Numerically exact generalized Green's function cluster expansions for electron-phonon problems

Matthew R. Carbone, David R. Reichman, and John Sous

Phys. Rev. B **104**, 035106 — Published 6 July 2021

DOI: 10.1103/PhysRevB.104.035106

Numerically Exact Generalized Green's Function Cluster Expansions for Electron-Phonon Problems

Matthew R. Carbone,[1,*] David R. Reichman,[1] and John Sous[2,†]

[1] *Department of Chemistry, Columbia University, New York, New York 10027, USA*
[2] *Department of Physics, Columbia University, New York, New York 10027, USA*

(Dated: June 8, 2021)

We generalize the family of approximate Momentum Average (MA) methods to formulate a numerically exact, convergent hierarchy of equations whose solution provides an efficient algorithm to compute the Green's function of a particle dressed by bosons, suitable in the entire parameter regime. We use this approach to extract ground state properties and spectral functions. Our approximation-free framework, dubbed the Generalized Green's function Cluster Expansion (GCCE), allows access to exact numerical results in the extreme adiabatic limit, where many standard methods struggle or completely fail. We showcase the performance of the method, specializing to three important models of charge-boson coupling in solids and molecular complexes: the molecular Holstein model, which describes coupling between charge density and local distortions, the Peierls model, which describes modulation of charge hopping due to inter-site distortions, and a more complex Holstein+Peierls system with couplings to two different phonon modes, paradigmatic of charge-lattice interactions in organic crystals. The GGCE serves as an efficient approach that can be systematically extended to different physical scenarios, thus providing a tool to model the frequency-dependence of dressed particles in realistic settings.

I. INTRODUCTION

The interaction of a particle with its environment is central to the study of many physical systems. One classic problem of this type is that of the polaron, which describes a mobile carrier dressed by bosonic fluctuations.[1] Originally predicted by Landau,[2] expanded upon by Pekar[3,4] and cemented into condensed matter canon by Lee, Low and Pines,[5] Fröhlich, Pelzer and Zienau,[6,7] Feynman,[8] and Holstein,[9,10] a polaron forms when a particle such as an electron or hole moves in a deformable medium. The motion of the particle induces a local polarization cloud, which is dragged along with the particle as it moves, renormalizing its effective mass and yielding a non-zero quasiparticle weight. Polarons arise in a variety of physical contexts beyond that of electron-phonon systems,[11] such as excitons in photoexcited molecular crystals,[12–15] hole-doped magnets,[16] light-matter systems,[17–19] impurities embedded in ultracold gases[20–23] and in other more exotic physical settings.[24–26]

Over the last two and a half decades, many (in principle) exact numerical methods have been devised to study polaronic problems. One can broadly classify these approaches into two main categories: real- and imaginary-frequency methods. Approaches in the former class include Variational Exact Diagonalization,[27] and its variants,[28,29] Limited Phonon Basis Exact Diagonalization[30] and Matrix-Product-State techniques.[31–34] Methods in the latter class are most prominently Monte Carlo methods, such as Diagrammatic,,[35–37] Path-integral[38] and Continuous-time[39] Monte Carlo. While Monte Carlo techniques are well suited for the study of finite-temperature systems over the complete range of polaronic model parameters, they require ill-conditioned analytic continuation to the real-frequency axis in order to study dynamics.[40] In contrast, direct real-time methods face a daunting challenge in several parameter regimes, including the so-called adiabatic limit where the lattice response is slow, as well as the strong-coupling limit, where a large number of bosons is excited in the system and the size of basis states becomes too large to efficiently manage.

In this work, we introduce the Generalized Green's function Cluster Expansion (GGCE), a non-perturbative approach that enables an exact, efficient numerical computation of real-frequency Green's functions of polaronic models even in regimes challenging for related real-frequency approaches. We restrict ourselves to the limiting case of a single carrier in an otherwise unoccupied band,[1,41] reserving an attempt to formulate a cluster expansion approach for the real-frequency properties of polaron models at finite concentrations[42] for future work. In particular, we show that the GGCE provides access to exact spectra in the portions of the adiabatic and strong-coupling limits inaccessible to more standard Variational Exact Diagonalization approaches, while converging more rapidly in accessible regimes. Our method builds on the Momentum Average (MA) Approximation,[43] proposed by Berciu in 2006,[44] which has since been adapted to describe realistic materials.[45,46] Our procedure is applicable to any form of particle-boson coupling, and proceeds via efficient generation of an equation of motion (EOM) in orders of the spatial extent of bosonic clusters that arise in the dynamics. We show that this approach variationally recovers the exact infinite boson Hilbert space, provided that one converges the computation with respect to the cluster size, and we find that this is achieved with a high level of efficiency when compared against standard numerical approaches, even in the adiabatic limit. In addition to providing access to quasiparticle spectra over

a wide frequency range, the GGCE comes with several strengths. In particular, it is formulated in the *infinite* system size limit, and thus provides access to exact spectra in the thermodynamic regime. It affords sufficient flexibility that permits extensions to finite-ranged models at finite temperatures and in higher dimensions, as well as to studies of bipolarons, and systems with different boundary conditions. Additionally, it allows the study of dynamics of non-equilibrium initial states. Lastly, since existing linear algebra solvers represent the only computational bottleneck in the approach, the GGCE serves as an easy-to-implement, methodologically unconstrained technique whose performance is limited only by access to computational resources such as large-scale parallel computing or GPU technology.

Our manuscript is organized as follows. In Section II, we review the foundations of the MA methods and devise a generalized formalism we use in the GCCE approach (Subsection II.A). We briefly discuss our computational implementation of the method (Subsection II.B) and highlight the relationship to and differences between our and other methods (Subsection II.C). In Section III, we demonstrate the power and scope of this implementation and present a combination of numerically exact and quasi-converged results on the Holstein,[9,10] Peierls[47–50] and mixed-boson mode Holstein+Peierls[51] models. Finally, in Section IV, we conclude and discuss possible future work.

II. METHODOLOGY AND GENERAL CONSIDERATIONS

Consider a mobile particle (e.g. electron, hole, etc.) coupled to a bosonic field

$$\hat{H} = \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} \hat{c}^\dagger_{\mathbf{k}} \hat{c}_{\mathbf{k}} + \sum_{\mathbf{q}} \hbar\Omega_{\mathbf{q}} \hat{b}^\dagger_{\mathbf{q}} \hat{b}_{\mathbf{q}} + \sum_{\mathbf{k},\mathbf{q}} g(\mathbf{k},\mathbf{q}) \hat{c}^\dagger_{\mathbf{k}+\mathbf{q}} \hat{c}_{\mathbf{k}} (\hat{b}^\dagger_{-\mathbf{q}} + \hat{b}_{\mathbf{q}}). \quad (1)$$

Here, the carrier (boson) has dispersion $\varepsilon_{\mathbf{k}}$ ($\hbar\Omega_{\mathbf{q}}$), and the interaction $\hat{V}$ contains a vertex $g(\mathbf{k},\mathbf{q})$ that in general depends on both $\mathbf{k}$ and $\mathbf{q}$. We use a compact notation $\sum_{\mathbf{k}}$ to imply a discrete sum for a problem formulated on the lattice or a d-dimensional integral $\frac{l^d}{(2\pi)^d}\int \mathrm{d}^d k$ with l^d the system volume for a problem in the continuum.

The goal of our approach is to derive the EOM of the one-electron Green's function at zero temperature,[1]

$$G(\mathbf{k},\omega) = \langle 0|\hat{c}_{\mathbf{k}} \hat{G}(\omega) \hat{c}^\dagger_{\mathbf{k}}|0\rangle . \quad (2)$$

For Hamiltonians of the form in Eq. (1), only the retarded component of $G(\mathbf{k},t)$ contributes,[44] and the propagator, in real frequency, takes the form

$$\hat{G}(\omega) = \left[\omega - \hat{H} + \mathtt{i}\eta\right]^{-1}, \quad (3)$$

where $\eta = 0^+$ is an artificial broadening parameter. Repeated application of Dyson's equation,

$$\hat{G}(\omega) = \hat{G}_0(\omega) + \hat{G}(\omega)\hat{V}\hat{G}_0(\omega), \quad (4)$$

with $\hat{H}_0 = \hat{H} - \hat{V}$ yields an infinite hierarchy of equations,[52] which we compute in the basis states $|k,n\rangle$, labeling a delocalized state of the carrier with definite momentum quantum number k in the presence of n bosons in the system. The first application of Dyson's equation yields

$$G(\mathbf{k},\omega) = G_0(\mathbf{k},\omega)\left[1 + \langle 0|\hat{c}_{\mathbf{k}}\hat{G}(\omega)\hat{V}\hat{c}^\dagger_{\mathbf{k}}|0\rangle\right], \quad (5)$$

and the second gives

$$\langle 0|\hat{c}_{\mathbf{k}}\hat{G}(\omega)\hat{V}\hat{c}^\dagger_{\mathbf{k}}|0\rangle = \langle 0|\hat{c}_{\mathbf{k}}\hat{G}(\omega)\hat{V}\hat{G}_0(\omega)\hat{V}\hat{c}^\dagger_{\mathbf{k}}|0\rangle , \quad (6)$$

where $\hat{G}_0(\omega)$ is the free particle propagator, and

$$\hat{G}_0(\omega)\,|k,n\rangle = G_0(k,\omega - n\hbar\Omega)\,|k,n\rangle . \quad (7)$$

Note this expansion can be indexed by the number of bosons contained in the created states. A coupling $\hat{V}$ that is linear in boson operators either creates or annihilates a boson, thus coupling states with n bosons to states with $n \pm 1$ bosons.

