Structure and Dynamics in Electron-Phonon Coupled Materials

Paul Joseph Pagano Robinson

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2023

Abstract

Structure and Dynamics in Electron-Phonon Coupled Materials

Paul Joseph Pagano Robinson

Electron-phonon interactions (EPIs) are ubiquitous in condensed matter physics and materials science. They are crucial for understanding numerous phenomena, including conventional superconductivity, charge-transport and, most pertinent for this thesis, polaron formation. A polaron is a charge carrier (electron or hole) dressed with a "cloud" of phonons. The polaronic quasiparticle may have vastly different ground- and excited-state properties from that of the bare, constituent charge carrier.

While polarons are well-studied and largely understood in canonical model Hamiltonians, recent advances have made it possible to study more complex, fully *ab initio* systems. Here, the numerically exact methods which are available for some model systems become much more challenging to apply, so accurate approximate methods are a necessity. In this dissertation, we present several advancements in approximate but accurate methods for different polaronic problems and polaron observables.

With respect to polaron dynamics, we focus on low-scaling methods to produce wave vector-dependent single-particle spectral function. We present a thorough study of the accuracy of the second- and fourth-order cumulant expansions (CE) of the electronic Green's function by comparing them against numerically-exact reference data for the one-dimensional Holstein model. We find that the second-order CE is accurate at zero electronic-momentum across a wide range of temperatures, while for non-zero electronic momenta, the CE is only accurate at

high-temperatures. The fourth-order cumulant expansion improves on the dynamics at short times and can improve the spectra; however, it can also introduce non-physical divergences and negative spectral weight. The second-order cumulant expansion is thus a useful tool for determining spectral functions in some instances. However, increasing the order of the CE introduces pathologies that may persist at arbitrarily high-order.

As an alternate approach to improving the CE, we introduce a new self-consistent cumulant expansion (SC-CE) which remedies many of the deficits of the CE. We compare the results for this new approximation against those from the second-order cumulant expansion as well as reference data for the one-dimensional Holstein model. Unlike the CE, the SC-CE can produce accurate spectra across the entire Brillouin-zone, and captures non-perturbative features excellently. The trade-off for this increased accuracy is the introduction of some degree of negative spectral-weight and the potential for rapid divergences in time in some instances. We find that these problems can be minimized, but not completely eliminated in the thermodynamic limit and in more realistic cases where phonon dispersion exists. We also demonstrate how the SC-CE fits into the greater scheme of Green's function methods which approximate the self-energy non-diagrammatically as has recently been proposed by Pandey and Littlewood, and we note the potential applications of the SC-CE both in *ab initio* polaron problems and in general many-body problems.

We finally consider a new method to determine the ground-state structure of the polaron in *ab initio* materials, a topic which has only recently appeared in the literature. We present a new all-coupling variational method based on the Nagy-Markoš variational ansatz for the Fröhlich model. The ansatz is a projected unitary transform which naturally interpolates between the weak-coupling (Lee-Low-Pines) ansatz and the strong-coupling adiabatic ansatz by modulating the momentum conservation of the electron-phonon scattering processes. We demonstrate our *ab initio* Nagy-Markoš ansatz on the Holstein model and the Fröhlich model, and show that it always improves upon the better of the weak or strong coupling result. We consider the *ab initio* case of lithium fluoride (LiF), and find that the ansatz provides accurate polaron binding energies for both

the hole-polaron and the electron-polaron which are classical cases of small and large polarons, respectively. We note how our flexible variational ansatz is an ideal starting point for perturbative energy corrections and cumulant Green's function methods.

Future developments and applications of the efficient methodologies presented in this dissertation may enable quantitative calculations of polarons in large-intermediately coupled *ab initio* systems, such as the lead-halide perovskites and other systems where it has hitherto been difficult to fully understand the effects of the electron-phonon interactions.

Table of Contents

List of Figures

List of Tables

Acknowledgements

There are many people who I would like to thank for their help, encouragement, and support. I am incredibly fortunate to have so many exceptionally supportive mentors, colleagues, friends, and family, who have contributed to my life.

To my doctoral advisor Prof. Dave Reichman. Thank you for meeting with the undergraduate who stopped you in the hallway. That was the beginning of five wonderful years in your group, and you have been the best Ph.D. advisor I could have hoped to have. Whether it was our lockdown Zoom meetings, our office discussions, or our chats around campus, your advice and insight, both related to research and the world at-large, is invaluable.

To my undergraduate advisor Prof. Anastassia Alexandrova of UCLA. You gave me an unrivalable undergraduate research experience. I didn't know what theoretical chemistry was when you took me into your group, but my time there set the foundation for my continued career in the field, and I am deeply grateful.

Prof. Garnet Chan of Caltech and Prof. Eric Neuscamman of UC Berkeley, the summers I spent with your groups were my first experiences in theory development. Getting to know your welcoming groups and your work during my time with you opened my eyes to new areas of research. Thank you both.

To my high-school chemistry teacher Stella Glogover of The Head-Royce School. The way you taught chemistry was unique. Your approach prepared me to learn quantum mechanics in college, and you even lent me Atkins' *Physical Chemistry*. Your classes and mentorship set me on

my path towards chemical physics.

Thank you also to my advisory committee Prof. Richard Friesner, and Prof. Timothy Berkelbach for their helpful discussions and insights during my time here at Columbia.

I would like to thank all of the other wonderful teachers and professors I have had throughout my academic career.

Thank you to the entire Reichman group for their camaraderie, conversation and help over the past five years. To all of the Ph.D. students: Matt Carbone, Ian Dunn, Benedikt Kloss, Haimi Nguyen, Pavan Ravindra, James Shee, Shu Fay Ung, and Jack Weber. And to all of the postdocs: Marco Baity-Jesi, Chitrak Bhadra, Yao-Wen Chang, Zhi-Hao Cui, Jaehyeok Jin, Dmytro Khomenko, Joonho Lee, Martina Lessio, Lachlan Lindoy, Chen Liu, Ankit Mahajan, Arkajit Mandal, Debora Monego, Nathan Ng, Hung Pham, Sambuddha Sanyal, John Sous, and Roel Tempelaar. Whether we were having a group lunch, group bagels, or impromptu discussions, I enjoyed working with each of you.

Matt Carbone, you were my host when I visited as a prospective student and I am happy to have been able to continue asking you questions about science and graduate school for another few years. Chitrak Bhadra, our wide-ranging discussions and excursions to Shakespeare in the Park are fond memories. Ian Dunn, thank you for spending so much time discussing cumulants with me, and helping me learn about polarons initially. Joonho Lee, thank you for all the discussions and insight into *ab initio* systems, I have enjoyed working with you this past year and I look forward to our continued work together.

Thank you to all the other members of the Alexandrova, Neuscamman, and Chan groups whom I had the pleasure of working alongside. David Reilley, thank you for planning excursions across LA and game-nights at your apartment. Mike Nechay, I look back with pleasure at the time we spent baking cakes in the office together and the late night espresso; you are missed.

I would also like to thank the numerous Columbia chemistry staff members for making the department run smoothly. In particular, thank you to Alison Doyle, Betty Cusack, Sheila Skaff, Shaina Johnson, Vivian Polanco, Dwayne Smith, and Ahmed Syed.

Thank you to the National Science Foundation Graduate Research Fellowship for financial support under Grant No. DGE-2036197, and thank you to XSEDE for computational resources under allocation TG-PHY210108.

I would now like to acknowledge all of the friends and colleagues who have enriched my life. Pierre Devlaminck, my companion to some 20-odd Broadway, off-Broadway, or off-off-off Broadway shows. You've become an amazing friend, travel companion, events co-chair and roommate. Jack Weber, the apple picking, trips upstate, and meals with you and Sara were a delight. Spending time with you two and Pierre has made Columbia feel like home. Arden Lee, whether it was chatting out the window during lockdown or spending two years together on the ASGC executive board, I always appreciated your humor and friendship. Allyson Li, you've been a great class rep and friend. I enjoyed almost seeing Bruce Springsteen with you; maybe next time we will get to instead of getting caught in the rain and dodging lightning on the way out. Davida Rios, you are the hostess with the mostess and your evenings of charcuterie, Mario Kart, and Jackbox games at your apartment were always such a fun way to spend time together. Having such a close group in the cohort has been a true blessing.

Bonnie Ko, you've been there for me since high school. Thank you for the many phone calls, trips to the LA and Met Operas, and excursions to visit me in whichever metro area we happen to both live in. Veronica Timpane, since we met backstage in 2010 you have been a constant in my life and my go-to texting companion. You also introduced me to Columbia for the first time. Thank you to the rest of the Thespian Troupe and D&D group: Jonathan Bianchini, Yurie Murayama, Brandon Callender, Kristy Brambila, and Alex Mangus. Thank you for the countless hours of joyful conversation and game-play.

To Connor Sears and Steph Jones and all the other members of The Young New Yorkers' Chorus, the two hours a week I spent singing with you over the past five years were often the highlight of my week and a source of joy.

Jacob D'Eustachio, thank you for taking me out of my comfort zone and providing me with tremendous support and comfort. Your genuine kindness, infectious smile, and sense of adventure

have brought me around the city far-and-wide, into circus performances, and onto the ice. Thank you.

To my aunts, uncles, cousins, and extended family throughout the country, thank you for your encouragement and support; it has meant a great deal to me. To my brother Michael Robinson, I have been so proud and happy to talk about science with you. Thank you for telling me what to think about baseball's new rules, and more importantly, congratulations on your recent B.S. in physics from UCSB! To my grandmother Chris Pagano, thank you for indulging my lengthy and sometimes incoherent explanations of my work and myriad topics during our frequent and extensive phone calls. Who else would I spend hours each week talking about philosophy and constitutional law? To my grandfather Tony Pagano, you were the first professor I met and I aspire to live my life the way you lived yours. I love and miss you. Finally, to my parents Liz Pagano and Ken Robinson. Mom, thank you for the daily phone calls from the car, for visiting to share in breathtaking (for one reason or another) theatrical productions, and for always helping me keep things in perspective. Dad, thank you for making sure I am, like the proverbial boy scout, always prepared with the tools I need whether they be actual tools, like when you helped me move to NYC in the August heat, or metaphorical tools and life skills including how to get the best deals on travel. I would not be here without the endless love and support both of you have given me throughout my life and your dedication to instilling a love of learning in me from the beginning.

For my parents and my grandparents

Chapter 1: Cumulant methods for electron-phonon problems. I. Perturbative expansions

The description of the dynamics of electrons interacting with phonons is a cornerstone topic in condensed matter physics due to its ubiquity and the importance of electron-phonon interactions (EPIs) in determining the properties of solids. Indeed, EPIs are crucial for understanding a wide range of phenomena in solids, including superconductivity, transport properties, and the vibronic satellite structure in emission and absorption spectra, to name just a few.[2–8] Unfortunately, even for simplified canonical EPI models, such as the Holstein, Fröhlich, and Su-Schrieffer-Heeger models which were introduced many decades ago, exact dynamical solutions are largely out of reach.[9–12]

There is a plethora of methods for extracting accurate properties of EPI models which may be useful under different circumstances. Exact ground state and low-lying excited state properties for the Holstein and other models are attainable through diagonalization in a variational Hilbert space (VD).[13–15] Focusing on the electron-phonon dynamics, the one-particle Green's function $\mathcal{G}(k,t)$ has been extensively studied at zero-temperature using exact diagonalization,[16–19] cluster perturbation theory,[20,21] a variational approach,[22] the momentum-averaged approximation,[23,24] and diagrammatic quantum Monte Carlo (DQMC) in conjunction with numerical analytic continuation.[25]

The frontier of finite-temperature dynamical calculations remains less well explored. The increased occupation of higher-lying phonon states at nonzero temperatures renders Fock space methods harder to converge. Only recently has the spectral function ($A(k,\omega) = -\pi^{-1}\Im\mathcal{G}(k,\omega)$) been reported at finite temperature for the Holstein model using VD with the finite-temperature

Lanczos method on 6- and 12-site systems.[26–28] DQMC has provided the temperature dependent mobilities for the Holstein and Fröhlich models; however, as with spectral information this approach is restricted by an ill-conditioned analytic continuation procedure.[5,29] More recently, numerically exact dynamical methods based on DMRG+VD,[30] a generalized cluster expansion,[31] and the Hierarchical Equations of Motion (HEOM)[32–40] approach have been introduced for real-time dynamics in lattice models with EPIs.

For realistic *ab initio* modeling[41–48] of systems with EPIs, many of the exact methods mentioned previously are infeasible. Instead, perturbative approaches are usually employed. However, since each order of perturbation theory exponentially increases the number of self-energy diagrams, it is not practical or computationally efficient to directly compute high-order diagrams in large, realistic systems, and approximate resummations of higher-order terms become essential. The cumulant expansion (CE) approach has been used for this purpose for many years, and was recently combined with density functional perturbation theory to calculate the finite temperature photoemission spectra of MgO, LiF[44], and TiO_2.[43] While the utility of the CE for the calculation of $\mathcal{G}(k,t)$ at finite temperature has been known for years, only a few papers have systematically explored its validity.[49–51]

Motivated by recent exact dynamical results in the finite-temperature Holstein model,[28] we systematically explore the CE in this system as proposed by Dunn in the context of Fröhlich insulators.[49] A similar expansion was also used in conjunction with the Matsubara formalism by Gunnarsson *et al* for describing *zero-temperature* spectral properties of the half-filled Holstein model.[51] In Section 1.1 we introduce the model, as well as the definition of $\mathcal{G}(k,t)$, and the framework of the CE. In Section 1.2 we provide a detailed comparison of the CE in the 6-site Holstein model with exact VD results. This comparison highlights a number of interesting features which demand more detailed investigation. In Section 1.2.1 we discuss errors of the CE that are associated with finite lattice size. In Section 1.2.2 we demonstrate several useful and problematic features of the fourth-order CE. In Section 1.2.3 we analyze the short-time convergence of the CE as well as the limitations of the CE in capturing fine spectral features and long-time behavior. In Section 1.2.4 we

will present results for the spectral function of an infinite system. We conclude with a summary of the main results and the implications of our findings for the use of the CE in the *ab initio* modeling of materials.

1.1 Model and Perturbative Cumulant Expansion

In this section we provide background information needed for the remainder of the paper. While all of the information in this section is well-known, this information is useful for setting notation and for providing a self-contained discussion of the results that follow. Throughout this work we focus only on a very specific model, namely the one-dimensional Holstein model with Einstein phonons and periodic boundary conditions.[4,9,10] We consider only the single particle case, that is a single electron promoted into an otherwise empty band. The model is defined by a system-bath Hamiltonian

$$H = H_e + H_p + V, \tag{1.1}$$

where the kinetic energy term

$$H_e \equiv -t_0 \sum_n \left(a_n^\dagger a_{n+1} + a_n^\dagger a_{n-1} \right) = \sum_k \varepsilon_k a_k^\dagger a_k, \tag{1.2}$$

$$\varepsilon_k = -2t_0 \cos k, \tag{1.3}$$

describes the purely electronic system, and

$$H_p \equiv \omega_0 \sum_n b_n^\dagger b_n = \omega_0 \sum_k b_k^\dagger b_k, \tag{1.4}$$

describes the bath. Lastly,

$$\begin{aligned} V &\equiv g\omega_0 \sum_n a_n^\dagger a_n \left(b_n + b_n^\dagger\right) \\ &= \frac{g\omega_0}{\sqrt{N}} \sum_{kq} a_{k+q}^\dagger a_k \left(b_q + b_{-q}^\dagger\right), \end{aligned} \tag{1.5}$$

accounts for the EPI, which is linear in the bath coordinates. The Holstein model describes the deformation of a discrete lattice,[9,10] reflecting the decoupled nature of sites in a molecular crystal by including only strictly local electron-phonon coupling. In addition, the model further isolates the effects of intermolecular relaxation by ignoring Peierls-like coupling.[12,52–54] For an excellent review that discusses the relation between the Holstein model and continuum models such as the Fröhlich model, see the work of Devreese and Alexandrov.[55]

1.1.1 One-Particle Green's Function

We will focus on the calculation of the finite temperature one-particle (causal [1]) Green's function,[4,56]

$$\mathcal{G}(k,t) \equiv -i\Theta(t) \frac{\mathrm{Tr}\left[e^{-\beta(H-\mu N)} a_k(t) a_k^\dagger(0)\right]}{\mathrm{Tr}\left[e^{-\beta(H-\mu N)}\right]}. \tag{1.6}$$

This quantity is directly related to experimentally measurable quasi-particle spectra as probed by, e.g., photoemission spectroscopy, and can be used to infer transport properties such as charge mobilities in an approximate manner.[4,47,48,57] In addition, the one-particle Green's function provides a testbed for the comparison of numerical methods ranging from the approximate to the exact which may be applied to general electron-phonon problems.[25,29,58,59]

As mentioned above, we study an insulator where the chemical potential μ satisfies $\mu \ll -2|t_0|$ and there is a single electron placed in the conduction band.[49] When this is the case, it is simple to demonstrate that the trace over the many-electron Fock space in Eq. 1.6 can be exactly replaced

[1] In the model studied here with a single electron, this is equivalent to the retarded Green's function

by a trace over zero-electron states and non-interacting phonon states weighted by the canonical density operator for an uncoupled phonon bath,[4,49] namely

$$\mathcal{G}(k,t) = -i\Theta(t)\,\frac{\mathrm{Tr}\left[e^{-\beta H_p} a_k(t) a_k^\dagger(0)\right]}{\mathrm{Tr}\left[e^{-\beta H_p}\right]},$$
$$\equiv -i\Theta(t)\,\langle a_k(t) a_k^\dagger(0)\rangle. \tag{1.7}$$

As will be useful in the next subsection, we also define the quantity

$$\Phi(k,t) \equiv \log\frac{\mathcal{G}(k,t)}{\mathcal{G}_0(k,t)}, \tag{1.8}$$

where

$$\mathcal{G}_0(k,t) = -i\Theta(t)\,\langle e^{iH_e t} a_k e^{-iH_e t} a_k^\dagger\rangle. \tag{1.9}$$

Finally, most comparisons with exact calculations will be made via consideration of the spectral function, defined as

$$A_k(\omega) = -\frac{1}{\pi}\Im\left[\int_{-\infty}^{\infty} dt e^{i\omega t}\mathcal{G}(k,t)\exp(-\gamma t)\right], \tag{1.10}$$

which most closely connects the one-particle Green's function to angle resolved photoemission experiments.[4] Here, γ is a broadening parameter which is used to enable comparison with VD calculations, serves to dampen recurrences for calculations with a small number of sites, and which may be considered as an effective "experimental" resolution for the spectral function itself.

1.1.2 Cumulant Expansion for $\mathcal{G}(k,t)$

Various perturbative approaches have been developed for the explicit calculation of Green's functions such as $\mathcal{G}(k,t)$. The standard approach, which we shall not follow here, follows the now well-established rules of quantum field theory.[4,56] Instead, we will follow the "linked-cluster"

or cumulant approach perhaps first used in the form we employ by Brout and Englert,[60,61] and applied to polaron models originally by Mahan, Dunn and others.[49,50]

First, consider the difference between expansions of the moment generating function (MGF) and the cumulant generating function (CGF). For a (classical) Gaussian random variable X, an expansion of the MGF truncated at second-order will only approximately describe the MGF,

$$\langle e^{-igX} \rangle = 1 - ig\langle X \rangle - \frac{g^2}{2}\langle X^2 \rangle + O\left(g^3\right). \tag{1.11}$$

However, the CGF in this case is exactly described by a second-order expansion,

$$\log\langle e^{-igX} \rangle = -ig\langle X \rangle - \frac{g^2}{2}\left(\langle X^2 \rangle - \langle X \rangle^2\right). \tag{1.12}$$

By taking the logarithm of the Gaussian MGF before expanding, one effectively resums an infinite number higher-order terms in the MGF. The inclusion of even approximate terms of higher-order in the perturbation expansion leads one to expect that a cumulant method can be accurate, especially if the expanded quantity is "nearly Gaussian" in the sense of having small cumulants of order higher than second.

The CE is readily adaptable for perturbative calculation of both thermodynamics and quantum dynamics, where, like the MGF, both the Boltzmann factor and the propagator are exponential functions to be averaged, albeit in time-ordered form.[62] Analogous to a MGF, the one-particle Green's function

$$\mathcal{G}(k,t) = \mathcal{G}_0(k,t)\left\langle e_T^{-i\int_0^t d\tau \hat{V}(\tau)} \right\rangle_k \tag{1.13}$$

$$\mathcal{G}_0(k,t) = -i\Theta(t)e^{-i\varepsilon_k t}, \tag{1.14}$$

can also be calculated approximately via a perturbative calculation of $\left\langle e_T^{-i\int_0^t d\tau \hat{V}(\tau)} \right\rangle_k$ in powers of a coupling constant. Here, the k subscript denotes the average over all one-electron states with electronic momentum k, the T subscript denotes time-ordering, and the hat designates $\hat{V}(\tau)$ as an

operator in the interaction picture.

