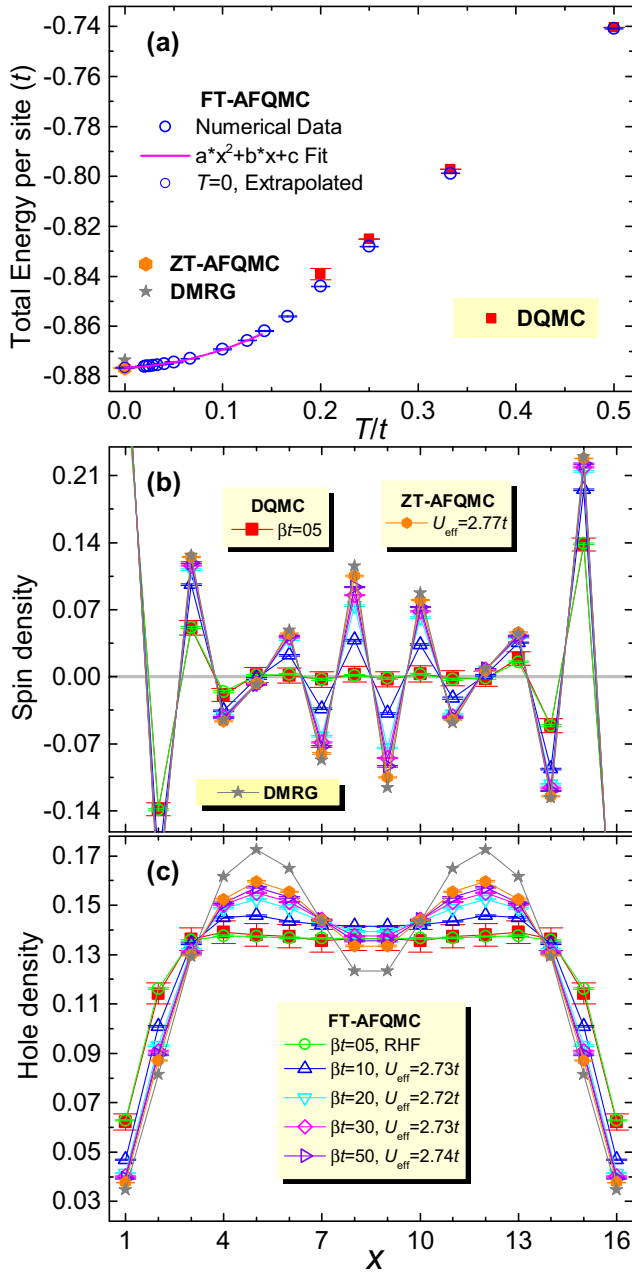
FIG. 6. Development of antiferromagnetic correlations and charge order as temperature is lowered. The system is 16×4, $U = 6t$, and $1/8$ hole doping, with AFM pinning fields of $h = 0.10t$ applied on both edges along the y direction. The AFQMC results are from self-consistent calculations. Energies are shown in (a), and spin density and hole density are plotted in panels (b) and (c), which share the same symbols, legends, and scales of the x axis.

discussed aims to generate a $\hat{B}_T$ by optimizing $e^{-\beta \hat{H}_T}$ with respect to the many-body calculation at the desired temperature β. The alternative we also tested, of using a ground-state $\hat{H}_T$, takes the view of optimizing an effective *Hamiltonian*. No significant difference was observed in the performance of these in the Hubbard model. Another approach which will be interesting to test is to, as random walk proceeds (varying ℓ), refine the constraint $e^{-\tau \hat{H}_T}$ "dynamically" with τ, which

corresponds to $\prod_n^{M-\ell} \hat{B}_T$ in the constraining conditions in Eq. (11), as the random walk proceeds (varying ℓ). Applying the computed single-particle density matrix as a constraint to form a self-consistent procedure is also possible, as discussed in Ref. [40] for the ground state. We will leave these to a future study.

