Constrained-path auxiliary-field quantum Monte Carlo for coupled electrons and phonons

Joonho Lee,[1,*] Shiwei Zhang,[2,3,†] and David R. Reichman[1,‡]
[1]*Department of Chemistry, Columbia University, New York, New York 10027, USA*
[2]*Center for Computational Quantum Physics, Flatiron Institute, New York, New York 10010, USA*
[3]*Department of Physics, College of William and Mary, Williamsburg, Virginia 23187, USA*

(Received 28 December 2020; accepted 18 February 2021; published 15 March 2021)

We present an extension of constrained-path auxiliary-field quantum Monte Carlo (CP-AFQMC) for the treatment of correlated electronic systems coupled to phonons. The algorithm follows the standard CP-AFQMC approach for description of the electronic degrees of freedom while phonons are described in first quantization and propagated via a diffusion Monte Carlo approach. Our method is tested on the one- and two-dimensional Holstein and Hubbard-Holstein models. With a simple semiclassical trial wave function, our approach is remarkably accurate for $\omega/(2dt\lambda) < 1$ for all parameters in the Holstein model considered in this study where d is the dimensionality, ω is the phonon frequency, t is the electronic hopping strength, and λ is the dimensionless electron-phonon coupling strength. In addition, we empirically show that the autocorrelation timescales as $1/\omega$ for $\omega/t \lesssim 1$, which is an improvement over the $1/\omega^2$ scaling of the conventional determinant quantum Monte Carlo algorithm. In the Hubbard-Holstein model, the accuracy of our algorithm is found to be consistent with that of standard CP-AFQMC for the Hubbard model when the Hubbard U term dominates the physics of the model, and is nearly exact when the ground state is dominated by the electron-phonon coupling scale λ. The approach developed in this work should be valuable for understanding the complex physics arising from the interplay between electrons and phonons in both model lattice problems and *ab initio* systems.

DOI: 10.1103/PhysRevB.103.115123

I. INTRODUCTION

The coupling of electrons to nuclear lattice distortions is responsible for myriad important physical phenomena in bulk materials [1]. In particular, the thermodynamic and transport properties of solids are crucially influenced by electron-phonon (el-ph) interactions. Perhaps the most spectacular consequence of el-ph interactions is the emergence of superconductivity as described by the Bardeen-Cooper-Schrieffer (BCS) theory. Here, the el-ph interaction mediates an effective electron-electron (el-el) attraction which results in the Cooper pairing of electrons of opposite spin [2]. The BCS theory provides a quantitative framework for the description of conventional superconductivity such as that found at low temperatures in simple metals.

A simple microscopic picture is unfortunately not available for unconventional superconductors such as the cuprates, whose critical temperature (T_c) can be above 90 K at ambient pressure [3,4]. It is believed that the el-ph interaction alone cannot give rise to these high-T_c values [5]. However, experimental evidence exists which indicates that non-negligible el-ph interactions are present in these materials [6–24]. It remains unclear what role el-ph interactions play in the cuprates and related materials, and if a potentially delicate interplay between el-el and el-ph interactions may influence their superconducting properties.

The canonical model Hamiltonian used to capture the physics of the cuprates is the two-dimensional (2D) repulsive Hubbard model [25]. The ground state of the hole-doped 2D Hubbard model has been thought to support d-wave superconductivity for many years [26]. A recent joint numerical study using two state-of-the-art approaches, density matrix renormalization group (DMRG) and constrained-path (CP) auxiliary-field quantum Monte Carlo (AFQMC), indicates that the ground state of the standard 2D repulsive Hubbard model with near-neighbor hopping supports modulated phases (e.g., stripes) that are not superconducting over a range of repulsion strengths and doping levels [27] expected to describe the cuprates. This suggests that features beyond those included in the simple Hubbard model, such as the effects of multiple bands, longer-ranged Coulomb interactions, and/or the role of el-ph interactions, may be needed to tip the balance of the ground state towards superconductivity for realistic values of doping levels and the magnitude of el-el repulsions.

Our work is motivated by precisely these considerations, namely, the development of a scalable and accurate numerical approach that can treat el-ph effects on the same footing as el-el correlations. This is a challenging task, as treating the complex electronic degrees of freedom in the pure 2D Hubbard model is already difficult, even with state-of-the-art numerical approaches [27–29]. The addition of el-ph effects, as contained in, e.g., the 2D Hubbard-Holstein model, thus requires nontrivial extensions of these approaches in order to treat electrons and phonons on an equal footing.

*jl5653@columbia.edu
†szhang@flatironinstitute.org
‡drr2103@columbia.edu

Several methods have been formulated or extended to coupled el-ph problems, including DMRG [30–35], variational exact diagonalization [36], variational Monte Carlo [37–39], dynamical mean-field theory [40–44], density matrix embedding theory [45,46], and coupled-cluster theory [47–49]. There are difficulties facing each approach. For example, large el-ph couplings and/or small phonon frequencies are challenging to handle in most methods based on a second-quantized representation of phonons because of the necessity of truncating the phonon Hilbert space. When a large number of phonons per site are required, the computational cost associated with treating them can grow prohibitively expensive. In addition to the demand of treating the phononic Hilbert space, there is of course the interacting many-electron problem. Clearly, the treatment of correlated el-ph coupled systems in two and higher dimensions over a wide range of the parameter space in an exact or near-exact manner is a forefront challenge.

The method that we propose here is an extension of the CP-AFQMC method developed and popularized by Zhang and coworkers [50,51]. For purely electronic problems, the CP-AFQMC approach is similar to the determinant quantum Monte Carlo (DQMC) method [52–59] in the sense that the two-body propagation is aided by the Hubbard-Stratonovich transformation [60] and is formulated in the space of determinants. There, however, are several key differences. CP-AFQMC reformulates the imaginary-time propagation by working with open-ended random walks. An exact boundary condition is introduced in auxiliary-field space, which can be approximately imposed using a trial wave function, to avoid the notorious fermion sign problem. The open-ended random-walk approach allows easy access to zero-temperature results, and is often much less prone to ergodicity problems in the Monte Carlo sampling. Moreover, CP-AFQMC can be naturally extended to *ab initio* Hamiltonians while coping with the fermionic phase problem associated with these more complex models using the phaseless approximation instead of the constrained-path approximation [61,62]. Because of the constraint imposed on walker trajectories, CP-AFQMC is no longer exact, unlike DQMC. Furthermore, due to the constraint, the ground-state energy computed via the usual mixed estimator is not variational [63]. On the other hand, CP-AFQMC can be used to access a wider range of interaction strengths and doping regimes in which DQMC cannot be used due to the inherent sign problem. It should be noted that, in addition to its flexibility, CP-AFQMC has been shown to yield excellent accuracy for strongly correlated electrons [27,29].

In this work we devise an extension of CP-AFQMC to treat both electrons and phonons on an equal footing, while retaining its benefits for electrons. Our framework is similar to the extension of Green's function Monte Carlo (GFMC) as formulated by McKenzie and others [64], where the phonons are treated in a first quantized space. We present the formulation of this new CP-AFQMC approach, provide thorough benchmark results on the one-dimensional (1D) and 2D Holstein and Hubbard-Holstein models for various phonon frequencies and el-ph couplings, and discuss the current scope and limitations of the proposed approach.

The paper is organized as follows: In Sec. II we outline the model we study and the important parameters that control its physics. In Sec. III we outline our algorithm. Section IV is devoted to a discussion of trial wave functions. Sections V A and V B discuss distinct perturbative approaches to the problem outlined in Sec. II. Sections VI and VII present results for the Holstein and Hubbard-Holstein models, respectively. Section VIII discusses the extension of our approach to realistic *ab initio* problems. In Sec. IX we conclude.

II. MODEL

A. Hubbard-Holstein Hamiltonian

Although the approach we outline is general, we focus on a paradigmatic model of a correlated system coupled to phonons, namely, the Hubbard-Holstein model [25,65]. The Hubbard-Holstein model is defined by the following Hamiltonian:

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_{\text{el}}^{(1)} + \hat{\mathcal{H}}_{\text{el}}^{(2)} + \hat{\mathcal{H}}_{\text{ph}} + \hat{\mathcal{H}}_{\text{el-ph}}, \tag{1}$$

where

$$\hat{\mathcal{H}}_{\text{el}} = \hat{\mathcal{H}}_{\text{el}}^{(1)} + \hat{\mathcal{H}}_{\text{el}}^{(2)}, \tag{2}$$

$$\hat{\mathcal{H}}_{\text{el}}^{(1)} = -t \sum_{\sigma\in\{\uparrow,\downarrow\}} \sum_{\langle ij\rangle} \hat{a}_{i_\sigma}^\dagger \hat{a}_{j_\sigma}, \tag{3}$$

$$\hat{\mathcal{H}}_{\text{el}}^{(2)} = U \sum_i \hat{n}_{i_\uparrow} \hat{n}_{i_\downarrow}, \tag{4}$$

$$\hat{\mathcal{H}}_{\text{ph}} = \omega \sum_i \hat{b}_i^\dagger \hat{b}_i, \tag{5}$$

and

$$\hat{\mathcal{H}}_{\text{el-ph}} = -g \sum_i \hat{n}_i (\hat{b}_i + \hat{b}_i^\dagger) \tag{6}$$

with

$$\hat{n}_i = \sum_{\sigma\in\{\uparrow,\downarrow\}} \hat{n}_{i_\sigma}. \tag{7}$$

$\hat{a}_{i\sigma}$ is the annihilation operator for electrons with spin σ on site i, $\hat{b}_i$ is the annihilation operator for bosons on site i, and $\hat{n}_i$ is the electronic number operator on site i. The nearest-neighbor electronic hopping is controlled by t and the onsite repulsion is characterized by the parameter U. The phonons are treated as harmonic oscillators with a single frequency ω. The electronic density is coupled to the phonon degrees of freedom characterized by a coupling constant g.

There are three relevant dimensionless parameters to define. The first is the adiabaticity ratio in units of the hopping parameter

$$\alpha = \frac{\omega}{t}. \tag{8}$$

The second is the effective onsite repulsion in units of the hopping parameter

$$\frac{U}{t}. \tag{9}$$

Lastly, we define the dimensionless el-ph coupling λ,

$$\lambda = \frac{g^2}{2dt\omega}, \tag{10}$$

where d is the dimensionality of the system. When U is the dominant parameter, a spin density wave (SDW) phase similar to that found in the Hubbard model is expected to arise. When λ dominates, a charge density wave (CDW) phase similar to that found in the Holstein model arises. A metallic or superconducting phase can arise when the system transitions between these two phases [58,66].

B. Phonons in first quantization

Since the Hamiltonian in Eq. (1) does not commute with the phonon number operator $\hat{b}_i^\dagger \hat{b}_i$, the number of phonons in the system is not conserved. Therefore, one needs to work with an infinitely large phonon Hilbert space in order to study eigenstates of the Hubbard-Holstein model. Methods working in a second-quantized space such as DMRG [30] generally require a specification of the maximum number of phonons *a priori* for the sake of computational tractability. Limiting the maximum number of phonons effectively truncates the infinite Hilbert space, which may introduce significant errors, particularly when α is small and/or λ is large.

For this reason we work within the framework of first quantization, namely, with position and momentum operators on each site i,

$$\hat{X}_i = \sqrt{\frac{1}{2m\omega}}(\hat{b}_i^\dagger + \hat{b}_i), \tag{11}$$

$$\hat{P}_i = i\sqrt{\frac{m\omega}{2}}(\hat{b}_i^\dagger - \hat{b}_i), \tag{12}$$

and thus reexpress

$$\hat{\mathcal{H}}_{\text{ph}} = \sum_i \left(\frac{m\omega^2}{2}\hat{X}_i^2 + \frac{1}{2m}\hat{P}_i^2 - \frac{\omega}{2}\right), \tag{13}$$

$$\hat{\mathcal{H}}_{\text{el-ph}} = -g\sqrt{2m\omega}\sum_i \hat{n}_i \hat{X}_i \tag{14}$$

in the Hubbard-Holstein Hamiltonian in Eq. (1). Here, we introduced a fictitious mass m and throughout this work we use $m = 1/\omega$. Working in a first-quantized space allows one to work directly at the complete basis set limit for the phonons and avoids the issues posed by a truncated phonon Hilbert space.

III. CONSTRAINED-PATH AUXILIARY-FIELD QUANTUM MONTE CARLO

AFQMC for mixed fermions and bosons was first formulated and studied by Rubenstein, Zhang, and Reichman [67]. In their formulation, bosons are treated within a second-quantized framework. Therefore, their approach would naturally suffer from the truncation of the infinite bosonic Hilbert space if applied to the Hubbard-Holstein model. In this work, we will reformulate the procedure to treat fermions in a second-quantized space and bosons in a first-quantized space. Such a formulation is closely related to that of Ref. [64], however, our work allows the control of the sign problem and introduces the full advantage of the CP-AFQMC approach in treating the electronic degrees of freedom.

In AFQMC, as in other projector QMC methods [68], we obtain the ground state via

$$|\Psi_0\rangle \propto \lim_{\tau\to\infty} e^{-\tau\hat{\mathcal{H}}}|\Phi_0\rangle, \tag{15}$$

where $|\Psi_0\rangle$ is the true ground state, τ denotes imaginary time, and $|\Phi_0\rangle$ is a trial wave function with nonzero overlap with the true ground state. Since $\hat{\mathcal{H}}$ involves both fermions and bosons and so do the wave functions $|\Psi_0\rangle$ and $|\Phi_0\rangle$, we represent these global vibronic wave functions as a function of imaginary time τ in a mixed basis

$$|\Psi(\tau)\rangle = \sum_k \omega_k |\psi_k(\tau), \mathbf{X}_k(\tau)\rangle, \tag{16}$$

where $|\psi_k\rangle$ is the electronic wave function and $|\mathbf{X}_k\rangle$ is a set of coordinates that represents the phonon degrees of freedom. In our algorithm, these basis states each take a product form

$$|\psi_k(\tau), \mathbf{X}_k(\tau)\rangle \equiv |\psi_k(\tau)\rangle \otimes |\mathbf{X}_k(\tau)\rangle\,, \tag{17}$$

where $|\psi_k\rangle$ is a single Slater determinant. We will show below that the projection process in Eq. (15) can be turned into a random walk in the space of product states of the form defined in Eq. (17). We note that it is also possible to work in momentum space ($|\mathbf{P}_k\rangle$) [69], however, it is more convenient to work in position space here since it makes the application of the e-ph coupling term straightforward.

We write the propagator for a finite time step $\Delta\tau$ as

$$\exp(-\Delta\tau\hat{\mathcal{H}}) = e^{-\Delta\tau\hat{\mathcal{H}}_{\text{el}}^{(2)}} e^{-\Delta\tau(\hat{\mathcal{H}}_{\text{el}}^{(1)}+\hat{\mathcal{H}}_{\text{el-ph}})} e^{-\Delta\tau\hat{\mathcal{H}}_{\text{ph}}} + O(\Delta\tau^2) \tag{18}$$

using the standard first-order Trotter approximation. By virtue of the Thouless theorem [70], $|\psi_k(\tau)\rangle$ remains a single Slater determinant after propagation via $\hat{\mathcal{H}}_{\text{el}}^{(1)}$ and $\hat{\mathcal{H}}_{\text{el-ph}}$ (note the latter is diagonal in $|\mathbf{X}_k\rangle$ space). $\hat{\mathcal{H}}_{\text{el}}^{(2)}$ is represented as a one-body operator coupled to Ising variables and, therefore, a single Slater determinant remains in the same manifold after propagation by $e^{-\Delta\tau\hat{\mathcal{H}}_{\text{el}}^{(2)}}$. The phonon propagation generated by $\hat{\mathcal{H}}_{\text{ph}}$ follows a commonly used diffusion MC (DMC) algorithm [71,72].