A key development made by Berciu[44,53] is to recast the EOM as a hierarchical "expansion" in orders of the spatial extent of the bosonic cloud, M, rather than treating it as a direct expansion in the number of bosons. Making use of the spatial structure of the Green's functions generated through repeated application of Dyson's identity allows one to derive a scheme in which states corresponding to clouds larger than a certain spatial extent M are excluded. To illustrate the idea, consider the example of $M = 2$. At this level of approximation, only states with bosons localized on single and first-neighbor sites are retained in the hierarchy. Note that this imposes no restriction on the distance between the carrier and the boson cloud. We can view this approximation as a variational ansatz in the space of Green's functions in which one allows the carrier anywhere in the system, but with bosons clustered in a cloud of a maximum length M.

Before delving into the details, we provide a brief summary of the convergence parameter space employed in our method. As discussed above, M indexes the maximum extent (in units of lattice sites) of the bosonic cloud contained in the set of linear equations generated through repeated application of Dyson's identity. Any equation in the closure must have cloud extent L such that $0 \leq L \leq M$. The total number of bosons, N, allowed in any cloud provides a second convergence parameter. Similarly to M, only equations with a total number of bosons $0 \leq n_{\mathrm{T}} \leq N$ are allowed. An example of the configuration of a single *auxiliary* Green's function (AGF), a Green's function describing the overlap between the single carrier and a configuration of the single carrier + some boson distribution, is given in Fig. 1. Converging

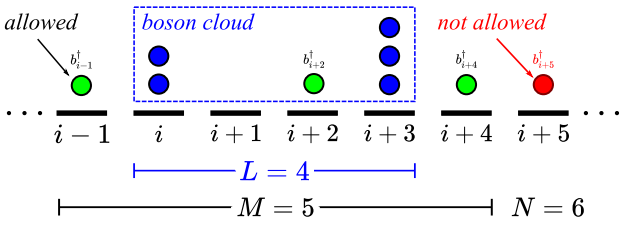

FIG. 1. Cartoon of a $L = 4$, $\mathbf{n} = [2, 0, 0, 3]$ boson cloud (blue), such that $n_{\mathrm{T}} = 5$, contained within a variational space specified by a maximum cloud extent $M = 5$ and maximum number of bosons $N = 6$. In this example, the constraint $M = 5$ spans sites $i-1$ to $i+4$ so that bosons can be created only on these sites (e.g., green circles), and are not allowed outside of the M-site cloud (e.g., red circle). $N = 6$ implies that states with two more bosons that those depicted in the figure (blue circles) are omitted from the variational space. Note that the carrier (not shown) is allowed to be anywhere on the chain.

M and N in numerical calculations allows us to approach the infinite Hilbert space limit.

Below, we detail the approach we use to construct and solve the linear system of equations in the EOM. Specifically, in Subsection II.A we derive a generalized expression for $G(k, \omega)$ for arbitrary models. Then, in Subsection II.B, we explain how to systematically generate and solve the system of equations in computer simulations. Finally, in Subsection II.C, we discuss the relation of the GGCE to other methods.

II.A. A Generalized Equation of Motion

We now specialize our construction to the case of one-dimensional (1D) lattice models described by Hamiltonians of the form

$$\hat{H} = -t \sum_{\langle ij \rangle} \hat{c}_i^\dagger \hat{c}_j + \Omega \sum_i \hat{b}_i^\dagger \hat{b}_i + \hat{V}, \tag{8}$$

where $\langle ij \rangle$ denotes nearest neighbors, for which numerical results are available, t is the hopping amplitude and Ω is the frequency of dispersionless Einstein phonons. This 1D Hamiltonian allows us to both benchmark GGCE against exact numerics and to tackle regimes that are typically difficult to study or inaccessible by related techniques even in the well-studied 1D limit. In what follows we set $\hbar = 1$ and the lattice constant $a = 1$.

Beginning with Eq. (5), we derive a generalized EOM (GEOM). Here the free particle Green's function is given by

$$G_0(k, \omega) = \left[\omega - \varepsilon_k + \mathtt{i}\eta\right]^{-1}, \tag{9}$$

with free particle dispersion $\varepsilon_k = -2t \cos k$.

Consider a generalized representation of $\hat{V}$ for models that describe coupling between a carrier and a single bosonic mode,

$$\hat{V} = \sum_{(g,\psi,\phi,\xi)} g \sum_i \hat{c}_i^\dagger \hat{c}_{i+\psi} \hat{b}^\xi_{i+\phi}. \tag{10}$$

Here g is the coupling constant, $\psi, \phi \in \mathbb{Z}$ encode the spatial dependence of the coupling, and $\xi = \{-, +\}$ labels bosonic operators as either annihilation ($b^- \equiv b$) or creation ($b^+ \equiv b^\dagger$). Specifically, ψ is an integer that indexes the structure of the carrier hopping in the coupling term, and ϕ is an integer that determines at which site relative to i (the site the fermion hops to) a phonon is created. This generalized notation completely specifies $\hat{V}$ for a given arbitary finite-ranged model. We present examples of such models in Appendix B. For clarity, let us specialize to the Holstein model as an example:

$$\hat{V}_{\mathrm{H}} = \alpha \sum_i \hat{c}_i^\dagger \hat{c}_i (\hat{b}_i^\dagger + \hat{b}_i) \tag{11}$$

can be represented in this notation as follows

$$\begin{aligned} \hat{V}_{\mathrm{H}} &= \alpha \sum_i \hat{c}_i^\dagger \hat{c}_i \hat{b}_i^\dagger + \alpha \sum_i \hat{c}_i^\dagger \hat{c}_i \hat{b}_i \\ &\leftrightarrow \{(\alpha, 0, 0, +), (\alpha, 0, 0, -)\}. \end{aligned} \tag{12}$$

We allow for an arbitrary but finite number of interaction terms, which need not be equal and can thus be used to model, for example, a long-ranged coupling of a carrier to a bosonic mode.

Using Eq. (5), we arrive at the GEOM for $G(k, \omega)$,

$$f_0(0) = G_0(k, \omega) \left[1 + \sum_{(g,\psi,\phi,\xi)} g e^{\mathtt{i}kR_{\psi-\phi}} f_1(\phi)\right]. \tag{13}$$

Here, we have defined an AGF[44,54] given by

$$f_n(\delta) = \mathcal{N}^{-1/2} \sum_i e^{\mathtt{i}kR_i} \langle 0 | \hat{c}_k \hat{G}(\omega) \hat{c}_{i-\delta}^\dagger \hat{b}_i^{\dagger n} | 0 \rangle, \tag{14}$$

where $\mathcal{N}$ is the number of lattice sites, $R_m \equiv m$ and $f_n(\delta) \equiv f_n(k, \delta, \omega)$. The AGFs can be thought of as higher-order propagators of an electron in a spatial cloud composed of multiple bosonic excitations. Further, we note the identity $f_0(\delta) = e^{\mathtt{i}kR_\delta} G(k, \omega)$, c.f. Eq. (14).

It is now necessary to introduce additional notation for describing how AGFs with greater than zero phonons couple. Since bosons can in general be created anywhere on the lattice, we define an occupation number vector $\mathbf{n}$, which labels the number of boson excitations starting from site i on a cloud embedded within the infinite lattice,

$$\mathbf{n} \equiv [n^{(i)}, n^{(i+1)}, ..., n^{(i+L-1)}], \tag{15}$$

where $L \leq M$ is the length of $\mathbf{n}$. This vector serves as a device for labeling the bosonic Hilbert space in the following way: $\mathbf{n} \leftrightarrow B_{i,\mathbf{n}}^\dagger |0\rangle$, where $B_{i,\mathbf{n}}^\dagger \equiv \hat{b}_i^{\dagger n_0} \hat{b}_{i+1}^{\dagger n_1} \cdots \hat{b}_{i+L-1}^{\dagger n_{L-1}}$.

This allows us to write a generalized version of Eq. (14), where n becomes a vector,

$$f_{\mathbf{n}}(\delta) = \mathcal{N}^{-1/2} \sum_i e^{\mathrm{i}kR_i} \langle 0|\hat{c}_k \hat{G}(\omega) \hat{c}^\dagger_{i-\delta} \hat{B}^\dagger_{i,\mathbf{n}}|0\rangle \,. \quad (16)$$

Upon Fourier transforming to reciprocal space and substituting Dyson's equation, we obtain

$$f_{\mathbf{n}}(\delta) = \mathcal{N}^{-1} \sum_i e^{\mathrm{i}kR_i} \sum_q e^{-\mathrm{i}qR_{i-\delta}} G_0(q, \omega - n_\mathrm{T}\Omega) \times \langle 0|\hat{c}_k \hat{G}(\omega) \hat{V} \hat{c}^\dagger_q \hat{B}^\dagger_{i,\mathbf{n}}|0\rangle \,, \quad (17)$$

where n_T is the total number of bosons in the configuration labeled by $\mathbf{n}$. Here we used the fact that when $n_\mathrm{T} > 0$, $\langle 0|\hat{c}_k \hat{G}_0(\omega) \hat{c}^\dagger_q \hat{B}^\dagger_{i,\mathbf{n}}|0\rangle = 0$. Defining $\tilde{\omega} \equiv \omega - n_\mathrm{T}\Omega$ and adopting a combined real/momentum-space representation, we have

$$f_{\mathbf{n}}(\delta) = \mathcal{N}^{-3/2} \sum_i e^{\mathrm{i}kR_i} \sum_q e^{-\mathrm{i}qR_{i-\delta}} G_0(q, \tilde{\omega}) \times \sum_m e^{\mathrm{i}qR_m} \langle 0|\hat{c}_k \hat{G}(\omega) \hat{V} \hat{c}^\dagger_m \hat{B}^\dagger_{i,\mathbf{n}}|0\rangle \,. \quad (18)$$