As mentioned, the self-consistent approach produces a rigorous procedure to identify an optimal independent-particle treatment of the many-body system. In the Hubbard model, this takes the form of an effective interaction. In a more general context, one can imagine this procedure as a way to optimize an exchange-correlation functional within the context of density functional theory (DFT). The optimized parametrization of a particular DFT flavor can then be applied to larger systems or other materials. Similarly, one could imagine using the approach to help determine a "U" parameter for DFT $+ U$ type calculations in correlated materials.

The finite-temperature extension of AFQMC allows many potential applications. Among these are direct calculations to study temperature dependence and obtain thermodynamic information in correlated electron models and materials. It could also potentially be combined with embedding methods such as dynamical mean-field theory [63], acting as an impurity solver.

To summarize, we have presented a self-consistent finite-temperature AFQMC approach to study many-fermion systems. The method shares the basic formalism of DQMC, but controls the sign or phase problem with a constraint. The calculations are formulated as branching random walks with importance sampling. The constraint is applied with a trial Hamiltonian in an independent-particle form, or a density matrix, and is optimized by a self-consistent feedback from the AFQMC calculation. The approach complements the ZT-AFQMC which has been widely applied in lattice models and in realistic systems in condensed matter and quantum chemistry. In this paper, we presented the finite-temperature algorithm in detail, including many technical and numerical aspects, which will help make the implementation of the FT-AFQMC method to other correlated fermion systems more straightforward. With our method and with improved numerical stabilization procedures (further described in Appendix B), calculations can reach very low temperatures (as low as $1/80$ in the Hubbard model).

In addition to presenting the method, we also carried out a careful benchmark, and studied its behavior in detail. Different forms of the trial Hamiltonian were tested. Comparisons were made with DQMC where the latter could be performed, as well as with ZT-AFQMC and ED and DMRG in the ground state. The FT-AFQMC method is indistinguishable from DQMC at high temperatures, and connects smoothly to the ZT-AFQMC at very low temperatures. The benchmark results show that the accuracy of FT-AFQMC method for thermodynamic properties is comparable to or even better than that of the ZT-AFQMC method for ground-state properties. The results in the Hubbard model demonstrate that the approach opens up regimes in temperature where interesting physics takes place and which are inaccessible by DQMC. It is hoped that applications to study thermodynamic properties in a variety of correlated fermion models will now be possible with developments in this work.

APPENDIX A: SOME FURTHER DETAILS IN FT-AFQMC METHOD

In this Appendix, we further discuss several technical details of the FT-AFQMC implementation.

In Sec. III B, we mentioned that there are two somewhat different ways to implement FT-AFQMC depending on the connection between $\hat{H}_T$ and $\hat{H}_0$. If $\hat{H}_T$ is closely related to the free fermion part $\hat{H}_0$, e.g., with the RHF or UHF form, we need to insert the $e^{-\Delta\tau(\hat{H}_I+\hat{H}_0-\hat{H}_T)} = \sum_{\mathbf{x}} p(\mathbf{x})\hat{B}_I(\mathbf{x})$ operator at every time slice. In this case we have, for the ℓth time slice,

$$\frac{\mathcal{P}_\ell^T}{\mathcal{P}_{\ell-1}^T} = p(\mathbf{x}_\ell)\frac{\text{Tr}\big[\big(\prod_{n=1}^{M-\ell}\hat{B}_T\big)\hat{B}_I(\mathbf{x}_\ell)\hat{B}_T\big(\prod_{m=1}^{\ell-1}\hat{B}_m\big)\big]}{\text{Tr}\big[\big(\prod_{n=1}^{M-\ell}\hat{B}_T\big)\hat{B}_T\big(\prod_{m=1}^{\ell-1}\hat{B}_m\big)\big]} = p(\mathbf{x}_\ell)\frac{\det\big[\mathbf{I}_{2N_s} + \big(\prod_{n=1}^{M-\ell}\mathbf{B}_T\big)\mathbf{B}_I(\mathbf{x}_\ell)\mathbf{B}_T\big(\prod_{m=1}^{\ell-1}\mathbf{B}_m\big)\big]}{\det\big[\mathbf{I}_{2N_s} + \big(\prod_{n=1}^{M-\ell}\mathbf{B}_T\big)\mathbf{B}_T\big(\prod_{m=1}^{\ell-1}\mathbf{B}_m\big)\big]}, \quad \text{(A1)}$$