Before elaborating on the propagation more concretely, let us introduce importance sampling, using a trial vibronic wave function $|\Psi_T\rangle$. We rewrite the global vibronic wave function in Eq. (16) in the following form [61]:

$$|\Psi(\tau)\rangle = \sum_k \omega_k \frac{|\psi_k(\tau), \mathbf{X}_k(\tau)\rangle}{\langle\Psi_T|\psi_k(\tau), \mathbf{X}_k(\tau)\rangle} \tag{19}$$

to perform imaginary-time propagation, namely,

$$|\Psi(\tau+\Delta\tau)\rangle = e^{-\Delta\tau\hat{\mathcal{H}}_{\text{el}}^{(2)}} e^{-\Delta\tau(\hat{\mathcal{H}}_{\text{el}}^{(1)}+\hat{\mathcal{H}}_{\text{el-ph}})} e^{-\Delta\tau\hat{\mathcal{H}}_{\text{ph}}}|\Psi(\tau)\rangle. \tag{20}$$

In the propagation of the phonon degrees of freedom we sample from the distribution

$$f[(\mathbf{X}(\tau+\Delta\tau)] \equiv \frac{\langle\Psi_T|\psi(\tau), \mathbf{X}(\tau+\Delta\tau)\rangle}{\langle\Psi_T|\psi(\tau), \mathbf{X}(\tau)\rangle} \times \langle \boldsymbol{X}(\tau+\Delta\tau)|e^{-\Delta\tau\hat{\mathcal{H}}_{\text{ph}}}|\boldsymbol{X}(\tau)\rangle\,, \tag{21}$$

where we have omitted the walker index k in the subscript. One can derive the following MC move for the updating of

the variable $\mathbf{X}(\tau) \equiv \mathbf{X}$:

$$\mathbf{X}(\tau+\Delta\tau) = \mathbf{X}(\tau) + \mathcal{N}\left(\mu=0, \sigma=\sqrt{\frac{\Delta\tau}{m}}\right) + \frac{\Delta\tau}{m}\frac{\nabla_{\mathbf{X}}\langle\Psi_T|\psi(\tau),\mathbf{X}(\tau)\rangle}{\langle\Psi_T|\psi(\tau),\mathbf{X}(\tau)\rangle}. \tag{22}$$

Here $\mathcal{N}(\mu,\sigma)$ is a normally distributed random number with mean μ and variance σ^2, and the last term is the so-called drift term. Updates for the walker weights are carried out as

$$w(\tau+\Delta\tau) = w(\tau)e^{-\frac{\Delta\tau}{2}[E_{\rm ph}(\tau+\Delta\tau)+E_{\rm ph}(\tau)-2E_{\rm shift}]}, \tag{23}$$

where $E_{\rm shift}$ is a constant shift that can be adjusted to control walker weight fluctuations, and we define

$$E_{\rm ph}(\tau) = \frac{\langle\Psi_T|\hat{\mathcal{H}}_{\rm ph}|\psi(\tau),\mathbf{X}(\tau)\rangle}{\langle\Psi_T|\psi(\tau),\mathbf{X}(\tau)\rangle}. \tag{24}$$

This algorithm is the same as the standard diffusion Monte Carlo algorithm [71–74].

Propagation arising from $\hat{\mathcal{H}}_{\text{el-ph}}$ is straightforward to implement since

$$\sum_i \hat{n}_i\hat{X}_i|\psi(\tau),\mathbf{X}(\tau)\rangle = \sum_i \hat{n}_i X_i(\tau)|\psi(\tau),\mathbf{X}(\tau)\rangle. \tag{25}$$

$\sum_i \hat{n}_i X_i(\tau)$ is thus a diagonal matrix in the single-particle space with its ith entry being $X_i(\tau)$. It is then straightforward to exponentiate this matrix and apply it along with $\hat{\mathcal{H}}_{\rm el}^{(1)}$ to the Slater determinant.

Lastly, propagation generated by $\hat{\mathcal{H}}_{\rm el}^{(2)}$ is the same as that for the standard AFQMC algorithm for the Hubbard model. We employ the discrete Hirsch spin decomposition for the two-body propagator [60]:

$$e^{-\Delta\tau U\hat{n}_{i_\uparrow}\hat{n}_{i_\downarrow}} = \frac{1}{2}e^{-\Delta\tau U(\hat{n}_{i_\uparrow}+\hat{n}_{i_\downarrow})/2}\sum_{x_i=\pm1} e^{\gamma x_i(\hat{n}_{i_\uparrow}-\hat{n}_{i_\downarrow})}, \tag{26}$$

where the constant γ is determined by

$$\cosh(\gamma) = e^{-\Delta\tau U/2}. \tag{27}$$

For a given $\mathbf{x}$, the action of Eq. (26) on a single Slater determinant keeps the Slater determinant in the single determinant manifold. In AFQMC we keep track of the overlap between the walker wave function and a chosen trial wave function. More specifically, we measure the overlap ratio of the kth walker

$$r_k = \frac{\langle\Psi_T|\psi_k(\tau+\Delta\tau),\mathbf{X}_k(\tau+\Delta\tau)\rangle}{\langle\Psi_T|\psi_k(\tau),\mathbf{X}_k(\tau+\Delta\tau)\rangle}. \tag{28}$$

If r_k is negative, the constraint condition is invoked and we set the weight w_k to zero, which then causes the walker to be removed from the simulation. Furthermore, we apply heat-bath sampling [51] using this ratio to importance sample the Ising variables. This completes the description of our algorithm for the Hubbard-Holstein Hamiltonian.

The local energy evaluation at τ with the Hubbard-Holstein model is straightforward via the one-body walker Green's function

$$G_{i_\sigma j_\sigma}(\tau) \equiv \frac{\langle\Psi_T|\hat{a}_{i_\sigma}^\dagger\hat{a}_{j_\sigma}|\psi(\tau),\mathbf{X}(\tau)\rangle}{\langle\Psi_T|\psi(\tau),\mathbf{X}(\tau)\rangle}, \tag{29}$$

and the two-body walker Green's function

$$\Gamma_{i_\uparrow i_\downarrow} = \frac{\langle\Psi_T|\hat{a}_{i_\uparrow}^\dagger\hat{a}_{i_\uparrow}\hat{a}_{i_\downarrow}^\dagger\hat{a}_{i_\downarrow}|\psi(\tau),\mathbf{X}(\tau)\rangle}{\langle\Psi_T|\psi(\tau),\mathbf{X}(\tau)\rangle}. \tag{30}$$

We will also need the mixed estimator for the phonon displacement

$$\langle\hat{X}_i\rangle(\tau) \equiv \frac{\langle\Psi_T|\hat{X}_i|\psi(\tau),\mathbf{X}(\tau)\rangle}{\langle\Psi_T|\psi(\tau),\mathbf{X}(\tau)\rangle} = X_i(\tau), \tag{31}$$

and for the squared phonon momentum

$$\langle\hat{P}_i^2\rangle(\tau) = -\frac{\langle\Psi_T|\nabla_{X_i}^2|\psi(\tau),\mathbf{X}(\tau)\rangle}{\langle\Psi_T|\psi(\tau),\mathbf{X}(\tau)\rangle}, \tag{32}$$

where $\nabla_{X_i}^2$ can be applied to the left on $|\Psi_T\rangle$. Using these mixed estimators, the local energy can be evaluated as

$$E_L = -t\sum_\sigma\sum_{\langle ij\rangle} G_{i_\sigma j_\sigma} + U\sum_i \Gamma_{i_\uparrow i_\downarrow} + \sum_i\left(\frac{m\omega^2}{2}X_i^2 + \frac{1}{2m}\langle\hat{P}_i^2\rangle - \frac{\omega}{2}\right) - g\sqrt{2m\omega}\sum_i(G_{i_\uparrow i_\uparrow}+G_{i_\downarrow i_\downarrow})X_i. \tag{33}$$

IV. TRIAL WAVE FUNCTIONS

The choice of the trial wave function can affect the quality of the CP approximation in treating the electronic degree of freedom. It can also affect the computational efficiency in treating the electronic and especially the phononic degrees of freedom; in particular, a poor choice of the importance function can magnify or even introduce additional ergodicity issues, especially in an el-ph system when multiple phonon modes are pronounced. It is highly advantageous if an accurate trial wave function allows the overlap ratio in Eq. (28), and the local energy in Eq. (33), to be efficiently evaluated.

A. Semiclassical state

The simplest variational trial wave function that we employ in this work takes a simple product form between electronic and bosonic degrees of freedom,

$$|\Psi_T\rangle = |\psi_T\rangle\otimes|\phi_T\rangle, \tag{34}$$

where $|\psi_T\rangle$ is a single determinant and $|\phi_T\rangle$ is a coherent state (or a shifted harmonic oscillator state). This wave function has been referred to as a "semiclassical state" in literature [75,76]. Due to its simple product form, there is no explicit entanglement between electrons and phonons. The electronic trial wave function $|\psi_T\rangle$ is parametrized by orbital rotation θ,

$$|\psi_T(\theta)\rangle = e^{\hat{\kappa}}|\psi_0\rangle, \tag{35}$$

where

$$\hat{\kappa} = \sum_{ij}[\theta_{ij} - (\theta_{ji})^*]\hat{a}_i^\dagger\hat{a}_j, \tag{36}$$

and $|\psi_0\rangle$ is some initial determinant (normally obtained by diagonalizing the one-body electronic Hamiltonian). Single determinant trial wave functions have been widely used in previous AFQMC studies of the Hubbard model [28,29,77,78].

The phonon trial wave function $|\phi_T\rangle$ is parametrized by coherent state displacements β,

$$|\phi_T(\beta)\rangle = e^{\sum_i \beta_i \hat{b}_i^\dagger - \beta_i^* \hat{b}_i}|0\rangle \equiv \hat{\mathcal{D}}(\beta)|0\rangle, \tag{37}$$

where $\hat{\mathcal{D}}(\beta)$ is the displacement operator. We optimize the energy of $|\Psi_T\rangle$ in Eq. (34) variationally over θ and β and use this as the final trial wave function. $|\phi_T(\beta)\rangle$ technically contains infinitely many bosons, but it has a convenient property which allows for an efficient AFQMC algorithm

$$\hat{b}_i|\phi_T(\beta)\rangle = \beta_i|\phi_T(\beta)\rangle. \tag{38}$$

Using this fact, one can show that the projection of $\langle\phi_T(\beta)|$ on to $|\mathbf{X}\rangle$ is

$$\langle\phi_T(\beta)|\mathbf{X}\rangle = \prod_i \left(\frac{m\omega}{\pi}\right)^{\frac{1}{4}} e^{-\frac{m\omega}{2}(X_i - \sqrt{\frac{2}{m\omega}}\beta_i)^2}. \tag{39}$$

Similarly, the numerator of Eq. (32) is straightforward to evaluate as well using

$$-\langle\phi_T(\beta)|\nabla_{X_i}^2|\mathbf{X}\rangle = -\nabla_{X_i}^2\langle\phi_T(\beta)|\mathbf{X}\rangle. \tag{40}$$

This semiclassical trial wave function therefore can be efficiently combined with the AFQMC algorithm.

We elaborate here on formal properties of this state. We first note that the operator $\hat{\kappa}$ is anti-Hermitian, and thus the orbital rotation operator $e^{\hat{\kappa}}$ is a unitary operator. Because $\hat{\kappa}$ is an anti-Hermitian operator, the variational parameters in θ are not all independent. Formally, one can represent the same wave function with only $M(M-1)/2$ parameters, where M is the number of lattice sites. Similarly, the displacement operator is also a unitary operator. While θ and β are in principle complex valued, we assume them to be real valued for the rest of this paper. A complex valued θ can be useful for the description of certain strongly correlated systems [79–81], but we do not focus on these cases here. A complex valued β introduces an average momentum to the coherent state through its imaginary component. However, with variational optimization the imaginary component of β is found to be zero.

The variational energy of the semiclassical state can be obtained within the Born-Oppenheimer (BO) approximation. After some algebra, it can be shown that the lowest energy of a semiclassical state can be obtained by minimizing

$$\langle\hat{H}_{\rm el}\rangle - \frac{g^2}{\omega}\left\langle\sum_i \hat{n}_i^2\right\rangle \tag{41}$$

over the variational parameters in $|\psi_T\rangle$. For a fixed λ, variations in α do not change the energy of the semiclassical state.

The semiclassical state is exact in (1) the limit $g \to \infty$, (2) the adiabatic limit $\omega \to 0$ (for a fixed λ) with $U \to 0$, and (3) the atomic limit $U \to \infty$. When $\lambda \to \infty$ (or $g \to \infty$ for a fixed ω), the use of a single semiclassical state is not problematic even though the BO potential develops into a well-separated double-well potential. This situation is physically similar to that of the atomic limit of the Hubbard model ($U \to \infty$) where spin flips do not cause an energy penalty and a degeneracy occurs amongst all possible 2^N spin flips where N is the number of electrons. Similarly, in the Holstein model, charge swapping does not cause an energy penalty and the same macroscopic degeneracy occurs. In other words, any one of the degenerate semiclassical states is equally well suited as an approximate wave function.

Aside from these limits, the semiclassical state itself can be inaccurate, but we find that the subsequent AFQMC calculation with the semiclassical trial wave function is often numerically exact. The most difficult parameter regime for our AFQMC framework is when the Holstein coupling strength g takes an intermediate value. That is, g is large enough that the el-ph correlation is strong but is small enough that the macroscopic degeneracy does not occur. A straightforward way to probe these situations is to increase ω for a fixed λ value. In this case, g can be much larger than t but is always smaller than ω as long as $2dt\lambda < \omega$. In these situations, the semiclassical state can be a poor choice of a trial wave function in AFQMC, as we shall see. This is because no correlation between electronic and bosonic degrees of freedom is built into this trial wave function.

From a different point of view, the difficulty of semiclassical states was noted in the work of Proville and Aubry, who defined the "quantumness" of the phonons as [82]

$$\zeta = \frac{\omega^2}{g^2} = \frac{\omega}{2dt\lambda}. \tag{42}$$

As ζ increases, the semiclassical state qualitatively fails [82]. This is consistent with the picture described above in that for a given λ, both g and ζ increase as ω increases. We attribute the difficulties associated with semiclassical states to the increase in correlation between electrons and phonons instead of the quantal effect associated with the phonons alone.

B. Multivibronic state

A linear combination of multiple semiclassical states can be used to correlate electrons and phonons

$$|\Psi_{\rm MS}\rangle = \sum_k c_k|\psi_T(\theta_k)\rangle \otimes |\phi_T(\beta_k)\rangle, \tag{43}$$

where one may determine $\mathbf{c}$, θ, and β variationally. We refer to this wave function as a multivibronic (MV) wave function, similar to the multi-Slater determinant trial states employed to study purely electronic systems [77]. The MV wave functions of this form would need exponentially many states for large systems. Nevertheless, due to their simplicity, multivibronic states can be valuable for the study of small systems.