The $\mathcal{M}^{\text{th}}$-order CE (linked-cluster) for $\mathcal{G}(k,t)$ is given by

$$\mathcal{G}_{\mathcal{M}}(k,t) = \mathcal{G}_0(k,t) \exp\left[\Phi_{\mathcal{M}}(k,t)\right], \tag{1.15}$$

and likewise $A_{\mathcal{M}}(k,\omega)$ is calculated via the Fourier transform of $\mathcal{G}_{\mathcal{M}}(k,t)$. Here, $\Phi_{\mathcal{M}}(k,t)$ is the sum of the cumulants C_μ up to order $\mathcal{M}$. The procedure for constructing the cumulants $C_\mu(k,t)$ from the moments $M_\mu(k,t)$ is well-known.[4] For models of the form given in Eqns. (1.1-1.5), the first few cumulants (up to the fourth-order cumulant) are explicitly given by

$$C_1 = 0, \tag{1.16}$$

$$C_2 = e^{i\varepsilon_k t} M_2, \tag{1.17}$$

$$C_3 = 0, \tag{1.18}$$

$$C_4 = e^{i\varepsilon_k t} M_4 - \frac{1}{2} C_2^2, \tag{1.19}$$

where

$$M_\mu(k,t) = \frac{(-i)^\mu}{\mu!} \int_0^t dt_1 \ldots \int_0^t dt_\mu \times \left\langle T\left\{\hat{a}_k(t)\hat{V}(t_1)\ldots\hat{V}(t_\mu)\hat{a}_k^\dagger(0)\right\}\right\rangle. \tag{1.20}$$

Here $T\{\ldots\}$ is the time-ordering operator which places later times to the left.

Calculation of the second-order and fourth-order CE for the Holstein model requires computing M_2 and M_4. These moments depend on ε_k and the form of the EPI vertex, which is a momentum-independent constant for the Holstein model. Expressions for M_2 and M_4 where we evaluate the time integrals and leave the momentum sums explicit are given in Appendix A.

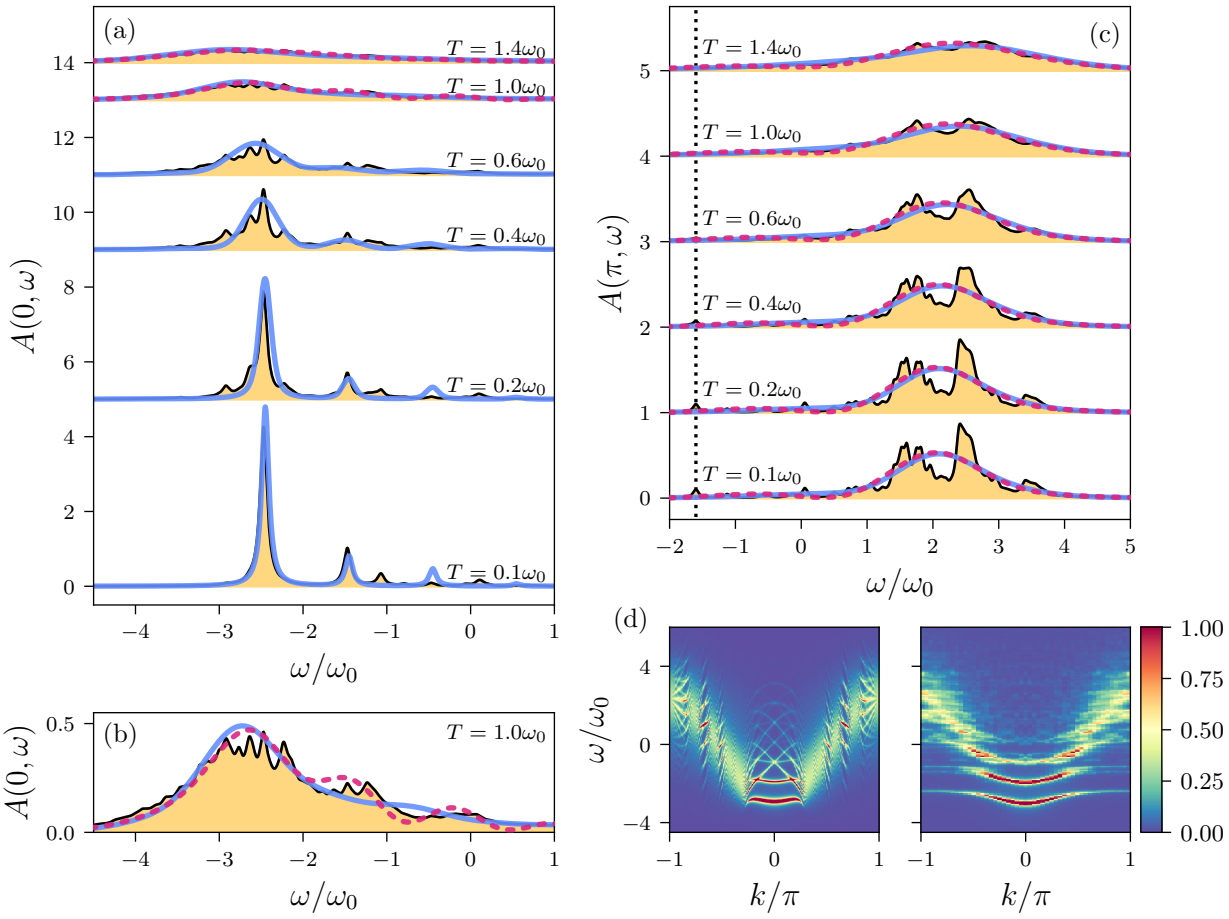

Figure 1.1: Spectral functions $A(\omega)$ for the 6-site Holstein model calculated via the second-order CE (blue solid line), fourth-order CE (pink dashed line), and VD (thin black line and gold shading). Model parameters: $\omega_0 = t_0 = g = 1$. We use $\gamma = 0.05$. VD results are the same as those presented in from Fig. 1(a) of Bonča *et al.*[28] (a) $k = 0$ for a range of temperatures. The fourth-order CE is only presented for the two highest temperatures because it is divergent at lower temperatures. Both orders of the CE capture the most prominent structures of the VD result. (b) $k = 0$ at only $T = \omega_0$. When the fourth-order CE is convergent it slightly corrects the quasiparticle energy and adds additional structure to the peak which better approximates the VD result. (c) $k = \pi$ for a range of temperatures. The VD result demonstrates that there is significant structure in the spectral function and a quasiparticle peak at $\sim -1.5\omega_0$ (dotted black line) while the second- and fourth-order CE only broadly model the structured features in the spectral function. In the fourth-order CE there is an addition peak centered around $-\omega_0$. (d) Heat maps for second-order CE (left) and VD (right, from Ref. 28) at all momenta for $\omega_0 = t_0 = 1$, $T = 0.1$ and $g^2 = 2$. Note how in the CE the extra bright peaks at $k = 0$ disperse into a series of shifted peaks which eventually coalesce into the incoherent polaron peak at $k = \pm\pi$. In contrast, the VD bands extend all the way from $k = 0$ to $k = \pm\pi$.

1.1.3 Convergence of Cumulant Expansion

Let us now examine some aspects of the convergence of the CE for the Holstein model as a function of temperature and EPI coupling strength. Following the definition of the CE in Eq. 1.15, a sufficient condition for the break down of the expansion occurs when successive higher-order cumulants C_μ are not relatively small. Therefore we examine the magnitude C_μ.

For $T \to 0$, the phonon occupation numbers N_0 vanish, such that we may order the terms in Appendix A as functions of the coupling strength for the finite, even cumulants

$$C_{2n} \sim g^{2n}. \tag{1.21}$$

The high temperature limit is slightly more subtle. Before performing the time integrations in Eq. (1.20), contracting the phonon operators yields

$$C_{2n} \sim g^{2n} \prod_{i=1}^{n} \left[\coth\left(\frac{\beta\omega_0}{2}\right) \cos\left(\omega_0\tau_i\right) - i \sin\left(\omega_0\tau_i\right) \right]. \tag{1.22}$$

Taking the high temperature limit,

$$C_{2n} \sim g^{2n} \prod_{i=1}^{n} \frac{2}{\beta\omega_0} \cos\left(\omega_0\tau_i\right) \sim \left(\frac{2g^2}{\beta\omega_0}\right)^n. \tag{1.23}$$

Thus, we see from these two limits that the CE in the time domain breaks down for large T and g such that, schematically, the expansion is governed by

$$\sim \max\left[g^2, \frac{2g^2}{\beta\omega_0}\right]. \tag{1.24}$$

According to Dunn,[49] the CE should also give a reasonable description of $A\left(k,\omega\right)$ at high enough temperatures and strong enough coupling such that ($N_0 = \frac{1}{e^{\beta\omega_0}-1}$)

$$gN_0 \gtrsim 1, \tag{1.25}$$

and/or

$$g\,(N_0 + 1) \gtrsim 2, \tag{1.26}$$

which supersedes the condition for convergence given above. In this regime the long time behavior of $\mathcal{G}\,(k,t)$ is quickly damped and $A\,(k,\omega)$ is broadened to such a degree to as to wash out all sharp spectral features. We shall see evidence of these behaviors in the following sections.

1.2 Results

Recently, Bonča *et al.* published the first exact temperature-dependent spectral function for the single-particle Holstein model using VD.[28] Since the VD and finite temperature Lanczos methods are well detailed in the literature, we do not review them here.[13–15,26–28] Due to the expense of the approach, the finite-temperature calculations of Bonča *et al.* were limited to small system sizes of 6 and 12 sites. It should be noted that at the level of heat maps of the k-dependent spectral function, the 6 and 12 site results differ by only a small amount. On closer inspection, small finite size effects are apparent, as will be discussed below. Recently, new techniques have been developed that are capable of providing exact finite temperature spectra in the single-particle polaron models for larger systems.[30] However comprehensive results for larger lattices in the Holstein model have not yet been published, and thus we compare only to the work of Ref.28.

To judge the accuracy of the CE and compare its performance to the exact results of Bonča *et al.*,[28] we now consider the CE approximation for spectral functions for a 6-site Holstein model at the band bottom ($k = 0$) and the band edge ($k = \pi$). We restrict the comparison to the intermediate coupling regime where $g = t = \omega_0$. In Fig. 1.1 we plot the $k = 0$ and $k = \pi$ second- and fourth-order CE spectral functions along with data from Ref. 28. For both momenta, the second-order CE captures the broad structure of the spectral function reasonably well, and for $k = 0$ the results are quantitatively accurate at both $T = 0$ and at high temperatures $T \geq \omega_0$. In particular, the second-order CE captures the quasi-particle peak and the first vibronic satellite peak at $T = 0$ in

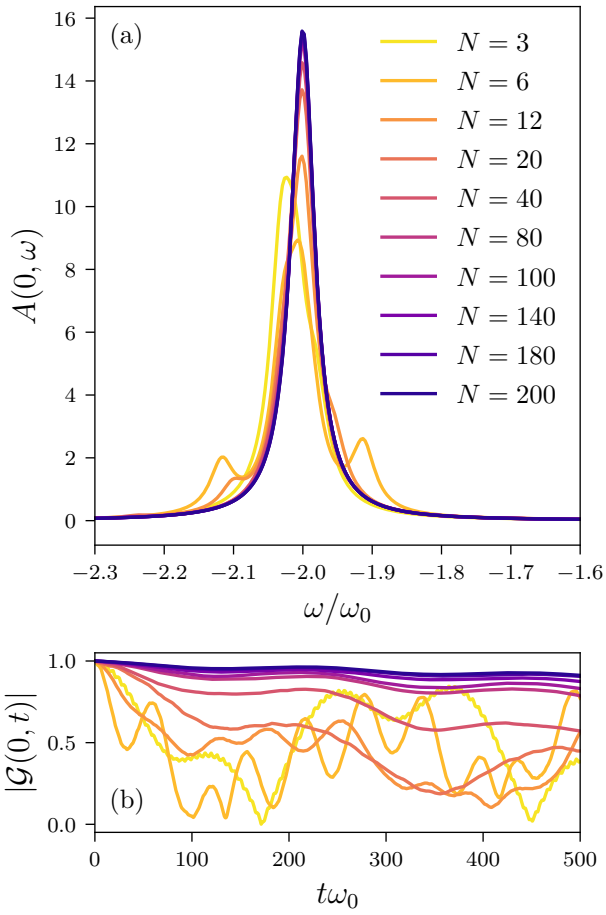

Figure 1.2: One-phonon exact diagonalization approximation results of Holstein model at increasing system sizes. $\omega_0 = t_0 = T = 1; g = 0.25$ (a) Spectral function as a function of system size demonstrates disappearing fine structure (b) Magnitude of the Green's function in time as a function of system size demonstrates disappearance of sharp beats.

excellent agreement with VD. At higher temperatures, the central features of these peaks are well captured, however the fine structure superposed on the quasi-particle peak exhibited by the exact VD spectra is absent in the CE spectra. We will see below that this fine structure is a consequence of the small lattice size, and thus the CE approximation does not properly capture this type of finite lattice effect.

At the band edge ($k = \pi$) the results produced by the second-order CE are not as encouraging, as illustrated in in Fig. 1.1(c). The VD data has two important features: a quasi-particle peak at low energy ($\sim -1.5\omega_0$), and a broad vibronic wing with a split peak structure centered around $2\omega_0$.

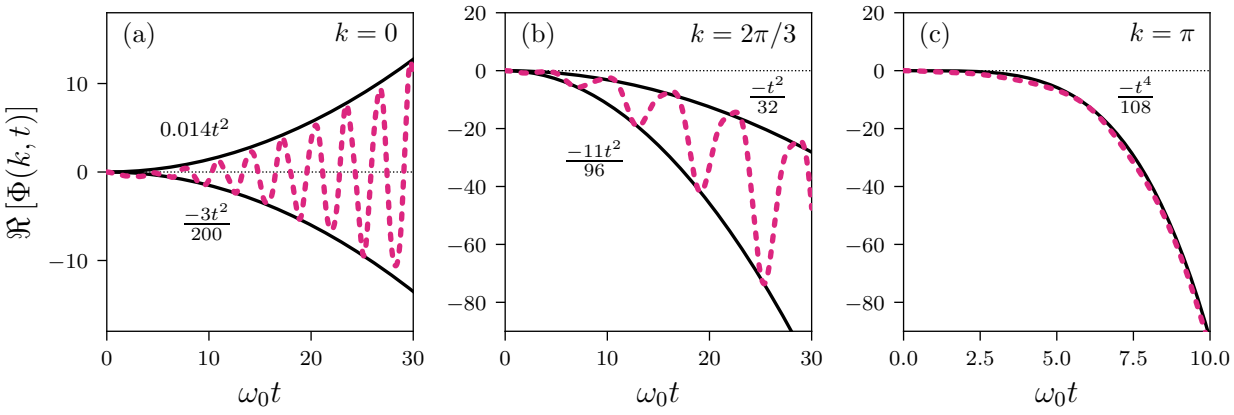

Figure 1.3: The fourth-order CE numerical results (dashed red) for $N = 6$ at $g = \omega_0 = t_0 = 1$ and $T = 0$ compared with the leading order envelope functions (solid black) at (a) $k = 0$, (b) $k = 2\pi/3$, and (c) $k = \pi$. The functional form of the envelope is given in each figure and describes the general shape of the numerical results. A discussion of the origin of these envelope functions is found in the main text and Appendix A.

The second-order CE misses the peak structure of the exact spectral function entirely, and instead can be described as a single broad peak centered near the average value of the peak intensity found in the exact VD result. Again, as temperature increases and the features of the spectral function broaden, the CE result becomes more and more accurate, reflecting the fact that the CE properly accounts for the spectral bandwidth even for $k = \pi$. The fact that the CE is accurate away from $k = 0$ for temperatures $T \geq \omega_0$ has important practical implications for the use of the CE to study transport phenomena, a topic we will return to before concluding.

The difference in accuracy of the CE between the $k = 0$ and the $k = \pi$ cases is seen generally across the full range of wave vectors. More specifically, we find that the $k = 0$ case is the only case for which the CE is in quantitative agreement with exact VD results for low temperatures. A full comparison of the exact and approximate CE spectral functions across the entire band can be found in Fig. 1.1(d). Here, several features are notable. The fact that for $k = 0$ the CE predicts a prominent series of small peaks beyond the first satellite spaced by ω_0, in reasonable agreement with the exact VD results, is actually the result of an incorrect intensity crossing structure which renders the satellite behavior for all $k \neq 0$ inaccurate. As we will discuss below, this behavior is the result of the manner in which the CE approximates higher-order multi-phonon scattering terms.

Note as well that for $k \neq 0$ there is fine structure in the high intensity band. This behavior is a finite size effect of the CE which vanishes when the number of lattice sites tends to infinity, as we will discuss later in this work. Such finite size effects are distinct from the true finite sized behavior exhibited in the VD results discussed above, and do not reflect the correct formation of structure exhibited in the satellite region of the exact spectra.

We next turn to a discussion of the corrections to the second-order CE provided by the fourth-order CE. In general, when the fourth-order CE is well-defined for the parameter regime of the Holstein model studied here, it only subtly alters the behavior found from the second CE. In Fig. 1.1(b) we provide a close-up of the $k = 0$, $T = \omega_0$ case found in Fig. 1.1(a). It can be observed that in general the fourth-order CE indeed redistributes spectral weight correctly, with the exception of a small region of negative spectral weight for $\omega > 0$. The fact that the fourth-order CE does not guarantee positivity of the spectra has been discussed in several previous works.[51,63,64] Gunnarsson *et al.* attribute this problem to the particular analytical form of the terms retained at fourth-order in the CE. More problematic is the fact that for some parameter regimes the fourth-order CE is not well defined due to unbounded growth in the time domain of some of the terms in the expansion.[51] We will see below that these terms take a similar form to those pointed out as contributing to negative spectral weight by Gunnarsson *et al.* Thus, these two issues appear to be connected. In Fig. 1.1(a) fourth-order CE results are not shown for $k = 0$ and $T \leq 0.6\omega_0$ due to the divergence in the time domain of the fourth cumulant. In the next three subsections we will investigate more deeply several of the features exposed here for the finite-sized Holstein chain before turning to the CE in the thermodynamic limit.

1.2.1 Finite Size Effects

In the discussion of results for the 6-site lattice presented above, we mentioned several aspects of both the exact VD results as well as the results of the CE that warrant further discussion. In this subsection we focus on one such feature, namely the role played by the small lattice size, and the implications for the failure of the CE to capture these effects. In particular, we focus now on the

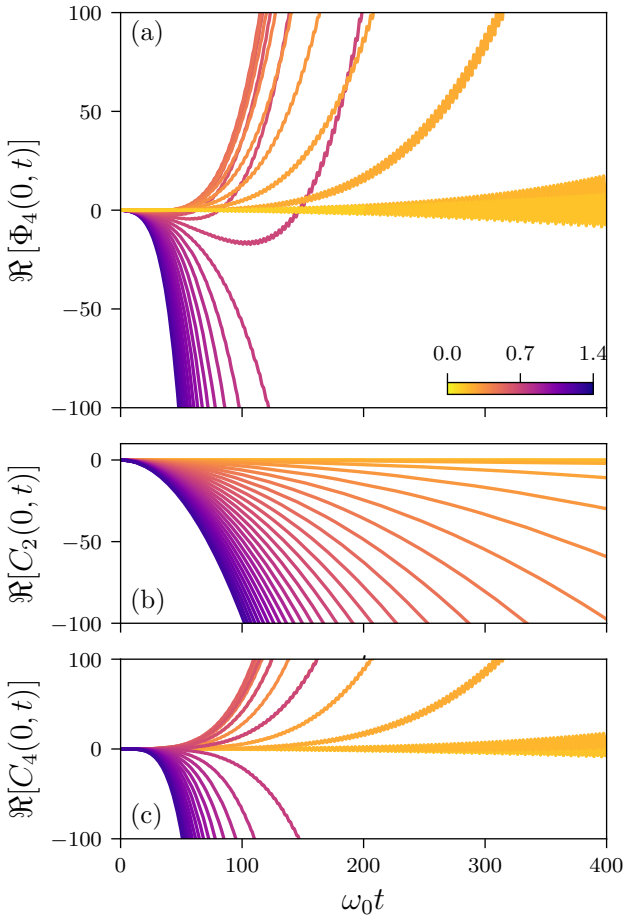

Figure 1.4: Breakdown of the time dependence of (a) $\Phi_4(0,t)$ and its dependence on (b) $C_2(0,t)$ and (c) $C_4(0,t)$, highlighting the temperature driven transition from divergent to non-divergent $\mathcal{G}(0,t)$. Here, $\omega_0 = t_0 = 1, g = 0.25$, $k = 0$ and $T = [0, 1.4]$. $\Phi_4(0,t)$ transitions from divergent to non-divergent between $T = 0.7$ and $T = 0.75$, and because $C_2(0,t)$ does not predict a divergent $\mathcal{G}(0,t)$ for any temperature, the transition from divergence or non-divergence is dictated by the transition in $C_4(0,t)$.

small ripples that appear in the main quasi-particle region of the VD spectra for a 6-site system in the regime $0.1\omega_0 \geq T \geq 0.6\omega_0$. We explicitly demonstrate that these features are due to the small lattice size, and thus the failure of the CE to capture this type of finite size effect is not relevant in the thermodynamic limit. Indeed, the expected change in the spectral function in transitioning from small finite size systems to the $N \to \infty$ limit plotted in the manner of Fig. 1(e) will largely appear confined to smoothing the intensity modulation of the most prominent spectral features.