where the notations are the same as the main text. In the second way, the trial Hamiltonian $\hat{H}_T$ is assumed to be quite different from $\hat{H}_0$, and we need to replace $\hat{B}_T$ by $\hat{B} = \hat{B}_I(\mathbf{x})\hat{B}_K$ for every time slice. In this case, we first replace $\hat{B}_T$ by $\hat{B}_K$ and then insert the $e^{-\Delta\tau\hat{H}_I} = \sum_{\mathbf{x}} p(\mathbf{x})\hat{B}_I(\mathbf{x})$ operator to evaluate $\mathcal{P}_\ell^T/\mathcal{P}_{\ell-1}^T$ during the random walk:

$$\frac{\mathcal{P}_\ell^T}{\mathcal{P}_{\ell-1}^T} = p(\mathbf{x}_\ell)\frac{\text{Tr}\big[\big(\prod_{n=1}^{M-\ell}\hat{B}_T\big)\hat{B}_I(\mathbf{x}_\ell)\hat{B}_K\hat{B}_{\ell-1}\ldots\hat{B}_2\hat{B}_1\big]}{\text{Tr}\big[\big(\prod_{n=1}^{M-\ell}\hat{B}_T\big)\hat{B}_T\hat{B}_{\ell-1}\ldots\hat{B}_2\hat{B}_1\big]}$$

$$= p(\mathbf{x}_\ell)\frac{\text{Tr}\big[\big(\prod_{n=1}^{M-\ell}\hat{B}_T\big)\hat{B}_I(\mathbf{x}_\ell)\hat{B}_K\hat{B}_{\ell-1}\ldots\hat{B}_1\big]}{\text{Tr}\big[\big(\prod_{n=1}^{M-\ell}\hat{B}_T\big)\hat{B}_K\hat{B}_{\ell-1}\ldots\hat{B}_1\big]} \cdot \frac{\text{Tr}\big[\big(\prod_{n=1}^{M-\ell}\hat{B}_T\big)\hat{B}_K\hat{B}_{\ell-1}\ldots\hat{B}_1\big]}{\text{Tr}\big[\big(\prod_{n=1}^{M-\ell}\hat{B}_T\big)\hat{B}_T\hat{B}_{\ell-1}\ldots\hat{B}_1\big]}$$

$$= p(\mathbf{x}_\ell)\frac{\det\big[\mathbf{I}_{2N_s} + \big(\prod_{n=1}^{M-\ell}\mathbf{B}_T\big)\mathbf{B}_I(\mathbf{x}_\ell)\mathbf{B}_K\mathbf{B}_{\ell-1}\ldots\mathbf{B}_1\big]}{\det\big[\mathbf{I}_{2N_s} + \big(\prod_{n=1}^{M-\ell}\mathbf{B}_T\big)\mathbf{B}_K\mathbf{B}_{\ell-1}\ldots\mathbf{B}_1\big]} \cdot \frac{\det\big[\mathbf{I}_{2N_s} + \big(\prod_{n=1}^{M-\ell}\mathbf{B}_T\big)\mathbf{B}_K\mathbf{B}_{\ell-1}\ldots\mathbf{B}_1\big]}{\det\big[\mathbf{I}_{2N_s} + \big(\prod_{n=1}^{M-\ell}\mathbf{B}_T\big)\mathbf{B}_T\mathbf{B}_{\ell-1}\ldots\mathbf{B}_1\big]}, \quad \text{(A2)}$$