The goal of the procedure is to extract a relationship between AGFs with n_T and $n_\mathrm{T} \pm 1$ bosons. This depends on the specific form of $\hat{V}$. It is thus advantageous to express $\hat{V}$ as defined in Eq. (10) to obtain

$$\hat{V} \hat{c}^\dagger_m \hat{B}^\dagger_{i,\mathbf{n}} |0\rangle = \sum_{(g,\psi,\phi,\xi)} g \sum_j \hat{c}^\dagger_j \hat{b}^\xi_{j+\phi} \hat{B}^\dagger_{i,\mathbf{n}} |0\rangle \, \delta_{m,j+\psi}. \quad (19)$$

Consider the case when $\xi = -$, implying the boson operator removes a boson from site $j + \phi$. Such a process can only have a non-zero contribution when a boson is removed from an occupied site, and the domain of sites where $\hat{b}_{j+\phi}$ can act in general is $j + \phi - i \in \Gamma^-_L = \{0, 1, ..., L-1\}$. In this case, an extra prefactor appears due to the boson commutation relations, $\hat{b}_j \hat{b}^{\dagger m}_i = \hat{b}^{\dagger m}_i \hat{b}_j + m\delta_{ij} \hat{b}^{\dagger m-1}_i$.

Up until now, this derivation has been exact. We now impose a limit on the maximum cloud extent, M, restricting the cluster of sites where bosons can be created to at most M connected sites, which are occupied with up to N bosons.[55] Thus, when $\xi = +$, we have $j + \phi - i \in \Gamma^+_L = \{L - M, L - M + 1, ..., M - 1\}$. This restriction requires that we replace the sum over j with a sum over the elements of the aforementioned set: $\sum_j \to \sum_{\gamma \in \Gamma^\xi_L}$.

To continue the derivation of the EOM, we introduce the notation: $\hat{B}^{(\xi,\gamma)\dagger}_{i,\mathbf{n}} |0\rangle$ as the state $\hat{B}^\dagger_{i,\mathbf{n}} |0\rangle$ with an extra boson created ($\xi = +$) or destroyed ($\xi = -$) on site $i + \gamma$ within the permitted variational space specified by the above restriction. We omit states indexed by $\mathbf{n}$ whose $n_\mathrm{T} > N$ from the space of AGFs. Fig. 1 demonstrates the variational space encoded in our notation.

Summing over m and q in Eq. (18) produces the following general form,

$$f_{\mathbf{n}}(\delta) = \sum_{(g,\psi,\phi,\xi)} g \sum_{\gamma \in \Gamma^\xi_L} n^{(\xi,\gamma)} g_0(\delta + \gamma - \phi + \psi, \tilde{\omega}) \times \mathcal{N}^{-1/2} \sum_i e^{\mathrm{i}kR_i} \langle 0|\hat{c}_k \hat{G}(\omega) \hat{c}^\dagger_{i+\gamma-\phi} \hat{B}^{(\xi,\gamma)\dagger}_{i,\mathbf{n}}|0\rangle \,, \quad (20)$$

where $n^{(\xi,\gamma)}$ is a prefactor associated with applying a boson creation or annihilation operator: it is equal to 1 if $\xi = +$, and is equal to the number of bosons on site $i + \gamma$ (before a boson is annihilated) if $\xi = -$.

Here, the free particle propagator in real space is given by[56]

$$g_0(\delta, \omega) = \frac{1}{\mathcal{N}} \sum_q e^{\mathrm{i}qR_\delta} G_0(q, \omega) = -\frac{\mathrm{i}\left[-\omega_\eta/2t + \mathrm{i}\sqrt{1 - (\omega_\eta/2t)^2}\right]^{|\delta|}}{\sqrt{4t^2 - \omega^2_\eta}}, \quad \omega_\eta \equiv \omega + \mathrm{i}\eta. \quad (21)$$

Observe that the second line in Eq. (20) is precisely an AGF with different arguments and with $n_\mathrm{T} \to n_\mathrm{T} \pm 1$ bosons. Indexing a new AGF in the same manner as before we have

$$f_{\mathbf{n}}(\delta) = \sum_{(g,\psi,\phi,\xi)} g \sum_{\gamma \in \Gamma^\xi_L} n^{(\xi,\gamma)} \times g_0(\delta + \gamma - \phi + \psi, \tilde{\omega}) f^{(\xi,\gamma)}_{\mathbf{n}}(\phi - \gamma). \quad (22)$$

Finally, we note that in order to abide by our labeling convention, certain "reduction rules" for the AGFs must be followed in order to produce a valid closure. When removing or adding bosons, as in the case $f_{\mathbf{n}} \to f^{(\xi,\gamma)}_{\mathbf{n}}$, additional phase prefactors may appear. The details of these rules are summarized in Appendix A (see also Ref. 54 for a specific example).

II.B. Implementation

Together, Eqs. (13) and (22), along with the rules in Appendix A, contain all information necessary to solve for $G(k, \omega)$ for some chosen values of M, N. In this section, we describe the *computational* approach for representing these equations and solving them numerically.

Every possible combination of $1 \le n \le N$ bosons on $1 \le L \le M$ sites will contribute to the calculation of $G(k, \omega)$. In the first step, we systematically generate all combinations, noting the only requirement that the first and last sites for some cloud extent L must be at least singly occupied. This amounts to symbolically constructing and storing representations of these objects, e.g.

$$\mathcal{G} = \{f_{[0]}(\delta), f_{[1]}(\delta), f_{[1,1]}(\delta), f_{[1,0,2]}(\delta), ...\}, \quad (23)$$

such that all possible AGFs corresponding to a given configuration are generated. This can be thought of precisely as the classic combinatorics problem of N indistinguishable balls in M distinguishable bins, with the added constraint of requiring at least one boson on each end of the cloud. In this way, the total number of equations generated at this step (the total number of elements in $\mathcal{G}$, defined as $|\mathcal{G}|$) has a straightforward representation,

$$|\mathcal{G}| = 1 + \sum_{L=1}^{M}\sum_{n=1}^{N} \begin{cases} 1 & \text{if } L=1 \text{ or } n=2 \\ \binom{L+n-3}{n-2} & \text{otherwise} \end{cases}, \quad (24)$$

where the one extra term reflects the first equation in the set of equations (for $G(k,\omega)$).

The second step consists of finding the values for δ each function $f_{\mathbf{n}}$ requires. Observing that the only δ-dependence on the RHS of Eq. (22) is contained in g_0 (and importantly not in $f_{\mathbf{n}}$), we obtain the full closure of equations by finding, for every $f_{\mathbf{n}}$, the values of δ prescribed by the indices $\phi - \gamma$ on the RHS. This set is informally denoted as $\mathcal{S}$, e.g.,

$$\mathcal{S} = \{f_{[0]}(-1), f_{[0]}(0), f_{[1]}(-1), ...\}. \quad (25)$$

The terms contained in $\mathcal{S}$ are determined by a nontrivial function of M, N and depend on the model type. Every term in $\mathcal{S}$ is simply a specific case of the LHS of Eq. (22). To further clarify, the generalized equations in the set $\mathcal{G}$ leave δ unfixed. The equations in the set $\mathcal{S}$ *fix* the allowed values of δ based on the indices $\phi - \gamma$. We note this does not constitute an approximation to the EOM since the conditions that fix $\mathcal{S}$ arise naturally within the hierarchy.

In the final step, we formulate this as a inhomogenous linear system of equations and aim to find the solution for all $f_{\mathbf{n}}(\delta)$ for some values of k, ω, M, N,

$$A\mathbf{f} = \mathbf{b}. \quad (26)$$

Above, A is a matrix of coefficients which can be read from the aforementioned equations, and $\mathbf{b}$ is proportional to the unit vector and inherits the inhomogeneity of Eq. (13). This matrix equation can be solved in one of two ways. The solution for $\mathbf{f}$ can be obtained in a single step, which amounts to applying some direct solver to the $|\mathcal{S}| \times |\mathcal{S}|$ matrix A. However, this approach is either inefficient (using a sparse solver) or intractable using a dense solver due to the large size of A in cases such as the extreme adiabatic limit. Alternatively, we find that a continued fraction approach using dense linear algebra provides the optimal middle ground. Formally, the continued fractions (here we suppress the M and N dependence) $V_n = \mathcal{A}_n(\mathbf{k},\omega)V_{n-1} + \mathcal{B}_n(\mathbf{k},\omega)V_{n+1}$, where $\mathcal{A}_n(\mathbf{k},\omega)$ and $\mathcal{B}_n(\mathbf{k},\omega)$, are sparse matrices read off directly from the EOM, and V_n is a vector of AGF's with $n \leq N$ bosons.[53,54] The matrix inversions required are much smaller in this approach, although there are $\Theta(N)$ of them. We note that using this more efficient approach, the calculations become challenging in our current implementation only around $(M, N) \sim (10, 7)$, which produces ∼60k equations. Adding one more boson balloons the calculation to ∼150k equations, which are in principle within reach on large supercomputer architectures with sufficient memory capacity.