A particular flavor of MV wave function that we focus on in this work is closely tied to the underlying order encoded in the semiclassical states themselves. Let us consider the two-electron, two-site Holstein model. It is well known that for the Holstein model at large coupling λ, the BO surface develops into a double-well potential [65,83]. For the two-site problem, the BO potential energy surfaces (PESs) are characterized by

$$\hat{H}_{\rm BO}(X_1, X_2) = \hat{H}_{\rm el} - g\sqrt{2m\omega}(\hat{n}_1 X_1 + \hat{n}_2 X_2) + \frac{m\omega^2}{2}\left(X_1^2 + X_2^2\right), \tag{44}$$

where X_1 and X_2 are constant scalars denoting the coordinates of the classical phonons. We can find the ground-state electronic wave function of the Hamiltonian in Eq. (44) by exactly

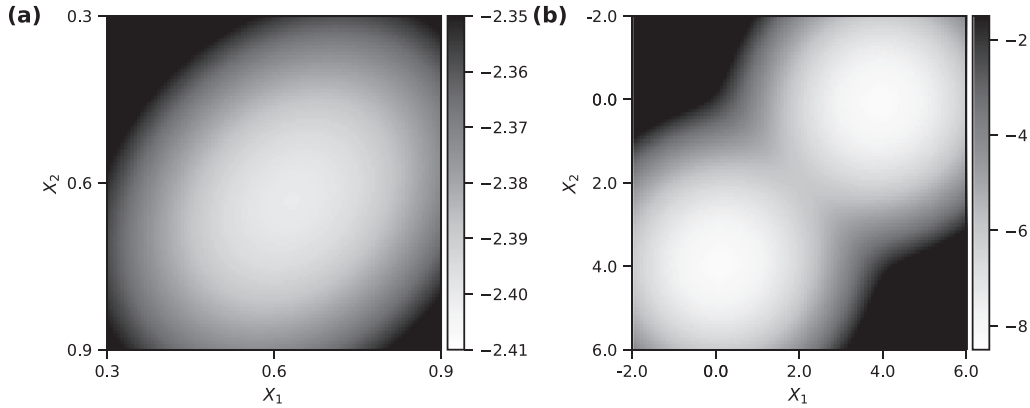

FIG. 1. Born-Oppenheimer potential energy surfaces in units of t for the two-electron and two-site Holstein model: (a) $\omega = t$, $\lambda = 0.1$, and $g = 0.447t$ and (b) $\omega = t$, $\lambda = 1$, and $g = 1.414t$. The minimum in (a) is $E = -2.40t$ at $(X_1 = 0.63, X_2 = 0.63)$, while the two minima in (b) are $E = -8.25t$ at $(X_1 = 3.90, X_2 = 0.10)$ and $(X_1 = 0.10, X_2 = 3.90)$, respectively.

diagonalizing it and forming a potential energy surface for each combination of (X_1, X_2). In Fig. 1, a representative example of the BO PESs is given. Figure 1(a) illustrates an example of the weak coupling case, where the coherent states have the same centers for all sites and no charge modulation occurs. In such cases, the minimum BO state (i.e., the semiclassical state) is an excellent variational wave function. In Fig. 1(b) there are two distinct minima with equal BO energies. Here, a wave function of a single semiclassical state with a Gaussian function centered at one of the two BO minima in position space would not provide a good description of the system. When used as an importance function, it can introduce or exacerbate ergodicity problems in the Monte Carlo sampling and induce a large or even infinite variance in the energy fluctuations.

We propose the following improved variational wave function in this situation for the Holstein model on a bipartite lattice. At half-filling there are two *exactly* degenerate semiclassical states. In particular, one state is characterized by

$$\beta_i^{(1)} = \begin{cases} \beta_e, & \text{if } i \text{ on A sublattice} \\ \beta_o, & \text{if } i \text{ on B sublattice} \end{cases} \tag{45}$$

where i is a site index. β_e and β_o are determined by variationally optimizing semiclassical states. The pertinent orders in 2D are stripe orders, checkerboard orders, etc. One can easily identify the bipartite sublattice sites in such orders. The other degenerate solution is given by switching the A and B sublattices. One can smoothly interpolate between the two states by defining a convex combination

$$\beta(\alpha) = \alpha\beta^{(1)} + (1-\alpha)\beta^{(2)} \tag{46}$$

for $\alpha \in [0, 1]$. For each $\beta(\alpha)$, we find a single determinant that minimizes the energy of a single semiclassical state. One can take a linear combination of all of these semiclassical states along the line that interpolates two solutions to form a MV wave function. We refer to this as the Thouless path (TP) wave function $|\Psi_{\mathrm{TP}}\rangle$,

$$|\Psi_{\mathrm{TP}}\rangle = \sum_\alpha c_\alpha |\Psi_T(\beta(\alpha))\rangle, \tag{47}$$

where c_α is determined by variationally minimizing the energy. The construction of the TP wave function can be generalized to arbitrary filling fractions and number of sites because different filling fractions simply give rise to ordered states with different wavelengths. The cost for its construction is negligible compared to the optimization of a semiclassical state. Its use in AFQMC as a trial wave function simply introduces a prefactor depending on the number of states included in Eq. (47). We will refer to a TP wave function with n interpolation values of α as TP(n). While TP wave functions provide a simple and accurate importance function for double-well potential energy surfaces, they also become inaccurate when the correlation between electrons and phonons becomes strong.

C. Variational Lang-Firsov trial wave functions

A simple, widely used way to incorporate correlation effects between electrons and phonons is to use the polaron transformation or the Lang-Firsov (LF) transformation [84]

$$|\Psi_{\mathrm{LF}}\rangle = \hat{U}_{\mathrm{LF}}(\xi)|\psi_T(\theta)\rangle \otimes |\phi_T(\beta)\rangle, \tag{48}$$

where

$$\hat{U}_{\mathrm{LF}}(\xi) = e^{\frac{1}{\sqrt{2}}\sum_i \xi_i \hat{n}_i (\hat{b}_i^\dagger - \hat{b}_i)}, \tag{49}$$

and the set ξ are referred to as the LF amplitudes which are variational parameters along with θ and β. From the wave-function viewpoint, (48) provides a way to explicitly build a wave function with nonperturbative el-ph correlation on top of semiclassical states via a unitary transformation. Typical LF implementations involve the phonon vacuum state as opposed to the coherent state in (48). We find that having the coherent state provides additional variational flexibility and thereby yields lower energies compared to those that use the vacuum state. Since it does not complicate the underlying optimization problem, we use the coherent state as written in (48).

While the details of the LF transformation and its variational optimization have been well documented [84–87], we briefly summarize them to provide a self-contained description. Our goal is to simultaneously determine ξ, θ, and β variationally. To carry this out, we find that it is simpler to work with the unitary-transformed Hamiltonian $\hat{H}^{\mathrm{LF}}$, based on $\hat{U}_{\mathrm{LF}}$, and optimize the variational energy of $\hat{H}^{\mathrm{LF}}$ evaluated

with the semiclassical wave function. We start from

$$\hat{U}_{\rm LF}(\xi)^\dagger a_{i_\sigma}^\dagger a_{j_\sigma} \hat{U}_{\rm LF}(\xi) = a_{i_\sigma}^\dagger a_{j_\sigma} e^{[-\xi_i(b_i^\dagger - b_i)+\xi_j(b_j^\dagger - b_j)]} \tag{50}$$

and

$$\hat{U}_{\rm LF}(\xi)^\dagger b_i^\dagger \hat{U}(\xi) = b_i^\dagger - \xi_i \hat{n}_i, \tag{51}$$

$$\hat{U}_{\rm LF}(\xi)^\dagger b_i \hat{U}(\xi) = b_i - \xi_i \hat{n}_i. \tag{52}$$

The LF transformed Hamiltonian reads as

$$\hat{\mathcal{H}}_{\rm el}^{\rm LF} = -t \sum_\sigma \sum_{\langle ij \rangle} a_{i_\sigma}^\dagger a_{j_\sigma} e^{[-\frac{\xi_i}{\sqrt{2}}(b_i^\dagger - b_i)+\frac{\xi_j}{\sqrt{2}}(b_j^\dagger - b_j)]} + U \sum_i \hat{n}_{i_\uparrow} \hat{n}_{i_\downarrow}, \tag{53}$$

$$\hat{\mathcal{H}}_{\rm ph}^{\rm LF} = \omega \sum_i \left(\hat{b}_i^\dagger + \frac{\xi_i}{\sqrt{2}}\hat{n}_i\right)\left(\hat{b}_i + \frac{\xi_i}{\sqrt{2}}\hat{n}_i\right), \tag{54}$$

and

$$\hat{\mathcal{H}}_{\text{el-ph}}^{\rm LF} = -g \sum_i \hat{n}_i (\hat{b}_i + \hat{b}_i^\dagger + \sqrt{2}\xi_i \hat{n}_i). \tag{55}$$

All of the energy terms are straightforward to evaluate with semiclassical trial wave functions. The electronic kinetic energy is more complex than its bare Hamiltonian counterpart due to the presence of exponential bosonic operators, so we provide more details here. To utilize Eq. (38), we write the exponential term in the kinetic energy operator as

$$e^{\frac{1}{\sqrt{2}}[\xi_j(\hat{b}_j^\dagger - \hat{b}_j) - \xi_i(\hat{b}_i^\dagger - \hat{b}_i)]} = e^{\frac{1}{\sqrt{2}}(\xi_j \hat{b}_j^\dagger - \xi_i \hat{b}_i^\dagger)} e^{\frac{1}{\sqrt{2}}(\xi_i \hat{b}_i - \xi_j \hat{b}_j)} e^{-\frac{1}{4}(\xi_i^2 + \xi_j^2)}, \tag{56}$$

where we have used

$$e^{\frac{1}{\sqrt{2}}\xi_j(\hat{b}_j^\dagger - \hat{b}_j)} = e^{\frac{1}{\sqrt{2}}\xi_j \hat{b}_j^\dagger} e^{-\frac{1}{\sqrt{2}}\xi_j \hat{b}_j} e^{-\frac{1}{4}\xi_j^2}. \tag{57}$$

The expectation value of Eq. (56) is simple to evaluate with the semiclassical state of Eq. (34):

$$\langle \Psi_T | e^{\frac{1}{\sqrt{2}}(\xi_j \hat{b}_j^\dagger - \xi_i \hat{b}_i^\dagger)} e^{\frac{1}{\sqrt{2}}(\xi_i \hat{b}_i - \xi_j \hat{b}_j)} | \Psi_T \rangle e^{-\frac{1}{4}(\xi_i^2 + \xi_j^2)} = e^{\frac{1}{\sqrt{2}}(\xi_j \beta_j - \xi_i \beta_i)} e^{\frac{1}{\sqrt{2}}(\xi_i \beta_i - \xi_j \beta_j)} e^{-\frac{1}{4}(\xi_i^2 + \xi_j^2)}. \tag{58}$$

The variational LF wave function is expected to be more accurate than the semiclassical state due to the explicit correlation between electrons and phonons. Furthermore, the limit of $\omega \to \infty$ which is difficult for simple semiclassical wave functions to treat, can be exactly treated by the LF wave function, because the el-ph coupling term in $\hat{H}^{\rm LF}$ can be removed by setting $\xi_i = \sqrt{2}g/\omega$. Due to the fact that phonon displacements are significantly penalized in this limit, the variational optimization over β naturally yields $\beta = 0$. Therefore, the bosonic operators in the hopping amplitude in Eq. (53) all vanish. Provided that one can handle the remaining electronic Hamiltonian terms exactly, the variational LF wave function should be exact in this limit. We note that for many-electron systems in the $\omega \to \infty$ limit, the LF Hamiltonian takes the same form as the attractive Hubbard model, which is another sign-free lattice model that can be efficiently simulated in AFQMC [88,89].

Despite these desirable properties, there seems to be no simple and general way to use this wave function in AFQMC without invoking a major increase in scaling. As an exception to this, we mention here the work of Hohenadler and coworkers [69] where a QMC algorithm with the LF Hamiltonian was formulated for single-electron problems. It was demonstrated, however, that the transformed electronic Hamiltonian in Eq. (53) creates a complex phase problem.

Therefore, we briefly investigate a simpler linearized LF (LLF) wave function of the form

$$|\Psi_{\rm LLF}\rangle = \left(1 + \frac{1}{\sqrt{2}} \sum_i \xi_i \hat{n}_i \hat{b}_i^\dagger\right) |\psi_T(\theta)\rangle \otimes |\phi_T(\beta)\rangle, \tag{59}$$

where we have omitted a term that is proportional to $\hat{n}_i \hat{b}_i$ since the action of $\hat{b}_i$ on $|\phi_T(\beta)\rangle$ is trivial due to Eq. (38). We variationally optimize ξ in Eq. (59) to maximize the accuracy of the LLF trial wave function. The AFQMC algorithm presented in Sec. III can be efficiently implemented for Eq. (59).

It is possible to formulate a simple extension of the LLF wave function in the spirit of the TP wave function:

$$|\Psi_{\text{TP-LLF}}\rangle = \sum_{\alpha=1}^{n} c_\alpha |\Psi_{\rm LLF}^{(\alpha)}\rangle, \tag{60}$$

where each of the $|\Psi_{\rm LLF}^{(\alpha)}\rangle$ terms has its own variational parameters. Following the discussion of the TP trial wave function, it may be possible to determine these variational parameters via a convex interpolation of β and ξ as in Eq. (46). We refer to this wave function as the TP-LLF(n) wave function which goes beyond both the TP(n) and the LLF wave functions in sophistication.

In contrast with the LF form, a trial wave function with an el-ph Jastrow factor can be used more straightforwardly in AFQMC since $\hat{X}_i$ operators are involved in the exponent instead of $\hat{P}_i$ as in LF. The unitary transformation in the LF wave function can be thought of as a simple Jastrow factor that encodes correlation between the electronic density and the phonon momentum on a site. However, $\hat{U}_{\rm LF}(\xi)$ is unitary and we thus expect this transformation to behave differently from Jastrow factors in el-ph problems [37–39]. (It is also different from the coupled-cluster operators considered in recent studies of el-ph problems [47–49].) Nonetheless, when linearized both Jastrow and $\hat{U}_{\rm LF}(\xi)$ yield identical results. Given the performance improvement with the LLF trial wave function (as discussed below), we expect an el-ph Jastrow trial wave function will greatly reduce the difficulties in parameter regimes with strong el-ph coupling, and result in a major improvement in our AFQMC approach. We leave the implementation and systematic studies using an el-ph Jastrow trial wave function in AFQMC for future work.

D. Additional details

The semiclassical state in Eq. (34) can describe two competing phases, SDW and CDW. To obtain the variational wave function for these two distinct states, we employ the following protocol:

(1) For a CDW state, we perform a variational optimization of a semiclassical state with spin restriction. Due to the spin restriction, any states that arise from minimization are not capable of describing SDW order.

(2) For an SDW state, we perform a variational optimization of a spin-unrestricted Hartree-Fock (UHF) wave function to minimize the electronic energy. Once a UHF state is obtained, we determine the shift vector β variationally while fixing the electronic degrees of freedom. As long as the UHF state exhibits SDW order, such a coupled el-ph semiclassical state will exhibit the same SDW order. We have used an *ad hoc* effective repulsion strength (U_{eff}/t) of 0.5 [77] in our UHF calculations to obtain SDW trial wave functions for the Hubbard-Holstein model in this work. The CP-AFQMC results are not sensitive to this particular choice. (We note that it is possible to determine this effective repulsion strength via a self-consistent procedure with CP-AFQMC [90].)