where the second and first ratios in the last line correspond to the insertion of $\mathbf{B}_I(\mathbf{x}_\ell)$ matrix and replacement of $\mathbf{B}_T$ by $\mathbf{B}_K$ matrix. To evaluate $\mathcal{P}_\ell^T/\mathcal{P}_{\ell-1}^T$ in Eq. (A2) numerically, we need to calculate the determinant $\det[\mathbf{I}_{2N_s} + \big(\prod_{n=1}^{M-\ell}\mathbf{B}_T\big)\mathbf{B}_K\mathbf{B}_{\ell-1}\ldots\mathbf{B}_1]$ and the static single-particle Green's function at every time slice, which results in extra computational cost compared to the first way in Eq. (A1).

We next describe our numerical stabilization procedures. In the DMQC and FT-AFQMC methods, the numerical stabilization procedures are similar and they both involves the static single-particle Green's-function matrix $\mathbf{G}(\tau,\tau) = \{G_{i\sigma,j\sigma'} = \langle c_{i\sigma}c_{j\sigma'}^+\rangle_\tau\}$ at $\tau = \ell\Delta\tau$. Generally, its matrix element takes the form

$$G_{i\sigma,j\sigma'} = \frac{\text{Tr}[\hat{B}(\beta,\tau)c_{i\sigma}c_{j\sigma'}^+\hat{B}(\tau,0)]}{\text{Tr}[\hat{B}(\beta,\tau)\hat{B}(\tau,0)]} = \big\{\big[\mathbf{I}_{2N_s} + \mathbf{B}(\tau,0)\mathbf{B}(\beta,\tau)\big]^{-1}\big\}_{i\sigma,j\sigma'}. \quad \text{(A3)}$$

For DQMC, $\hat{B}(\beta,\tau) = \hat{B}_M\ldots\hat{B}_{\ell+2}\hat{B}_{\ell+1}$ and $\hat{B}(\tau,0) = \hat{B}_\ell\ldots\hat{B}_2\hat{B}_1$, with $\mathbf{B}(\beta,\tau) = \mathbf{B}_M\ldots\mathbf{B}_{\ell+2}\mathbf{B}_{\ell+1}$ and $\mathbf{B}(\tau,0) = \mathbf{B}_\ell\ldots\mathbf{B}_2\mathbf{B}_1$. For the FT-AFQMC algorithm, $\hat{B}(\beta,\tau) = \prod_{n=1}^{M-\ell}\hat{B}_T$ and $\hat{B}(\tau,0) = \hat{B}_\ell\ldots\hat{B}_2\hat{B}_1$, with $\mathbf{B}(\beta,\tau) = \prod_{n=1}^{M-\ell}\mathbf{B}_T$, $\mathbf{B}(\tau,0) = \mathbf{B}_\ell\ldots\mathbf{B}_2\mathbf{B}_1$, instead.

We apply the column-pivoted QR algorithm [54,55] to perform the following decompositions for $\mathbf{B}(\beta,\tau)$ and $\mathbf{B}(\tau,0)$ matrices: $\mathbf{B}(\beta,\tau) = \mathbf{V}_L\mathbf{D}_L\mathbf{U}_L$ and $\mathbf{B}(\tau,0) = \mathbf{U}_R\mathbf{D}_R\mathbf{V}_R$, where $\mathbf{U}_R$, $\mathbf{U}_L$, $\mathbf{V}_R$, $\mathbf{V}_L$ are $2N_s \times 2N_s$ matrices and $\mathbf{U}_R$, $\mathbf{U}_L$ are unitary, and $\mathbf{D}_R$, $\mathbf{D}_L$ are $2N_s \times 2N_s$ real diagonal matrices. To calculate the Green's function matrix in Eq. (A3), we further separate $\mathbf{D}_R$ as $\mathbf{D}_R^{\max}\mathbf{D}_R^{\min}$, which are both $2N_s \times 2N_s$ real diagonal matrices satisfying the following condition: if $|(\mathbf{D}_R)_{ii}| \geqslant 1$, then $(\mathbf{D}_R^{\max})_{ii} = |(\mathbf{D}_R)_{ii}|$ and $(\mathbf{D}_R^{\min})_{ii} = \text{Sgn}[(\mathbf{D}_R)_{ii}]$; if $|(\mathbf{D}_R)_{ii}| < 1$, then $(\mathbf{D}_R^{\max})_{ii} = 1$ and $(\mathbf{D}_R^{\min})_{ii} = (\mathbf{D}_R)_{ii}$. Similarly, we have $\mathbf{D}_L = \mathbf{D}_L^{\min}\mathbf{D}_R^{\max}$ following the same definitions. We can then calculate the single-particle Green's-function matrix in Eq. (A3) in a numerically stable manner [54,55] as