To shed light on the type of finite size effects expected to arise in small lattice systems, and to

reveal why these effects show up prominently only at low to intermediate values of the temperature, we turn to exact diagonalization for finite sized systems in the one-phonon sector. This approach is outlined in Appendix C. Due to the strong restriction on the phonon excitations allowed, we do not aim for quantitative results and merely expose the qualitative nature of the spectral features associated with the quasi-particle peak as the system is tuned from finite to infinite lattice size.

In Fig. 1.2 we show the behavior of the spectral function and the real-time behavior of the one-particle Green's function for $k = 0$ for a weakly coupled electron-phonon system ($g = 0.25$) with parameters $\omega_0 = t = T = 1$. The behavior of the Green's function in the time domain reveals the existence of higher frequency beating behavior superposed on lower frequency oscillations. The high frequency behavior is related to recurrences due to transitions associated with the discreteness of the spectrum in the small N limit. Such behavior will manifest most strongly at intermediate temperatures, where thermally-populated low-lying states can participate in producing the observed beating behavior but where the temperature is not so high that damping effects dominate the decay of the Green's function. We note that already by $N = 20$ the erratic high frequency behavior vanishes, although finite size effects are still present. In the frequency domain, spectral functions of finite size systems with $N \leq 12$ exhibit small secondary peaks similar to the behavior exhibited in Fig. 1(a).

The finite size behavior and the inability of the CE to capture it is similar to that seen in purely electronic systems. In particular, McClain *et al.* have studied the spectral function of the electron gas with coupled-cluster and cumulant-based techniques in finite sized systems.[65] Here the CE also shows a relative inability to reproduce structure associated with the discrete nature of finite sized systems. We emphasize that the structure of the spectral function seen in small systems in the Holstein model discussed in this section are distinct from larger scale features for $k \neq 0$ such as that seen in Fig. 1(c) which are also absent in low-order CE calculations. The more important failure to reproduce these larger scale features is expected to persist in the $N \to \infty$ limit.

N	g	T/ω_0	$\omega_0 t_{\Re[\Phi_4]>0}$
6	1.00	0.40	7.3
6	1.00	0.60	10.7
6	1.00	0.70	23.2
6	1.00	0.72	80.0
6	1.00	0.73	$> 10^4$
6	0.75	0.00	7.3
6	0.50	0.00	10.7
6	0.25	0.00	20.2
6	0.10	0.00	51.6
6	1.00	0.00	7.1
12	1.00	0.00	10.2
50	1.00	0.00	24.9
100	1.00	0.00	49.0
150	1.00	0.00	73.3

Table 1.1: $t_{\Re[\Phi_4]>0}$ for varying system sizes, coupling strengths and temperatures. Increasing the system size or decreasing the coupling can push the onset of the divergence to longer times. Raising the temperature past some transition temperature fully removes the divergence. Model parameters: $t_0 = \omega_0 = 1; k = 0$

1.2.2 Divergences in the fourth-order CE

Fig. 1(b) illustrates that, aside from the unphysical appearance of regions with a (small) negative spectral weight for $\omega \gtrsim 1.1\omega_0$ (not shown), the fourth-order CE improves upon the second-order CE for the spectral function at $k = 0$ at higher temperatures. However, as mentioned above, for the same k value at low temperatures, the fourth-order CE is divergent at longer times, and thus truncated higher-order CEs cannot always be used to systematically improve upon low-order results. Here we focus on the factors which can shift the fourth-order CE between well-behaved and divergent at long times to better understand where corrections to the second-order CE are applicable. We will demonstrate that in the Holstein model, the divergence of the fourth-order CE depends intimately on the wave vector, system size and temperature under consideration, and is closely connected to the issue of negative spectral weight first pointed out for this model by Gunnarsson *et al.*[51]

The CE is an exponential function of the quantity $\Phi(k,t)$ defined in Sec. 1B for which physical results require $\Re\left[\Phi(k,t)\right] \leq 0$ for all times. In addition, at finite temperatures the requirement

$\lim_{t\to\infty} \Re\left[\Phi(k,t)\right] \to -\infty$ must hold, reflecting the finite lifetime of quasi-particles. It is easily checked that the second-order CE always satisfies these requirements. In particular, $-g^2t^2(2N_0 + 1) \leq \Re[\Phi_2(k,t)] \leq 0$, and thus the second-order CE never diverges.

With these considerations in hand, we focus on the fourth-order CE, characterizing the divergence of the fourth-order term $\Phi_4(k,t)$ by the quantity $t_{\Re[\Phi_4]>0}$, which marks the earliest time where $\Re[\Phi_4] > 0$. In Table 1.1 we compile $t_{\Re[\Phi_4]>0}$ for the Holstein model with the same parameters as found in Fig. 1 for the spectral function at $k = 0$ as a function of temperature and the number of lattice sites. Several aspects of the data are worthy of note. We focus first on the fact that as T approaches a temperature between $T = 0.72$ and $T = 0.73$, the divergence is abruptly pushed from a finite time to infinite time for all practical purposes. This behavior is consistent with the results plotted in Fig. 1.1(a), where only the cases $T = 1.0\omega_0$ and $T = 1.4\omega_0$ have non-divergent fourth-order CE results.

The root of this abrupt behavioral change in the long-time limit of $\Phi_4(k,t)$ becomes manifest upon examining the analytical forms of the individual cumulant terms. Details may be found in Appendix A. Direct examination of $\Phi_4(k,t)$ at zero temperature reveals terms of the form $te^{i\alpha t}$ where α is some real number. Care must be taken with the evaluation of $C_4(k,t)$, as it contains many apparent singularities which are actually well-defined when appropriate limits are taken. Depending on the particular limit and the values of t_0 and ω_0, $C_4(k,t)$ contains real-valued terms which may diverge linearly, quadratically, or quartically in time. Some of the seemingly divergent terms of order t^2 in $M_4(k,t)$ are exactly cancelled by the transformation from moments to cumulants in Eq. 1.19.

Here, we present the leading-order contributions to the fourth cumulant for $N = 6, t_0 = \omega_0, T = 0$ and $k = 0$,

$$\Phi_4(0,t) = \frac{-g^4t^2\left(2e^{-5it\omega_0} + 25e^{-2it\omega_0}\right)}{1800\omega_0^2} + o(t^2). \tag{1.27}$$

Since the exponential functions in Eq. 1.27 contain no real damping, $\Phi_4(0,t)$ diverges quadratically in time. The envelope growth rapidly becomes the only significant term in the expansion. This is illustrated in Fig. 1.3(a) where the case of $g = \omega_0 = 1$ is explicitly shown, and $\Re\left[\Phi(0,t)\right]$

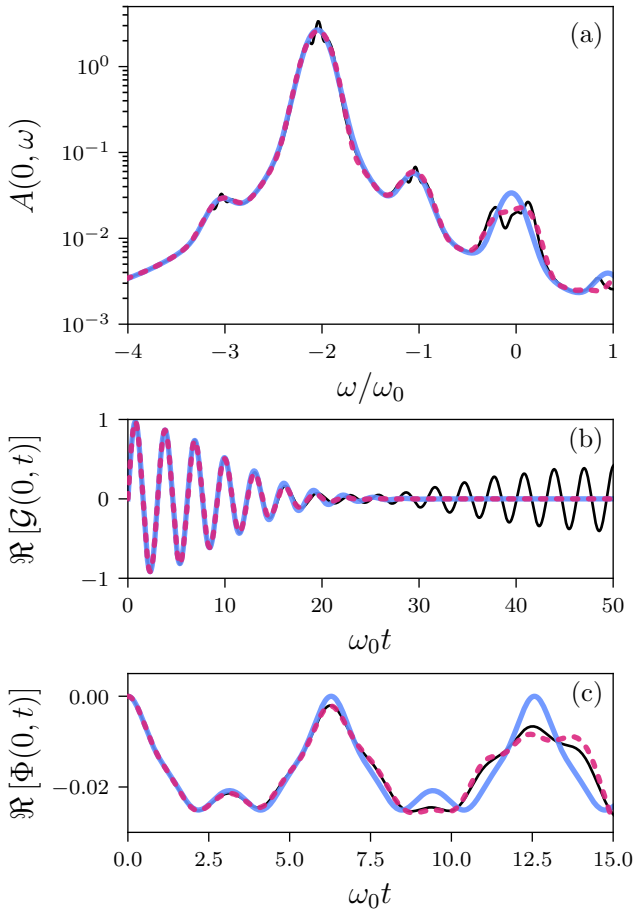

Figure 1.5: (a): Comparison of A_2 (blue solid line), A_4 (pink dashed line) and $A_{\infty/\mathrm{HEOM}}$ (thin black line and gold fill) for $T = 1.0\omega_0$. (b): Comparison of the real parts of $\mathcal{G}_2(t)$ and $\mathcal{G}_4(t)$ with $\mathcal{G}_\infty(t)$ for $T = 1.0\omega_0$. (c): Comparison of $\Phi_2(t)$ and $\Phi_4(t)$ with $\Phi_\infty(t)$ for $T = 0.0\omega_0$. Model parameters: $N = 6; k = 0; \omega_0 = t_0 = 1$, $g = 0.25$, $\gamma = 0.04$ (spectral function only). HEOM calculations are performed using modified versions of *PHI*[66] and *pyrho*.[67]

indeed grows along the upper and lower bounds of Eq. 1.27. Although the case presented here is quadratically divergent, this is specific to $t_0 = \omega_0$. A more general version of Eq. 1.27 contains only linear divergences (albeit many of them), while the proper evaluation of the limit $t_0 \to \omega_0$ introduces quadratically growing terms. It is worth noting that terms of this general form were also found by Gunnarsson *et al.* in their zero-temperature CE study of similar polaron models.[51] These authors determined that such terms give rise problematic negative spectral weight, however they did not report a divergent behavior in $G(k,t)$. We will return to the issue of negative spectral

weight below.

Inspection of Eq. 1.27 also makes clear why decreasing the value of the electron-phonon coupling delays the onset of the divergence, as seen in Table 1. The divergent portion of the fourth-order cumulant is scaled by g^4, so it is quite expected that smaller g decreases the time scale of divergence. This is demonstrated in Table 1.1 with the example of a six site system where decreasing the value of g indeed increases the divergence time.

In a similar vein, we can examine the terms in $\Phi_4(\pi, t)$ and $\Phi_4(2\pi/3, t)$ to understand why, for $N = 6$ and $\omega_0 = t_0$, the fourth-order CE is non-divergent at zero temperature. The case of $k = \frac{2\pi}{3}$ is very similar in form to $k = 0$ but critically contains an extra constant in the expression for the leading term, namely

$$\Phi_4\left(\frac{2\pi}{3}, t\right) = \frac{-g^4 t^2 \left(4e^{it\omega_0} + e^{-2it\omega} + 6\right)}{96\omega_0^2} + o(t^2). \tag{1.28}$$

Because of the $-6t^2g^4/96\omega_0^2$ term in Eq. 1.28, $\Re[\Phi_4(2\pi/3, t)]$ tends towards negative infinity quadratically, which corresponds to a strongly damped $\mathcal{G}(2\pi/3, t)$. This is shown in Fig. 1.3(b) for the case of $g = \omega_0 = t_0 = 1$. As in the case of $k = 0$, the avoidance of an unphysical divergence is specific to the choice that $t_0 = \omega_0$, so it merely serves as a demonstration of one way in which the terms in the fourth-order CE conspire to avoid divergences issues for a specific set of parameters at a specific wave vector.

While these examples demonstrate how terms quadratic in time in the expansion of $\Phi_4(k, t)$ can lead to a convergent or divergent approximation to $\mathcal{G}(k, t)$, leading terms of even higher-order in time are possible. An example of this occurs at $k = \pi$, where in Fig. 1.1(c) we observe that the fourth-order CE is always well behaved. Once again taking $N = 6, \omega_0 = t_0$ and $T = 0$, we can show that the leading-order divergence for $k = \pi$ is

$$\Phi_4(\pi, t) = \frac{-g^4 t^4}{108} + o(t^4). \tag{1.29}$$

Unlike the behaviour expressed by Eq. 1.27, the leading-order divergence of $\Phi_4(\pi, t)$ is not oscil-

latory; rather, it is strictly negative and quartic, rendering irrelevant any oscillating and increasing terms of slower growth. This qualitative difference in behavior between the $k = 0$ and $k = \pi$ cases ensures a non-divergent Green's function at the band edge even at low temperatures.

The divergent terms discussed above are related to the double pole structure described by Gunnarsson *et al.*[51] Analyzing the behavior of the fourth cumulant in the short-time limit, a function of the form $\exp\left(\alpha t^n e^{-zt} + \dots\right)$ can be linearized to give $1 + \alpha t^n e^{-zt} + \dots$. Here, α is a complex coefficient, n is a positive integer, and z is purely imaginary. The Fourier transform of a function of this form will be proportional to the n^{th} derivative of a delta function centered at z, and this feature will be present in the spectra even if the overall CE is convergent. While this argument is approximate, as it relies on the short-time dynamics, it nonetheless makes clear the connection between negative spectral weight and the potential for divergent behavior in the fourth- (and presumably higher)-order CE. Since the CE to all orders is exact, the cancellation of these problematic terms at high-orders must occur, albeit clearly in a complicated manner which likely obviates the possibility of removing such terms in lower-order versions of the the expansion in a reliable way. In most applications we are interested in the $N \to \infty$ limit, and here, as shown in Sec. 1.2.4, we note that for some wave-vectors divergences are suppressed with increased system size. As empirically demonstrated in Table 1.1 for $k = 0$, the onset time of the divergence grows linearly with the system size. To understand this behavior, we again consider which terms are present in the summation of the expressions for the fourth-order cumulant. As the system size increases, the number of terms in the momentum sums over q_1 and q_2 grows as N^2, while the weight of each individual term decreases in magnitude as N^{-2}. Singularities in $\Phi_4(k, t)$ that produce quadratic growth in time only occur when specific energetic conditions are met. A few examples of these conditions are $\epsilon_{k+q_1} - \epsilon_k + \omega_0 = 0$, $\epsilon_{k+q_1} - \epsilon_{k+q_1+q_2} + \omega_0 = 0$, and $\epsilon_{k+q_1+q_2} - \epsilon_k + 2\omega_0$=0. Crucially, these conditions exist only on one-dimensional lines in the space of q_1 and q_2. Thus, the ratio of the non-singular evaluations to the total number of evaluations falls of at least as $1/N$, as we observe numerically. In particular, for $k = 0$ the non-singular term occurs in 61.1% of evaluations of the momentum sum for $N = 6$, 99.7% of the evaluations for $N = 600$, and 99.8% of the evaluations

for $N = 1200$. Thus, for very large system sizes we can drop all of the singular cases of q_1 and q_2 by recognizing that the ratio of singular cases to non-singular disappears as $\sim 1/N$. The true thermodynamic limit of $C_4(k,t)$ corresponds to a principle value integral over momentum space with real terms at most linearly divergent in t.

While the preceding argument justifies why the non-linear in time divergences present in small systems disappear as $N \to \infty$ for $k = 0$, it does not explain why linear time divergences do not appear. We now heuristically argue that a distinct type of behaviour suppresses divergent growth in time as $N \to \infty$ for some wave vectors.

With the remaining linear terms proportional to $te^{i\alpha t}$ where $\alpha \in \mathfrak{R}$, the momentum sum in the fourth-order cumulant becomes one of many oscillating exponential functions, each with weight N^{-2}. As the frequencies in the exponential become continuously distributed, interference of the many out of phase components can delay the onset of divergence to arbitrarily long times. It must be noted that this cancellation depends on specific properties of the unperturbed energy dispersion which are not trivially satisfied at all k. Nonetheless, we find numerically that as N tends towards an infinite number of sites for both $k = 0$ and $k = \pi$, the first constructive beat is pushed to $t = \infty$, hence the results in section 1.2.4 are well-behaved at all temperatures for those values of k unlike for the case $N = 6$. We have not been able to uncover a deeper analytical argument for this behaviour, and must appeal to numerical heuristics, which are presented in detail in Sec. 1.2.4.

Since the infinite-order CE provides an exact representation of the dynamics, it must be true that even higher-order cumulant terms eventually conspire to remove the divergent terms at lower-orders. However, because $C_n(k,t) \propto g^n$, we know that the higher-order terms can not directly cancel the lower-order divergences, and instead must form the series representation of a well-behaved exact $\Phi(k,t)$. There are a number of approaches one could attempt to remove these divergences, but all presume some knowledge of the higher-order terms in the CE. For an approximate means of resumming higher-order cumulants, we refer the reader to the self-consistent cumulant approximation in the companion paper.[68]

1.2.3 Convergence to Exact Result: Short-Time Analysis

The results in the previous subsection illustrate that the use of the fourth-order CE can improve agreement with exact benchmarks (Fig. 1(b)) but can also lead to unphysical results associated with instabilities and negative spectral weight. While carrying out the CE to infinite-orders yields exact results, it is clear that the manner in which convergence occurs is complicated. Here, we focus on the time domain, explicitly illustrating how higher-order expansions *always* systematically improve the accuracy of the short-time behavior. To carry out this comparison, we employ the numerically exact "Hierarchical Equations-of-Motion" (HEOM) method.[32–39] This approach provides rapid convergence to the exact result for models such as the spin-boson model. For the one-dimensional Holstein model, exact convergence for finite times is attainable for weak-to-moderate coupling strengths in moderately-sized chains.[39] Since this method may be unfamiliar to some readers, a brief description is provided in Appendix B.

We first work at weak coupling ($\omega_0 = t_0 = 1, g = 0.25$) and high temperatures ($T = 1.0$) where we can easily converge the exact HEOM results for times sufficient to provide the full spectral function with minimal artificial damping. In Fig. 1.5(a) we show results for $N = 6$ and $k = 0$ which are consistent with the behavior found in Fig. 1. In particular, the fourth-order CE improves subtly on the second-order result, bringing the theory into quantitative agreement with exact spectral function, with the exception of very small secondary peak structure visible at $\omega \sim -2$ and $\omega \sim 0$. As expected from the discussion in Sec.IIA, this behavior is due to finite size recurrences which are expected to vanish as $N \to \infty$. Fig. 1.5(b) illustrates the behavior underlying the spectral function in the frequency domain. In particular, a large visible recurrence starting at $\omega_0 t \sim 20$, which is missed by the second- and fourth-order CEs, can be observed.