To approach the infinite phonon Hilbert space limit using the continued fraction approach, we set $V_{N+1} = 0$, solving the set of equations until we obtain $G(\mathbf{k},\omega)$, which corresponds to V_0. In the $N \to \infty$ limit, this represents a sensible boundary condition because it becomes energetically expensive to generate clouds with larger than N bosons. In practice we treat N as a convergence parameter. All results shown in this work appear to be converged with respect to N to desirable accuracy, unless otherwise stated.

II.C. Comparison to Other Methods

1. *Comparison to related methods: Momentum Averge (MA) and Limited Phonon Basis Exact Diagonalization (LPBED) methods*

The GGCE method combines advantages from the MA and Limited Phonon Basis Exact Diagonalization[30] (LPBED) methods. In the MA approach, one makes an educated guess of the value of M needed to obtain accurate results, in essence employing a variational ansatz to the EOM. One then derives the EOM in MA(M) analytically "by hand" and solves for $G(k,\omega)$ numerically. LPBED is a more general ED analog of MA, and in principle also relies on a variational ansatz, albeit one different from that of MA. Another successful version of LPBED[57] discussed in the literature included clouds of size $M = 5$ whilst allowing for two extra bosons anywhere on the lattice even away from the cloud, but with a more restricted total number of bosons.[58]

We can roughly view MA and LPBED methods as specific variational cases of the GGCE, which benefits from allowing an arbitrary systematic choice of maximal cloud extent, M, in the $N \to \infty$ limit. The GGCE thus serves as a systematically exact method which allows one to tailor resources based on the underlying physics of the problem, and is limited only by computational resources. This provides the potential to access regimes that are difficult to quantitatively describe by other approaches, as we show below.

2. *Comparison to Variational Exact Diagonalization (VED) methods*

Variational Exact Diagonalization (VED)[27] represents another class of successful approaches to the polaron problem. In VED, a variational Hilbert space is iteratively generated by applying the off-diagonal parts of the Hamiltonian to a reference state taken to be a Bloch state of an electron and zero bosons in an infinite system. After N_h iterations, one diagonalizes the Hamiltonian in

the generated basis using standard Lanczos techniques. Convergence with respect to N_h, when possible, guarantees access to the exact ground state and a small manifold of low-lying excited states.[28] There are at least two main differences between GGCE and VED.

First, VED naturally imposes a restriction on the distance between the electron and phonon configurations, which can be at most $\sim N_h$ sites (the precise value depends on the coupling), while GGCE (and MA[54]) includes states with the electron arbitrarily far away from the phonon clouds with no restriction (this can be seen from the application of $\hat{G}_0(\omega)$ in the EOM on states in AGFs with both an electron and phonons, which moves the electron arbitrarily in the system without regard to the location of the bosonic cloud, c.f. Eq. (20)). We note that VED is capable of describing the ground and low-lying excited states in the weak- and lower-intermediate regimes of coupling in the adiabatic limit.[27] We suspect that the restriction on the distance between the electron and the phonons in VED prohibits access to very strong couplings in the adiabatic regime and to continuum states since these are generally delocalized states (see discussion below). In contrast, as we show below, GGCE can tackle strong coupling in the adiabatic regime.

Second, GGCE is formulated as an expansion in terms of cloud sizes, and the computation must be converged with respect to the cloud size, while VED imposes no restriction on cloud sizes (a cloud in VED can extend over, at most, $\sim N_h$ sites). For example, $N_h = 11$ implies clouds extended over $\sim$ 10-11 sites (the exact number depends on the specific model of the electron-boson coupling). Such a value of N_h represents a rough lower bound within what is typically used in VED in the intermediate adiabatic limit. These values imply clouds with sizes that are much larger than those used in GGCE in the current work. This suggests that GGCE may benefit in terms of efficiency by employing a smaller number of states resulting from smaller clouds without compromising accuracy. We believe this is a direct result of using an EOM formulation of propagators, which ensures we keep only those states generated in the dynamics and nothing further. Comparing, empirically, to Ref. 27, we note that the number of states needed in GGCE appears to be two orders of magnitude smaller than those in VED in order to achieve convergence in similar parameter regimes.

Finally, we note that other variants of VED with extra restrictions on the variational space have been used with great success.[28,59,60] These, however, are either not formulated in a general enough manner to be applied to a generic form of electron-boson coupling[59] or involve further constraints that, while variational, are not completely motivated physically especially at strong couplings. In contrast, GGCE in its current form follows naturally from the EOM and has no restrictions beyond the cloud size, which is taken to the infinite limit sequentially and in an efficient manner. In principle further restrictions of this type can be imposed in our GGCE, but we do not explore this direction in the current manuscript.

The preliminary analysis presented here suggests that GGCE may perform more favorably than related approaches, at least in some parameter regimes and for some quantities. Future work must be devoted to address these issues and compare the range of variational restricted-basis approaches over the full range of parameter space for both ground-state energies and spectral functions to fully access the utility and efficiency of each approach.

III. RESULTS

In this section, we show results for a variety of 1D lattice models described by the Hamiltonian defined by Eqs. (8) and (10). This allows us to both benchmark GGCE against exact numerics, and to tackle regimes typically inaccessible even in the well-studied 1D limit. In what follows, we characterize the interaction strength via the dimensionless coupling constant

$$\lambda = E_{\rm GS}(t=0)/E_{\rm GS}(\alpha=0), \tag{27}$$

which is the ratio of the ground state (GS) energy in the atomic limit to that in the free particle limit, and the adiabaticity ratio

$$\Lambda = \Omega/W, \tag{28}$$

where $W = 4t$ is the carrier's bandwidth.

While DMC and other quantum Monte Carlo methods may access the GS in the adiabatic limit, dynamics are generally difficult to obtain due to uncertainties associated with analytical continuation to the real-frequency axis. We showcase the ability of the GGCE to simulate dynamics in the low-energy regime for the Holstein[9,10] (H) and Peierls (P) (also known as the Su-Schrieffer-Heeger[47]) models. Finally, we study an experimentally motivated mixed Holstein+Peierls (HP) model in which the carrier couples to two different boson modes, one describes a Holstein coupling and the other a Peierls coupling.

III.A. Holstein Model

We first consider the prototypical Holstein model[9,10] for which

$$\hat{V} = \alpha \sum_i \hat{c}_i^\dagger \hat{c}_i (\hat{b}_i^\dagger + \hat{b}_i), \quad \lambda_{\rm H} \equiv \alpha^2/2\Omega t. \tag{29}$$

In Fig. 2, we compute the GS energy of a Holstein polaron for $\Lambda \in [0.0025, 2.5]$. For $\Omega/t = 0.1$ and 0.5, we compare our results to those obtained via Diagrammatic Monte Carlo (DMC).[61] Not only does GGCE converge to the exact result for $\lambda \in [0, 1.2]$, but it also yields slightly lower GS energies than DMC in the strong-coupling regime $\lambda \gtrsim 1$, although the differences are likely

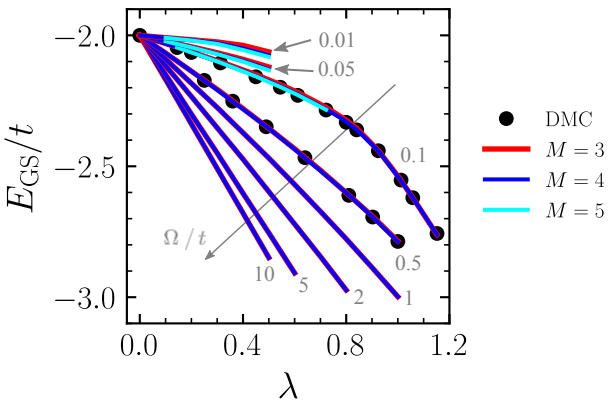

FIG. 2. Ground state energy E_{GS}/t as a function of coupling strength $\lambda = \alpha^2/2\Omega t$ for the Holstein model in adiabaticity limits extending from anti- ($\Lambda \gg 1$) to extreme-adiabatic ($\Lambda \ll 1$). Values for Ω/t are shown in grey, and GGCE results for $\Omega/t = 0.1$ and 0.5 are compared to DMC data (symbols) obtained from Ref. 61. Ground state peak locations are converged with respect to N, and generally require ≈ 10 bosons at small couplings, but up to ≈ 30 at large couplings.

due to statistical errors in DMC.[62] Importantly, we are able to converge our results to the exact limit even at extremely small $\Omega/t \in [0.01, 0.1]$ for intermediate coupling strengths $\lambda \gtrsim 0.5$, overcoming previous limitations of momentum average methods. Beyond demonstrating GGCE's ability to simulate the adiabatic limit of massive bosons, our results show a trend at intermediate couplings of the polaron binding energy $|E_{GS}(\lambda) - E_{GS}(0)|$ that monotonically decreases with Ω.

To demonstrate the ability of the GGCE method to converge spectral functions and probe a broad range of physical regimes, we present an array of spectral functions in Fig. 3. These results cover all combinations of $\Omega/t \in \{0.1, 0.5\}$, $k \in \{0, \pi/2, \pi\}$ and $\lambda \in \{0.2, 0.5, 0.8\}$ and highlight the potential of the method. For example, in both cases treated in Fig. 3, we find excellent convergence of the ground state peak location and structure. The first excited state, which for the values of λ considered, lies in the polaron + one boson continuum, proves more difficult to converge. Nonetheless, we show reasonable convergence of this second peak for a wide range of parameters. However, convergence becomes more challenging for $\Omega/t = 0.5$ at $\lambda = 0.5$, as seen in the second column of Fig. 3(a), even when using extremely large cloud sizes ($M = 10$), and as a result this peak is not sufficiently converged. Difficulty in resolving excitations above $E_{GS} + \Omega$ is not surprising, since the nature of these continuum states involves scattering between a delocalized electronic state and an extended cloud of phonons that is generally not small. As such, a sufficiently large cloud and therefore a bigger variational space is needed for convergence. Thus, with increasing computational resources, convergence of the spectral function proceeds naturally from low to high energy. This implies that one can readily achieve convergence of lower-energy states with much ease.