$$\mathbf{G}(\tau,\tau) = (\mathbf{U}_L)^{-1}[(\mathbf{U}_R)^{-1}(\mathbf{U}_L)^{-1} + \mathbf{D}_R\mathbf{V}_R\mathbf{V}_L\mathbf{D}_L]^{-1}(\mathbf{U}_R)^{-1} = (\mathbf{U}_L)^{-1}[(\mathbf{U}_L\mathbf{U}_R)^{-1} + \mathbf{D}_R^{\max}\mathbf{D}_R^{\min}\mathbf{V}_R\mathbf{V}_L\mathbf{D}_L^{\min}\mathbf{D}_L^{\max}]^{-1}(\mathbf{U}_R)^{-1}$$

$$= (\mathbf{U}_L)^{-1}\big(\mathbf{D}_L^{\max}\big)^{-1}\big[\big(\mathbf{D}_R^{\max}\big)^{-1}(\mathbf{U}_L\mathbf{U}_R)^{-1}\big(\mathbf{D}_L^{\max}\big)^{-1} + \mathbf{D}_R^{\min}\mathbf{V}_R\mathbf{V}_L\mathbf{D}_L^{\min}\big]^{-1}\big(\mathbf{D}_R^{\max}\big)^{-1}(\mathbf{U}_R)^{-1}. \quad \text{(A4)}$$
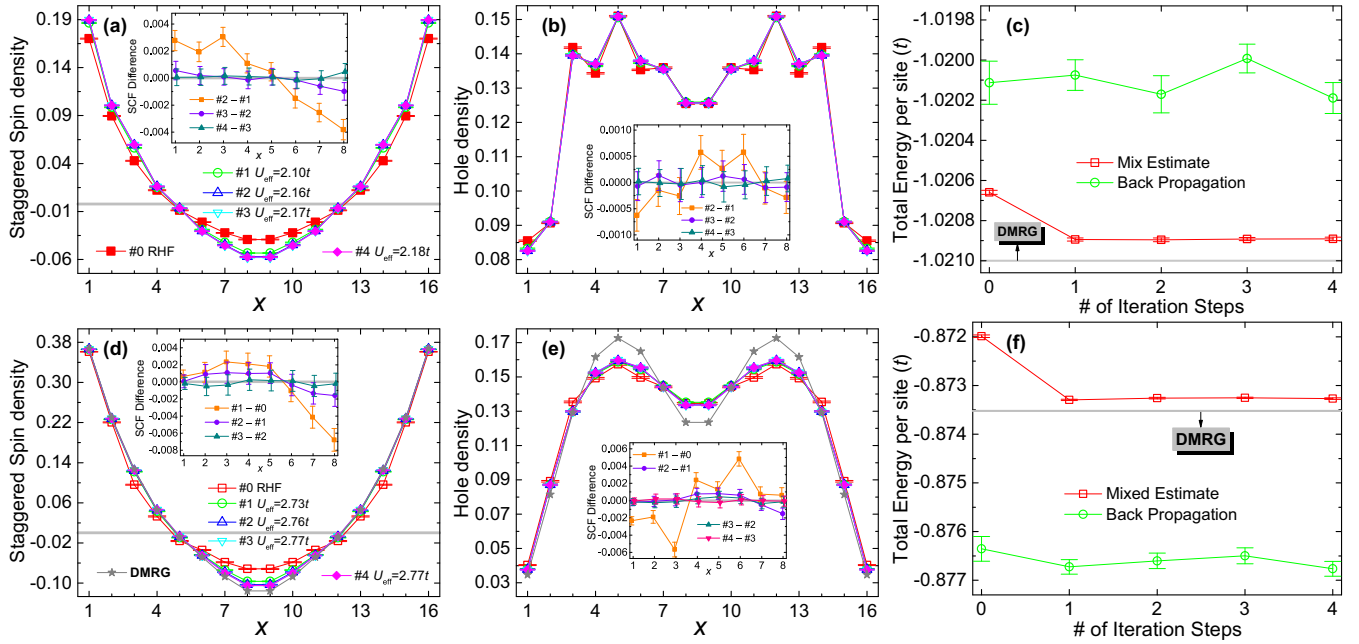
FIG. 7. Numerical results from self-consistent ZT-AFQMC calculations at $T = 0$ for 16×4 system with $U = 4t$ and $U = 6t$, with 1/8 hole doping. AFM pinning fields are applied on both edges along y direction, with $h = 0.10t$ for $U = 4t$ and $h = 0.25t$ for $U = 6t$. (a)–(c) Results of staggered spin density, hole density, and total energy per site are plotted in three panels in each row, with the top row for the $U = 4t$ system and the bottom row (d)–(f) for $U = 6t$. Panels (a),(b) and (d),(e) share the same symbols and legends, respectively. The insets in the first two columns show the difference between successive iterations.