On the scale of Fig. 1.5(b), it is nearly impossible to parse what the fourth-order CE provides over the second-order CE to improve the distribution of spectral weight as seen in Fig. 1.5(a). However, by focusing on the function $\Phi(k, t)$ directly, one clearly observes the systematic improvement provided by the fourth-order CE over the second-order CE. In Table 1.2, we consider two temperatures and two coupling strengths, along with the time, $t_{\mathcal{M}}$, after which $|\mathcal{G}_{\mathcal{M}}(t) - \mathcal{G}_{\infty}(t)| > 5\times10^{-4}$,

g	T/ω_0	$t_2\omega_0$	$t_4\omega_0$
0.25	0.0	2.90	8.70
0.25	1.0	1.35	5.55
1.0	0.0	0.50	0.85
1.0	1.0	0.35	0.65

Table 1.2: $t_{\mathcal{M}}$ at two EPI strengths and temperatures. In all cases, raising either the temperature or the coupling decreases $t_{\mathcal{M}}$ and $t_2 < t_4$, indicating that the fourth-order CE improves upon the second-order CE at short times. Model parameters: $N = 6; \omega_0 = t_0 = 1$. HEOM calculations were performed using modified versions of *PHI*[66] and *pyrho*.[67]

where $\mathcal{M}$ is the order of the CE, and $\mathcal{G}_\infty$ is given by the exact HEOM result. This improvement in short-time behavior is manifest in Fig. 1.5(c), which illustrates the improved description of $\Phi(0, \omega)$ for parameters such that the long-time limit of the fourth-order CE is divergent. In particular, in all cases, $t_4 > t_2$ indicating that the fourth-order CE improves upon second-order CE. Clearly, the long-time pathological behavior of the fourth-order CE does not corrupt the increase in accuracy of the short-time behavior of the cumulant generating function.

The above discussion suggests that convergence of the CE occurs in the time domain such that the short-time behavior can be systematically converged for longer and longer times, while concomitantly longer-time anomalies in $\Phi(k, t)$ must resum into functions which behave in a non-singular manner. It is difficult to guess the form taken by such functions from just the first two terms in the expansion. In this sense the fourth-order CE does not appear to be generically useful. In the companion paper, we will present a self-consistent cumulant scheme that, while still suffering from some of the ill-effects introduced by the fourth-order CE, does provide access to non-perturbative behavior that appears to be completely out of reach of low-order CEs.[68]

1.2.4 Thermodynamic Limit

Inspired by the possible suppression of physical fine structure and the elimination of poorly behaved spurious oscillations in the infinite-system limit, we now continue in the spirit of Dunn's continuum calculation on the Fröhlich model[49] to treat the finite-temperature infinite Holstein model in the thermodynamic limit using the CE.

We start by investigating the finite-k behaviour of $A_2(k, \omega)$ and $A_4(k, \omega)$ in Fig. 1.6 where

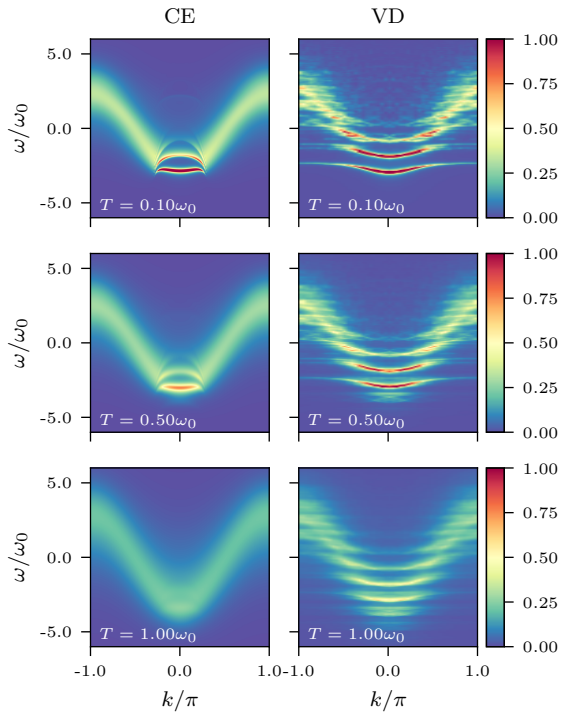

Figure 1.6: Heat maps of $A_2(k,\omega)$ for $N = 600$ (left), and $A_{\infty/\mathrm{VD}}(k,\omega)$ for $N = 6$ (right, from Ref. 28) for several temperatures. Note that for the CE there is a sharp transition at around $\pi/3$ from a clear quasiparticle peak to a incoherent spectrum. Additionally, note that compared to the VD results that the band curvature is qualitatively incorrect for the vibronic peaks at small k. Model parameters: $\omega_0 = t_0 = 1$, $g^2 = 2$, and $\gamma = 0.05$.

$A_n(k,\omega)$ denotes the n^{th}-order CE approximation. Here, $N = 600$ for the CE calculations while $N = 6$ for the VD results. We expect that on the scale of these plots, finite size effects in the VD results are small as discussed in Ref. 28. Note, however, the large changes that appear within $A_2(k,\omega)$ as N is increased. While $A_2(0,\omega)$ is accurate compared to the exact result, even near $k = 0$ the curvature of the bands that represent satellite peaks are described in a qualitatively incorrect manner. Further, while the finite size effects described in section 1.2.1 do vanish after an abrupt change of behaviour which occurs at $\epsilon_k = \omega_0$, they are replaced with a single incoherent band centered on ϵ_k. Thus, the prominent satellite structure for $k > \omega_0$ in the exact results is completely absent in $A_2(k,\omega)$.

The fourth-order CE (not plotted in Fig. 1.6) produces an $A(k,\omega)$ broadly similar to that of the

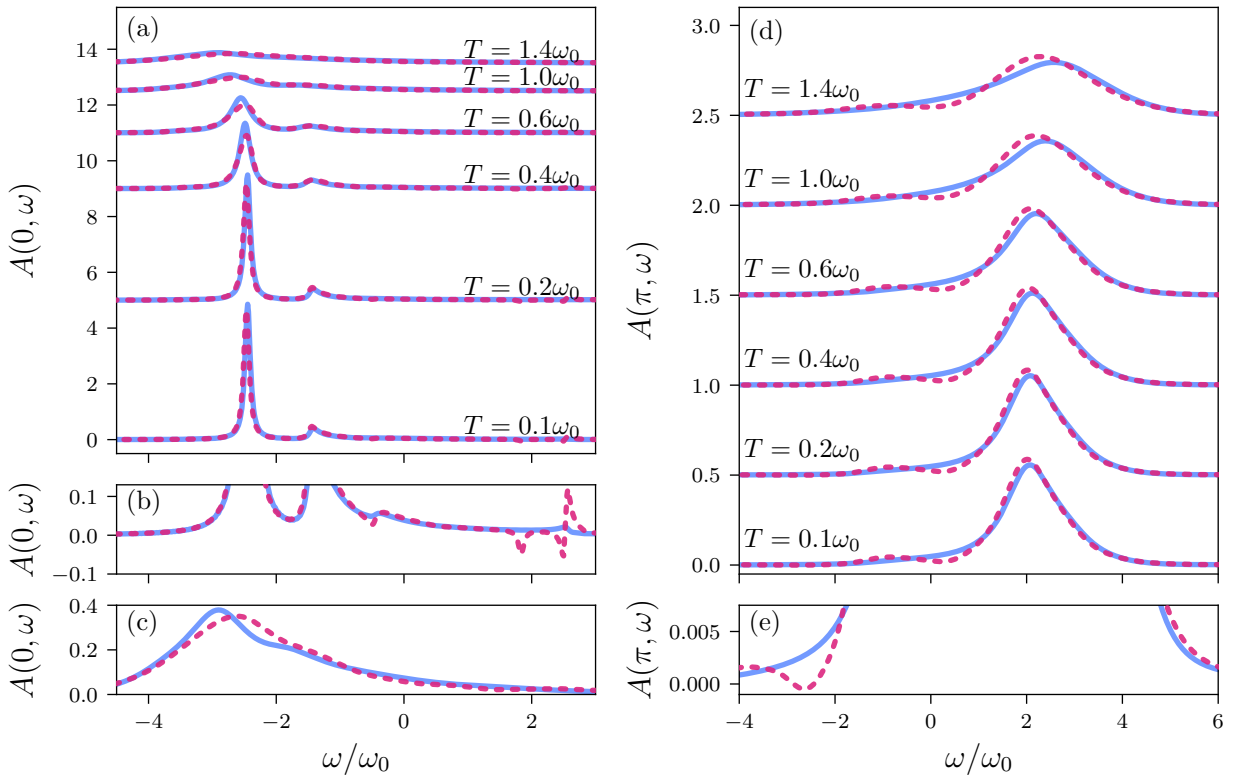

Figure 1.7: Spectral functions from the second-order CE (solid blue line) and the fourth-order CE (dashed pink line) for $N = 600$ Holstein model for a range of temperatures. (a) Results for the band bottom ($k = 0$) for $T = [0.1, 1.4]$. At low temperatures fourth-order CE slightly lowers the quasiparticle energy while also inserting regions of negative spectral weight at high frequencies. (b) $k = 0, T = 0.1$ spectral functions enlarged to emphasize the negative spectral weight predicted by the fourth-order cumulant. (c) $k = 0, T = 1.4$ spectral functions enlarged to emphasize the shift in peak location between the second- and fourth-order CE. (d) Results for the band edge ($k = \pi$) for $T = [0.1, 1.4]$. The fourth-order CE prediction lowers the energy of the main peak predicted by second-order CE, and additionally adds a broad weak intensity peak at $-\omega_0$. Both orders of the CE broaden similarly with increasing temperature. (e) $k = \pi, T = 0.1$ spectral functions enlarged to emphasize the added peak and the region of negative spectral weight predicted by the fourth-order CE compared to the second-order CE. Model parameters: $\omega_0 = t_0 = g = 1; \gamma = 0.05$ ($k = 0$ only).

second-order CE for all k which produce non-divergent fourth-order results in our approximation of the thermodynamic limit (N=600). However, $A_4(k,\omega)$ does produce some notable differences from $A_2(k,\omega)$. The fourth-order CE correctly predicts positive curvature of the satellite peaks at low temperatures near $k = 0$, marking an improvement over the second-order cumulant result. Along with this improvement, one obvious feature that emerges is small regions of negative spectral weight. The appearance of negative spectral weight, discussed previously by Gunnarsson *et*

al.[51] and in the previous subsection, has also been noted in other studies.[63,64] Though only present in high energy regions of the spectrum near $k = 0$, the negative spectral weight in $A_4(k, \omega)$ appears at much lower energies, approximately at the location of VD quasiparticle energy, for $k = \pi$. Finally, there are large regions of k space for which divergent behaviour occurs. It is possible that these regions become well-behaved for even larger system sizes, however, we have no evidence that the fourth-order CE is globally well behaved for $N = \infty$.

In Fig. 1.7 we closely examine the temperature dependence of $A_2(k, \omega)$ and $A_4(k, \omega)$ at $k = 0$ and $k = \pi$. At low temperatures, the $k = 0$ CE appears nearly converged, with $A_2(0, \omega)$ and $A_4(0, \omega)$ showing nearly identical behaviour around the quasiparticle peak. On the other hand, at $k = \pi$ the low temperature fourth-order CE does not appear converged with respect to the second-order CE, and $A_2(\pi, \omega)$ deviates notably from $A_4(\pi, \omega)$ for $\omega > 0$. This distinction in performance at $k = 0$ and $k \neq 0$ is consistent with the overall comparison of $A_n(k, \omega)$ with results from VD. In the companion paper, we will discuss how $A(k, \omega)$ for $k \neq 0$ can be more accurately calculated from a self-consistent cumulant approach.[68]

For the $k = 0$ case, the apparent convergence of the CE for some temperatures warrants more consideration, and thus we devote the remainder of this subsection to a more detailed discussion of this case. The second- and fourth-order CE results match best for high and low T. The low T convergence of the main spectral features is supported by the analysis in Sec. 1.1.3, which shows how the CE in the time-domain $\mathcal{G}(0, t)$ breaks down only at high T, where the phonon occupation numbers N_0 contribute to a growth in the magnitude of higher-order cumulants. Meanwhile, for very high temperatures, Dunn's argument that the long time behavior of $\mathcal{G}_2(0, t)$ and $\mathcal{G}_4(0, t)$ may markedly differ with differences hidden by rapid damping such that the resulting extremely broad spectral functions may appear converged, is borne out.[49] This extreme damping is seen for $T = 1.4$ highlighted in Fig. 1.7(c), where although the centroids of the main second- and fourth-order CE peaks are displaced, the broadening makes the high T results appear converged. Such misleading convergence behavior was also seen for the 6-site system in Sec. 1.2.3. At intermediate temperatures such as $T = 0.6$, apparently neither the low temperature real-time convergence illustrated in

Sec. 1.1.3, nor the high temperature damping behaviour discussed above is operative, such that the CE results shown in Fig. 1.7 display a lack of convergence for intermediate temperatures, where the second-order CE polaron peak is considerably shifted from the fourth-order CE polaron peak. The shift of the polaron peak to lower energies exhibited by $A_2(0,\omega)$ seems to be an artificial feature that is corrected in $A_4(0,\omega)$ where the center of the polaron peak appears fixed in location with respect to temperature. The origin of the distinction between $A_2(0,\omega)$ and $A_4(0,\omega)$ is subtle. An important approximation that distinguishes the second-order and fourth-order CEs is the second-order assumption[49]

$$\varepsilon\left(k+q_1+q_2\right) \approx \varepsilon\left(k+q_1\right)+\varepsilon\left(k+q_2\right)-\varepsilon\left(k\right). \tag{1.30}$$

This approximation implies that the second-order CE overestimates the energy of multiphonon processes, for example the consecutive emission of two phonons with $q_1 = q_2 = \pi$. This is likely the origin of the unphysical shift in $A_2(0,\omega)$.

1.3 Conclusion

In this paper we have presented a comprehensive study of the behavior and properties of the CE method for one paradigmatic model of an electron interacting with phonons, namely the one-dimensional Holstein model. The motivation for this choice is the fact that it is in this case where the most extensive numerically exact results are available for comparison. Some of the conclusions we draw may be generic and connect to other polaron models, but future work is necessary before such a conclusion can be drawn.

Although of great interest for applications to realistic systems, the CE in higher dimensional systems is not directly explored in this work. The formalism for the CE presented here is dimensionality agnostic, so we expect that many of our conclusions should hold in higher dimensions.

Within the confines of the second-order CE, we find that the spectral function is rather well described for up to intermediate coupling strengths at both low and high ($T > \omega_0$) temperatures

for $k = 0$ but is quantitatively accurate for other wave vectors in the high temperature regime only. Finite lattice effects are present in the numerically exact simulations which are not captured by low-order CE methods. These features are small, and are not expected to be present in the infinite lattice limit. The correct placement of satellite peaks for $k = 0$ is revealed in the structure of the second-order CE in part to be the result of a spurious intensity crossing structure. In the infinite size limit this structure is converted into a satellite region which exists only for wave vectors such that $(\epsilon_k + 2t_0 - \omega_0) < 0$ with *negative* band curvature for the higher order satellites. The effect of other models and parameters on the size of this region is a topic of future study. Both the sharp change of behavior at $\epsilon_k + 2t_0 = \omega_0$ and the sign of the curvature contrast with the exact finite lattice results which are expected to semi-quantitatively describe the infinite lattice behavior.

We have also explored the properties of the fourth-order CE. At fourth order, we find that the short-time real-time evolution of the Green's function is always systematically improved, while the long-time behavior may become pathological depending on the parameters of the model and the wave vector in question. When the fourth-order CE is well-behaved, improved spectral features are noted even at relatively low frequencies. We have explored the origins of the ill-behaved fourth-order CE. The general structure of the problematic terms take the algebraic form noted by Gunnarsson *et al.* to also give rise to negative spectral weight.[51] In addition, classifying the divergent contributions for fixed lattice size N, we illustrate the subtle balance of terms that conspire to render the fourth-order CE either useful in correcting the second-order CE or pathological. We note that in general the fourth-order CE does not generally appear capable of producing stable and sizable corrections to the second-order CE, even for intermediate electron-phonon coupling values. In the companion paper, we formulate and study a self-consistent version of the CE which is capable of accurately capturing features beyond that of the low order CE.[68]

Lastly, we comment on the recent use of the CE for the study of transport behavior in real materials. Specifically, Bernardi et al. have used the second-order CE, in conjunction with the "bubble" approximation to the current-current correlation function, to compute mobilities in both $SrTiO_3$ and in organic crystals. This approach has the advantage of capturing incoherent relaxation

channels which are not described in the simplest semi-classical theories based on the Boltzmann equation. While we cannot comment on the accuracy of the CE for systems like $SrTiO_3$ for which the Fröhlich model is most appropriate, nor can we comment on models with sizable Peierls coupling such as organic crystals, our results do suggest that the second-order CE should reasonably accurately model the full wave vector dependent one-particle spectral function for the situation $T \gtrsim \omega_0$, while likely becoming significantly less accurate for all but $k \sim 0$ at lower temperatures. This of course does not imply that the independent bubble approximation is itself accurate. Further work will be devoted to testing this approach in model systems where a controlled assessment of the various approximations is possible.

Chapter 2: Cumulant methods for electron-phonon problems. II. The self-consistent cumulant expansion

The standard tools of quantum field theory enable the rigorous calculation of Green's functions as an order-by-order expansion in terms of the interaction between excitations. The use of Dyson's equation and its associated self-energy, along with the Feynman diagram technique, provides a systematic means of partially resumming this infinite order expansion to all orders.[4,56,69] If the self-energy is expressed in terms of the interacting Green's function, then a self-consistent theory results in which additional diagrams are included. Perhaps the simplest example of such a self-consistent resummation is the Hartree-Fock theory for interacting electrons, where the self-energy is calculated to lowest non-trivial order with the fully dressed electron Green's function. More elaborate resummations also exist which are capable of including effects such as screening beyond the Hartree-Fock level.

One prominent approximation which includes such screening effects is the GW approach, which is analogous to Hartree-Fock theory with an RPA-screened interaction.[70–74] The suite of GW approximations also illustrates the distinctions between bare and self-consistent expansions. If the self-energy is expressed entirely in terms of the bare Green's function, then the so-called G_0W_0 theory results. This approach has been widely used in the computational condensed matter literature due to its accuracy and affordability. Unfortunately, for the treatment of realistic systems, the myriad distinct choices for input forms of G_0 renders the accuracy of this approach sensitive to this input. In principle this may be solved via a self-consistent solution to the equations.[75–77] However, in addition to the expense of self-consistency, there is no guarantee that the effort is

rewarded with increased accuracy.

For the important benchmark of the electron gas, both the G_0W_0 approximation and fully self-consistent version fail to reproduce important features of the spectral function such as satellite peaks in the spectral function near $k = 0$. To remedy this shortcoming, a cumulant expansion version of the GW approach has been formulated.[78,79] By construction, the cumulant version of GW is a non-self-consistent approach, formulated by exponentiating the low order GW expansion. While the cumulant expansion greatly increases the accuracy of the theory with regard to spectral information, it suffers from the same input sensitivity as the standard G_0W_0 approach. A systematic expansion technique, akin to the standard Dyson expansion with self-consistent resummations, appears lacking in the cumulant Green's function literature. [1]

In this work we attempt to fill this gap by developing the self-consistent cumulant Green's function theory and test its ability to describe spectral properties in models of an electron coupled to a lattice.[44] Such models are of more than circumscribed interest, as they also serve as analogs for purely electronic systems such as the electron gas. The approach we take was actually formulated nearly five decades ago, but never implemented or studied in any context.[49] When compared to the bare cumulant expansion carried out in the previous paper on the same class of models, we find that the simplest self-consistent cumulant theory has the ability to greatly increase the accuracy of spectral features in a non-perturbative manner away from $k = 0$, but comes with some deficits, which include ease of solution and the generation of small regions of negative spectral weight.[1,51] Promisingly, we find that addition of realistic model features such as finite phonon dispersion enhance the stability of the self-consistent equations. Thus we believe the self-consistent cumulant method may be a viable approach to accurately study realistic many-body models in condensed matter physics in a manner independent of input information.

This paper is organized as follows: In Sec. 2.1 we introduce the models which are the focus of this study. In Sec. 2.2 we present a generalized derivation of the self-consistent cumulant expansion (SC-CE) for the electron-phonon problem. Some details of the approach are confined to

[1]While this work was being written, Ref. 80 appeared, which presents a promising self-consistent approach to electron-phonon problems. The relationship, if any, to the work presented here remains to be explored.

the appendices. In Sec 2.3 we comment on the connection between the SC-CE and the Modified Direct Interaction Approximation (MDIA), a classical closure for fluid turbulence. In Sec. 2.4.1 we study the SC-CE in finite lattice situations, while in Secs. 2.4.2 and 2.4.3 we investigate the Holstein model without and with phonon dispersion in the thermodynamic limit. In the last section we conclude with future perspectives.