III.B. Peierls Model

In Fig. 4, we present exact spectral functions of a polaron in the Peierls model[47,49,50,63] defined by

$$\hat{V} = \alpha \sum_{\langle ij \rangle} (\hat{c}_i^\dagger \hat{c}_j + \text{h.c.})(\hat{b}_i^\dagger + \hat{b}_i - \hat{b}_j^\dagger - \hat{b}_j),$$

$$\lambda_P = 2\alpha^2/\Omega t, \quad (30)$$

for a variety of different dimensionless couplings. Although in principle no more difficult than for the case of the Holstein model, we reserve exploring the extreme-adiabatic limit ($\Omega/t \ll 1$) to future work and show results only for $\Omega/t = 1$.

The Peierls model exhibits distinct polaron physics when compared with the Holstein model. A Peierls polaron exhibits a sharp transition from a state with $k_{GS} = 0$ to one with $k_{GS} \neq 0$ for $\lambda > \lambda_c(\Omega/t)$,[64] while a Peierls bipolaron exhibits a significantly smaller mass than its Holstein counterpart[65] and can exhibit transitions under certain conditions.[66] We observe the transition to a band minimum at a finite wave vector in Fig. 4 as λ changes from $\lambda = 0.8$ to $\lambda = 1$, consistent with Ref. 64. Importantly, we are able to resolve the spectrum above the ground state within sufficient accuracy. The excited states of this model play an important role in presence of other perturbations, as will become apparent next.

III.C. Mixed-Boson Mode Holstein + Peierls Model

We now consider a realistic model applicable to organic crystals, molecular complexes, etc., in which the charge carrier couples to both Holstein and Peierls phonon modes, each with its own frequency.[51,67–69] The Hamiltonian is given by

$$\begin{aligned}\hat{H} = &-t \sum_{\langle ij \rangle} \hat{c}_i^\dagger \hat{c}_j + \Omega_H \sum_i \hat{h}_i^\dagger \hat{h}_i + \Omega_P \sum_i \hat{p}_i^\dagger \hat{p}_i \\ &+ \alpha_H \sum_i \hat{c}_i^\dagger \hat{c}_i (\hat{h}_i^\dagger + \hat{h}_i) \\ &+ \alpha_P \sum_{\langle ij \rangle} (\hat{c}_i^\dagger \hat{c}_j + \text{h.c.})(\hat{p}_i^\dagger + \hat{p}_i - \hat{p}_j^\dagger - \hat{p}_j),\end{aligned} \quad (31)$$

where $\hat{h}_i \equiv \hat{b}_i^H$, $\hat{p}_i \equiv \hat{b}_i^P$ and the Holstein and Peierls boson operators act on *different* boson Hilbert spaces. We note that the combinatorics of multi-phonon models require vastly more resources than single mode cases. Here, $\lambda_H = \alpha_H^2/2\Omega_H t$ and $\lambda_P = 2\alpha_P^2/\Omega_P t$, as before.

First, we detail the differences between this HP model and that presented in Ref. 70. The latter model represents a toy model of a carrier coupled to *one boson*

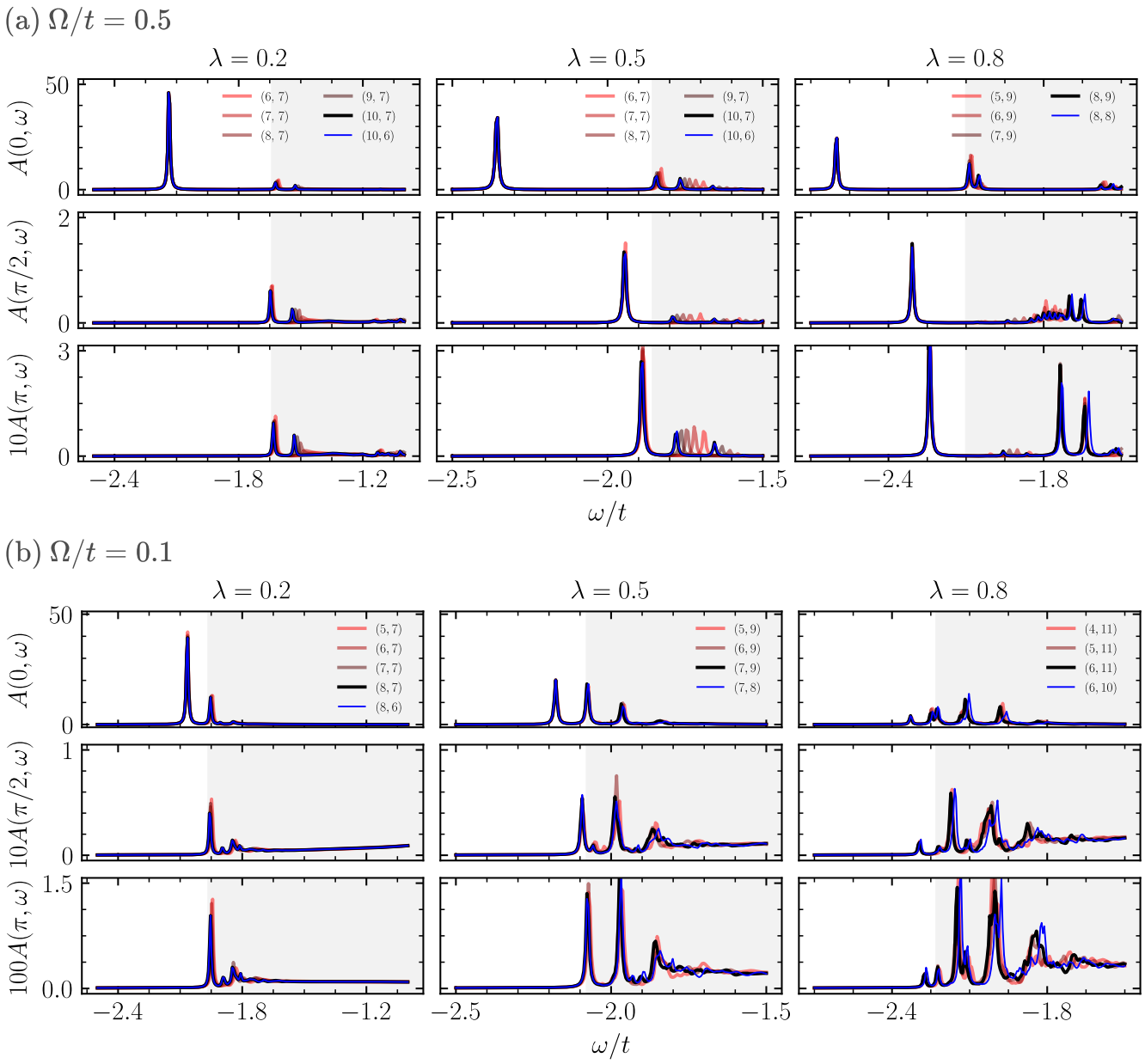

FIG. 3. Spectral function $A(k, \omega)$ at $k = 0, \pi/2$ and π for the Holstein model in (a) the mildly adiatabic ($\Omega/t = 0.5$) and (b) adiabatic limits ($\Omega/t = 0.1$), for dimensionless couplings $\lambda = 0.2, 0.5$ and 0.8. In this figure we use $\eta = 0.005$. Various values of (M, N) are shown in the legend. Specifically, the gradient from red to black shows convergence with respect to M for fixed N. We demonstrate convergence with respect to N via the blue line, which shows results that use the largest used M, but with one less boson than the largest-used N. The onset of the continuum is shown in shaded gray, and is defined as $E_{\rm GS} + \Omega$, where $E_{\rm GS}$ is the polaron ground-state energy.

type, with two coupling contributions: diagonal (Holstein) and off-diagonal (Peierls). Computations for this type of model possess the same scaling complexity as that for H or P models, making it much easier to converge. However, a realistic calculation requires modeling couplings to multiple phonon modes, typically of vastly different energies, characteristic of experimental systems. A straightforward generalization of our implementation allows us to treat the boson modes as explicitly distinguishable, even when $\Omega_{\rm P} = \Omega_{\rm H}$. We simply introduce two types of bosonic clouds, one for Holstein bosons with $M_{\rm H}$ and one for Peierls bosons with $M_{\rm P}$. These can overlap, and we thus need an extra variational parameter to constraint the absolute extent, A, over which the com-