From Eq. (A4), we can observe that the matrix $[(\mathbf{D}_R^{max})^{-1}(\mathbf{U}_L\mathbf{U}_R)^{-1}(\mathbf{D}_L^{max})^{-1} + \mathbf{D}_R^{min}\mathbf{V}_R\mathbf{V}_L\mathbf{D}_L^{min}]$, whose inverse matrix needs to be calculated, should have all its matrix elements around -10^0–10^0. For time-displaced single-particle Green's-function matrices $\mathbf{G}(\tau, 0)$ and $\mathbf{G}(0, \tau)$, there are similar formulas to calculate them in a numerically stable manner. Practically, the numerical stabilization is carried out with a suitably chosen number of imaginary time slices as an interval.

APPENDIX B: NUMERICAL RESULTS FROM SELF-CONSISTENT ZT-AFQMC CALCULATIONS

In Sec. V B, we presented $T = 0$ results from ZT-AFQMC calculations for comparisons with the self-consistent FT-AFQMC results on 16×4 systems at 1/8 hole doping. Here we include the ZT-AFQMC results for completeness, summarized in Fig. 7.

The self-consistent calculations started from free-electron trial wave functions. For both systems, five steps of iterations

reached the converged results for the densities, while the total energy converged within two steps. The converged effective interaction strengths are $U_{eff} = 2.18t$ and $U_{eff} = 2.77t$ for those two systems, respectively. The convergence of staggered spin density and hole density are illustrated by the differences of results between successive iterations as shown in Figs. 7(a) and 7(b) for the $U = 4t$, $h = 0.10t$ system and Figs. 7(d) and 7(e) for the $U = 6t$, $h = 0.25t$ system. For the $U = 6t$ system, the converged results of staggered spin density are very close to the DMRG results, highlighting the high accuracy. The hole density still has some deviations, similar to the results shown in Ref. [40]. The results of total energy per site from both the mixed estimate and back propagation measurements are shown in the last column. As expected, the error bars of results of total energy per site from mixed estimate measurements is much smaller than those from back propagation measurements as shown in Figs. 7(c) and 7(f). For $U = 6t$, the converged results of total energy per site from those two different ways of measurement have relative errors of 0.028% and 0.355% from DMRG, respectively.