V. PERTURBATION THEORY

A. Coherent state Møller-Plesset perturbation theory

It is instructive to consider low-order perturbation theory for comparison to numerically exact approaches. We first note that

$$\hat{\mathcal{D}}(\beta)^\dagger b_i^\dagger \hat{\mathcal{D}}(\beta) = b_i^\dagger + \beta_i, \tag{61}$$

$$\hat{\mathcal{D}}(\beta)^\dagger b_i \hat{\mathcal{D}}(\beta) = b_i + \beta_i. \tag{62}$$

Using this property, we write

$$\hat{\tilde{\mathcal{H}}}_{\text{ph}} \equiv \hat{\mathcal{D}}(\beta)^\dagger \hat{\mathcal{H}}_{\text{ph}} \hat{\mathcal{D}}(\beta) = \omega \sum_i (b_i^\dagger + \beta_i)(b_i + \beta_i) \tag{63}$$

and

$$\hat{\tilde{\mathcal{H}}}_{\text{el-ph}} \equiv \hat{\mathcal{D}}(\beta)^\dagger \hat{\mathcal{H}}_{\text{el-ph}} \hat{\mathcal{D}}(\beta) = -g \sum_i \hat{n}_i (\hat{b}_i + \hat{b}_i^\dagger + 2\beta_i). \tag{64}$$

Thus, we have

$$\hat{\tilde{\mathcal{H}}} = \hat{\mathcal{H}}_{\text{el}} + \hat{\tilde{\mathcal{H}}}_{\text{ph}} + \hat{\tilde{\mathcal{H}}}_{\text{el-ph}}. \tag{65}$$

We note that the following zeroth-order Hamiltonian naturally has the semiclassical state of Eq. (34) as its ground state

$$\hat{\mathcal{H}}_0 = \hat{\mathcal{F}} + \omega \sum_i \left(\beta_i^2 + \hat{b}_i^\dagger \hat{b}_i\right), \tag{66}$$

where $\hat{\mathcal{F}}$ is the Fock operator defined as (for spin $\sigma = \uparrow$ or $\downarrow$)

$$\hat{\mathcal{F}}_\sigma = \hat{\mathcal{F}}_\sigma^{\text{el}} - 2g \sum_i \hat{n}_{i_\sigma} \beta_i, \tag{67}$$

with the electronic Fock operators

$$\hat{\mathcal{F}}_\uparrow^{\text{el}} = -t \sum_{\langle ij \rangle} \hat{a}_{i_\uparrow}^\dagger \hat{a}_{j_\uparrow} + U \sum_i \hat{n}_{i_\uparrow} \langle \hat{n}_{i_\downarrow} \rangle_{\psi_T}, \tag{68}$$

$$\hat{\mathcal{F}}_\downarrow^{\text{el}} = -t \sum_{\langle ij \rangle} \hat{a}_{i_\downarrow}^\dagger \hat{a}_{j_\downarrow} + U \sum_i \hat{n}_{i_\downarrow} \langle \hat{n}_{i_\uparrow} \rangle_{\psi_T}, \tag{69}$$

where

$$\langle \hat{n}_{i_\sigma} \rangle_{\psi_T} = \frac{\langle \psi_T | \hat{n}_{i_\sigma} | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle}. \tag{70}$$

It is straightforward to show that $|\Psi_T\rangle$ is an eigenstate of $\hat{\mathcal{H}}_0$. From this starting point, one can develop an order-by-order perturbation theory to capture all of the correlation effects among electrons and between electrons and phonons built through

$$\begin{aligned} \hat{V} &= \hat{\tilde{\mathcal{H}}} - \hat{\mathcal{H}}_0 \\ &= (\hat{\mathcal{H}}_{\text{el}} - \hat{\mathcal{F}}_{\text{el}}) + \sum_i (\omega \beta_i - g \hat{n}_i)(\hat{b}_i + \hat{b}_i^\dagger). \end{aligned} \tag{71}$$

We note that such a partitioning of the Hamiltonian resembles the widely used Møller-Plesset (MP) perturbation theory in quantum chemistry [91]. We refer to this perturbation theory as "coherent state Møller-Plesset perturbation theory" (CSMP) since a coherent state (or a semiclassical state) is an eigenstate of the zeroth-order Hamiltonian. This was also recently discussed in the work of White and coworkers in the context of coupled-cluster theory [49].

Similar to MP, CSMP recovers the energy of the semiclassical state with the first-order perturbation correction

$$E^{(0)} + E^{(1)} = \langle \hat{\mathcal{H}}_{\text{el}} \rangle_{\psi_T} + \omega \sum_i \beta_i^2 - 2g \sum_i \langle \hat{n}_i \rangle_{\psi_T} \beta_i, \tag{72}$$

where we have defined

$$\langle \hat{n}_i \rangle_{\psi_T} \equiv \langle \psi_T | \hat{n}_{i_\uparrow} + \hat{n}_{i_\downarrow} | \psi_T \rangle. \tag{73}$$

In this work, we are interested in comparing the second-order perturbation theory (CSMP2) with AFQMC. The evaluation of the CSMP2 energy is most natural in the molecular orbital (MO) basis rather than in the site basis. The MO basis is defined by a set of orbitals $\{\psi_{i_\sigma}\}$ that satisfy

$$\hat{\mathcal{F}}_\sigma \psi_{p_\sigma} = \epsilon_{p_\sigma} \psi_{p_\sigma}, \tag{74}$$

where ϵ_{p_σ} is the pth MO energy and the pth MO, ψ_{p_σ}, is expanded via a set of site orbitals $\{\phi_{\mu_\sigma}\}$:

$$\psi_{p_\sigma} = \sum_\mu C_{\mu p_\sigma} \phi_{\mu_\sigma}. \tag{75}$$

We then transform Eq. (71) from the site basis to the MO basis using the coefficient matrix $\mathbf{C}$ for each spin

$$\begin{aligned} \hat{\tilde{V}} = &\sum_{\sigma \in \{\uparrow,\downarrow\}} \sum_{pq} \sum_\mu \left((\omega \beta_\mu - g_{pq}) \hat{a}_{p_\sigma}^\dagger \hat{a}_{q_\sigma} \hat{b}_\mu^\dagger + \text{H.c.} \right) \\ &+ \sum_{pqrs} U_{p_\uparrow q_\downarrow r_\downarrow s_\uparrow} \hat{a}_{p_\uparrow}^\dagger \hat{a}_{q_\downarrow}^\dagger \hat{a}_{r_\downarrow} \hat{a}_{s_\uparrow}, \end{aligned} \tag{76}$$

where

$$g_{pq} = g \sum_\mu (C_{\mu p})^* C_{\mu q}, \tag{77}$$

$$U_{p_\uparrow q_\downarrow r_\downarrow s_\uparrow} = U \sum_\mu (C_{\mu p_\uparrow})^* (C_{\mu q_\downarrow})^* C_{\mu r_\downarrow} C_{\mu s_\uparrow}. \tag{78}$$

The CSMP2 energy expression follows in a spin-orbital MO basis

$$\begin{aligned} E^{(2)} = &- \sum_{\mu=1}^{M} \frac{\left[\omega \beta_\mu - g \sum_i (C_{\mu i})^* C_{\mu i}\right]^2}{\omega} - \sum_{ia} \frac{-|g_{ai}|^2}{\omega + \epsilon_a - \epsilon_i} \\ &- \sum_{i_\uparrow a_\uparrow} \sum_{j_\downarrow b_\downarrow} \frac{|U_{i_\uparrow j_\downarrow b_\downarrow a_\uparrow}|^2}{\epsilon_{a_\uparrow} + \epsilon_{b_\downarrow} - \epsilon_{i_\uparrow} - \epsilon_{j_\downarrow}}, \end{aligned} \tag{79}$$

where the orbital energies $\{\epsilon_p\}$ are eigenvalues of the Fock operator $\hat{\mathcal{F}}$. We note that the first term in Eq. (79) is zero if the semiclassical reference state is fully optimized.

B. Lang-Firsov perturbation theory

It may be useful to develop a second-order perturbation theory from a reference state given by the Lang-Firsov transformation [84] in (48). In the spirit of the original LF transformation [84], we set

$$\xi_i = \frac{\sqrt{2}g}{\omega}, \tag{80}$$

which removes the Holstein coupling in the transformed framework. With this choice of the LF amplitudes instead of the variational LF amplitudes, the transformed Hamiltonian is simplified to

$$\hat{\mathcal{H}}^{\mathrm{LF}} = \omega \sum_i \hat{b}_i^\dagger \hat{b}_i - \frac{g^2}{\omega} \sum_i \hat{n}_i + \left(U - \frac{2g^2}{\omega}\right) \sum_i \hat{n}_{i_\uparrow} \hat{n}_{i_\downarrow} - t \sum_\sigma \sum_{\langle ij \rangle} \hat{a}_{i_\sigma}^\dagger \hat{a}_{j_\sigma} e^{[-\frac{g}{\omega}(\hat{b}_i^\dagger - \hat{b}_i) + \frac{g}{\omega}(\hat{b}_j^\dagger - \hat{b}_j)]}. \tag{81}$$

For convenience, we rewrite

$$e^{[-\frac{g}{\omega}(\hat{b}_i^\dagger - \hat{b}_i) + \frac{g}{\omega}(\hat{b}_j^\dagger - \hat{b}_j)]} = e^{-\frac{g^2}{\omega^2}} e^{-\frac{g}{\omega}(\hat{b}_i^\dagger - \hat{b}_j^\dagger)} e^{\frac{g}{\omega}(\hat{b}_i - \hat{b}_j)}. \tag{82}$$

Based on the approach of Bonca, Trugman, and coworkers [92,93], we choose the zeroth-order Hamiltonian as

$$\hat{\mathcal{H}}_0^{\mathrm{LF}} = \omega \sum_i \hat{b}_i^\dagger \hat{b}_i - \frac{g^2}{\omega} \sum_i \hat{n}_i + \left(U - \frac{2g^2}{\omega}\right) \sum_i \hat{n}_{i_\uparrow} \hat{n}_{i_\downarrow}, \tag{83}$$

with the perturbing Hamiltonian

$$\hat{V}^{\mathrm{LF}} = -te^{-\frac{g^2}{\omega^2}} \sum_\sigma \sum_{\langle ij \rangle} \hat{a}_{i_\sigma}^\dagger \hat{a}_{j_\sigma} e^{-\frac{g}{\omega}(\hat{b}_i^\dagger - \hat{b}_j^\dagger)} e^{\frac{g}{\omega}(\hat{b}_i - \hat{b}_j)}. \tag{84}$$

For concreteness, we consider the specific case of second-order perturbation theory for a two-electron, two-site problem. The ground state of $\hat{\mathcal{H}}_0^{\mathrm{LF}}$ is either $|\uparrow_1\downarrow_1\rangle \otimes |0,0\rangle$ or $|\uparrow_2\downarrow_2\rangle \otimes |0,0\rangle$ with an energy $U - \frac{4g^2}{\omega}$. We start from an unperturbed state,

$$\left|\Psi_0^{(0)}\right\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_1\rangle \otimes |0,0\rangle, +|\uparrow_2\downarrow_2\rangle \otimes |0,0\rangle), \tag{85}$$

noting that an excited state which can be connected to the unperturbed ground state via $\hat{V}^{\mathrm{LF}}$ takes the form

$$\left|\Psi_{mn}^{(0)}\right\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle) \otimes |m,n\rangle. \tag{86}$$

It is then easy to show that the first-order energy correction to the unperturbed state is zero. The second-order energy reads as

$$E_{\mathrm{LFPT}}^{(2)} = -2t^2 e^{-\frac{2g^2}{\omega^2}} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(g/\omega)^{2(m+n)}}{m!n!} \times \frac{1 + (-1)^{m+n}}{(m+n)\omega - U + 2\frac{g^2}{\omega}}. \tag{87}$$

We numerically evaluate this expression in a brute-force manner, observing that a maximum n of 200 is sufficient to converge the energy to machine precision. We note that this expression differs from that of Bonca, Trugman, and coworkers [92,93] since in their work a single bipolaron was considered in the continuum limit (infinite lattice) whereas in our work we focus on a two-site problem. This approach is referred to as the second-order Lang-Firsov perturbation theory (LFPT2).

VI. HOLSTEIN MODEL

To study the behavior of the proposed AFQMC algorithm with simple trial wave functions such as the semiclassical, LLF, and TP wave functions, we shall investigate the 1D and 2D Holstein models first, namely, we set $U = 0$ in Eq. (1).

For the Holstein model, the sign problem is absent, as is well known in the determinant quantum Monte Carlo (DQMC) approach [94]. Similar to DQMC, the overlap function $\langle \Psi_T | \psi_k(\tau), \mathbf{X}_k(\tau) \rangle$ in Eq. (19) remains non-negative throughout the imaginary-time propagation since the phonon component $\phi_T(\mathbf{X}(\tau))$ (omitting the walker index again) is positive everywhere, and the electronic component $\langle \psi_T | \psi \rangle = |\langle \psi_{T,\uparrow} | \psi_\uparrow \rangle|^2$ with a spin-restricted form is also non-negative. Thus, in the Holstein model the difference between our approach and DQMC is primarily in the way the Monte Carlo sampling is conducted. AFQMC uses a branching random walk with a population of walkers to construct the imaginary-time path iteratively, as we have described, while DQMC treats the entire path as a path integral or world line, and updates it by sweeping different imaginary-time locations via a Metropolis-like algorithm. A second difference is the introduction of an importance function in our approach via the similarity transformation, as indicated in Eqs. (21) and (28). These factors can affect the behavior of the Monte Carlo sampling, and yield different performances in different regimes of the parameter space, including efficiency, autocorrelation time, and possibly different levels of difficulty with ergodicity. The examples in the Holstein model below serve as a first test of the AFQMC method in this context.

A. Two-electron two-site model

We start with this small problem where we easily can compare results against exact diagonalization (ED). Since there are only two sites in our model, we compute energies with open boundary conditions (OBCs). In Fig. 2, we present the error in the total energy of CSMP2, LFPT2, and AFQMC compared to ED. Understanding the behaviors of the two flavors of perturbation theory helps gauge nonperturbative effects in our system. In Fig. 2(a), it is clear that the CSMP2 energy becomes more inaccurate as we increase ω. This is because the zeroth-order wave function, a semiclassical trial wave function, starts to degrade when increasing el-ph correlation. Perhaps the most striking behavior to note concerning CSMP2 is that this approach performs worst for intermediate λ values (e.g., $\lambda = 0.5$) and is in fact more accurate for larger λ values such as $\lambda = 2.0$. This can also be understood in terms of the increase in el-ph correlation as explained in Sec. IV A. On the other hand, LFPT2 in Fig. 2(b) is comparatively more accurate than CSMP2 for λ values larger than $\lambda = 0.5$. The LF reference state (namely, two electrons occupying one site and with a phonon vacuum state) is qualitatively incorrect when the el-ph coupling is small. In such cases, we cannot treat

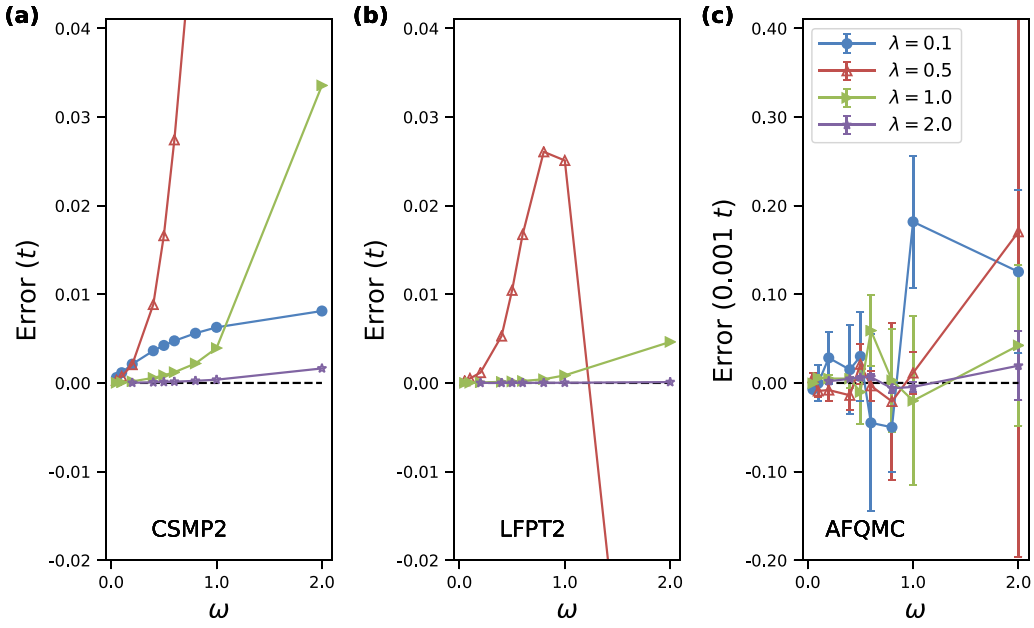

FIG. 2. Error in the total energy in units of t for the two-site two-electron Holstein model as a function of ω for various λ values: (a) CSMP2, (b) LFPT2, and (c) AFQMC results. For $\lambda = 0.1$, LFPT2 energy errors lie outside the plotted range. In (c), for $\lambda \geqslant 0.5$ AFQMC/TP(11) results are shown, while for $\lambda = 0.1$ we present AFQMC/S results. Note the different vertical scales in (c).

the hopping term perturbatively. This is clearly reflected in Fig. 2(b) as LFPT2 exhibits large errors for small λ values. As LFPT2 is well known to produce accurate results for strong coupling, it is remarkable that a weak coupling perturbation theory, CSMP2, performs equally well even at $\lambda = 2.0$.