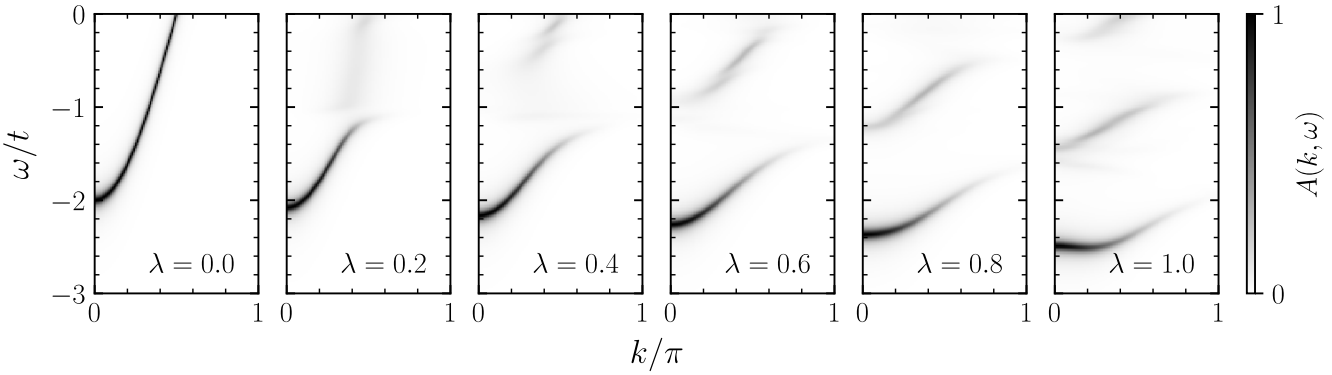

FIG. 4. Spectral function $A(k,\omega)$ (scaled to a maximum of 1) of the Peierls model for $\Omega/t = 1$, $\eta = 0.05$ and various values of the dimensionless coupling strength, λ. $M = 5$ and $N = 10$ used here are sufficient for convergence of the bands on the scale of the plot. It is worth noting that fine structure in states above the lowest energy band can be resolved on a finer grid and smaller value of η, although the intensity of these states is at most roughly an order of magnitude smaller than that of the lowest energy band.

bined clouds extend. We detail this construction in Appendix C. This approach allows us to explore this more experimentally relevant model. As previously mentioned, this comes with the downside of increased computational complexity. However, as we show below, we are able to semi-quantitatively converge the lowest-energy band for reasonably large couplings, and, with modest computational resources, we resolve the spectrum in the experimentally relevant regime (see Fig. 5), $\Omega_{\rm P} < \Omega_{\rm H}$, as well as in a hypothetical scenario with the frequencies reversed. This requires modest choices of $M_{\rm H}$, $M_{\rm P}$ and A. In Fig. 5), in the more experimentally relevant case, with $\Omega_{\rm H}/t = 2.5$ and $\Omega_{\rm P}/t = 0.5$, we see a non-negligible bandwidth and thus significant P-like character, an important observation for experiment. In our simulations of this model, we also find interesting behavior in the second peak in the spectrum involving a minimum away from $k = 0$ (not shown), which we leave to a future detailed analysis.

We quantify the ground state convergence as a function of the individual maximum cloud extents $M_{\rm H}, M_{\rm P}$, the absolute cloud extent A, and maximum number of bosons in the variational space, $N_{\rm H}, N_{\rm P}$ in Fig. 6. This analysis suggests that an increase of computational resources, within reach on large computers, will permit complete convergence.

IV. CONCLUSIONS

We have presented an exact, general approach to solving the EOM of a Green's function of a particle dressed by bosons, suitable for treating difficult regimes such as the adiabatic limit, and have demonstrated the power of the approach by calculating the polaron ground state and spectral functions in coupling regimes ranging from weak to strong, and adiabaticity limits ranging from ex-

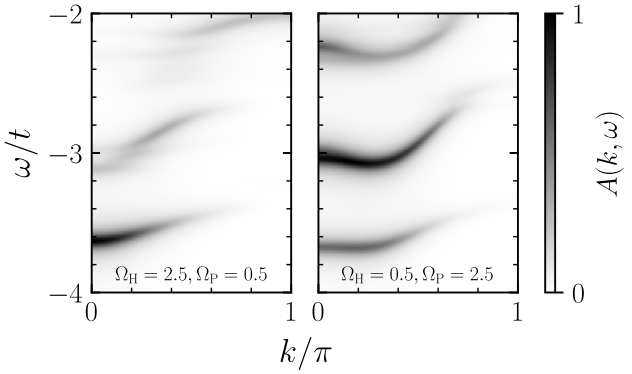

FIG. 5. Spectral function $A(k,w)$ (scaled to a maximum of 1) of the mixed-boson mode Holstein+Peierls model for various values of $\Omega_{\rm H}$ and $\Omega_{\rm P}$, $\lambda_{\rm H} = \lambda_{\rm P} = 1$, $t = 1$ and $\eta = 0.05$. For these calculations, we use $M_{\rm H} = M_{\rm P} = 3$, and a maximum total cloud length, or absolute extent, $A = 3$ (see Appendix C), and $N_{\rm H} = N_{\rm P} = 5$, for which semi-quantitive convergence is achieved.

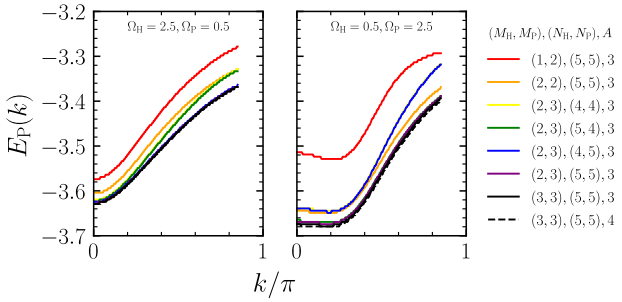

FIG. 6. Convergence of the energy of the ground-state polaron band, $E_{\rm P}(k)$, for parameters shown in Fig. 5 against combinations of $M_{\rm H}, M_{\rm P}, A, N_{\rm H}$ and $N_{\rm P}$.

treme anti-adiabatic to extreme adiabatic. We note that at large couplings, the GGCE achieves ground state energies in agreement with DMC (Fig. 2), without the introduction of stochastic error. Exact simulated spectra for $\Lambda \ll 1$ are, in general, difficult to achieve with Monte Carlo methods due to the reliance on analytic continuation, and inaccessible to most Exact Diagonalization methods due to the large basis size needed for convergence.

We emphasize the success achieved by the MA method in characterizing polarons and bipolarons in various systems under different physical conditions of experimental relevance. In most of these cases, verification of the accuracy of the method against an exact approach was needed to justify *a posteriori* its utility and potential in limits where exact numerics are difficult to obtain, e.g., in higher dimensional systems. The GGCE method systematically makes use of the MA hierarchy, resulting in an exact yet efficient approach, and a new physically motivated expansion in orders of the boson cluster size, thus expanding the horizon of possibilities in characterizing dressed quasiparticles in previously challenging regimes. Finally, the GGCE computational framework is well-suited for future practical extensions, including higher-dimensional systems, finite-temperature studies, the computation of observables connected to higher-order Green's functions, such as optical spectra and polaron mobilities,,[71] as well as studies of the dynamics of bipolarons,[72] and in other contexts we plan to address in future work.

ACKNOWLEDGMENTS

We acknowledge helpful discussions with Kipton Barros, Robert Blackwell, Denis Golez, Kyle T. Mandli, Riccardo Rossi, David Stein, Kwangmin Yu and especially Mona Berciu and Andrew J. Millis. We thank Alexandru Macridin for providing Quantum Monte Carlo data for the ground state energy of the Holstein polaron. M. R. C. acknowledges support from the U.S. Department of Energy through the Computational Sciences Graduate Fellowship (DOE CSGF) under grant number: DE-FG02-97ER25308. D. R. R. and J. S. acknowledge support from the National Science Foundation (NSF) Materials Research Science and Engineering Centers (MRSEC) program through Columbia University in the Center for Precision Assembly of Superstratic and Superatomic Solids under Grant No. DMR-1420634. J. S. also acknowledges the hospitality of the Center for Computational Quantum Physics (CCQ) at the Flatiron Institute. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We acknowledge computing resources from Columbia University's Shared Research Computing Facility project, which is supported by NIH Research Facility Improvement Grant 1G20RR030893-01, and associated funds from the New York State Empire State Development, Division of Science Technology and Innovation (NYSTAR) Contract C090171, both awarded April 15, 2010.

Appendix A: Reduction Rules for AGFs

In this Appendix, we detail the reduction rules the AGFs follow in order to produce a valid EOM.

Annihilating or creating a boson to the right of the last occupied site does not come with any additional rule for re-indexing:

$$f_{[n,n',\ldots,n'',0,\ldots,0,1]}(\delta) \xrightarrow{\hat{b}} f_{[n,n',\ldots,n'',0,\ldots,0]}(\delta) = f_{[n,n',\ldots,n'']}(\delta), \quad \text{(A1)}$$

$$f_{[n,n',\ldots,n'']}(\delta) \xrightarrow{\hat{b}^\dagger} f_{[n,n',\ldots,n'',0,\ldots,0,1]}(\delta) = f_{[n,n',\ldots,n'',0,\ldots,0,1]}(\delta), \quad \text{(A2)}$$

where here $n, n'' > 0$.

However, when creating or annihilating a boson to the left of the first occupied site on the chain, we must re-index the state such that the label i always references the first occupied site:

$$f_{[1,0,\ldots,0,n,n',\ldots,n'']}(\delta) \xrightarrow{\hat{b}} f_{[0,\ldots,0,n,n',\ldots,n'']}(\delta) \to e^{-\mathrm{i}kR_z} f_{[n,n',\ldots,n'']}(\delta + z), \quad \text{(A3)}$$

$$f_{[n,n',\ldots,n'']}(\delta) \xrightarrow{\hat{b}^\dagger} f_{[1,0,\ldots,0,n,n',\ldots,n'']}(\delta) \to e^{\mathrm{i}kR_z} f_{[1,0,\ldots,0,n,n',\ldots,n'']}(\delta - z), \quad \text{(A4)}$$

where z is the number of shifted sites $i \to i \pm 1 \to i \pm 2, \ldots$ in the phase incurred.