2.1 Models

We focus here on the Holstein model,[9,10] and a modified version of the Holstein model which includes dispersed phonons.[81] For both of these models exact VD data exist for comparison.[28,81] Work applying the non-self-consistent CE to other systems is common in the literature, both in model systems[44,49,50,82] and, more recently, in ab initio descriptions of realistic materials.[47] Self-consistent approaches based on the standard self-energy approach in conjunction with the Dyson equation in model[17,83] and more elaborate *ab initio* systems have been explored.[42,84,85] Relative to this large body of work, self-consistent cumulant approaches have received little to no attention. One notable example is the application of a self-consistent cumulant approach to the electron gas.[86] This approach differs from the one we investigate here, and the connections between and differences of these approaches will be explored in a future publication. Our study of two well-known models is presented as a proxy for assessment of the application of such approaches to realistic materials including those with complicated band structure, *ab initio*-determined electron-phonon coupling constants, and multiple phonon modes.

We exclusively consider the case of a single electron in an otherwise empty band coupled linearly to a bath of phonons

$$H = H_e + H_p + V. \tag{2.1}$$

The electronic Hamiltonian is given by the tight-binding form

$$H_e = \sum_k \epsilon_k a_k^\dagger a_k, \tag{2.2}$$

$$\epsilon_k = -2t_0 \cos k, \tag{2.3}$$

and the bath of longitudinal phonons is given by

$$H_p = \sum_n \omega_0 b_n^\dagger b_n + t_1 \left(b_n^\dagger b_{n+1} + \text{h.c.} \right) = \sum_q \omega_q b_k^\dagger b_k, \tag{2.4}$$

$$\omega_q = \omega_0 + 2t_1 \cos q. \tag{2.5}$$

When $t_1 = 0$ the model is precisely the standard dispersionless Holstein model. Here, $a_k(b_k)$ is an annihilation operator for electrons (phonons), ϵ_k is the bare electronic energy, and ω_q is the phonon dispersion. The electron-phonon interaction is given by

$$V = \frac{g}{\sqrt{N}} \sum_{kq} a_{k+q}^\dagger a_k \left(b_q + b_{-q}^\dagger \right). \tag{2.6}$$

Throughout this work we characterize our systems with the dimensionless effective electron-phonon coupling strength $\lambda = g^2/2t_0\sqrt{\omega_0^2 - 4t_1^2}$.[81,87]

2.2 Self-Consistent Cumulant Expansion

The primary quantity of interest in this work is the one-electron (causal) Green's function.[2] In the case of a single electron on the lattice, the thermal trace over all states is replaced with a trace

[2]For single particle models such as ours, this is equivalent to the retarded Green's function.

over zero-electron states

$$\mathcal{G}(k,t) = -i\Theta(t)\,\frac{\mathrm{Tr}\left[e^{-\beta H_p} a_k(t) a_k^\dagger(0)\right]}{\mathrm{Tr}\left[e^{-\beta H_p}\right]},$$
$$\equiv -i\Theta(t)\,\langle a_k(t) a_k^\dagger(0)\rangle. \tag{2.7}$$

We additionally define the spectral function as

$$A(k,\omega) = -\pi^{-1}\Im\left[\int_{-\infty}^{\infty} dt\; e^{i\omega t}\mathcal{G}_k(t)e^{-\gamma t}\right], \tag{2.8}$$

including an exponential broadening factor γ.

In the preceding paper, we followed the standard route of calculating cumulants order-by-order from moments,[4] and we refer the reader to that work for a detailed derivation and discussion of the CE.[1] This approach, however, does not lend itself to deriving self-consistent methods. Here, we construct the SC-CE through an alternative derivation first formulated by Dunn,[49] and fully presented in Appendix D.

Starting with the time-ordered exponential form of Eq. 2.7,

$$\mathcal{G}_k(t) = \mathcal{G}_k^{(0)}(t)\langle a_k T e^{-i\int_0^t d\tau V(\tau)} a_k^\dagger\rangle, \tag{2.9}$$
$$\mathcal{G}_k^{(0)}(t) \equiv -i\Theta(t)e^{-i\epsilon_k t}, \tag{2.10}$$

we directly evaluate the thermal trace, expressed in terms of a differential operator S, such that

$$\mathcal{G}_k(t) = \mathcal{G}_k^{(0)}(t) T e^S = -i\Theta(t)e^{C_k(t)}. \tag{2.11}$$

Here, T the time-ordering operator places operators in order of increasing time from right to left, and $C_k(t)$ is the sum of the cumulants of all orders. This contrasts with the standard approach, where the thermal trace is evaluated for each order individually rather than for all orders at once. Finding S is not practical since determining a closed form for S is tantamount to solving the full

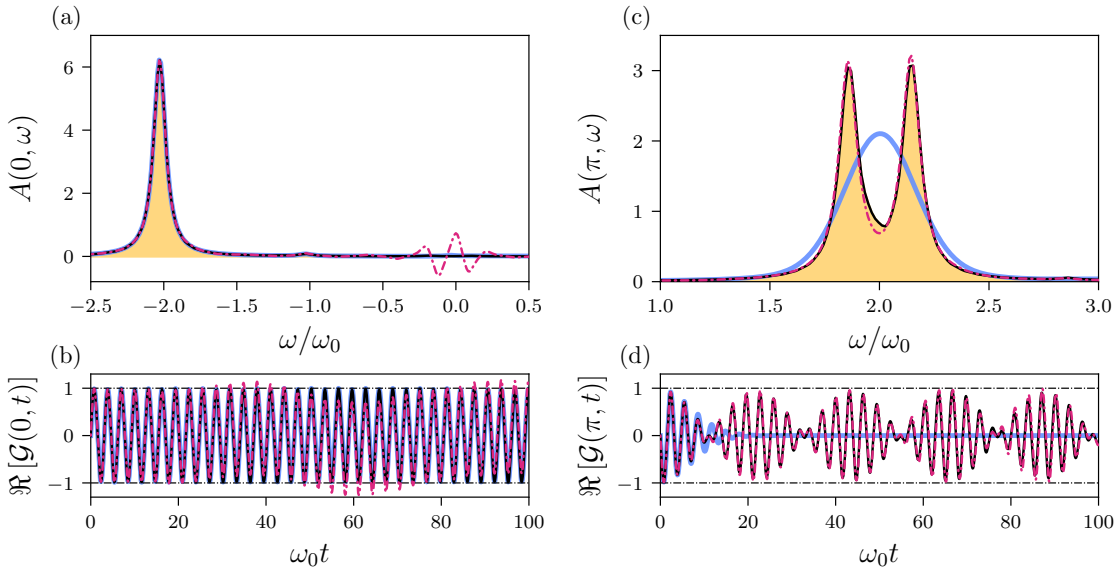

Figure 2.1: Comparison of the SC-CE (dot-dashed red line), CE (solid blue line) and ED (solid black line with gold shading) for the Holstein model with $N = 6$, $\lambda = 1/32$ and $T = 0.1\omega_0$ (a) Spectral function for $k = 0$. Note the negative spectral weight predicted by the SC-CE. (b) $\mathcal{G}(0,t)$ for the three methods at $k = 0$. (c) The spectral function at $k = \pi$, note that the SC-CE improves upon CE by correctly predicting a large peak splitting. (d) Demonstration of the fact that the peak splitting predicted by the SC-CE is a product of recurrences in $\mathcal{G}(\pi,t)$ which the CE fails to capture. The spectra in (a) and (c) are calculated with a broadening of $\gamma = 0.05$.

problem.

The advantage of introducing S is that it can be transformed to other bases before the expansion and resummation. The details of this transformation are included in Appendix D, the result of which is that occurrences of $\epsilon_k t$ in the integrand of the cumulant can be replaced with a nearly arbitrary function of time and momentum. The most intriguing use of this transformation is the creation of self-consistent versions of the cumulant expansion.

Dunn explored S in the context of the Fröhlich model with what he called the "improved cumulant" approach, where one sets $\epsilon_k t \rightarrow E_k t$, where E_k is a renormalized energy.[49] He also suggested, but did not explore, a fully self-consistent version of the cumulant-expansion, where $i\epsilon_k t \rightarrow C_k(t)$. This formulation is the focus of the work presented here.

To leading order, the resulting self-consistent equation for the cumulant is (see Appendices D

and E)

$$C_k(t) = -i\epsilon_k t - i\frac{g^2}{N}\sum_q \int_0^t d\tau \int_0^\tau d\tau' D_q^{(0)}(\tau - \tau') \times e^{-C_k(\tau)+C_{k-q}(\tau)-C_{k-q}(\tau')+C_k(\tau')}. \tag{2.12}$$

Here, $D_q^{(0)}(t)$ is the free phonon propagator, $D_q^{(0)}(t) = -i\left[(1+N_q)e^{-i\omega_q|t|} + N_q e^{i\omega_q|t|}\right]$, and N_q is the Bose occupation factor, $N_q = (\exp(\beta\omega_q) - 1)^{-1}$. Eq. 2.12 is a difficult integral equation to solve, coupling the solution for a fixed wave vector at time t to all other Green's functions for other momenta and all previous times. It is, however, readily transformed into a self-consistent expression for the interacting Green's functions, which is more convenient for calculations

$$\frac{d\mathcal{G}_{\vec{k}}(t)}{dt} = -i\epsilon_k \mathcal{G}_{\vec{k}}(t) - i\frac{g^2}{N}\sum_{\vec{q}} \int_0^t d\tau D_{\vec{q}}^{(0)}(t-\tau)\frac{\mathcal{G}_{\vec{k}}(\tau)\mathcal{G}_{\vec{k}-\vec{q}}(t)}{\mathcal{G}_{\vec{k}-\vec{q}}(\tau)}. \tag{2.13}$$

The properties of the solutions to Eqs. 2.12-2.13 are not readily apparent; however, it can be shown by expressing Eq. 2.12 recursively (Appendix E) that the weak-coupling limit of this SC-CE returns to that of the standard CE. The recursive SC-CE solution demonstrates that the SC-CE approximates higher-order cumulants by screened second-order cumulants.

The recursive expression Eq. 2.13 suggests that SC-CE might suffer from some of the same problems as the fourth-order CE which introduces, e.g., negative spectral weight. However, it is not clear whether terms which appear to be problematic in the recursive formulation are suppressed when Eq. 2.13 is solved directly. Before taking up the question of the performance of the SC-CE, we first take a detour, exploring Eq. 2.13 from a different perspective.

2.3 Connection to a Theory of Turbulence

In this section we briefly discuss how the formalism developed above and in the appendices is connected to a seemingly unrelated technique in the classical theory of turbulence which has been employed to study the closure problem in passive scalar advection models. As is well-known,

one of the earliest successful self-consistent closures in field theory is Kraichnan's Direct Interaction Approximation (DIA), which when applied to the Navier-Stokes equation or to stochastic models of turbulent flow, makes detailed predictions for a variety of observable quantities.[88,89] In the context of the Holstein model studied here, the DIA is tantamount to a simple one-loop approximation to the self-energy without vertex renormalization, namely the self-consistent Migdal approximation.[90]

Models of turbulence can also be studied in a systematic way within the Mori-Zwanzig formalism.[91] Vanden-Eijnden and collaborators have demonstrated that for passive scalar models of turbulence, the DIA arises from a simple self-consistent manipulation of the standard time-retarded memory function.[92,93] However it is well-known that an exact time-local framework also exists for the description of equilibrium and non-equilibrium dynamical phenomena in statistical mechanics.[94] Time-retarded and time-local methods essentially correspond to the resummation of perturbation theory in distinct classes of cumulants, and both approaches are formally exact if all cumulants are included.[95] Appealing to the time-local framework, Vanden-Eijnden and collaborators developed an analog of the DIA that is local in time, which they called the Modified Direct Interaction Approximation (MDIA), and demonstrated that the MDIA was the most accurate simple closure for the passive scalar models of turbulence they studied.[93] It is important to note that in the theory of open quantum systems, the time-local approach has been employed for many years to study the relaxation of a quantal system connected to a bath.[96] However these calculations are never formulated and employed in a self-consistent manner.

One can heuristically follow the reasoning of Kramer *et al.* to see the connection between the MDIA and Eq. 2.13. Consider the Dyson equation in the time domain for $t > 0$,

$$i\frac{d\mathcal{G}_{\vec{k}}(t)}{dt} - \epsilon_k \mathcal{G}_{\vec{k}}(t) - \int_0^t d\tau \Sigma(\vec{k}, t-\tau)\mathcal{G}_{\vec{k}}(\tau) = 0, \tag{2.14}$$

where for the Holstein model in one dimension $\Sigma(\vec{k}, t) = i\frac{g^2}{N}\sum_{\vec{q}} D^0_{\vec{q}}(t)\mathcal{G}^{(0)}_{\vec{k}-\vec{q}}(t)$ to lowest order in perturbation theory. Next, one may use the relationship $\mathcal{G}^{(0)}_{\vec{k}-\vec{q}}(t-\tau) = -i[\mathcal{G}^{(0)}_{\vec{k}-\vec{q}}(\tau)]^{-1} \cdot \mathcal{G}^{(0)}_{\vec{k}-\vec{q}}(t)$ in

the self-energy term. Finally, partial resummation may be affected by replacing all bare Green's functions with dressed ones, which converts the Dyson equation precisely into Eq. 2.13. Note that while this provides the lowest order in a series of self-consistent cumulant equations, systematic higher-order terms may be constructed order by order. Dunn has discussed higher-order terms from the standpoint of the self-consistent expansion detailed in the appendices. At higher orders, the procedure outlined by Dunn and the renormalization technique discussed here will differ. Such higher-order terms are cumbersome and will not be considered further in this work.

The connection to the classical MDIA provides an interesting perspective on the expected behavior of the self-consistent cumulant expansion. For example, it was noted in Paper I that inclusion of the 4th order cumulant may improve the accuracy of the approximate Green's function, but comes at the expense of potentially unphysical behavior, such as the generation of regions of negative spectral weight for some wave vectors.[1] One may ask if this deficit is cured within the lowest order self-consistent version of the theory. Within the classical theory of turbulence it is well-known that the DIA approximation is "weakly realizable" in the sense that it is exact for certain well-defined models which preserve some physical symmetries and properties. The MDIA is not even weakly realizable, and thus despite its superior accuracy, the price to be paid may be some degree of unphysical behavior, such as negative spectral weight.[97] As we shall see in the following sections, the same is true of the SC-CE for quantum models.

2.4 Results

For our investigations of both the dispersive and non-dispersive Holstein models, we compare the SC-CE to exact spectral functions calculated using VD and Exact Diagonalizaton (ED), as well as with the standard CE. For the dispersionless Holstein model at intermediate coupling, we compare against the finite-temperature VD data of Bonča, Trugman, and Berciu.[28] As in the companion paper, we can compare with these limited sized systems because on the scale of the heat maps the spectral functions appear to change only slightly as the system size increases.[1] For the zero-temperature dispersive Holstein model we compare with the VD data of Bonča and

λ	T/ω_0	$\max(\lvert\mathcal{G}(0,t)\rvert)$	$\max(\lvert\mathcal{G}(\pi,t)\rvert)$
1/32	0.10	1.49	1.00
1/8	0.10	1.69	1.00
1/2	0.10	1.41×10^1	5.61×10^1
1	0.10	2.11×10^4	1.52×10^5
1/32	1.00	2.01	1.00
1/8	1.00	3.06	1.33
1/2	1.00	4.23×10^4	2.33×10^4
1	1.00	3.88×10^6	8.21×10^6

Table 2.1: Demonstration of the growth of SC-CE divergences for the six-site Holstein model. For weak couplings ($\lambda = 1/32$ and $1/8$) and $T = 0.1\omega_0$ the divergences are suppressed (thought the Green's function does still display non-physical magnitude > 1). For $T = 1\omega_0$, only $\lambda = 1/32$ has a small enough magnitude at $t = 100$ to be useful for spectral analysis. For both temperatures $\lambda = 1/2$ and 1 rapidly diverge. For all calculations, the integration time-step is $0.001\omega_0^{-1}$.

Trugman[81] which are computed in the thermodynamic limit. We do not review the numerical benchmark approaches here and refer the reader to extensive literature on these methods.[13–15,26–28] Details of the CE and truncated-ED ("K-phonon approximation") methods are presented in the companion paper.[1]

2.4.1 Six-Site Holstein Model

We begin with the six-site Holstein model at very weak coupling ($\lambda = 1/32$) and low-temperature ($T = 0.1\omega_0$). One may expect the standard CE to perform well in this parameter regime because the coupling is so weak. However, the perturbative CE is a short-time expansion as illustrated in the preceding work. Thus, there is no guarantee of its success even for very small g away from short times. As we will see, even for such small coupling values the standard second order cumulant fails for $k \neq 0$.

In Fig. 2.1, we compare the spectral functions and $\mathcal{G}(k,t)$ for $k = 0$ and $k = \pi$ using SC-CE, CE, and ED at $T = 0.1\omega_0$. The ED results are truncated at the three-phonon level which should be sufficient given the coupling strength. Fig. 2.1(a) shows that the spectral function at $k = 0$ is nearly exactly described by CE (as expected for weak coupling and $k = 0$). The SC-CE predicts the main features approximately as accurately, but inserts a region of non-physical negative spectral weight

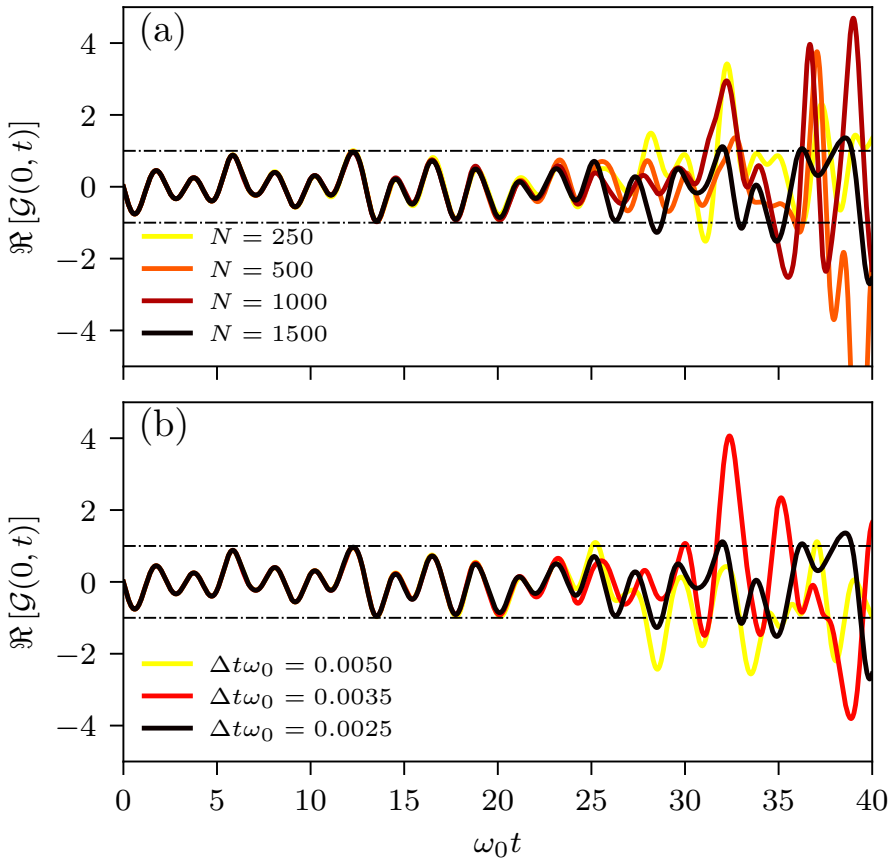

Figure 2.2: Convergence of the real part of $\mathcal{G}(0,t)$ as calculated by the SC-CE with respect to (a) system size, and (b) time-step. In all cases, $T = 0$, $\lambda = \omega_0 = 1$, and the horizontal dashed-black lines mark ± 1 and serve as a guide of when non-physical behaviour sets in. In (a) the time-step is uniformly $0.0025\omega_0^{-1}$. $\mathcal{G}(0,t)$ is nearly converged with system size up to approximately $25\omega_0^{-1}$ for $N > 250$, but for longer times the functions diverge. For the case of $N = 1500$ (b) demonstrates that very small time steps are necessary to converge the SC-CE at later times, and large time steps worsen the longer-time behaviour. However, continued decrease of the time step cannot control the instabilities at long times within our integration scheme.

around $\omega = 0$.

In Fig. 2.1(b), the time dependence of $\mathcal{G}(0,t)$ is plotted, revealing the origin of the negative spectral weight. The SC-CE result deviates from ED via the addition of a long wavelength oscillation peaking for the first time at $\sim 40\omega_0^{-1}$. This additional structure introduces non-physical behaviour into $\mathcal{G}(0,t)$ with a norm that exceeds unity.