We also show the performance of AFQMC for the same two-site Holstein dimer. The error of AFQMC is shown on a much smaller (100 times) scale. We have tested both the single semiclassical state and TP(11), i.e., the TP state with a superposition of 11 semiclassical states, as trial wave function. These are referred to as AFQMC/S and AFQMC/TP(11), respectively. Results are shown in the figure, with AFQMC/S for smaller λ and AFQMC/TP(11) for $\lambda \geqslant 0.5$. Near-exact energies are obtained for all parameters examined here. We observed that results can become severely biased with AFQMC/S for large λ, as a consequence of a poor importance function causing large, or even diverging, variances. Even with an improved importance function TP(11), small residual effects can be present (via underestimation of the statistical error, or bias from population size). We also note a large statistical error at $\omega = 2t$ which is maximized at an intermediate value of $\lambda = 0.5$ (or $g = \sqrt{2}t$). Nevertheless, with TP(11), the bias (if any) is smaller than 0.001 t for the Holstein dimer, which highlights the accuracy and sampling efficiency of AFQMC/TP(11). We discuss the issue of bias in AFQMC in sign-problem-free models further in Secs. VI B and VI E.

B. 1D four-site model at half-filling

To further investigate the effect of the importance function on the sampling result and any potential bias, we consider a 1D four-site Holstein model employing periodic boundary conditions (PBCs), at half-filling with $\lambda = 0.5$, $\omega = 4t$, and $g = 2t$. We compute the ground-state energies with AFQMC using the following trial wave functions: a single semiclassical state (S); the TP wave function with 13 semiclassical states [TP(13)]; the LLF wave function; and the TP-LLF wave function with two LLF states [TP-LLF(2)]. With AFQMC/TP-LLF(2), a ground-state energy of $-10.293(2)$ is obtained, compared to the exact result of -10.292 (obtained from DMRG using ITENSOR [95], although ED can also be done here). In contrast, a biased result is seen with each of the other forms of the trial wave function. The bias is about 0.4% relative to the exact result, using the computational parameters specified in Sec. X, and is essentially independent of whether bifurcation is accounted for or not in the trial wave function. These results suggest that, to remove the sampling bias in this parameter regime, it is critical to have in the importance function a means to both overcome the adiabatic potential bifurcation issue and treat el-ph correlation.

C. 1D 20-site model at half-filling

Next, we discuss a 20-site 1D Holstein model at half-filling employing PBCs. ED is no longer feasible for systems of this size so we used the ITENSOR [95] package to perform DMRG calculations [30]. The DMRG calculations were carried out by placing alternating fermionic and bosonic sites on a 1D lattice so that overall there are twice the number of sites compared to the physical lattice problem. While it is possible to use an optimized phonon basis [96] to handle larger el-ph coupling cases, here we employed the most primitive version of DMRG for simple comparisons. The bond dimension we used was fixed at 1000 and the maximum number of bosons for each site was taken to be 60.

We compare the total energy per site within DMRG, CSMP2, variational LF, AFQMC/S, and AFQMC/TP(11) in

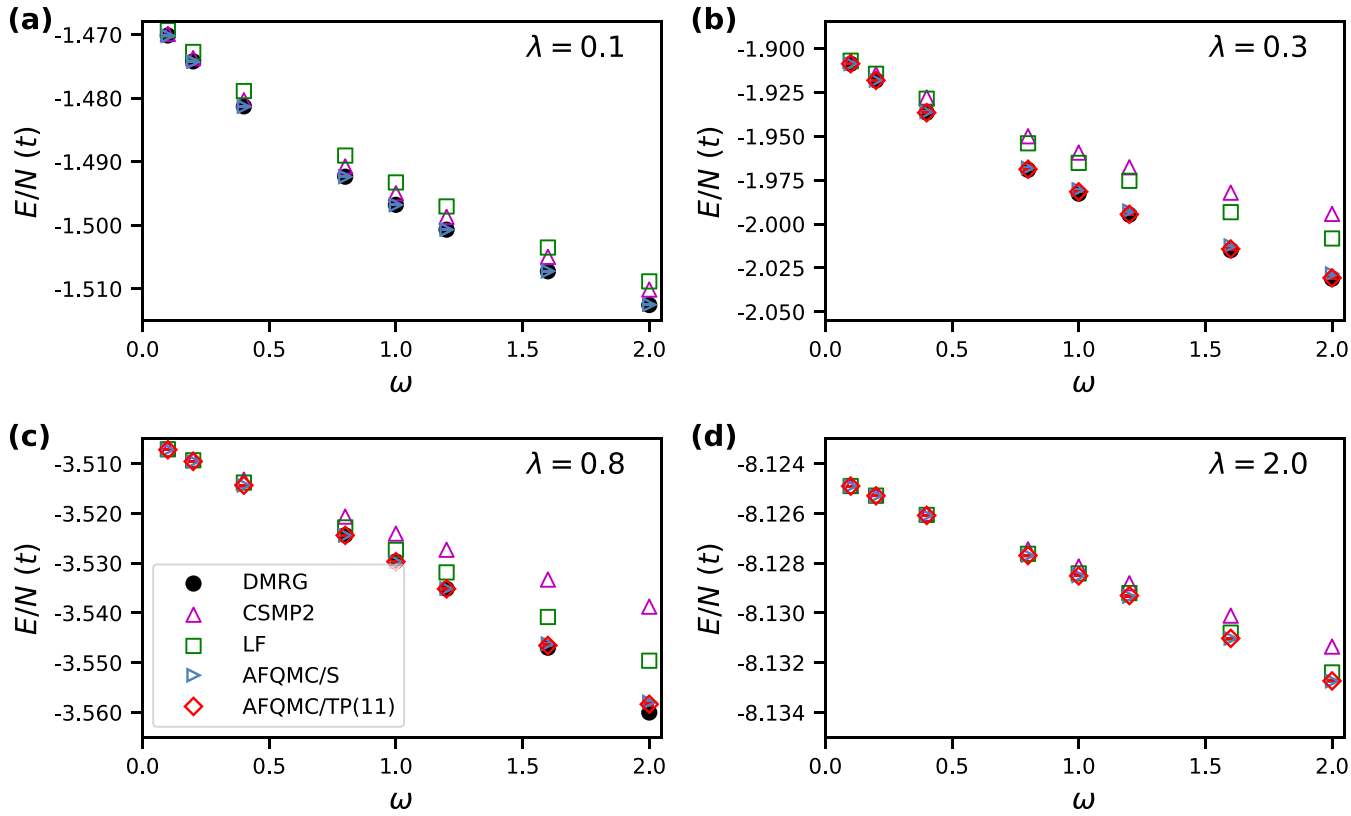

FIG. 3. Total energy per site in units of t for the 20-site 20-electron 1D Holstein model as a function of ω for various values of λ: (a) $\lambda = 0.1$ results, (b) $\lambda = 0.3$ results, (c) $\lambda = 0.8$ results, and (d) $\lambda = 2.0$ results. Note that all error bars for AFQMC are too small to be seen on the plotted scales, DMRG reference values are unavailable for $\omega < 0.8$ when $\lambda = 0.8$ and all values of ω when $\lambda = 2.0$, and AFQMC/TP(11) is not presented for $\lambda = 0.1$ because the results for this trial wave function are nearly identical to those of AFQMC/S. The inset shows energy differences from DMRG for AFQMC/S (blue) and AFQMC/TP(11) (red) ($E_{\mathrm{AFQMC}} - E_{\mathrm{DMRG}}$).

Fig. 3 for various λ and ω. Given the discussion of Sec. VI B, it is desirable to employ the TP-LLF(n) wave functions in general, but we leave a more detailed study with this trial wave function for a future study. Here, we focus on AFQMC with simpler and less accurate trial wave functions [AFQMC/S and AFQMC/TP(11)].

Similarly to the two-site problem, we observe that CSMP2 follows the (near-exact) answers given by DMRG and AFQMC closely at small λ (e.g., $\lambda = 0.1$) as the frequency of the phonon mode is varied. However, we see a clear quantitative deviation of CSMP2 from the other curves as ω increases. The deviation is again maximized at an intermediate coupling $\lambda = 0.8$ and is smaller at weak and strong couplings. The variational LF wave function works better than CSMP2 for all $\lambda > 0.1$. Its strength over CSMP2 is highlighted as ω increases. This clearly suggests that the variational LF wave function includes el-ph correlation beyond the second-order contribution provided in CSMP2.

The performance of AFQMC/S is very good at all coupling strengths considered here for $\omega \leqslant 2.0$. Similar issues with biased final estimates from poor importance functions are seen at $\lambda = 0.3$ and 0.8. AFQMC/TP(11) shows improvement over the simplest semiclassical importance function in the case of $\lambda = 0.3$. However, its improvement for $\lambda = 0.8$ as ω becomes larger is very small. For example, the residual bias is still visible at $\omega = 2.0$. This points to the need to improve the importance function over the forms we have used. We expect that incorporating el-ph correlation directly, as in the LF-type wave functions, will ameliorate this sampling bias greatly, as observed in Sec. VI B. For the rest of the paper, we focus on AFQMC with the simplest trial wave function, namely, a single semiclassical trial wave function, because the observed sampling bias is small enough that it does not affect the conclusions of this work.

D. 2D 4 × 4 model at half-filling

We have established the expected behavior of AFQMC with semiclassical wave functions as importance functions from studying one-dimensional problems such as the Holstein dimer and a 1D chain. Here, we explore higher dimensions by investigating a 2D square lattice problem with a 4×4 geometry. We employed PBCs along the x axis and OBCs along the y axis. The main reason for choosing this boundary condition is to ease the convergence of the DMRG calculations. We were able to converge DMRG calculations only for $0.8 \geqslant \omega$ and $\lambda \leqslant 0.5$ where we used a bond dimension of 2500 and a maximum number of bosons of 25.

In Fig. 4, the energy per site as a function of ω for various λ values is presented for this 2D model. We observe conclusions

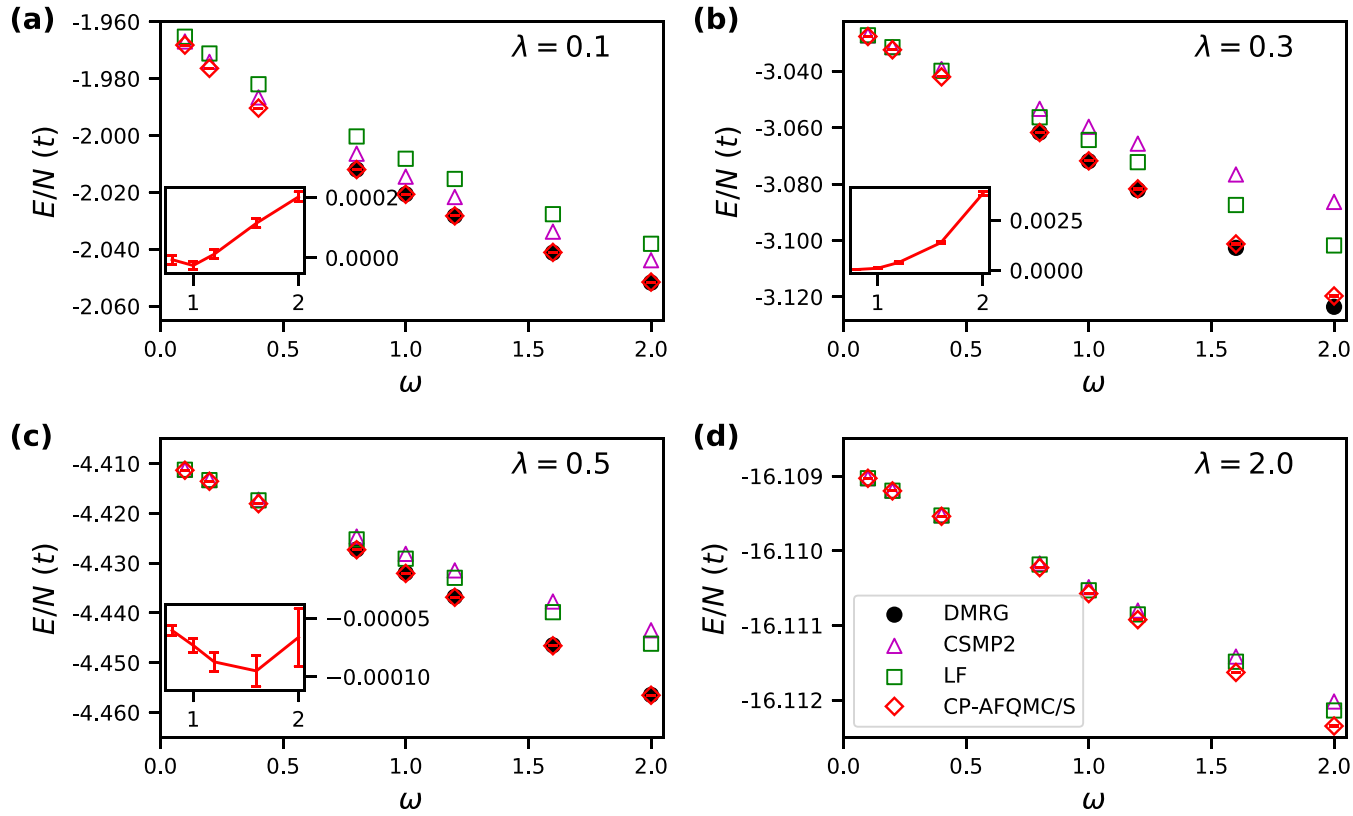

FIG. 4. Total energy per site in units of t for the 4×4 2D Holstein model at half-filling as a function of ω for various λ values: (a) $\lambda = 0.1$ results, (b) $\lambda = 0.3$ results, (c) $\lambda = 0.5$ results, and (d) $\lambda = 2.0$ results. Note that all error bars of CP-AFQMC/S are too small to be seen on the plotted scales, The inset shows energy differences between CP-AFQMC/S and DMRG ($E_{\mathrm{AFQMC}} - E_{\mathrm{DMRG}}$), for $\omega \geqslant 0.8$ when $\lambda < 2.0$ where DRMG results are available.

similar to our previous one-dimensional examples. CSMP2 quantitatively fails as ω increases. Furthermore, for a fixed ω, CSMP2 performs worst for intermediate λ values and is more accurate for small and large λ values. Similarly to the 1D 20-site case, the variational LF energy is more accurate than CSMP2 for $\lambda > 0.1$ and its improvement over CSMP2 becomes larger as ω increases. AFQMC/S is well behaved in the range $\omega \in [0.1, 2.0]$. Its maximum error occurs at $\lambda = 0.3$ and $\omega = 2.0$, where clear indications of sampling bias arises. Nevertheless, the range of parameters where AFQMC/S can be reliably performed with the simplest possible semiclassical trial wave function is quite broad even in 2D, highlighting the utility and potential of this approach.