[1] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278 (1981).

[2] G. Sugiyama and S. Koonin, Ann. Phys. **168**, 1 (1986).

[3] S. Hands, A. Kocic, and J. Kogut, Ann. Phys. **224**, 29 (1993).

[4] S. Chandrasekharan, Phys. Rev. D **82**, 025007 (2010).

[5] J. Carlson, S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. E. Schmidt, and R. B. Wiringa, Rev. Mod. Phys. **87**, 1067 (2015).

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

[7] D. J. Scalapino and R. L. Sugar, Phys. Rev. B **24**, 4295 (1981).

[8] S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar, Phys. Rev. B **40**, 506 (1989).

[9] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. Lett. **74**, 3652 (1995).

[10] S. Zhang, Phys. Rev. Lett. **83**, 2777 (1999).

[11] F. Assaad and H. Evertz, in *Computational Many-Particle Physics*, Lecture Notes in Physics Vol. 739, edited by H. Fehske, R. Schneider, and A. Weiße (Springer, Heidelberg, 2008), pp. 277–356.

[12] S. Zhang, in *Emergent Phenomena in Correlated Matter: Modeling and Simulation* (Verlag des Forschungszentrum, Julich, 2013), Chap. 15, pp. 15.1–15.31.

[13] J. Hubbard, Phys. Rev. Lett. **3**, 77 (1959).

[14] F. Becca and S. Sorella, in *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, UK, 2017), pp. 214–232.

[15] K. E. Schmidt and M. H. Kalos, *Applications of the Monte Carlo Method in Statistical Physics* (Springer-Verlag, Heidelberg, 1984).

[16] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B **41**, 9301 (1990).

[17] R. R. d. Santos, Braz. J. Phys. **33**, 36 (2003).

[18] J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).

[19] C. Wu and S.-C. Zhang, Phys. Rev. B **71**, 155115 (2005).

[20] E. Berg, M. A. Metlitski, and S. Sachdev, Science **338**, 1606 (2012).

[21] E. F. Huffman and S. Chandrasekharan, Phys. Rev. B **89**, 111101 (2014).

[22] L. Wang, P. Corboz, and M. Troyer, New J. Phys. **16**, 103008 (2014).

[23] Z.-X. Li, Y.-F. Jiang, and H. Yao, Phys. Rev. B **91**, 241117 (2015).

[24] L. Wang, Y.-H. Liu, M. Iazzi, M. Troyer, and G. Harcos, Phys. Rev. Lett. **115**, 250601 (2015).

[25] Z.-X. Li, Y.-F. Jiang, and H. Yao, Phys. Rev. Lett. **117**, 267002 (2016).

[26] Z. C. Wei, C. Wu, Y. Li, S. Zhang, and T. Xiang, Phys. Rev. Lett. **116**, 250601 (2016).

[27] E. Huffman and S. Chandrasekharan, Phys. Rev. E **94**, 043311 (2016).

[28] S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003).

[29] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. B **55**, 7464 (1997).

[30] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. Lett. **78**, 4486 (1997).

[31] M. Guerrero, J. E. Gubernatis, and S. Zhang, Phys. Rev. B **57**, 11980 (1998).

[32] J. Carlson, J. E. Gubernatis, G. Ortiz, and S. Zhang, Phys. Rev. B **59**, 12788 (1999).

[33] M. Enjalran, F. Hébert, G. G. Batrouni, R. T. Scalettar, and S. Zhang, Phys. Rev. B **64**, 184402 (2001).

[34] C.-C. Chang and S. Zhang, Phys. Rev. B **78**, 165101 (2008).

[35] C.-C. Chang and S. Zhang, Phys. Rev. Lett. **104**, 116402 (2010).

[36] H. Shi and S. Zhang, Phys. Rev. B **88**, 125132 (2013).

[37] H. Shi, C. A. Jiménez-Hoyos, R. Rodríguez-Guzmán, G. E. Scuseria, and S. Zhang, Phys. Rev. B **89**, 125129 (2014).