Appendix B: Examples of the Generalized Notation Used in Eq. (10)

In this work, we considered H, P and HP models, each of which have different carrier-boson couplings, $\hat{V}$. Within the framework of the GGCE, these differences amount to a simple change in input parameters. The fully expanded coupling terms $\hat{V}$, and their representation in terms of the notation defined in Eq. (10), are shown here. We present the three models used and reference the derivation as performed in Section II. First, recall that the vectors which represent the coupling are notated as (g, ψ, ϕ, ξ).

In the H model, this notation translates to

$$\hat{V}_{\mathrm{H}} = \underbrace{\alpha \sum_i \hat{c}_i^\dagger \hat{c}_i \hat{b}_i^\dagger}_{(\alpha,0,0,+)} + \underbrace{\alpha \sum_i \hat{c}_i^\dagger \hat{c}_i \hat{b}_i}_{(\alpha,0,0,-)}. \quad \text{(B1)}$$

In the P model, we have

$$\begin{aligned}\hat{V}_{\rm P} = &\underbrace{\alpha \sum_i \hat{c}_i^\dagger \hat{c}_{i+1} \hat{b}_i^\dagger}_{(\alpha,1,0,+)} + \underbrace{\alpha \sum_i \hat{c}_i^\dagger \hat{c}_{i+1} \hat{b}_i}_{(\alpha,1,0,-)} \\ &\underbrace{-\alpha \sum_i \hat{c}_i^\dagger \hat{c}_{i+1} \hat{b}_{i+1}^\dagger}_{(-\alpha,1,1,+)} \underbrace{-\alpha \sum_i \hat{c}_i^\dagger \hat{c}_{i+1} \hat{b}_{i+1}}_{(-\alpha,1,1,-)} \\ &\underbrace{+\alpha \sum_i \hat{c}_i^\dagger \hat{c}_{i-1} \hat{b}_{i-1}^\dagger}_{(\alpha,-1,-1,+)} + \underbrace{\alpha \sum_i \hat{c}_i^\dagger \hat{c}_{i-1} \hat{b}_{i-1}}_{(\alpha,-1,-1,-)} \\ &\underbrace{-\alpha \sum_i \hat{c}_i^\dagger \hat{c}_{i-1} \hat{b}_i^\dagger}_{(-\alpha,-1,0,+)} \underbrace{-\alpha \sum_i \hat{c}_i^\dagger \hat{c}_{i-1} \hat{b}_i}_{(-\alpha,-1,0,-)} .\end{aligned} \tag{B2}$$

The case of the HP model is a bit more elaborate, since the model involves different boson operators: $\hat{h}_i \equiv \hat{b}_i^{(\Omega_{\rm H})}$ and $\hat{p}_i \equiv \hat{b}_i^{(\Omega_{\rm P})}$. Thus, we have

$$\begin{aligned}\hat{V}_{\rm HP} = &\underbrace{\alpha_{\rm H} \sum_i \hat{c}_i^\dagger \hat{c}_i \hat{h}_i^\dagger}_{(\alpha_{\rm H},0,0,+)} + \underbrace{\alpha_{\rm H} \sum_i \hat{c}_i^\dagger \hat{c}_i \hat{h}_i}_{(\alpha_{\rm H},0,0,-)} \\ &\underbrace{+\alpha_{\rm P} \sum_i \hat{c}_i^\dagger \hat{c}_{i+1} \hat{p}_i^\dagger}_{(\alpha_{\rm P},1,0,+)} + \underbrace{\alpha_{\rm P} \sum_i \hat{c}_i^\dagger \hat{c}_{i+1} \hat{p}_i}_{(\alpha_{\rm P},1,0,-)} \\ &\underbrace{-\alpha_{\rm P} \sum_i \hat{c}_i^\dagger \hat{c}_{i+1} \hat{p}_{i+1}^\dagger}_{(-\alpha_{\rm P},1,1,+)} \underbrace{-\alpha_{\rm P} \sum_i \hat{c}_i^\dagger \hat{c}_{i+1} \hat{p}_{i+1}}_{(-\alpha_{\rm P},1,1,-)} \\ &\underbrace{+\alpha_{\rm P} \sum_i \hat{c}_i^\dagger \hat{c}_{i-1} \hat{p}_{i-1}^\dagger}_{(\alpha_{\rm P},-1,-1,+)} + \underbrace{\alpha_{\rm P} \sum_i \hat{c}_i^\dagger \hat{c}_{i-1} \hat{p}_{i-1}}_{(\alpha_{\rm P},-1,-1,-)} \\ &\underbrace{-\alpha_{\rm P} \sum_i \hat{c}_i^\dagger \hat{c}_{i-1} \hat{p}_i^\dagger}_{(-\alpha_{\rm P},-1,0,+)} \underbrace{-\alpha_{\rm P} \sum_i \hat{c}_i^\dagger \hat{c}_{i-1} \hat{p}_i}_{(-\alpha_{\rm P},-1,0,-)} .\end{aligned} \tag{B3}$$

Appendix C: Additional Notation for Mixed-Boson Mode HP Models

In Subsection III.C, and specifically Fig. 5, we introduced new notation required to define the configuration space of the HP model. First, the occupation number vector $\mathbf{n}$ is now a two-row matrix, $\overline{n}$, where as usual the columns index the site index starting with i, and the two rows correspond to the occupation numbers of the Holstein and Peierls bosons. For clarity, we label the first row $n_{\rm H}$ and the second $n_{\rm P}$. The logic presented in Section II still applies in for the HP model: $\hat{V}$ can still create or destroy only a single boson at a time, $\hat{B}_{i,\overline{n}}$ and corresponding objects now reference both sets of boson occupation numbers (and boson operators now carry a boson-type index), $\tilde{\omega} \equiv \omega - \Omega_{\rm H} \sum_j n_{\rm H}^{(j)} - \Omega_{\rm P} \sum_j n_{\rm P}^{(j)}$, etc. As before, the left-most occupied site is still the anchor for the entire cloud, and thus the same reduction rules in Appendix A apply.

In terms of the configuration space, we now limit the maximum number of Holstein and Peierls bosons individually, using $N_{\rm H}$ and $N_{\rm P}$, respectively, and the extent of the clouds individually, using $M_{\rm H}$ and $M_{\rm P}$, respectively. Given there are now two "overlapping" clouds of bosons which live in different Hilbert spaces, we must define yet another configuration space parameter, which we call the absolute extent, A. This is the maximum extent of the cloud measured from the site index of the left-most boson to the site index of the right-most boson, regardless of boson type. Note that we converged results in Fig. 5 with respect to A as well as the other four convergence parameters. We present an exemplary configuration space in Fig. 7 to further highlight the aforementioned definitions.

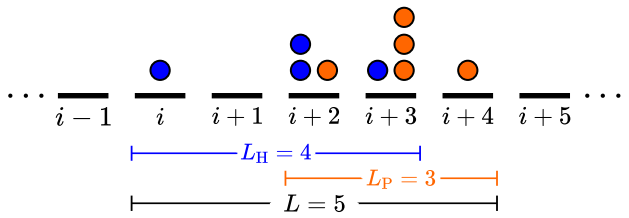

FIG. 7. Example of a configuration of HP bosons corresponding to $n_{\rm H} = [1, 0, 2, 1, 0]$ and $n_{\rm P} = [0, 0, 1, 3, 1]$. Similar to the single boson models, we require that $\sum_j n_{\rm H}^{(j)} \leq N_{\rm H}$, $\sum_j n_{\rm P}^{(j)} \leq N_{\rm P}$, $L_{\rm H} \leq M_{\rm H}$, $L_{\rm P} \leq M_{\rm P}$ and $L \leq A$.

* mrc2215@columbia.edu
† js5530@columbia.edu

[1] G. D. Mahan, *Many-particle Physics* (Springer Science & Business Media, 2013).
[2] L. D. Landau, Phys. Z. Sowjetunion **3**, 664 (1933).
[3] S. Pekar, Zh. Eksp. Teor. Fiz. **16**, 341 (1946).
[4] L. Landau and S. Pekar, Zh. Eksp. Teor. Fiz **18**, 419 (1948).
[5] T. Lee, F. Low, and D. Pines, Phys. Rev. **90**, 297 (1953).
[6] H. Fröhlich, H. Pelzer, and S. Zienau, Lond. Edinb. Dubl. Philos. Mag. J. Sci. **41**, 221 (1950).