E. Autocorrelation time and variance control

As we mentioned at the beginning of this section, in the Holstein model the difference between our AFQMC approach and DQMC is mainly in the details of the Monte Carlo sampling algorithm. The two methods can thus have different behavior in terms of efficiency in different regimes of the parameter space. Here, we look into this to help understand the domain of applicability. We note that this is not the focus of our study since in the most general case where electron interactions are present, the branching random-walk approach must be adopted in order to control the sign or phase problem.

The standard DQMC algorithm based on local updates exhibits a long autocorrelation time in the Holstein model for $\omega < 0.5$ and for low temperatures. It has been found that this is a consequence of an ergodicity problem. A careful mathematical analysis of the causes of this problem can be found in the work of Hohenadler and coworkers [69,97,98]. These authors have shown that the condition number of the bosonic action sampled in DQMC for small values of $\Delta\tau$ scales as $1/(\omega\Delta\tau)^2$. This poorly conditioned action leads to a long autocorrelation time that scales quadratically with increasing ω^{-1}. There have been attempts to ameliorate this problem based on global moves such as the Langevin dynamics approach [57,94,99] and the self-learning Monte Carlo approach [100]. We also mention that the work of Hohenadler and coworkers removed the autocorrelation problem using the Lang-Firsov transformation along with a principal component analysis [69].

Since AFQMC applies a projector to the entire set of electronic and phonon degrees of freedom using a population of random walkers, it is less prone to ergodicity problems. The Monte Carlo time coincides with the imaginary-time direction, with open-ended random walks evolving along the world lines, which makes them less likely to become trapped in particular configurations of the phonon paths. There is a deep connection between this fact and the necessity to resort to this sampling approach in order to impose a CP or phaseless gauge condition [51,61]. To quantify this, we directly compute

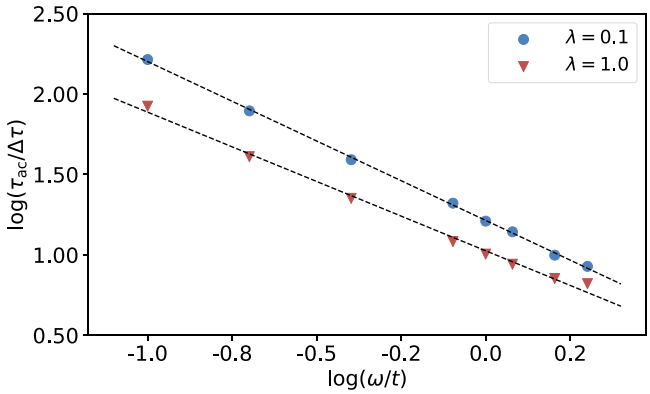

FIG. 5. Log-log plot of autocorrelation time $\tau_{ac}/\Delta\tau$ and phonon frequency ω/t for $\lambda = 0.1$ and 1.0 for the 4×4 2D Holstein model. The black dotted lines are linear fits for each curve. For $\lambda = 0.1$, the slope is -0.9876 with $R^2 = 0.9993$ and for $\lambda = 1.0$, the slope is -0.8617 with $R^2 = 0.9942$.

the energy autocorrelation function

$$c_E(\tau) = \frac{1}{N-\tau}\sum_{n=1}^{N-\tau}(E_n - \langle E\rangle)(E_{n+\tau} - \langle E\rangle), \tag{88}$$

which gives an estimate of the integrated autocorrelation time via

$$\tau_{ac} = \sum_{\tau=-\infty}^{\infty}\frac{c_E(\tau)}{c_E(0)}. \tag{89}$$

The summation in Eq. (89) needs to be performed within some window instead of over the entire set of samples since, for $\tau \gg \tau_{ac}$, the summed noise becomes comparable to the actual signal. We follow Sokal's prescription of the automated windowing procedure to handle this issue [101,102]. In Fig. 5, we present estimates of autocorrelation times for $\lambda = 0.1$ and 1.0 for the 4×4 2D Holstein model. For both values of λ, we observe a near linear behavior in the log-log scale correlation between τ_{ac} and ω as in Fig. 5. Empirically, we find that the autocorrelation time scales as $1/\omega^{0.9876}$ for $\lambda = 0.1$ and $1/\omega^{0.8617}$ for $\lambda = 1.0$. This scaling is a significant improvement over that of the standard DQMC algorithms where τ_{ac} scales as $1/\omega^2$ [69].

On the other hand, in AFQMC we use an importance function to guide the random walks. If the quality of the importance function is very poor, the variance can grow and even become infinite [88]. In cases where the trial wave function suppresses certain regions of the Hilbert space being sampled with a qualitatively incorrect functional form, the autocorrelation time and thus the variance can diverge, as mentioned earlier in this section. This situation was seen in the examples with the semiclassical wave function where there is a strong bifurcation of the adiabatic potential in the Holstein dimer. Another example occurs with the use of semiclassical trial wave functions where the lack of explicit el-ph correlation leads to large sampling biases. In extreme cases, calculations will be seemingly well behaved in "normal-sized" runs, as the Monte Carlo sampling is strongly biased by the wrong importance function and the autocorrelation time grows exponentially. These situations require careful analysis of the variance and study of the dependence on the details of the importance function to reveal the problem [88]. Separate but related to the quality of the importance function is the issue of efficiently sampling of multimodal landscapes in the el-ph models, as we have only incorporated local moves in our random walks. In the worst cases, AFQMC can, even with reasonable choices of importance functions, experience difficulties with long autocorrelations as occurs in DQMC. In AFQMC the use of a population of open-ended random walkers with branching can help avoid the sampling being stuck.

VII. HUBBARD-HOLSTEIN MODEL

The focus of our CP-AFQMC method is on doped systems and more realistic Hamiltonians, where the sign problem or phase problem will be present. In the previous section, we studied the 1D and 2D Holstein models with simple trial wave functions. In this section, we present benchmark data using the same approach for the Hubbard-Holstein model with $U/t = 4$. Because of the competition between U and g, we carefully study trial wave functions with both CDW and SDW order. Note that the CP-AFQMC algorithm is no longer exact because the el-el repulsion will lead to a sign problem. Karakuzu, Seki, and Sorella have presented an efficient QMC algorithm which is free of the sign problem at half-filling as long as $U > 2g^2/\omega$ [57]. Within CP-AFQMC, we can simulate any parameter regime efficiently by controlling the sign problem at the expense of introducing a constraint bias. Since CP-AFQMC has been extensively benchmarked for electronic systems in the past [78], we focus on any additional biases that may arise from the interplay between electrons and phonons in this section. We use an *ad hoc* $U_{\text{eff}} = 0.5\,t$ in the electronic mean-field part to generate all the SDW trial wave functions for CP-AFQMC in this section.

Below, we first examine the behavior of our algorithm in 1D, and then in 2D both at half-filling and $\frac{1}{8}$ doping. We will focus on benchmarking the accuracy of the computed ground-state energies. We note that for the purely electronic cases with $\lambda = 0$, all of our models in 1D and at half-filling in 2D are sign-problem free. However, CP-AFQMC can incur a systematic error in the energy in these cases because of an "artificial node" in auxiliary-field space [51]; this error can be removed straightforwardly [77,78]. Instead of invoking the scheme to remove this artificial bias, we will perform the CP-AFQMC calculation in the generic way as described above since our main focus in this work is the most general situation of a doped Hubbard-Holstein model where the sign problem is present.

A. 1D 20-site model at half-filling

We benchmark CP-AFQMC against DMRG for the 20-site 1D Hubbard-Holstein model at half-filling with PBCs. Unlike for the case of the pure Holstein model, CSMP2 and the variational LF approach are quantitatively and qualitatively inaccurate for all parameters examined here. This is not surprising because the onsite repulsion term for $U/t = 4$ is not small, so the failure of mean-field theories and a low-order perturbation theory on the el-el interaction is expected. For

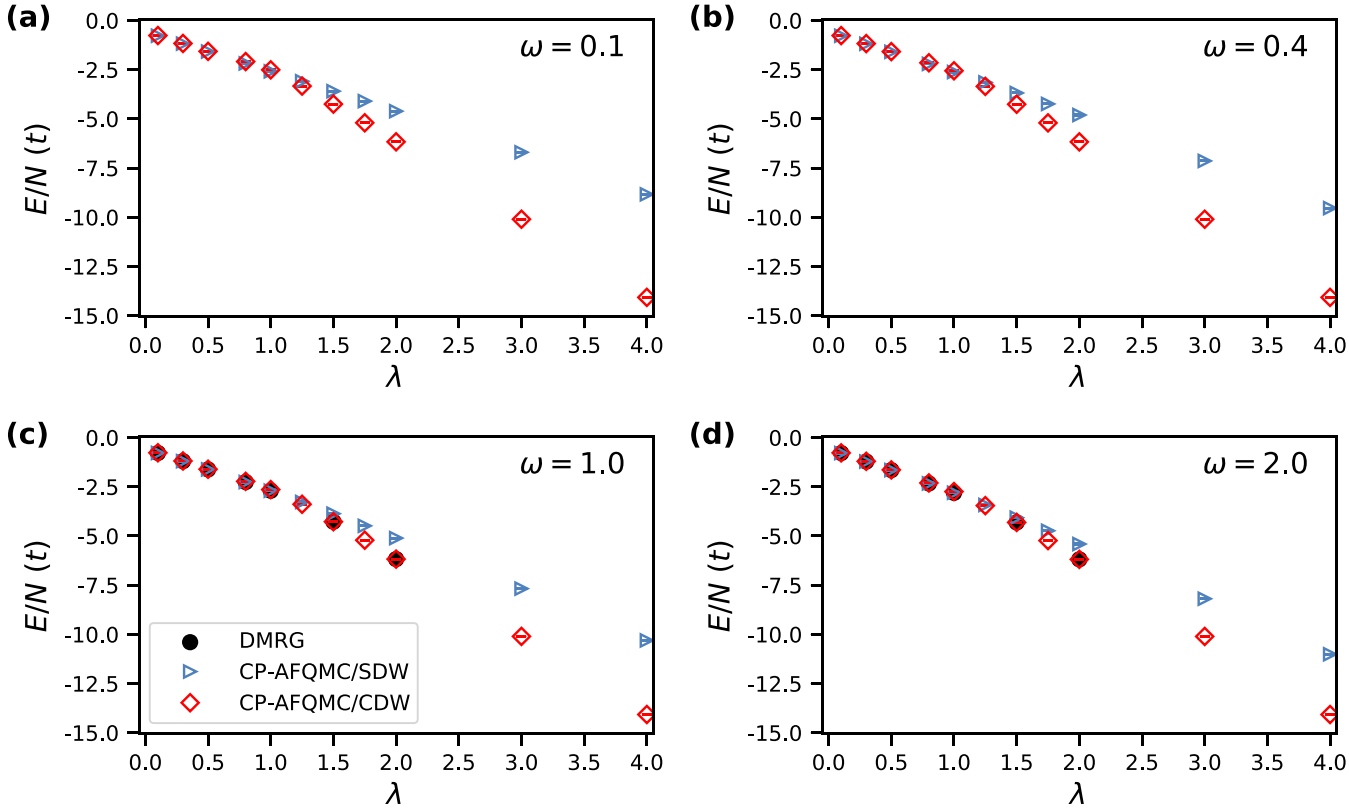

FIG. 6. The total energy per site in units of t for the 20-site 1D Hubbard-Holstein model at half-filling with $U = 4t$ as a function of λ for various ω values: (a) $\omega = 0.1$, (b) $\omega = 0.4$, (c) $\omega = 1.0$, and (d) $\omega = 2.0$. Note that all error bars of CP-AFQMC are too small to be seen on the plotted scales. Note that the DMRG results are unavailable for $\omega = 0.1$ and 0.4 as well as for $\lambda > 2$.

this reason we do not discuss CSMP2 and variational LF results here.

To study the CDW and SDW phases and the possibility of a phase transition between them, we carry out CP-AFQMC calculations using two different mean-field trial wave functions with the corresponding broken symmetry. Comparison of the computed energies indicates which one is the ground state at each Hamiltonian parameter choice as well as the existence and location of a transition, although the fact that our CP-AFQMC energies computed from the mixed estimate are not variational [63] adds a subtlety to this procedure. Here, the calculation leading to the higher energy can be thought of as the constraint acting to "hold" the projection to an excited state compatible with the broken symmetry of the trial wave function. In actual applications, we could use a self-consistent CP-AFQMC procedure [90] to tune the trial wave function and reduce its effect on the result, but for the purpose of benchmarks we will only perform one-shot calculations here using UHF trial wave functions generated with a fixed U_{eff}, and rely on comparison with DMRG results to gauge the accuracy.

In Fig. 6, we compare two sets of CP-AFQMC results at different ω values, one set with SDW trial wave functions (denoted by CP-AFQMC/SDW) and another with CDW trial wave functions (denoted CP-AFQMC/CDW). The trial wave functions themselves show a SDW to CDW transition at $\lambda \sim 0.5$, given our *ad hoc* choice of $U_{\text{eff}} = 0.5$. We see that CP-AFQMC/CDW leads to higher energies than CP-AFQMC/SDW for $\lambda \leqslant 1.0$. The energy differences are large enough to make it straightforward to identify the correct phase.

We next make more quantitative comparison of the total energy per site obtained from CP-AFQMC and DMRG in Fig. 6. DMRG calculations are performed with a bond dimension of 1000 and with the maximum number of bosons of 40. In Figs. 6(c) and 6(d), we observe that CP-AFQMC/SDW closely follows the DMRG energies from $\lambda = 0.1$ to 1.0. At $\lambda = 2.0$, the energy obtained from CP-AFQMC/SDW is significantly higher than that from CP-AFQMC/CDW, with the latter in good agreement with DMRG. The procedure described above of combining the lowest-energy curves between CP-AFQMC/SDW and CP-AFQMC/CDW produces quantitatively accurate results across the full range of parameters that we study. Variations in the value of the phonon frequency do not change the qualitative conclusions.

In Fig. 7 we show a magnified view of the absolute discrepancies between CP-AFQMC and DMRG energies. As a comparison, for the purely electronic Hubbard model ($\lambda = 0$), CP-AFQMC exhibits an error per site of $0.002\,45(8)t$ with respect to the DMRG reference values. Similarly in the Hubbard-Holstein model, CP-AFQMC/SDW energies exhibit an error per site of approximately 0.002–0.003t for $\lambda = 0.1, 0.3, 0.8$. When the system reaches values as large as $\lambda = 1.0$, we observe a small increase in the error as ω

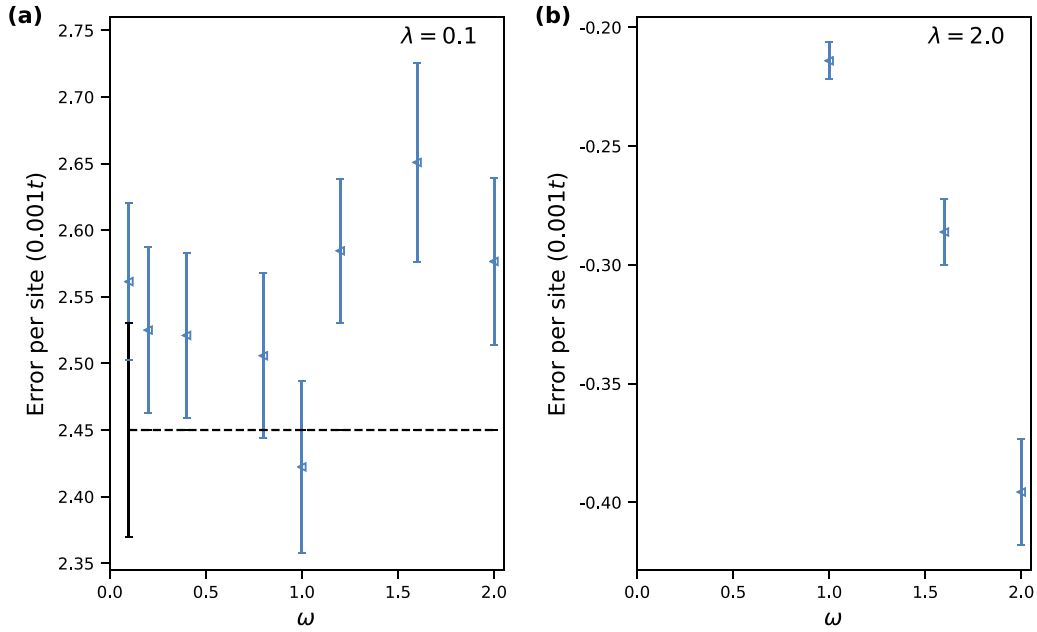

FIG. 7. Error in total energy per site compared to DMRG in units of $0.001t$ for the 20-site 1D Hubbard-Holstein model at half-filling with $U/t = 4$ as a function of ω for various λ values: (a) $\lambda = 0.1$ results and (b) $\lambda = 1.0$ results. The black dotted line in (a) indicates the electronic CP-AFQMC energy error (i.e., $\lambda = 0.0$). In (b), DMRG energies are only available for $\omega = 1.0, 1.5, 2.0$.

increases. The largest error found for $\lambda = 1.0$ is $0.0064(2)t$ at $\omega = 2.0$, which is slightly larger than the constraint bias found in the purely electronic problem. The point at which the largest error was observed coincides with the expected phase transition point between SDW and CDW (see below). At $\lambda = 2.0$, the discrepancy between CP-AFQMC and DMRG is an order of magnitude smaller, with a maximum deviation of CP-AFQMC per site of $-0.000\,40(2)t$ at $\omega = 2.0$.