The presence of this negative spectral weight is not surprising given the understanding of the

behaviour of higher-order cumulants. As discussed in the preceding companion paper, within the context of perturbative higher-order cumulant calculations, this problem has been known for some time.[1,51,63,64] Additionally, the M-DIA, a classical analogue of the SC-CE for turbulence discussed in Sec. 2.3, can also produce non-realizable negative spectral regions.[97]

While for the example of $k = 0$ at weak coupling the SC-CE is *worse* than the CE, the CE is less accurate away from $k = 0$, so we next consider the case of $k = \pi$ in Fig. 2.1(c). The most prominent feature of the ED spectrum is a split main peak with splitting centered at $\omega = 2\omega_0$. The CE, while generally placing the spectral intensity in the correct position does not include a peak splitting and thus provides a qualitatively incorrect description of the spectrum. The SC-CE, on the other hand, replicates the ED spectrum nearly exactly, including the prominent peak splitting. Fig. 2.1(d) demonstrates that the SC-CE captures the splitting because it accurately describes the long-time dynamics, specifically recurrences in the temporal evolution of $\mathcal{G}(\pi, t)$. While the initial damping of $\mathcal{G}(\pi, t)$ is well-described by both CE and SC-CE, the CE does not recur once it decays for the first time at $\sim 10\ \omega_0^{-1}$. The SC-CE, however, continues to recur periodically, in excellent agreement with the ED result, and without introducing any non-physical behaviour. It should be noted that even inclusion of the fourth-order CE does not capture this behaviour. Thus, the SC-CE can describe some non-perturbative effects beyond accessible low-order CEs.

While the SC-CE results for $\lambda = 1/32$ and $k = \pi$ demonstrate remarkable improvements over the CE, only a limited parameter region was accessible for this calculation because the SC-CE rapidly diverges for even slightly larger couplings when $N = 6$. In Table 2.1, we demonstrate that these divergences are exacerbated by increasing coupling strength when N is fixed at a small value, and, to a lesser extent, by increasing temperature. We characterize the severity of the divergence with the maximum of $|\mathcal{G}(k, t)|$ over the range $0 \leq t \leq 40$.

For $T = 0.1\omega_0$ or $1.0\omega_0$ and $\lambda = 1/32$ or $1/8$, the Green's function does not dramatically deviate from unity. Further, for several cases at $k = \pi$ the magnitude does not exceed one within the observed time window. In these cases, as well as the cases of weak deviations from unity, a modest damping enables us to calculate sharp spectral functions (e.g. Fig. 2.1(a) and (c)). This

is not the case for the larger couplings. These divergences render Green's functions for finite sized lattices unusable for spectral interpretation. The origin of these divergences is unclear, but they appear to be part of the SC-CE rather than numerical in nature. Finally, we note that the dependence of the divergent behaviour on temperature is a weaker effect.

While the divergences may be fundamental to finite sized SC-CE, they are reminiscent of the divergences we observed in the fourth-order CE, and there we were able to eliminate some of this problematic behaviour by increasing the system size towards the thermodynamic limit. Following this line of investigation, we devote the following two subsections to testing the SC-CE in systems approaching the thermodynamic limit, which is the case of most relevance for the description of experiments.

2.4.2 Thermodynamic Limit of the Holstein Model

Motivated by the difficulty of reaching the strong coupling regime in a six-site system, we explore the Holstein model with increasingly large numbers of sites as a means of reaching the thermodynamic limit and potentially suppressing instabilities. In Fig. 2.2(a) we study intermediate coupling ($\lambda = 1$), showing that for short times the SC-CE Green's function does become stable and appears to converge to the infinite lattice limit over some time scale as N is increased. However, beyond $t \approx 30\omega_0^{-1}$ the dynamics for different N values diverge dramatically, indicating that convergence has not been reached. Further, increasing the system size from $N = 1000$ to $N = 1500$ does not definitively push back the divergence, so it remains unclear how effective increasing the system size is at pushing back the divergence. Fig. 2.2(b) shows that decreasing the integration time step also eventually becomes ineffective in suppressing instability at long times. It remains unclear if the long-time behaviour we observe is an intrinsic failure of the SC-CE approach, as it is for the fourth-order CE in the same model, or if it is a numerical issue associated with the integration of a non-linear integro-differential equation. Either way, this behaviour presents a practical limit to how far we can currently converge results for the Holstein model. If the divergence can be overcome for $N \to \infty$, a different numerical approach which more effectively evaluates the

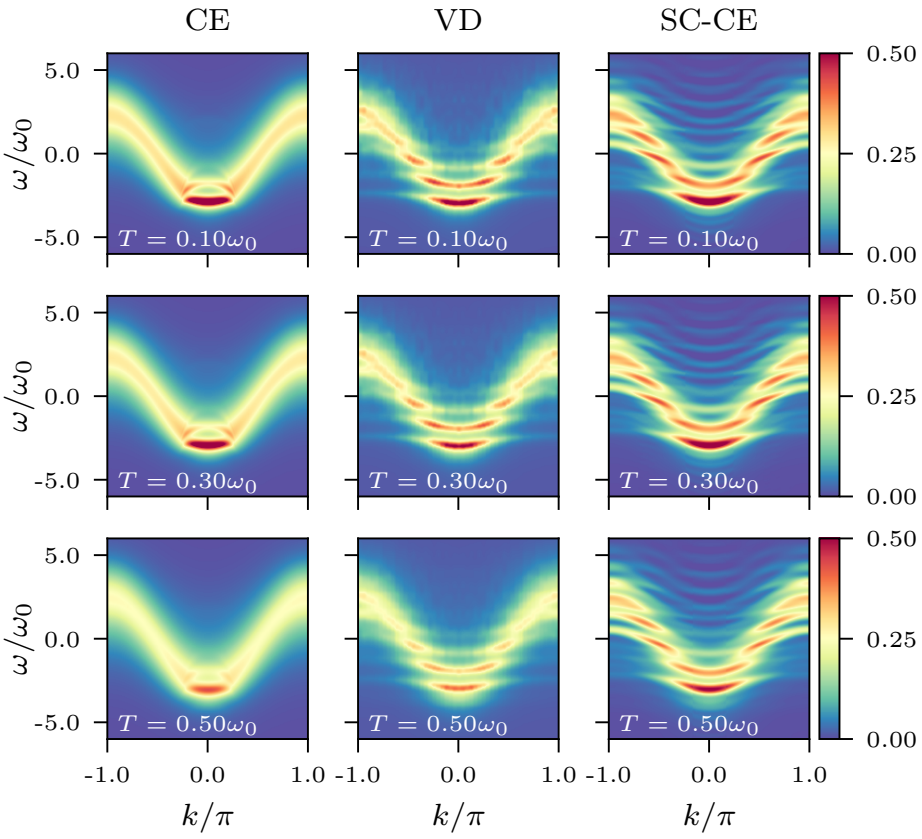

Figure 2.3: Heat maps of the spectral function for the finite temperature Holstein model ($\lambda = \omega_0 = 1$). Left: the CE with $N = 250$. The CE is converged on this scale for this value of N. Center: VD results of Bonča *et al.*[28] for six sites which reasonably approximates the thermodynamic limit on this scale. Right: the SC-CE with $N = 1500$. All three methods use a large broadening of $\gamma = 0.25$. The SC-CE was calculated with a time-step of $0.005\omega_0^{-1}$ and the Green's function was truncated beyond the first time its norm exceeded unity for each momentum. The SC-CE improves upon the CE and semi-quantitatively captures the VD spectra at all temperatures.

highly-oscillatory momentum integral in Eq. 2.13 might provide a better path to convergence.[98,99]

By using the times before which $|\mathcal{G}(k,t)| > 1$ and sufficient damping ($\gamma = 0.25$), we can compute spectra from SC-CE in a stable fashion. In Fig. 2.3 we compare the spectra of the CE and the SC-CE to the VD results of Bonča *et al.*[28] We find that the SC-CE and VD results are generally in agreement over the whole range of temperatures and momenta. The SC-CE correctly predicts that the quasiparticle peak flattens and diminishes in intensity from $k = 0$ to $k = \pi$. This contrasts with the CE which, while accurate for $k = 0$, has an abrupt transition where the quasiparticle peak vanishes and the structured features are replaced by a single broad peak. We noted in the previous

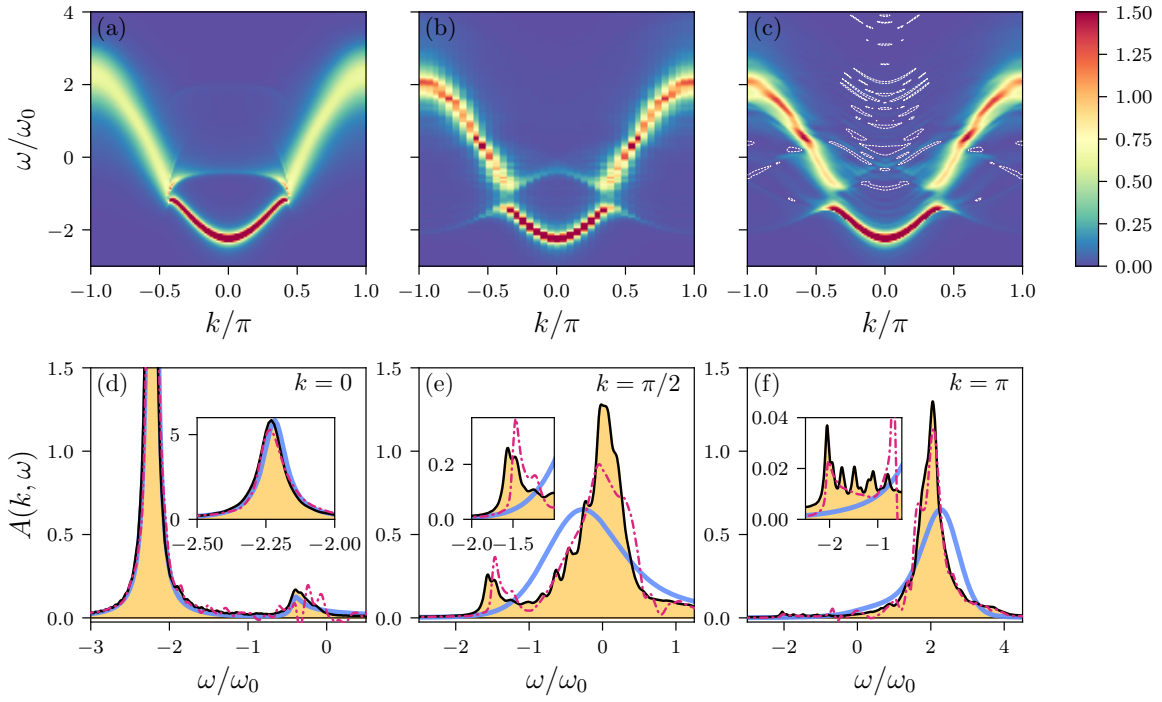

Figure 2.4: Spectral functions of the dispersive Holstein model for $T = 0, \omega_0 = t_0 = 1, \lambda = 0.5$ and $t_1 = 0.4$. The comparison between the (a) $N = 2000$ CE, (b) VD results of Bonča and Trugman,[81] and (c) $N = 2000$ SC-CE demonstrates that the SC-CE accurately captures the VD results across all momenta while CE is accurate only at $k = 0$. In (c), small regions of negative spectral weight are bounded with a white dashed line. Slices of the spectral functions at $k = 0, \pi/2$, and π are presented in (d), (e), and (f) respectively. Each slice compares the CE (solid blue line), VD (solid black line with gold shading), and the SC-CE (dot-dashed red line).

paper that the deficits of the second-order CE could not be generally corrected by inclusion of the fourth-order term.[1] By contrast, the curvature of the SC-CE satellite peaks is qualitatively correct and the appearance of the spectrum is correct across the entire range of k-space. Considering that CE predicts qualitatively incorrect satellite curvature, and predicts a nearly featureless spectrum for $k \geq \pi/3$, the SC-CE provides a vast improvement.

Due to the substantial broadening we have used, we cannot make quantitative statements concerning the fine structure of the spectral functions. It does appear, however, that the SC-CE generally over-weighs the intensity of the vibronic peaks. The SC-CE most resembles VD at lower temperatures, and most severely deviates from exact results at higher temperatures where we have less confidence in our ability to converge the Green's function in this non-dispersive case.

Increasing the system size towards the thermodynamic limit improves the stability of the SC-CE, and in light of the excellent spectra produced, the question becomes whether this method could be applicable to realistic systems. The simplicity of the Holstein model with its single, non-dispersed, phonon mode also potentially contributes to the unstable behaviour, and a more realistic model which naturally generates more internal damping may lead to better behaviour. The next subsection addresses this question with an investigation of the dispersive Holstein model.

2.4.3 Dispersive Holstein Model

We now consider a dispersive Holstein model at $T = 0$, with $t_1 = 0.4, \lambda = 0.5$ and $\omega_0 = t_0 = 1$. As we will see, when naturally dispersive phonons are present, the divergence seen for the SC-CE is strongly suppressed and we are able to propagate $\mathcal{G}(k, t)$ to long times, allowing for analysis of the fine structure as well as the main spectral peaks.

In Fig. 2.4 we compare (a) the CE, (b) the VD spectra of Bonča and Trugman,[81] and (c) the SC-CE. The SC-CE much more closely resembles the VD result than the CE. In particular, it correctly predicts the parabolic behaviour of the quasiparticle peak near $k = 0$, the gap in the spectral function close to $k \approx 0.3\pi$, and the loss in intensity in the spectral signatures as $k \to \pi$. Other features, such as the main satellite band, appear better described as well. For $k = 0$, the CE is more accurate than the SC-CE. As illustrate in Fig. 2.4(d), the SC-CE introduces noise and small regions of negative spectral weight around the first satellite and slightly underestimates the quasiparticle weight. At intermediate k, the SC-CE correctly predicts the formation of a gap between the quasiparticle peak and the first satellite while the CE predicts a single peak. Fig. 2.4(e) demonstrates this at $k = \pi/2$. There, the SC-CE predicts a quasiparticle and secondary peaks with positions, shapes and intensities in good agreement with exact results. The CE, in contrast, simply predicts a broad Gaussian-like spectrum centered in an incorrect location. At $k = \pi$ the SC-CE correctly predicts the very weak quasiparticle peak and the shape and intensity of the main satellite peak at $\omega = 2\omega_0$. However, the spectrum is somewhat noisy and contains some small regions of negative spectral weight.

Across the entire range of momenta, SC-CE correctly predicts the shapes and positions of the major peaks of the spectrum, and with the exception of $k = 0$, is a vast improvement over the CE. It appears that for more realistic dispersive models, the SC-CE is a very promising technique if one is accepting of physical defects, such as the prediction of small regions of negative spectral weight on the few percent level. It is possible these defects may be removed in higher-order versions of the theory, although the numerical solutions of the resulting equations may present a challenge. Future work investigating systems such as the electron gas should shed further light on the ability of the SC-CE to capture non-trivial physics in a wide range of distinct physical settings.

2.5 Conclusion

In this paper we have presented a self-consistent cumulant expansion (SC-CE) for the description of the dynamics of a single electron coupled to a bath of phonons. The models we have studied have direct relevance in polaron problems, and more broadly may be applied to purely electronic problems such as the interacting electron gas. Furthermore, the work outlined here takes a step toward placing the cumulant approach to Green's functions on an equal footing with the standard Dyson-equation based formulation of field theory via the introduction of a hierarchy of self-consistent renormalized equations. Via a link through the Dyson equation, we have connected the SC-CE at lowest order to a time-local self-consistent closure technique in the theory of classical turbulence. This connection gives insights into expectations for the successes and failures of the SC-CE.

We find that for the six-site Holstein model with a single phonon frequency at weak coupling, the SC-CE spectral function includes non-perturbative corrections beyond the second and fourth order bare cumulant expansions (CEs) and correctly describes the spectrum at all momenta while the CE is only accurate at $k = 0$. This is a consequence of the SC-CE capturing long time recurrences in the Green's function. The SC-CE also introduces a small degree of non-physical negative spectral weight for some wave vectors. More troubling, we find that the behavior of the SC-CE becomes unstable at long times. It remains unclear the degree to which this reflects the need for

better methods to solve the non-linear integro-differential equation that embodies the SC-CE, a failure of the SC-CE itself, or both.

We have studied the SC-CE for the Holstein model with a single phonon frequency and with dispersive phonons. Increasing the system size of the non-dispersed Holstein model towards the thermodynamic limit partially suppresses the divergences prominent in the finite size lattice models. It is unclear if a truly infinite system would have divergences since the SC-CE equations becomes challenging to converge at later times. Using a broadening parameter larger than that employed in recent VD studies along with a long-time truncation of the SC-CE time evolution, comparisons may be made with the standard CE and exact numerical results. We find that the SC-CE spectral function replicates the VD results well at all momenta and corrects many deficits of the CE. Since natural internal damping of the decay of the Green's function occurs in more realistic models with a finite phonon dispersion, we turn to the dispersive Holstein model both as a means to assess the SC-CE in more realistic settings as well as to see if stability of the SC-CE is enhanced in such cases. Indeed, for the dispersive Holstein model, we find that no additional damping or truncation is needed, and the results greatly improve many of the features of the CE, albeit with the defect of a small degree of unphysical negative spectral weight.

Given these findings, and the fact that the cost of the SC-CE is only $\approx O(N^2(t_{\max}/\Delta t)^2)$ where Δt is the time step, we believe the SC-CE is promising and should be explored further. An obvious next step would be the investigation of the homogeneous electron gas problem, where the screened Coulomb interaction would play the role of the damped phonon Green's function. In addition, higher-order SC-CE schemes can be explored, although the hierarchy of self-consistent cumulant equations rapidly becomes extremely numerically involved. These and other possibilities will be the subject of future investigations.

Chapter 3: Connections Between Green's Function Methods for Electron-Phonon Problems

3.1 Introduction

There are two main field-theoretical approaches to constructing approximations for Green's function. First, there is the standard Dyson equation approach with it's associated self-energy which permits successively more accurate approximations via the inclusion of higher-order diagrams.[56,69] Second there is is the linked-cluster or cumulant approach which approximates the entire Dyson-perturbative series via exponential resummation.[4,60,61] Both approaches see wide application in all corners of condensed matter dynamics and each confers a distinct set of advantages. Dyson methods are strictly realizable, meaning they always produce positive semi-definite spectra; however, they typically under-perform cumulant-based methods in approximating the quasiparticle (QP) satellite peaks. For example, fully self-consistent GW, a Dyson method, is known to incorrectly remove the satellite peaks from the electron gas, while the cumulant-based approach accurately predicts them.[82,100,101]

Self-consistent approximations can be constructed for both classes of methods, and hold the promise of improvement upon their non-self-consistent counterparts through implicitly including corrections up to infinite-order. Within the Dyson framework, many self-consistent methods have been developed and have become standard tools. To mention a few particularly famous examples, there is the aforementioned GW approximation,[70–74] and the ubiquitous Hartree-Fock approximation for electronic structure.[56,102] For electron-phonon problems the Migdal approximation, or alternately the Fan self-energy, is analogous to the GW self-energy, but with a phonon propagator in place of an electron propagator.[42,103,104] All three of these methods work by incorporating all self-similar diagrams which can be generated via insertions into the first-order self energy.

For cumulant-based methods the landscape of self-consistent quantum dynamics methods is barer. While the non-self-consistent second-order cumulant has become a workhorse in electron-phonon problems for model systems and *ab inito* calculations alike,[1,42–44,46,48,50,105] there are few examples of self-consistent cumulant methods. Indeed the only previous work we are aware of is a self-consistent cumulant approach for the electron gas by Holm and Aryasetiawan.[86] While the loss of strict realizablility for all cumulant-based methods may preclude some applications of resulting spectra, some cumulant methods, such as the second-order cumulant,[1] are realizable. [1] Even for the general case where the method is not realizable, the potentially vast improvements are worthy of more consideration.

Recently, two self-consistent methods, the self-consistent cumulant expansion (SC-CE) and the power series (PS) method were independently developed and presented in simple electron-phonon model systems.[68,80] Both of these methods go beyond the Dyson approach, and also incorporate self-consistency. In this chapter we will present the connections between these two methods and place them in larger context of self-consistent Green's function methods.