[38] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N.

V. Prokof'ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn *et al.* (Simons Collaboration on the Many-Electron Problem), Phys. Rev. X **5**, 041041 (2015).

[39] M. Qin, H. Shi, and S. Zhang, Phys. Rev. B **94**, 085103 (2016).

[40] M. Qin, H. Shi, and S. Zhang, Phys. Rev. B **94**, 235119 (2016).

[41] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, and G. K.-L. Chan, Science **358**, 1155 (2017).

[42] M. Suewattana, W. Purwanto, S. Zhang, H. Krakauer, and E. J. Walter, Phys. Rev. B **75**, 245123 (2007).

[43] M. Motta, D. M. Ceperley, G. K.-L. Chan, J. A. Gomez, E. Gull, S. Guo, C. A. Jiménez-Hoyos, T. N. Lan, J. Li, F. Ma, A. J. Millis, N. V. Prokof'ev, U. Ray, G. E. Scuseria, S. Sorella, E. M. Stoudenmire, Q. Sun, I. S. Tupitsyn, S. R. White, D. Zgid, and S. Zhang (Simons Collaboration on the Many-Electron Problem), Phys. Rev. X **7**, 031059 (2017).

[44] B. M. Rubenstein, S. Zhang, and D. R. Reichman, Phys. Rev. A **86**, 053606 (2012).

[45] Y. Liu, M. Cho, and B. Rubenstein, J. Chem. Theory Comput. **14**, 4722 (2018).

[46] E. Cocchi, L. A. Miller, J. H. Drewes, M. Koschorreck, D. Pertot, F. Brennecke, and M. Köhl, Phys. Rev. Lett. **116**, 175301 (2016).

[47] L. W. Cheuk, M. A. Nichols, K. R. Lawrence, M. Okan, H. Zhang, and M. W. Zwierlein, Phys. Rev. Lett. **116**, 235301 (2016).

[48] M. Boll, T. A. Hilker, G. Salomon, A. Omran, J. Nespolo, L. Pollet, I. Bloch, and C. Gross, Science **353**, 1257 (2016).

[49] L. W. Cheuk, M. A. Nichols, K. R. Lawrence, M. Okan, H. Zhang, E. Khatami, N. Trivedi, T. Paiva, M. Rigol, and M. W. Zwierlein, Science **353**, 1260 (2016).

[50] S. B. Fahy and D. R. Hamann, Phys. Rev. Lett. **65**, 3437 (1990).

[51] S. Zhang, in *Handbook of Materials Modeling* (Springer, Cham, 2018), pp. 1–27.

[52] H. Shi and S. Zhang, Phys. Rev. E **93**, 033303 (2016).

[53] S. Zhang, Comput. Phys. Commun. **127**, 150 (2000).

[54] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Int. J. Mod. Phys. C **16**, 1319 (2005).

[55] A. Tomas, C. Chang, R. Scalettar, and Z. Bai, in *2012 IEEE 26th International Parallel and Distributed Processing Symposium* (IEEE, Shanghai, 2012), pp. 308–319.

[56] E. Vitali, H. Shi, M. Qin, and S. Zhang, Phys. Rev. B **94**, 085140 (2016).

[57] D. J. Scalapino, *Handbook of High-Temperature Superconductivity: Theory and Experiment*, edited by J. R. Schrieffer and J. S. Brooks (Springer New York, NY, 2007), pp. 495–526.

[58] S. R. White and A. L. Chernyshev, Phys. Rev. Lett. **99**, 127004 (2007).

[59] S. R. White and D. J. Scalapino, Phys. Rev. B **79**, 220504 (2009).

[60] F. F. Assaad and I. F. Herbut, Phys. Rev. X **3**, 031010 (2013).

[61] J. E. Hirsch, Phys. Rev. B **28**, 4059 (1983).

[62] C.-C. Chang and M. A. Morales, arXiv:1711.02154.

[63] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).