It has been shown by several methods [37,46,57,58] that in the thermodynamic limit the Hubbard-Holstein model undergoes a transition between SDW and CDW at

$$U \approx \frac{2g^2}{\omega} = 4dt\lambda. \tag{90}$$

This value of U is where the effective onsite interaction changes sign as shown in Eq. (81). For $U/t = 4$ and $d = 1$, we expect the phase transition to occur at approximately $\lambda = 1.0$. We find that in the 20-site model, despite the expected finite-size effects, the onset of the phase transition is captured quite well. In particular, the crossover between CP-AFQMC/SDW and CP-AFQMC/CDW occurs roughly at $\lambda = 1.0$ in Fig. 6. While this is encouraging, detailed phase diagram studies with CP-AFQMC should be carried out in the future. We note that CP-AFQMC often restores the symmetry breaking of the underlying mean-field trial wave function [103,104], as would be expected of an exact many-body computation. Therefore, a proper phase diagram study with CP-AFQMC should involve a direct measurement of correlation functions [105] or order parameters with explicit symmetry breaking induced [27]. Furthermore, there may be intermediate phases such as metallic or superconducting phases near the onset of the phase transition between the SDW and CDW phases. Studying these putative intermediate phases is of great interest [58].

B. 2D 4 × 4 model at half-filling and $\frac{1}{8}$ hole-doping

Instead of comparing CP-AFQMC with other methods for 2D Hubbard-Holstein systems, we simply report the computed total energy per site using PBCs along both the x and y directions, as shown in Fig. 8. Based on the benchmark studies in the previous sections and on experience from the purely electronic model, we expect that our results will be of similar accuracy (or better because of effective reduction of the el-el interaction from the el-ph coupling) to that in the Hubbard model for most parameters considered in this work. In Fig. 8, we see that CP-AFQMC/CDW has a lower energy for $\lambda > 0.5$ at both frequencies ($\omega = 0.1$ and 2.0) at both half-filling and $\frac{1}{8}$ hole doping. Based on Eq. (90), it is expected that the onset of the crossover occurs around $\lambda = 0.5$ in 2D, consistent with our numerical results. (Note that our definition of λ includes dimensionality, hence the change in the crossover value from 1D to 2D.)

Consistent with our previous results, for $\lambda = 0.1$ and 0.3 at all frequencies up to $\omega = 2.0$, the CP-AFQMC error per site is approximately $0.009t$ or slightly larger. At the onset of the crossover between CP-AFQMC/CDW and CP-AFQMC/SDW ($\lambda \sim 0.5$), we expect the error to be maximized and larger than that of the CP-AFQMC bias for the electronic problem, similar to the 1D 20-site model at half-filling. For $\lambda = 0.8$ and 2.0, we expect that our results will be nearly exact for the ω values studied. For the purely electronic Hubbard model ($\lambda = 0$), CP-AFQMC exhibits an error per site of about $0.009\,01(9)t$ at half-filling and $0.004\,69(4)t$ at $\frac{1}{8}$ hole doping, using a UHF trial wave function. As mentioned, the error at half-filling is "artificial" and can be removed [77,78], but this is not done here. For $\lambda \leqslant 1$ at half-filling we expect an error of comparable size. Comparing two different fillings, we do not see qualitative differences in physical

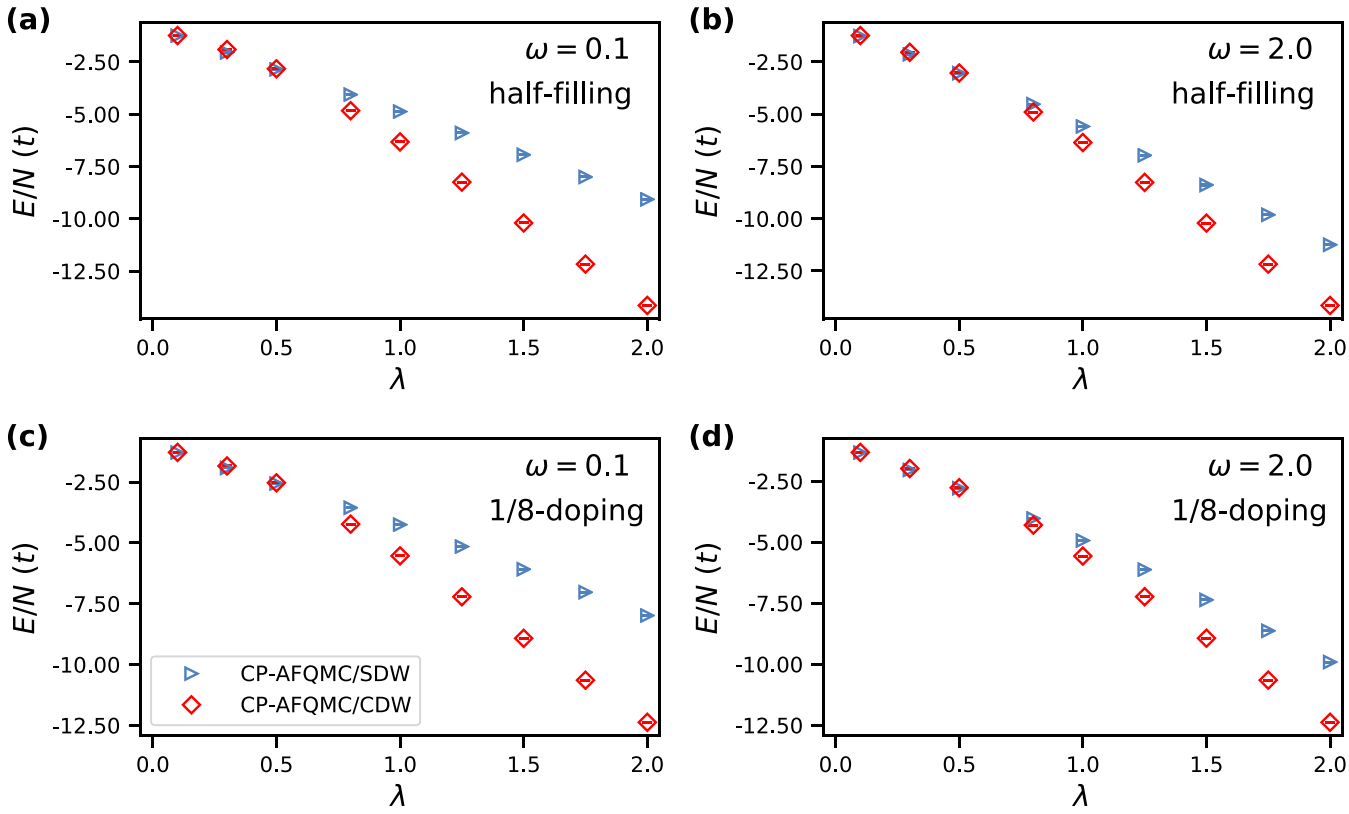

FIG. 8. Total energy per site in units of t for the 4×4 2D Hubbard-Holstein model with $U/t = 4$ as a function of ω: (a) $\omega = 0.1$ results at half-filling, (b) $\omega = 2.0$ results at half-filling, (c) $\omega = 0.1$ results at $\frac{1}{8}$ hole doping, and (d) $\omega = 2.0$ results at $\frac{1}{8}$ hole doping. Note that all error bars of CP-AFQMC are too small to be seen on the plotted scales.

behavior in our finite-sized lattice, and the value of the phonon frequency does not appear to make qualitative differences as well. We note that the energy difference between CP-AFQMC/SDW and CP-AFQMC/CDW noticeably shrinks as the phonon frequency ω increases.

VIII. TOWARDS *AB INITIO* HAMILTONIANS

We briefly discuss the extension of the presented algorithm for general *ab initio* Hamiltonians. The *ab initio* Hamiltonian that describes el-ph problems typically involves linear el-ph coupling. Therefore, the most widely used *ab initio* Hamiltonian has the same form as Eq. (1) with more general Hamiltonian matrix elements:

$$\hat{\mathcal{H}}_{\text{el}}^{(1)} = \sum_{\sigma\in\{\uparrow,\downarrow\}} \sum_{pq} h_{p_\sigma q_\sigma} \hat{a}_{p_\sigma}^\dagger \hat{a}_{q_\sigma}, \tag{91}$$

$$\hat{\mathcal{H}}_{\text{el}}^{(2)} = \frac{1}{2} \sum_{\sigma,\sigma'\in\{\uparrow,\downarrow\}} \sum_{pqrs} (p_\sigma r_\sigma | q_{\sigma'} s_{\sigma'}) \hat{a}_{p_\sigma}^\dagger \hat{a}_{q_{\sigma'}}^\dagger \hat{a}_{s_{\sigma'}} \hat{a}_{r_\sigma}, \tag{92}$$

$$\hat{\mathcal{H}}_{\text{ph}} = \sum_I \omega_I \hat{b}_I^\dagger \hat{b}_I, \tag{93}$$

and

$$\hat{\mathcal{H}}_{\text{el-ph}} = \sum_{\sigma\in\{\uparrow,\downarrow\}} \sum_{pqI} g_{p_\sigma q_\sigma I} \hat{a}_{p_\sigma}^\dagger \hat{a}_{q_\sigma} (\hat{b}_I^\dagger + \hat{b}_I), \tag{94}$$

where we have suppressed other quantum numbers such as **k**-point dependencies and have expressed everything in terms of the electronic ($\{p, q, r, s, \dots\}$) and phononic bands ($\{I, J, K, L, \dots\}$). The computation of these matrix elements at the level of density functional theory has been well documented [1,106,107] so here we focus on briefly describing the *phaseless* AFQMC (ph-AFQMC) algorithm [61] for these realistic el-ph problems.

The walkers take the same form as in Eq. (19). Therefore, the essence of the propagation algorithm remains unchanged. The only complication arises from the generalized form of $\hat{\mathcal{H}}_{\text{el}}^{(2)}$ which necessitates the use of a continuous Hubbard-Stratonovich transformation [108,109]. The continuous transformation leads to the fermionic phase problem which can be removed via the phaseless constraint [61]. The propagation is carried out the same way with appropriate modifications to the constraint to account for the phase problem. The *ab initio* generalization of semiclassical states used in this work is also straightforward. The trial wave function still takes the form of Eqs. (34), (35), and (37). A variational minimization of the total energy then leads to a trial wave function that can be used in ph-AFQMC. The projection of the trial wave function onto phonon displacements $|\mathbf{X}\rangle$ is identical to Eq. (39) except that the phonon mass and frequency now depend on band indices $\{I\}$. The *ab initio* generalization of the LF wave function may also be carried out straightforwardly by extending the LF generator in Eq. (49). We expect that

the *ab initio* ph-AFQMC approach will become a valuable tool for understanding polaronic physics in realistic correlated materials in the future.

IX. CONCLUSIONS AND OUTLOOK

In this work, we have introduced an extension of CP-AFQMC to describe correlated systems with el-ph coupling. Our approach utilizes a mixed first- and second-quantized representation where the phonons are propagated in first quantization following the commonly used diffusion MC algorithm, and the electronic degrees of freedom are handled in second quantization via AFQMC. The resulting algorithm is compared with numerically exact DMRG and low-order perturbation theories for the Holstein model as a first test of the basic algorithm. We have demonstrated that the autocorrelation time problems that arise in the commonly used DQMC methods are greatly ameliorated in AFQMC, with autocorrelation time that scales roughly as $1/\omega$.

While the Holstein model is sign-problem free, AFQMC with the simplest trial wave functions, namely, semiclassical states, is found to introduce a small bias when the underlying adiabatic surface develops into a double-well potential surface and/or g is larger than t but smaller than ω (e.g., $\omega \to \infty$ for a fixed λ). Based on a four-site model, we have shown that this bias can be removed by using an improved trial wave function where both bifurcations and increased el-ph correlation are accounted for. We have demonstrated the remarkable accuracy of AFQMC for both 1D and 2D Holstein models over a reasonably broad set of coupling and phonon frequency parameters via direct comparison with DMRG.

We have tested CP-AFQMC on the finite-sized versions of the 1D and 2D Hubbard-Holstein models with $U/t = 4$, using the simplest form of trial wave functions consisting of a semiclassical state with a single Slater determinant. For the 1D Hubbard-Holstein model, we have compared CP-AFQMC against numerically exact DMRG results. When λ is small and the ground state is dominated by the Hubbard U term, we find that the error of our algorithm is roughly the same as that expected from standard CP-AFQMC applied to the purely electronic Hubbard model. Furthermore, when the ground state is dominated by the el-ph coupling term and exhibits charge density wave order, we find that the overall error becomes remarkably small (smaller than that expected in purely electronic systems). These facts have motivated the production of what we believe are benchmark results for the finite-sized 2D Hubbard-Holstein model for various values of λ and ω at half-filling and $\frac{1}{8}$ hole doping.

For $U < 2g^2/\omega$ at half-filling and for all parameter regimes at any hole doping, standard QMC approaches suffer from the sign problem [57]. Therefore, our AFQMC approach should become an essential tool for producing accurate results scalable to large system sizes for this model. We have investigated the competing spin and charge density wave orders in the Hubbard-Holstein model. At the onset of the phase transition between these phases, we observe a crossover in the energies between two AFQMC calculations targeting the two phases. Lastly, we have briefly discussed the extension of the presented algorithm to *ab initio* Hamiltonians that can be easily formulated based on the phaseless AFQMC method for general electronic Hamiltonians [61].

Some immediate future directions include using this framework to provide a detailed study of the phase diagram of the Hubbard-Holstein model and other lattice models, and extending this framework to finite-temperature problems based on the constrained-path approximation [110]. As mentioned, a trial wave function with an el-ph Jastrow factor can be implemented straightforwardly in AFQMC, which is expected to further reduce the bias and improve sampling efficiency in antiadiabatic regimes with $2dt\lambda < \omega$. It will also be valuable to further investigate the implementation of the full LF trial wave function. Furthermore, application of the proposed AFQMC approach to *ab initio* systems will be of great interest as well. While there are several algorithmic aspects that can be further improved, including improved forms of importance functions and better sampling in large phonon frequency regimes, we believe that the algorithms and insights presented in this work will serve as stepping stones towards simulating model as well as *ab initio* systems with a nontrivial interplay between electronic correlation and el-ph couplings, which continue to be of great importance in modern condensed matter physics.