The SC-CE was first implemented and explored by us, but was originally derived by Dunn in 1975 as a further generalization to his Brillouin-Wigner perturbation theory-like self-consistent energy cumulant.[49,68] The SC-CE integro-differential equations do not permit solution in frequency space, thereby constraining the method to solution in the time-domain. Numerical investigation revealed that the method could produce a divergent-in-time Green's function, and regions of negative spectral weight. However, when available, the resulting spectral functions vastly improved upon the non-self-consistent cumulant, and captured the QP peak at all electronic momenta while additionally providing an excellent approximation to the satellite peaks including splittings and shoulders at all momenta.[68]

The PS method was derived by Pandey and Littlewood via a first-order self-consistent treatment with the non-interacting Green's function factored out and a novel temporal contraction relation applied. It was demonstrated to capture the non-perturbative peak splitting in the two-site Holstein

[1] Only cumulant expansions of second or infinite-order are guaranteed to be realizable because no distribution has a cumulant generating function that is a finite polynomial greater than degree two.

model for weak and intermediate coupling strengths.[80] It was additionally generalized for arbitrary system sizes.[106] In this chapter we will demonstrate that the PS method is a hybrid Dyson-cumulant method. Additionally we will show that it shares features with the SC-CE at small system sizes and converges to the SC-M as the system approaches the thermodynamic limit.

The structure of this chapter is as follows. In Sec. 3.2 we will introduce the Holstein model and the mathematical description of the methods. Following, in Sec. 3.3 we will present results for the methods in the intermediate coupling 1D Holstein model. We will compare not against numerically exact Variational Diagonalization (VD) and Truncated ED (ED).[1,28] Additionally, we will compare against the self-consistent Migdal (SC-M) approximation and its non-self-consistent counterpart the non-self-consistent Migdal approximation (NSC-M) as examples of Dyson-type methods.[42,103,104,107] Finally, we will conclude with comments about the general applicability and future developments of self-consistent approaches.

3.2 Formalism

The one-electron Green's function ($\mathcal{G}_k(t)$) can be determined via the Dyson equation, which when presented in its time domain form is

$$0 = i\frac{d\mathcal{G}_k(t)}{dt} - \epsilon_k \mathcal{G}_k(t) - \int_0^t d\tau \Sigma_k(t-\tau)\,\mathcal{G}_k(\tau)\,. \tag{3.1}$$

Here, $\Sigma_k(t)$ is the self-energy, and ϵ_k is the bare energy. The associated spectral function is given through the Fourier transform

$$A(k,\omega) = -\frac{1}{\pi}\mathrm{Im}\left[\int_{-\infty}^{\infty} dt e^{i\omega t}\mathcal{G}(k,t)\exp(-\gamma t)\right]. \tag{3.2}$$

The decay of γ is a Lorentzian broadening to permit the use of finite simulation times. [2]

[2]We note that the definition for the spectral function used in Ref. 80 and 106 cannot be used for methods that are not realizable. While the exact spectral function is positive semi-definite and thus is unaffected by the introduction of an absolute value, the prediction of some methods should in principle produce negative spectral weight. Specifically, non-realizable methods such as the higher-order perturbative cumulant expansions, the SC-CE or the PS fall in this category.

All of the methods we discuss here can be viewed as different renormalization schemes for the first-order self-energy. The derivation of the non-self-consistent Migdal approximation (NSC-M) applies the standard tools of diagrammatic perturbation theory[50,56,69] and produces the following self energy

$$\Sigma_k^{\text{NSC-M}}(t) = \frac{ig^2}{N}\sum_q \mathcal{D}_q^{(0)}(t)\,\mathcal{G}_{k-q}^{(0)}(t)\,, \tag{3.3}$$

$$\mathcal{G}_k^{(0)}(t) = -i\theta(t)e^{-i\epsilon_k t}. \tag{3.4}$$

Physically, this corresponds to the approximation that phonon emission/absorption events are statistically independent from one another. The corresponding frequency space version of this method (Eq. 157 of Ref. 42) has been used in the context of *ab initio* problems.[42,108]

The SC-M is derived by recognizing that the infinite class of non-crossing diagrams can be succinctly expressed by replacing the bare electronic Green's function with the interacting Green's function. Thus, the SC-M self energy is

$$\Sigma_k^{\text{SC-M}}(t) = \frac{ig^2}{N}\sum_q \mathcal{D}_q^{(0)}(t)\,\mathcal{G}_{k-q}(t)\,. \tag{3.5}$$

As it includes corrections up to infinite-order, the SC-M is valid for a much broader range of couplings than the NSC-M.

We now turn our attention to the SC-CE and PS which deal with a more general memory kernel ($\Sigma(t,\tau)$) in place of the self-energy in Eq. 3.1. As shown in Ref. 68, the SC-CE can be derived via either a general Mori-Zwanzig approach[68] or a more electron-phonon specific manipulation[49]. The memory kernel is thus given by

$$\Sigma_k^{\text{SC-CE}}(t,\tau) = \frac{g^2}{N}\sum_q \mathcal{D}_q^{(0)}(t-\tau)\,\frac{\mathcal{G}_{k-q}(t)}{\mathcal{G}_{k-q}(\tau)}. \tag{3.6}$$

The PS method is derived fully in Appendix G, and is the self-consistent Migdal approximation

where additionally it is assumed the temporal contraction for intra-band terms is

$$\mathcal{G}_k(\tau)\,\mathcal{G}_k(t-\tau) \rightarrow \mathcal{G}_k(t)\,\mathcal{G}_k(0)\,. \tag{3.7}$$

While originally presented as an iterative method it can be reworked into an equivalent real-time method. The full derivation is given in Appendix H, and leads to a memory kernel of

$$\Sigma_k^{\mathrm{PS}}(t,\tau) = \frac{g^2}{N}\sum_q \begin{cases} i\mathcal{D}^{(0)}_{k-q}(t-\tau)\,\mathcal{G}_q(t-\tau) & q \neq k \\ \mathcal{D}^{(0)}_0(t-\tau)\,\frac{\mathcal{G}_k(t)}{\mathcal{G}_k(\tau)} & q = k \end{cases}. \tag{3.8}$$

Through this presentation it is clear that the power series method applies the SC-M for all inter-band terms and the SC-CE for intra-band terms. In the limit of large N, the number of inter-band terms will grow as N^2 while the intra-band terms will only grow as N, so the PS will tend to the SC-M as $1/N$.

We note that this memory kernel may not be ideal to directly implement due to the unnecessary inclusion of $\mathcal{G}_k(\tau)$ in the denominator. Our numerical implementation (shown in Appendix I) is on the simplified expression.

3.3 Numerical Results

We begin our numerical analysis by considering the two-site Holstein model. We consider the case of intermediate coupling in Fig. 3.1 ($\lambda = \omega_0 = t_0 = 1$), and compare to a reference Exact Diagonalization (ED) calculation truncated at 15 phonons.[1] For the case of $k = 0$ in Fig. 3.1(a), both the SC-CE and the PS methods capture the QP peak at $\sim -3.2\omega_0$ and the satellite peaks at $\sim -2.2\omega_0, -1.2\omega_0$ and $-0.2\omega_0$. There is a slight ($< 0.1\omega_0$) positive energy shift in the PS and SC-CE compared to the ED results throughout. Though not visible on the scale of the plot there are a few difference between the SC-CE and PS. The SC-CE predicts a slightly lower QP energy ($\sim 0.01\omega_0$) and also continues to generates weak features lower in energy than the QP with peaks approximately at $-5\omega_0, -6\omega_0$ and $-7\omega_0$. The PS does not produce any features below

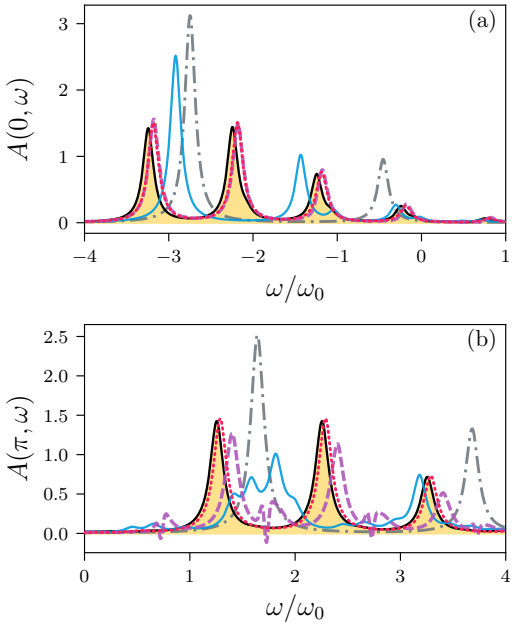

Figure 3.1: Comparison between the SC-M (blue solid line), PS (purple dashed line), SC-CE (red dotted line), NSC-M (dot-dashed grey line), and ED (black line and yellow fill) for $k = 0$ (a) and $k = \pi$ (b). Model parameters: $\lambda = t_0 = \omega_0 = 1$, and $T = 0$. N=2. $\gamma = 0.075$

the QP which shows the correct behaviour, however it does produce small regions of negative spectral weight at high frequencies ($\sim 2.75\omega_0, 3.75\omega_0,$ and $4.75\omega_0$). It appears there are several other points which would produce negative spectral weight but for the applied Lorentzian damping. Nonetheless, the near indistinguishably of the PS and SC-CE indicates that the on-site term is much larger than the inter-band term.

It is not surprising that we observe negative spectral weight in the PS since similar effects have been observed in higher-order cumulants, and the SC-CE.[51,68] A full analysis of the divergences in the SC-CE is given in Ref. 68. however negative spectral weight arises when the time-domain green's function exceeds a norm of one. For the SC-CE this could both occur in the form of divergences and not-apparently divergent recurrences with amplitude greater than one (e.g. Fig. 1 Ref. 68). For the PS method, the negative spectral weight is attributable to a divergence at longer times which appears to be exponential. In the case of $k = 0$ the PS divergence is weaker than our modest Lorentzian damping thus allowing spectra to be plotted. As demonstrated in Fig. 3.3,

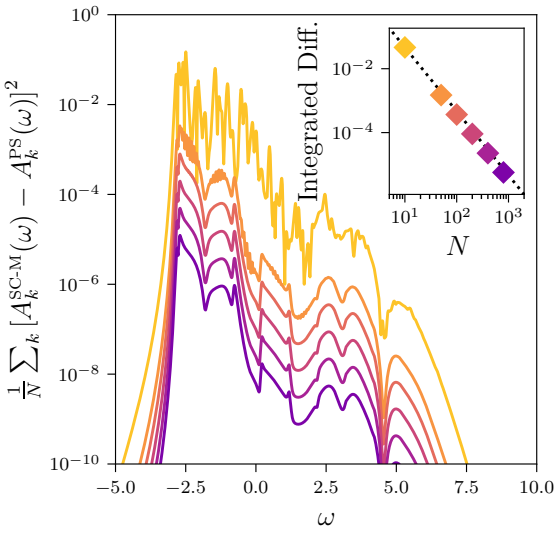

Figure 3.2: Demonstration of the equivalence of the SC-M and the PS methods in the thermodynamic limit. The outer panel shows the squared difference between the spectral functions of the two methods averaged over the whole Brillouin zone. The inset pane demonstrates the expected N^{-2} falloff of the integrated difference. The dotted line is the function N^{-2} scaled to match the integral at $N = 10$. Model Parameters: $\lambda = t_0 = \omega_0 = 1$ and $T = 0$. $N = [10, 50, 100, 200, 400, 800]$ and colors in the inset correspond to the curves in the outer panel. The spectra are broadened with a Lorentzian broadening of 0.075.

this divergence is stable with respect to decreasing the integration time step, so we believe it to be intrinsic to the SC-CE and PS methods; however, the lack of divergence in the SC-CE is likely due to a cancellation specific to the two-site system.

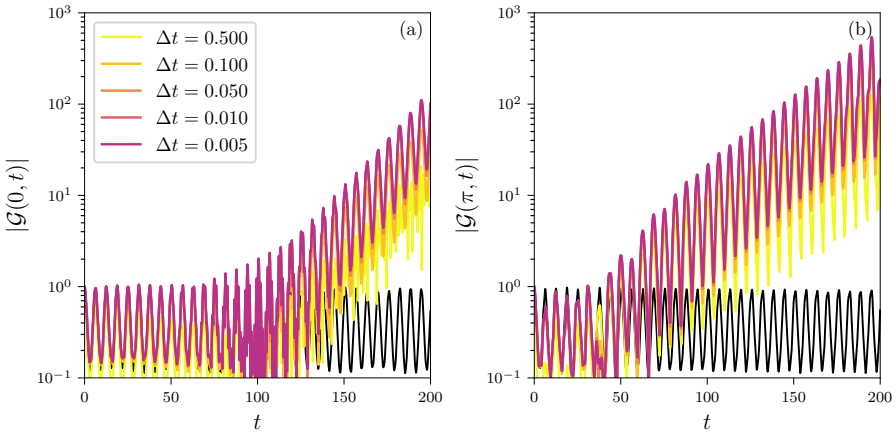

Figure 3.3: Demonstration of divergences in PS method for 2 site Holstein model ($\lambda = \omega_0 = t_0 = 1$). Decreasing the time-step does not obviously improve the divergence time. The ED result is shown in black.

In contrast to the SC-CE and PS, the NSC-M and SC-M provide qualitatively incorrect pictures of QP and satellite peaks. For both, the QP weight and energy are overestimated, and the first satellite at $-2.2\omega_0$ is missed completely. The NSC-M does not predict a satellite peak until $-0.5\omega_0$, while the SC-M predicts an overly prominent peak splitting for the satellite at $-1\omega_0$. While both self-energy methods are inaccurate, the SC-M is a clear improvement over the NSC-M.

At $k = \pi$ in Fig. 3.1(b), the SC-CE and the PS no longer resemble one another. The SC-CE retains nearly quantitative accuracy across the whole range of energies; it captures the large anti-bonding splitting and additionally captures the smaller satellite at $3\omega_0$. Again there is a small ($< 0.1\omega_0$) shift between the SC-CE and ED. Not visible on the scale of the plot, the SC-CE predicts very weak peaks below the QP energy like for $k = 0$.

The PS more prominently displays negative spectral weight at $k = \pi$ because the Green's function diverges faster in time than for the $k = 0$ case. This can be understood when we consider that the Green's function for $k = 0$ exceeds a magnitude of 10 for the first time around $t \approx 140$ while the Green's function for $k = \pi$ exceeds a magnitude of 10 near $t \approx 75$. The exponential

damping at these two times differ by a factor of 130, so the $k = 0$ negative spectral weight is washed out in the broadening. Because the exponential divergence is slow we can plot a spectrum with reasonably small broadening. Aside from the negative spectral weight the PS predicts under-weighted peaks with approximately the right inter-peak spacing but noticeably shifted absolute energies.

Here, the SC-M performs far worse than at $k = 0$ and does not capture the anti-bonding splitting but rather distorts the NSC-M anti-bonding peak into a broad peak with several shoulders and small splittings.

We note that for $\lambda \ll 1$ all three self consistent methods are nearly indistinguishable and capture the so-called "anti-bonding orbital splitting" at $k = \pi$. See for example Ref. 80 Fig. 3 which presents $\lambda = 1/32$ with $t_0 = 3$ and $\omega_0 = 6$. The SC-CE is nearly indistinguishable except for a small displacement of the weak satellite at $\omega \approx 3$.

While the PS and SC-M methods differ substantially for small system sizes, we know that as the system approaches the thermodynamic limit, the PS will converge into the SC-M. In Fig. 3.2 we demonstrate this convergence. The self-energy kernel in Eq. 3.8 predicates a $1/N$ falloff in the SC-CE contribution for each k-point. So, when we track the square of the difference between the two methods as a function of system size we expect a $1/N^2$ convergence. The inset in Fig. 3.2 numerically demonstrates this falloff. Hence for our consideration of the thermodynamic limit we will calculate the SC-M with the understanding that the PS is identical given a large enough system.

All of these methods are readily applicable to systems with hundreds/thousands of sites since forward integration of integral differential equations of the form in Eq. 3.1 scales as $O(N_t^2 N_k^2)$ where N_t and N_k are the number of time and grid points respectively. In the case of the pure Dyson methods (NSC-M and SC-M), it is possible and more common to carry out the problem in frequency space.

In Fig. 3.4 we compare the NSC-M, SC-M, and SC-CE against numerically exact VD results from Ref. 28 at a small but finite temperature. Data in Fig. 3.4(d) is taken from Ref. 68 and details

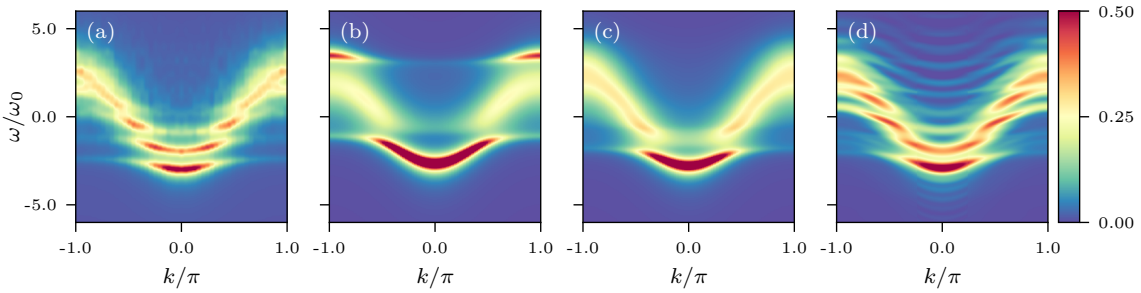

Figure 3.4: Comparison of large system sizes $\lambda = \omega_0 = t_0 = 1$ and $T = 0.1$. (a) VD data from Ref. 28 (b) NSC-M (c) SC-M (d) SC-CE

considering the truncation procedure can be found there.

The NSC-M does predict the QP peak at all momenta, though it is overly weighted. For the satellite peaks, it predicts an erroneous reflection of the QP near $5\omega_0$ at $k = \pi$, but otherwise predicts a featureless spectrum. The SC-M (and thereby the PS) largely corrects the over weighting of the QP and removes the extra peak at $k = \pi$; however it still predicts a featureless satellite region. These observations are consistent with those found in Mitrić et al. where the comparisons with Dynamical Mean Field Theory (DMFT) results in the thermodynamic limit revealed that the NSC-M (one-shot Migdal) was only valid at very weak couplings and the SC-M was of similar quality to the non-self-consistent cumulant expansion.[107] The SC-CE while truncated after approximately $35\omega_0^{-1}$, is very accurate at short times demonstrating an ability to predict the entire satellite structure.[68]

3.4 Conclusion

In this chapter we have explored the connections between several self-consistent time-domain Green's functions methods, namely the SCM approximation, the PS approach, and the SC-CE. For small systems we found that the SC-CE and PS both capture non-perturbative spectral features that neither N-SCM nor SCM predict. We additionally demonstrated the potential emergence of divergent and negative spectral weight in both the SC-CE and PS. In the thermodynamic limit the

SC-CE generates spectra that closely resemble the exact results; however, some of its pathologies may persist. On the other hand, we have demonstrated that in the thermodynamic limit the PS method becomes the SCM approximation.

We next consider the broader possibilities and future developments of these and similar methods. The novel approach taken in deriving the PS method has potential in the systematic search for new approximate methods. In particular, the introduction and application of approximate temporal contraction relationships of the Green's function at the heart of both the PS and the SC-CE is an exciting step forward for generating a class of methods distinct from standard self-energy-based approximations. However, there are trade-offs when departing from diagrammatic rigor in favor of a more general form. On the positive side, when we are willing to relax strict demands of realizability, we have seen that such methods can predict spectral features which standard low-order methods cannot. On the other hand, when we depart from situations where cumulant approaches are guaranteed to be positive definite, namely 2nd order theories, negative spectral weight may appear even if it can be greatly suppressed in the thermodynamic limit when phonon dispersion is present.

There has been comparatively little exploration on methods like the PS or SC-CE, yet there is potential for them to find broad application given that they are computationally inexpensive and potentially accurate. Further work should explore this frontier of approximations to better understand both the limitations and the potential advantages of these and other yet-to-be-derived post-Dyson Green's function methods.

Chapter 4: *Ab Initio* Polaron Structure via an All-Coupling Unitary Transformation

4.1 Introduction

The interaction between charge carriers and phonons is responsible for understanding condensed-phase phenomena from superconductivity to charge transport. One of the most ubiquitous effects is the formation of polarons, a quasiparticle composed of a charge carrier and a cloud of phonons. A polaron can have dramatically different properties from that of the underlying electron/hole. For example, the band gap, carrier mass, and optical conductivity can all be strongly renormalized by polaron formation. It is widely believed that the exceptional charge-transport properties of the lead-halide perovskites arise due to polaron formation,[109] so the development of *ab initio* methods to predict the properties of polarons is essential to predicting the properties of the next-generation of functional materials.