[7] H. Fröhlich, Adv. Phys. **3**, 325 (1954).
[8] R. P. Feynman, Phys. Rev. **97**, 660 (1955).
[9] T. Holstein, Ann. Phys. (NY) **8**, 325 (1959).
[10] T. Holstein, Ann. Phys. (NY) **8**, 343 (1959).
[11] A. Alexandrov and N. Mott, Rep. Prog. Phys. **57**, 1197 (1994).
[12] G. Scholes and G. Rumbles, Nat. Mater. **5**, 683696 (2006).
[13] S. H. Park, A. Roy, S. Beaupre, S. Cho, N. Coates, J. S. Moon, D. Moses, M. Leclerc, K. Lee, and A. J. Heeger, Nat. Photonics **3**, 297 (2009).
[14] F. C. Spano, Acc. Chem. Res. **43**, 429 (2010).
[15] K. Miyata, D. Meggiolaro, M. T. Trinh, P. P. Joshi, E. Mosconi, S. C. Jones, F. De Angelis, and X.-Y. Zhu, Sci. Adv. **3**, e1701217 (2017).
[16] S. Trugman, Phys. Rev. B **37**, 1597 (1988).
[17] D. Basov, M. Fogler, and F. G. De Abajo, Science **354** (2016).
[18] J. Feist, J. Galego, and F. J. Garcia-Vidal, ACS Photonics **5**, 205 (2018).
[19] B. Xiang, R. F. Ribeiro, M. Du, L. Chen, Z. Yang, J. Wang, J. Yuen-Zhou, and W. Xiong, Science **368**, 665 (2020).
[20] A. Schirotzek, C.-H. Wu, A. Sommer, and M. W. Zwierlein, Phys. Rev. Lett. **102**, 230402 (2009).
[21] M. Koschorreck, D. Pertot, E. Vogt, B. Fröhlich, M. Feld, and M. Köhl, Nature **485**, 619 (2012).
[22] N. B. Jørgensen, L. Wacker, K. T. Skalmstang, M. M. Parish, J. Levinsen, R. S. Christensen, G. M. Bruun, and J. J. Arlt, Phys. Rev. Lett. **117**, 055302 (2016).
[23] M.-G. Hu, M. J. Van de Graaff, D. Kedar, J. P. Corson, E. A. Cornell, and D. S. Jin, Phys. Rev. Lett. **117**, 055301 (2016).
[24] M. Baggioli and O. Pujolas, Phys. Rev. Lett. **114**, 251602 (2015).
[25] J. Sous and M. Pretko, npj Quantum Mater. **5**, 1 (2020).
[26] J. Sous and M. Pretko, Phys. Rev. B **102**, 214437 (2020).
[27] J. Bonča, S. Trugman, and I. Batistić, Phys. Rev. B **60**, 1633 (1999).
[28] J. Bonča, S. Maekawa, and T. Tohyama, Phys. Rev. B **76**, 035121 (2007).
[29] J. Bonča and S. A. Trugman, Phys. Rev. B **103**, 054304 (2021).
[30] G. De Filippis, V. Cataudella, A. Mishchenko, C. A. Perroni, and N. Nagaosa, Phys. Rev. B **80**, 195104 (2009).
[31] E. Jeckelmann and S. R. White, Phys. Rev. B **57**, 6376 (1998).
[32] C. Zhang, E. Jeckelmann, and S. R. White, Phys. Rev. B **60**, 14092 (1999).
[33] F. Dorfner, L. Vidmar, C. Brockt, E. Jeckelmann, and F. Heidrich-Meisner, Phys. Rev. B **91**, 104302 (2015).
[34] B. Kloss, D. R. Reichman, and R. Tempelaar, Phys. Rev. Lett. **123**, 126601 (2019).
[35] N. V. Prokof'ev and B. V. Svistunov, Phys. Rev. Lett. **81**, 2514 (1998).
[36] N. Prokofev and B. Svistunov, Phys. Rev. B **77**, 020408 (2008).
[37] A. Mishchenko, N. Prokofev, A. Sakamoto, and B. Svistunov, Phys. Rev. B **62**, 6317 (2000).
[38] J. T. Titantah, C. Pierleoni, and S. Ciuchi, Phys. Rev. Lett. **87**, 206406 (2001).
[39] P. Kornilovitch, Phys. Rev. Lett. **81**, 5382 (1998).
[40] We note that recent advances in Diagrammatic Quantum Monte Carlo enable the summation of graphs on the real-frequency axis, see e.g. Refs. 73–75. Such techniques may soon enable the direct calculation of spectral information in the type of models we consider here.
[41] D. Dunn, Can. J. Phys. **53**, 321 (1975).
[42] Determinant Quantum Monte Carlo and its variations[76–79] has been shown to perform well for these problems.
[43] The MA approach has been validated for a large number of systems, including, but not limited to, Holstein,,[44,53,80–83] Peierls,[64] Edwards,[54] and dual-coupled polarons,[70] Holstein[84] and Peierls bipolarons,[65] and has been applied to model experimental systems such as graphene[85] and cuprates,[45] for example.
[44] M. Berciu, Phys. Rev. Lett. **97**, 036402 (2006).
[45] H. Ebrahimnejad, G. A. Sawatzky, and M. Berciu, Nat. Phys. **10**, 951 (2014).
[46] M. M. Möller, G. A. Sawatzky, M. Franz, and M. Berciu, Nat. Commun. **8**, 1 (2017).
[47] W. Su, J. Schrieffer, and A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979).
[48] W.-P. Su, J. Schrieffer, and A. Heeger, Phys. Rev. B **22**, 2099 (1980).
[49] S. Barišić, Phys. Rev. B **5**, 932 (1972).
[50] S. Barišić, Phys. Rev. B **5**, 941 (1972).
[51] K. Hannewald, V. Stojanović, J. Schellekens, P. Bobbert, G. Kresse, and J. Hafner, Phys. Rev. B **69**, 075211 (2004).
[52] The connection to other approaches that utilize exact hierarchies for dynamics in polaron models (for example, the HEOM approach[86,87]) remains to be explored.
[53] M. Berciu and G. L. Goodvin, Phys. Rev. B **76**, 165109 (2007).
[54] M. Berciu and H. Fehske, Phys. Rev. B **82**, 085116 (2010).
[55] Note that the quasi-analytical formulation of the *approximate* MA methods represent a specific case of this general formalism in which $M = 1, 2$ or 3.
[56] E. N. Economou, *Greens Functions in Quantum Physics* (Springer-Verlag, Berlin, 1983).
[57] G. De Filippis, V. Cataudella, A. Mishchenko, and N. Nagaosa, Phys. Rev. B **85**, 094302 (2012).
[58] We note that it appears the largest values for M used in the MA method is 3, while the largest values for M used in the LPBED method is 5.
[59] A. Alvermann, H. Fehske, and S. A. Trugman, Phys. Rev. B **81**, 165113 (2010).
[60] M. Chakraborty and B. I. Min, Phys. Rev. B **88**, 024302 (2013).
[61] A. Macridin, *Ph. D. Thesis* (Rijksuniversiteit Groningen, 2003).
[62] We do not have access to statistical error bars in the DMC calculations.
[63] S. Barišić, J. Labbé, and J. Friedel, Phys. Rev. Lett. **25**, 919 (1970).
[64] D. Marchand, G. De Filippis, V. Cataudella, M. Berciu, N. Nagaosa, N. Prokofev, A. Mishchenko, and P. Stamp, Phys. Rev. Lett. **105**, 266605 (2010).
[65] J. Sous, M. Chakraborty, R. V. Krems, and M. Berciu, Phys. Rev. Lett. **121**, 247001 (2018).
[66] J. Sous, M. Chakraborty, C. Adolphs, R. Krems, and M. Berciu, Sci. Rep. **7**, 1 (2017).
[67] T. C. Berkelbach, M. S. Hybertsen, and D. R. Reichman, J. Chem. Phys. **138**, 114102 (2013).
[68] T. C. Berkelbach, M. S. Hybertsen, and D. R. Reichman, J. Chem. Phys. **138**, 114103 (2013).
[69] J. H. Fetherolf, D. Golež, and T. C. Berkelbach, Phys. Rev. X **10**, 021062 (2020).
[70] D. J. Marchand, P. C. Stamp, and M. Berciu, Phys. Rev. B **95**, 035117 (2017).

[71] G. L. Goodvin, A. S. Mishchenko, and M. Berciu, Phys. Rev. Lett. **107**, 076403 (2011).
[72] J. Sous, M. Berciu, and R. V. Krems, Phys. Rev. A **96**, 063619 (2017).
[73] A. Taheridehkordi, S. Curnoe, and J. LeBlanc, Phys. Rev. B **99**, 035120 (2019).
[74] A. Taheridehkordi, S. Curnoe, and J. LeBlanc, Phys. Rev. B **101**, 125109 (2020).
[75] J. Vučičević and M. Ferrero, Phys. Rev. B **101**, 075113 (2020).
[76] Y.-X. Zhang, W.-T. Chiu, N. Costa, G. Batrouni, and R. Scalettar, Phys. Rev. Lett. **122**, 077602 (2019).
[77] S. Li and S. Johnston, npj Quantum Mater. **5**, 40 (2020).
[78] Z.-X. Li, M. L. Cohen, and D.-H. Lee, Phys. Rev. B **100**, 245105 (2019).
[79] J. Lee, S. Zhang, and D. R. Reichman, Phys. Rev. B **103**, 115123 (2021).
[80] G. L. Goodvin, M. Berciu, and G. A. Sawatzky, Phys. Rev. B **74**, 245104 (2006).
[81] L. Covaci and M. Berciu, EPL **80**, 67001 (2007).
[82] G. L. Goodvin, L. Covaci, and M. Berciu, Phys. Rev. Lett. **103**, 176402 (2009).
[83] M. Berciu, A. S. Mishchenko, and N. Nagaosa, EPL **89**, 37007 (2010).
[84] C. P. J. Adolphs and M. Berciu, Phys. Rev. B **90**, 085149 (2014).
[85] L. Covaci and M. Berciu, Phys. Rev. Lett. **100**, 256405 (2008).
[86] Y. Tanimura and R. Kubo, J. Phys. Soc. Jpn. **58**, 101 (1989).
[87] L. Chen, Y. Zhao, and Y. Tanimura, J. Phys. Chem. Lett. **6**, 3110 (2015).