X. COMPUTATIONAL DETAILS

Our algorithm was implemented in a public open-source auxiliary-field quantum Monte Carlo package called PAUXY [111]. The blocking analysis was performed with PYBLOCK [112]. The pair-branching algorithm was used for population control [113]. Variational calculations were aided by automatic differentiation using JAX [114]. A total of 640 walkers and a time step of $0.005t^{-1}$ were used in all calculations except for the Holstein dimer and the four-site 1D Holstein model. For the Holstein dimer, we used a time step of $0.0005t^{-1}$ for $\omega \leqslant 1.6$ and $0.000\,25t^{-1}$ for $\omega > 1.6$ with 6400 walkers. For the four-site 1D Holstein model, we used 6400 walkers and a time step of $0.0005t^{-1}$. The population control bias and time-step error were found to be smaller than $0.001t$ in the absolute total energy per site. In calculations for $\omega = 0.1$, we used a very long β ($\sim 3000t^{-1}$) due to the long autocorrelation time in this regime. This yielded about 100 statistically independent samples after the blocking analysis. For the DMRG calculations, we gradually increased the number of bosons on each site to make sure that the reported DMRG energies are converged up to $0.001t$ per particle in that increasing the number of bosons by 10 showed an energy change smaller than $0.001t$ per particle. This is an energy scale that is small enough for the discussion in this work.

ACKNOWLEDGMENTS

We are grateful to Z. Li and H. Shi for their contributions during early stages of this work. We thank H. Shi, M. Lindsey, and F. Malone for fruitful discussions and M. Stoudenmire for help with ITENSOR calculations. D.R.R. thanks the National Science Foundation for funding under Grants No. CHE-1954791. This work was conducted using computational resources and services at the Flatiron Institute. The Flatiron Institute is a division of the Simons Foundation.

[1] F. Giustino, Rev. Mod. Phys. **89**, 015003 (2017).
[2] G. D. Mahan, *Many-particle Physics* (Springer, New York, 2013).
[3] R. Cava, R. Van Dover, B. Batlogg, and E. Rietman, Phys. Rev. Lett. **58**, 408 (1987).
[4] M.-K. Wu, J. R. Ashburn, C. Torng, P. H. Hor, R. L. Meng, L. Gao, Z. J. Huang, Y. Wang, and A. Chu, Phys. Rev. Lett. **58**, 908 (1987).
[5] D. J. Scalapino, Rev. Mod. Phys. **84**, 1383 (2012).
[6] K. A. Müller, Z. Phys. B **80**, 193 (1990).
[7] J. Song and J. F. Annett, Phys. Rev. B **51**, 3840 (1995).
[8] A. S. Alexandrov, J. Supercond. **13**, 985 (2000).
[9] T. Cuk, D. H. Lu, X. J. Zhou, Z. X. Shen, T. P. Devereaux, and N. Nagaosa, Phys. Status Solidi B **242**, 11(R) (2005).
[10] E. Khatami, A. MacRidin, and M. Jarrell, Phys. Rev. B **78**, 060502(R) (2008).
[11] O. Gunnarsson and O. Rösch, J. Phys.: Condens. Matter **20**, 043201 (2008).
[12] E. Van Heumen, E. Muhlethaler, A. B. Kuzmenko, H. Eisaki, W. Meevasana, M. Greven, and D. Van Der Marel, Phys. Rev. B **79**, 184512 (2009).
[13] C. Gadermaier, A. S. Alexandrov, V. V. Kabanov, P. Kusar, T. Mertelj, X. Yao, C. Manzoni, D. Brida, G. Cerullo, and D. Mihailovic, Phys. Rev. Lett. **105**, 257001 (2010).
[14] Y. He, M. Hashimoto, D. Song, S. D. Chen, J. He, I. M. Vishik, B. Moritz, D. H. Lee, N. Nagaosa, J. Zaanen, T. P. Devereaux, Y. Yoshida, H. Eisaki, D. H. Lu, and Z. X. Shen, Science **362**, 62 (2018).
[15] S. L. Yang, J. A. Sobota, Y. He, D. Leuenberger, H. Soifer, H. Eisaki, P. S. Kirchmann, and Z. X. Shen, Phys. Rev. Lett. **122**, 176403 (2019).
[16] Y. Hu, X. Chen, S. T. Peng, C. Lane, M. Matzelle, Z. L. Sun, M. Hashimoto, D. H. Lu, E. F. Schwier, M. Arita, T. Wu, R. S. Markiewicz, K. Shimada, X. H. Chen, Z. X. Shen, A. Bansil, S. D. Wilson, and J. F. He, Phys. Rev. Lett. **123**, 216402 (2019).
[17] B. Rosenstein and B. Y. Shapiro, Phys. Rev. B **100**, 054514 (2019).
[18] G. Grissonnanche, S. Thériault, A. Gourgout, M. E. Boulanger, E. Lefrançois, A. Ataei, F. Laliberté, M. Dion, J. S. Zhou, S. Pyon, T. Takayama, H. Takagi, N. Doiron-Leyraud, and L. Taillefer, Nat. Phys. **16**, 1108 (2020).
[19] B. Rosenstein and B. Y. Shapiro, arXiv:2003.03521.
[20] S. Banerjee, W. A. Atkinson, and A. P. Kampf, arXiv:2008.01401.
[21] E. I. Shneyder, S. V. Nikolaev, M. V. Zotova, R. A. Kaldin, and S. G. Ovchinnikov, Phys. Rev. B **101**, 235114 (2020).
[22] A. S. Mishchenko, N. Nagaosa, and N. Prokof'ev, arXiv:2007.09888.
[23] S. A. Sreedhar, A. Rossi, J. Nayak, Z. Anderson, Y. Tang, B. Gregory, M. Hashimoto, D. H. Lu, E. Rotenberg, R. J. Birgeneau, M. Greven, M. Yi, and I. M. Vishik, Phys. Rev. B **102**, 205109 (2020).
[24] Y. Y. Peng, A. A. Husain, M. Mitrano, S. X.-L. Sun, T. A. Johnson, A. V. Zakrzewski, G. J. MacDougall, A. Barbour, I. Jarrige, V. Bisogni, and P. Abbamonte, Phys. Rev. Lett. **125**, 097002 (2020).
[25] J. Hubbard, Proc. R. Soc. A **276**, 238 (1963).
[26] P. W. Anderson, Science **235**, 1196 (1987).
[27] M. Qin, C.-M. Chung, H. Shi, E. Vitali, C. Hubig, U. Schollwöck, S. R. White, S. Zhang *et al.*, Phys. Rev. X **10**, 031016 (2020).
[28] 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.*, Phys. Rev. X **5**, 041041 (2015).
[29] 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).
[30] E. Jeckelmann and S. R. White, Phys. Rev. B **57**, 6376 (1998).
[31] E. Jeckelmann, C. Zhang, and S. R. White, Phys. Rev. B **60**, 7950 (1999).
[32] M. Tezuka, R. Arita, and H. Aoki, Physica B (Amsterdam) **359-361**, 708 (2005).
[33] M. Tezuka, R. Arita, and H. Aoki, Phys. Rev. B **76**, 155114 (2007).
[34] H. Fehske, G. Hager, and E. Jeckelmann, Europhys. Lett. **84**, 57001 (2008).
[35] S. Ejima and H. Fehske, J. Phys.: Conf. Ser **200**, 012031 (2010).
[36] J. Bonča, S. A. Trugman, and I. Batistić, Phys. Rev. B **60**, 1633 (1999).
[37] T. Ohgoe and M. Imada, Phys. Rev. B **89**, 195139 (2014).
[38] T. Ohgoe and M. Imada, Phys. Rev. Lett. **119**, 197001 (2017).
[39] S. Karakuzu, L. F. Tocchio, S. Sorella, and F. Becca, Phys. Rev. B **96**, 205145 (2017).
[40] G. S. Jeon, T. H. Park, J. H. Han, H. C. Lee, and H. Y. Choi, Phys. Rev. B **70**, 125114 (2004).
[41] P. Paci, M. Capone, E. Cappelluti, S. Ciuchi, and C. Grimaldi, Phys. Rev. B **74**, 205108 (2006).
[42] P. Werner and A. J. Millis, Phys. Rev. Lett. **99**, 126405 (2007).
[43] Y. Murakami, P. Werner, N. Tsuji, and H. Aoki, Phys. Rev. B **88**, 125126 (2013).
[44] S. Li, E. Khatami, and S. Johnston, Phys. Rev. B **95**, 121112(R) (2017).
[45] B. Sandhoefer and G. K. L. Chan, Phys. Rev. B **94**, 085115 (2016).
[46] T. E. Reinhard, U. Mordovina, C. Hubig, J. S. Kretchmer, U. Schollwöck, H. Appel, M. A. Sentef, and A. Rubio, J. Chem. Theory Comput. **15**, 2221 (2019).
[47] M. Sibaev, I. Polyak, F. R. Manby, and P. J. Knowles, J. Chem. Phys. **153**, 124102 (2020).
[48] T. Dresselhaus, C. Bungey, P. J. Knowles, and F. R. Manby, J. Chem. Phys. **153**, 214114 (2020).
[49] A. F. White, Y. Gao, A. J. Minnich, and G. K. Chan, J. Chem. Phys. **153**, 224112 (2020).
[50] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. Lett. **74**, 3652 (1995).
[51] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. B **55**, 7464 (1997).
[52] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278 (1981).
[53] D. J. Scalapino and R. L. Sugar, Phys. Rev. Lett. **46**, 519 (1981).
[54] D. J. Scalapino and R. L. Sugar, Phys. Rev. B **24**, 4295 (1981).

[55] S. Johnston, E. A. Nowadnick, Y. F. Kung, B. Moritz, R. T. Scalettar, and T. P. Devereaux, Phys. Rev. B **87**, 235133 (2013).

[56] C. B. Mendl, E. A. Nowadnick, E. W. Huang, S. Johnston, B. Moritz, and T. P. Devereaux, Phys. Rev. B **96**, 205141 (2017).

[57] S. Karakuzu, K. Seki, and S. Sorella, Phys. Rev. B **98**, 201108(R) (2018).

[58] N. C. Costa, K. Seki, S. Yunoki, and S. Sorella, Commun. Phys. **3**, 1 (2020).

[59] N. C. Costa, K. Seki, and S. Sorella, arXiv:2009.05586.

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

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

[62] W. Al-Saidi, S. Zhang, and H. Krakauer, J. Chem. Phys. **124**, 224101 (2006).

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

[64] R. H. McKenzie, C. J. Hamer, and D. W. Murray, Phys. Rev. B **53**, 9676 (1996).

[65] T. Holstein, Ann. Phys. **8**, 325 (1959).

[66] Y. Wang, I. Esterlis, T. Shi, J. I. Cirac, and E. Demler, Phys. Rev. Research **2**, 043258 (2020).

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

[68] R. Blankenbecler and R. L. Sugar, Phys. Rev. D **27**, 1304 (1983).

[69] M. Hohenadler, H. G. Evertz, and W. von der Linden, Phys. Rev. B **69**, 024301 (2004).

[70] D. J. Thouless, Nucl. Phys. **21**, 225 (1960).

[71] C. Umrigar, M. Nightingale, and K. Runge, J. Chem. Phys. **99**, 2865 (1993).

[72] B. L. Hammond, W. A. Lester, and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, Vol. 1 (World Scientific, Singapore, 1994).

[73] J. W. Moskowitz, K. Schmidt, M. A. Lee, and M. H. Kalos, J. Chem. Phys. **77**, 349 (1982).

[74] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, Jr, J. Chem. Phys. **77**, 5593 (1982).

[75] G. Kalosakas, S. Aubry, and G. P. Tsironis, Phys. Rev. B **58**, 3094 (1998).

[76] A. H. Romero, D. W. Brown, and K. Lindenberg, J. Chem. Phys. **109**, 6540 (1998).

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

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

[79] D. W. Small, E. J. Sundstrom, and M. Head-Gordon, J. Chem. Phys. **142**, 024104 (2015).

[80] J. Lee and M. Head-Gordon, J. Chem. Phys. **150**, 244106 (2019).

[81] J. Lee, L. W. Bertels, D. W. Small, and M. Head-Gordon, Phys. Rev. Lett. **123**, 113001 (2019).

[82] L. Proville and S. Aubry, Eur. Phys. J. B **15**, 405 (2000).

[83] T. Holstein, Ann. Phys. **8**, 343 (1959).

[84] I. G. Lang and Y. A. Firsov, Zh. Eksp. Teor. Fiz. **43**, 1843 (1962).

[85] R. Silbey and R. A. Harris, J. Chem. Phys. **80**, 2615 (1984).

[86] R. A. Harris and R. Silbey, J. Chem. Phys. **83**, 1069 (1985).

[87] V. Pouthier, J. Chem. Phys. **138**, 044108 (2013).

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

[89] H. Shi and S. Zhang, Phys. Rev. B **95**, 045144 (2017).

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

[91] I. Shavitt and R. J. Bartlett, *Many-body Methods in Chemistry and Physics: MBPT and Coupled-cluster Theory* (Cambridge University Press, Cambridge, 2009).

[92] J. Bonča, T. Katrasnik, and S. A. Trugman, Phys. Rev. Lett. **84**, 3153 (2000).

[93] J. Bonča and S. Trugman, J. Supercond. **13**, 999 (2000).

[94] G. G. Batrouni and R. T. Scalettar, Phys. Rev. B **99**, 035114 (2019).

[95] M. Fishman, S. R. White, and E. M. Stoudenmire, arXiv:2007.14822.

[96] C. Zhang, E. Jeckelmann, and S. R. White, Phys. Rev. Lett. **80**, 2661 (1998).

[97] M. Hohenadler, H. G. Evertz, and W. Von Der Linden, Phys. Status Solidi B **242**, 1406 (2005).

[98] M. Hohenadler and W. von der Linden, *Polarons in Advanced Materials* (Springer, Dordrecht, The Netherlands, 2007), pp. 463–502

[99] F. Hébert, B. Xiao, V. G. Rousseau, R. T. Scalettar, and G. G. Batrouni, Phys. Rev. B **99**, 075108 (2019).

[100] C. Chen, X. Y. Xu, J. Liu, G. Batrouni, R. Scalettar, and Z. Y. Meng, Phys. Rev. B **98**, 041102(R) (2018).

[101] Autocorrelation time estimation, https://dfm.io/posts/autocorr/.

[102] A. Sokal, Monte Carlo methods in statistical mechanics: Foundations and new algorithms, in *Functional Integration: Basics and Applications*, edited by C. DeWitt-Morette, P. Cartier, and A. Folacci (Springer, Boston, 1997), pp. 131–192.

[103] W. Purwanto, W. A. Al-Saidi, H. Krakauer, and S. Zhang, J. Chem. Phys. **128**, 114309 (2008).

[104] J. Lee, F. D. Malone, and M. A. Morales, J. Chem. Theory Comput. **16**, 3019 (2020).

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

[106] S. Poncé, E. Margine, C. Verdi, and F. Giustino, Comput. Phys. Commun. **209**, 116 (2016).

[107] J.-J. Zhou, J. Park, I.-T. Lu, I. Maliyov, X. Tong, and M. Bernardi, arXiv:2002.02045.

[108] R. L. Stratonovich, Doklady Akad. Nauk S.S.S.R. **115**, 1097 (1957).

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

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

[111] See https://github.com/pauxy-qmc/pauxy for details on how to obtain the source code.

[112] See https://github.com/jsspencer/pyblock for details on how to obtain the source code.

[113] L. K. Wagner, M. Bajdich, and L. Mitas, J. Comput. Phys. **228**, 3390 (2009).

[114] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, and S. Wanderman-Milne, JAX: composable transformations of Python+NumPy programs.