Theoretical descriptions of polarons have evolved over the better part of a century, and have historically concerned a small number of canonical model Hamiltonians.[4] These are idealized limits representing different classes of polarons stripped of complications unnecessary for phenomenological understanding. Among the most influential models, the Fröhlich model describes the interaction of a free electron in a solid with a continuous polarizable medium.[11] The Fröhlich polaron has provided insight into ionic solids such as silver chloride (AgCl), lithium fluoride (LiF), etc.[110] The Holstein model, perhaps the simplest lattice polaron model, idealizes a molecular crystal with tight-binding electrons and a uniform on-site electron-phonon interaction (EPI).[9,10] Although we focus on the case of a single electron in this chapter, the Holstein model has been used to demonstrate the origin of net-attractive interactions between electrons. The Su-Schrieffer-Heeger (SSH) model describes electrons where the EPI modulates the hopping and originally provided a descrip-

tion of excitations in polyenes.[12] Each of these models is rich in phenomenology and the numerous studies of their solutions in different parameter regimes have been foundational to understanding polarons.

While these models are idealizations of the physics in *ab initio* systems, they have stayed the course because they provide physical insight into key aspects of polarons. More recently, numerically exact calculations on such models have been able to definitively answer many questions about the properties of polarons. Rather than obviating the models, these high quality solutions give them a new life as well-understood benchmarks for the development of methods capable of treating *ab initio* systems.

The literature of model electron-phonon systems is rich with variational ansatzes which can, in principle, be adapted to *ab initio* systems. The famous Lee-Low-Pines (LLP) transformation is a variational polaron transform which performs excellently at weak coupling (WC).[111] The adiabatic Landau-Pekar (LP) solution is a variational strong coupling (SC) approximation which treats the phonons via classical displacements.[4,112,113] The LP energy result is also recovered via the SC limit of Feynman's variational theory and by a simple unitary transformation.[114,115] Among the vast options of ansatzes we point out these two because they correctly describe the two limits of WC and SC, and because their forms are similar unitary transformations projected onto the phonon vacuum and single-electron state.

In fact, the difference between the LLP and adiabatic approaches melts away when we consider the WC and SC limits as two ends of a single continuum. Considered from a scattering perspective, WC is coherent scattering where electronic-momentum is conserved while SC involves (near-)impurity scattering, which correspond to an infinitely heavy electron. In the context of the Fröhlich model, Nagy and Markoš introduced an intriguing all-coupling ansatz which takes advantage of this continuum and is itself a generalization of an ansatz explored by Huybrechts in 1977.[116,117] The central insight behind both ansatzes is that the transition from WC-like to SC-like scattering can be represented by a parameter modulating the degree to which momentum is conserved in electron-phonon scattering events. The Huybrechts ansatz of Ref. 117 applied a

single variational parameter scaling the electronic momentum in all scattering processes by the same amount and has received considerable attention because it is a simple ansatz and avoids erroneously predicting an energy discontinuity in the Fröhlich model, which is a pitfall of many attempted variational ansatzes.[118] Far less attention has been given to the Nagy and Markoš (NM) ansatz of Ref. 116, which extends the Huybrechts ansatz by allowing the degree of momentum conservation to vary across the electronic band.

The bespoke nature of model Hamiltonians has long-inspired the development of computational tools to calculate EPI-renormalized quantities *ab initio*.[42] Such tools are desirable for quantitative understanding and prediction of new properties prior to experiments. Until recently, the main tools available to treat EPIs in *ab initio* systems were the Allen-Heine-Cardona (AHC)[119,120] theory and second-order cumulant expansions.[44] AHC theory incorporates the first-order Fan term and the second-order Debye-Waller terms in the many-body perturbation expansion and is routinely applied to calculate the temperature-driven effect of EPIs on the electronic bands; however, small polaron formation is a non-perturbative effect that AHC theory cannot capture. Similarly, the second-order cumulant expansion (CE), while capable of predicting polaron formation, is a perturbative theory.

Recently, several pioneering studies explored different approaches to calculating the polaron binding energy in *ab initio* systems.[45,46,121–124] Ref. 45 approaches the problem via a Landau-Pekar-like variational ansatz and Ref. 123 approaches it non-variationally via a Lang-Firsov-like canonical transformation.[125] Connections have been elucidated between the two methods which demonstrate that these methods both apply the same the same strong coupling ansatz to the problem and can both capture strongly localized polarons such as the hole-polaron in LiF.[124] The drawback for the Lang-Firsov-like transformation is that, unlike for the Holstein model the transformation cannot be closed exactly. Neither of these methods are able to accurately describe large polarons such as the electron-polaron in LiF.

A recent Green's function approach by Lafuente-Bartolome et al.[121,122] presented a robust many-body framework for solving for the polaron energy at all couplings. Their polaron equa-

tions extended beyond previous Green's function formalisms[126] by including not only the Fan-Migdal contribution, but also a polaron term which accounts for "static renormalization...due to localization effects."[122] As they noted, their resulting equations evoke the Hedin equations of GW theory.[70]

One avenue for development is certainly continuing down the path laid in Ref. 122, and we have no doubt that future studies which expand upon and explore that formalism will be of great consequence to the field. The direct description of spectra, as well as the wealth of avenues for systematic improvement are both major benefits of this approach. The drawback to their method is its complexity and non-variational nature.

The complications and expense of implementing the full set of these equations preclude a full self-consistent solution, so in the spirit of the successful G_0W_0 approximation, the authors of Refs. 121 and 122 make several approximations. Namely, applying only only the bare vertex, and replacing renormalized phonon propagators with the bare phonon propagator. They additionally approximate the polaron as composed of coherent states. Finally, they only self-consistently solve for the polaronic term, and apply the Fan-Migdal term perturbatively. This results in a method which, like those mentioned above, is based on SC but also includes first-order corrections.

All three of the aforementioned methods perform well when the system is well described by classical displacements as a starting point. However, the strong-coupling ansatz is not always a good description of the system and in such cases the methods presented in Refs. 45, 46, and 123 may dramatically under-bind the polaron while the method of Refs. 121, and 122 will capture the polaron energy purely through the Fan term.

In this chapter we develop a simple all-coupling ansatz for *ab initio* calculations. In Sec. 4.2 we will discuss the original NM ansatz and its properties in the context of the Fröhlich model. In Sec. 4.3, we present our *ab initio* Nagy-Markoš (*AI*-NM) ansatz and discuss it's properties. After a brief detour on the Holstein model in Sec. 4.3.1 to better understand *AI*-NM for band theories, we we will demonstrate the performance of the ansatz on *ab initio* calculations for LiF in Sec. 4.3.2. There we show that it always improves upon the better of the WC or SC approaches. Finally,

we conclude this chapter with general remarks about the broader applicability of the presented methods.

4.2 Original Nagy-Markoš Ansatz

To better understand the strengths and weaknesses of the original NM ansatz, we apply it to the Fröhlich model as was done in the original paper, and we refer the reader to Ref. 116 for a derivation and description. Thus, while the main method of this chapter is intended for application in *ab initio* systems, this section presents information about the general properties of the NM ansatz.

In Ref. 116 the authors considered a generalized Fröhlich model which allows for several types of polarons to be considered. The Hamiltonian we consider is

$$H = \sum_k \epsilon_k |k\rangle\langle k| + \omega_0 \sum_q b_q^\dagger b_q + \sum_{kq} g(q)|k+q\rangle\langle k|(b_q + b_{-q}^\dagger). \tag{4.1}$$

We restrict our discussion here to a dispersionless longitudinal optical (LO) phonon of frequency ω_0 and an EPI with $|g(q)|^2 = \frac{4\pi\alpha\omega_0^{3/2}}{\sqrt{2V}q^2}$ where α is the dimensionless coupling constant. The free-particle dispersion is given by $\epsilon_k = k^2/2m$.

Nagy and Markoš applied a unitary transformation which they defined in Eq. 6 of 116. They applied a form of the density operator in their transformation where the projection of the density is written $\sum_k |k + a_q q\rangle\langle k|$, where a_q is a variational parameter defined for all q to be between zero and one. They additionally consider an ansatz for the electronic wavefunction of a Gaussian form (an assumption which we will relax in our later *ab initio* version), with the k-space wavefunction given by $\sqrt{\frac{8}{V}}\left(\frac{\pi}{\lambda}\right)^{3/4} e^{\frac{-\mathbf{k}^2}{2\lambda}}$. They were able to write the energy as an integral equation dependent on λ.

The results of the NM ansatz for the Fröhlich model are summarized in Fig. 4.1(a). As mentioned previously, the Fröhlich model is challenging for variational approaches, and most incorrectly[127] predict a phase transition when transitioning from WC to SC in the form of a discontinuity

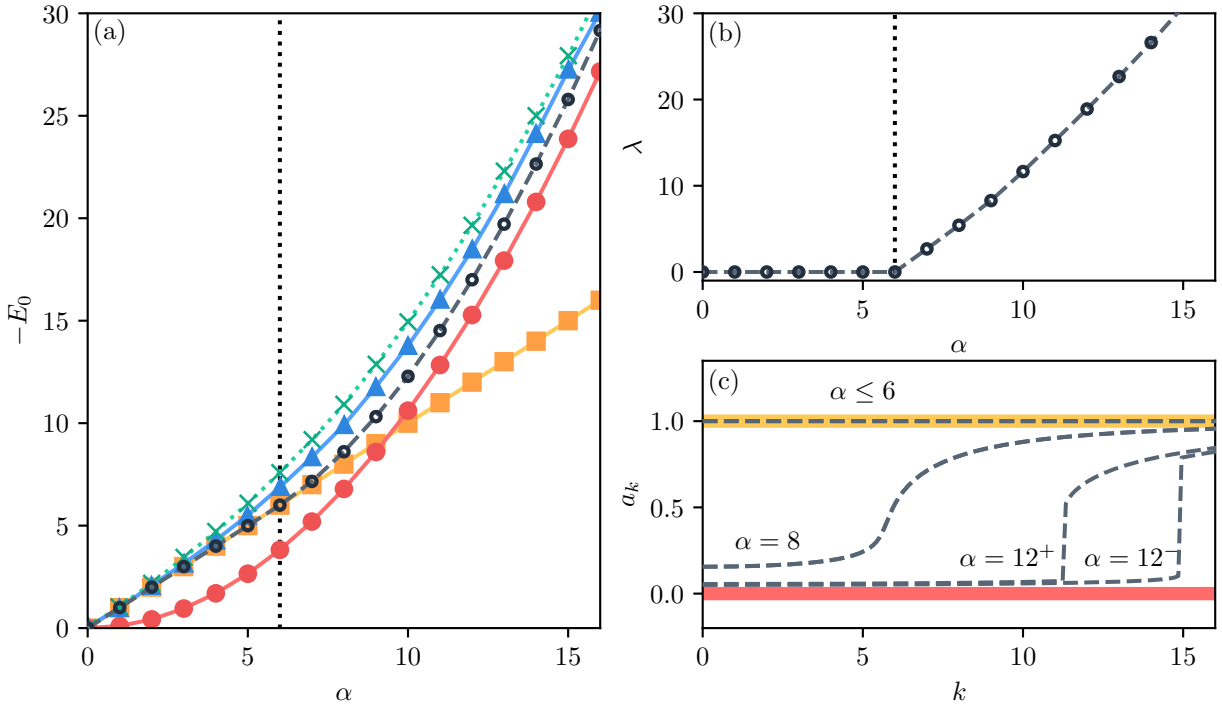

Figure 4.1: (a) The ground state energy of the Fröhlich model solved via Diagrammatic Quantum Monte Carlo from Ref. 25 (blue triangles), the Green's function approach from Ref. 121 (green X with dotted line), the WC ansatz (orange squares), the SC ansatz (red circles), and the NM ansatz (open black circles with dashed lines). We note that the NM ansatz is generally in good agreement with the exact result. (b) λ, the inverse width of the electronic wavefunction for the all-coupling ansatz acts as an order parameter for the transition from weak-coupling to strong coupling behaviour. (c) The diversity of behaviours for a_k in the model. For $\alpha \leq 6$ the a_k are uniformly 1, while for large couplings (e.g. $\alpha = 8$) the a_k tend towards 0 near $k = 0$. At the very large $\alpha = 12$, we display a_k optimized starting from 1 (12^+) and 0 (12^-). The red and orange lines at 0 and 1 indicate the strong and weak coupling ansatzes respectively.

in the energy. The NM ansatz predicts only a discontinuity in the derivative of the energy at $\alpha = 6$ which is an improvement over a discontinuity in the energy's value. The solution matches the perturbative WC result ($E = -\alpha$) up to $\alpha = 6$ where it then transitions not to the SC solution ($E = -\alpha^2/3\pi$), but to a solution that asymptotically approaches the SC solution.[4] Comparing with the numerically exact Diagrammatic Quantum Monte Carlo data of Ref. 25, the NM ansatz never errs by more than $\approx 14\%$. The Green's function approach of Ref. 121 has a slightly lower maximum error of $\approx 10\%$; however, it consistently overshoots the exact energy and does not do as well as the NM ansatz or the pure WC result at weak couplings. The added flexibility of the NM ansatz over both the SC and WC ansatzes allows it to closely match the exact result across all couplings with a relatively simple calculation.

As was also discussed in Ref. 116, we show in Fig. 4.1(b) that the optimized inverse width, λ, of the Gaussian electronic state acts as an order parameter and becomes non-zero at $\alpha = 6$. Fig. 4.1(c) demonstrates that the strong coupling limit is approached very slowly and that momenta far from $k = 0$ will always be treated in the weak coupling theory for any finite coupling. We also note that at very strong couplings the cubic equation for a_k admits multiple solutions, but the energy differences between them at even extreme coupling values ($\alpha = 20$) are small.

The Fröhlich model demonstrates that the NM ansatz provides an excellent approximation to the exact energies, and motivates its application to an arbitrary electron-phonon system.

4.3 *Ab Initio* Nagy-Markoš Ansatz

In both the Huybrechts and NM ansatzes, the formalism was limited to model Hamiltonians with forms like the Fröhlich model. In this section we develop a new *ab initio* Nagy Markoš (*AI*-NM) ansatz which is applicable to general electron-phonon problems.

The *ab initio* electron-phonon Hamiltonian is given to leading order in the EPI as

$$H = \sum_{iq} \epsilon_{iq} |i,q\rangle\langle i,q| + \sum_{q,\nu} \omega_{\nu q} b^\dagger_{\nu q} b_{\nu q} + \sum_{\substack{ij\nu \\ kq}} g^{ij}_\nu(k,q) |i,k+q\rangle\langle j,k| (b_{\nu q} + b^\dagger_{\nu -q}). \tag{4.2}$$

Here, the electronic bands, phonon frequencies and, electron-phonon matrix elements are given by ϵ_{iq}, $\omega_{\nu q}$ and $g_{\nu}^{ij}(k,q)$ respectively. The phonon annihilation/creation operators are given by $b_{\nu q}/b_{\nu q}^{\dagger}$. Because we are interested in applying this formalism to a single electron/hole, we work in the projector basis for electrons/holes, thus simplifying the application of the transformation itself.

We define our *AI*-NM ansatz via the trial wavefunction

$$|\Psi_{t_{ik}h_{\nu q}a_q}\rangle = U|0\rangle_{\text{ph}} \otimes |\psi\rangle_{\text{el}}. \tag{4.3}$$

The electronic portion of the wavefunction is given by the familiar configuration interaction (CI) singles form

$$|\psi\rangle_{\text{el}} = \sum_{ik} t_{ik}|i,k\rangle. \tag{4.4}$$

The phonon trial function is a product of a unitary transformation $U = e^S$ onto the phonon vacuum with

$$S = \sum_{\substack{i\nu \\ kq}} B_{\nu q}|i,k+a_q q\rangle\langle i,k|, \tag{4.5}$$

$$B_{\nu q} \equiv h_{\nu q}b_{\nu q} - h_{\nu -q}^{*}b_{\nu -q}^{\dagger}. \tag{4.6}$$

There are three sets of variational parameters to be minimized for this ansatz. The set of CI coefficients t_{ik} represent the electronic wavefunction, the $h_{\nu q}$ parameters are displacements of the phonon modes, and modulate the number of phonons in our ground state, and the a_k modify the momentum conservation at each wave-vector. In order to not allow momentum *increasing* processes, we constrain a_k to be a real number bounded between zero and one as in the NM ansatz. Additionally, in order for U to be unitary we must require that $a_k = a_{-k}$. From Eq. 4.5, it is apparent that the ansatz transitions from the LP-like SC approximation with classical phonons

($a_k = 0$) to a WC LLP-like transformation ($a_k = 1$).[111]

As is standard, we find an upper bound to the ground-state energy using the variational principle

$$E_0 \leq \min_{t_{ik} h_{\nu q} a_q} \frac{\langle \Psi_{t_{ik} h_{\nu q} a_q} | H | \Psi_{t_{ik} h_{\nu q} a_q} \rangle}{\langle \Psi_{t_{ik} h_{\nu q} a_q} | \Psi_{t_{ik} h_{\nu q} a_q} \rangle}. \tag{4.7}$$

In the evaluation of Eq. 4.7, we come across our first major difference between the NM and *AI*-NM approaches. Whereas in the NM ansatz the energy expression can be fully evaluated, it is not the case for the *AI*-NM because the application of the unitary transformation on Eq. 2.6 does not close. Thus, we must approximate the transformation.

While an approximate closure scheme necessarily makes the method non-variational, there are several important cases where the unitary transformation does close exactly and the method becomes rigorously variational. First, there is the WC limit where the electron-phonon coupling tends to zero. Second, there is the SC limit where a_k is uniformly zero. Finally, there is the special case where the $g_\nu^{ij}(k, q)$ does not depend on the electronic-momentum and ϵ_{ik} is parabolic. This last condition covers the case of the Fröhlich model and explains why the NM ansatz closed exactly. Full details of the transformation are given in Appendix J.

In the general *AI*-NM method, we approximate closure by retaining terms up to $O(h_{\nu q})^2$. This sacrifices the rigorous variational nature of our ansatz, but we can systematically approach a variational calculation by either including higher-order terms or constraining the maximum magnitude of $h_{\nu q}$ closer to zero. For this chapter we choose the simplest implementation of this constraint and apply a uniform cutoff which preferences against and discards results with $\max(|h_{\nu q}|) > 1.5$.

Within this approximation the energy given by our ansatz is

$$E = \frac{1}{\sum_{ik} |t_{ik}|^2} \Bigg\{ \sum_{iq} \left(\epsilon_{iq} - \frac{1}{2} \sum_{\nu k} |h_{\nu k}|^2 \mathcal{E}_{iqk} \right) |t_{iq}|^2 + \sum_{\nu q} \omega_{\nu q} |h_{\nu q}|^2 \sum_{iq} |t_{iq}|^2 - \sum_{\substack{ij\nu \\ kq}} \left(g_{\nu}^{ij}(k,q) h_{\nu -q} t^*_{ik+q} t_{jk+a_q q} + \text{c.c.} \right) \Bigg\}, \tag{4.8}$$

$$\mathcal{E}_{iqk} \equiv 2\epsilon_{iq} - \epsilon_{iq-a_k k} - \epsilon_{iq+a_k k}. \tag{4.9}$$

Up to this point we have formally assumed that we are in the thermodynamic limit; however, in order to evaluate Eq. 4.8 and its gradients we must define how to treat terms such as $t_{ik+a_q q}$ when momentum is discretized as is the case in any periodic DFT calculation. The presence of a continuous a_k which modifies the momentum arguments of the variational parameter brings us off of the grid to arbitrary points in the continuum. While this is hard to justify in a truly finite system, we are interested in properties extrapolated to the thermodynamic limit, so we develop a procedure which is applicable to systems of all sizes.

Taking inspiration from the Wannier interpolation methods routinely used to interpolate $g_{\nu}^{ij}(k,q)$ elements from a coarse to a fine grid, we define all of our 'off-grid' quantities via Fourier interpolation. This is formally similar to applying twisted boundary conditions, a well-known approach to approximating the thermodynamic limit.[28,128] This procedure not only allows us to have a well-defined procedure to evaluate our energy for any set of variational parameters, but as is detailed in Appendix K, we can also evaluate the gradients of Eq. 4.8 analytically, allowing efficient and direct minimization for Eq. 4.7.

4.3.1 Application to 1D Holstein Model

As a first step towards an arbitrary electron-phonon system, we consider the one-dimensional Holstein model. This is the first application of the numerical method developed in Sec. 4.3 and is a helpful first step because it isolates the complexities in the method from the inherent complexities of *ab initio* calculations while still allowing us to develop the method on a finite-sized lattice. The

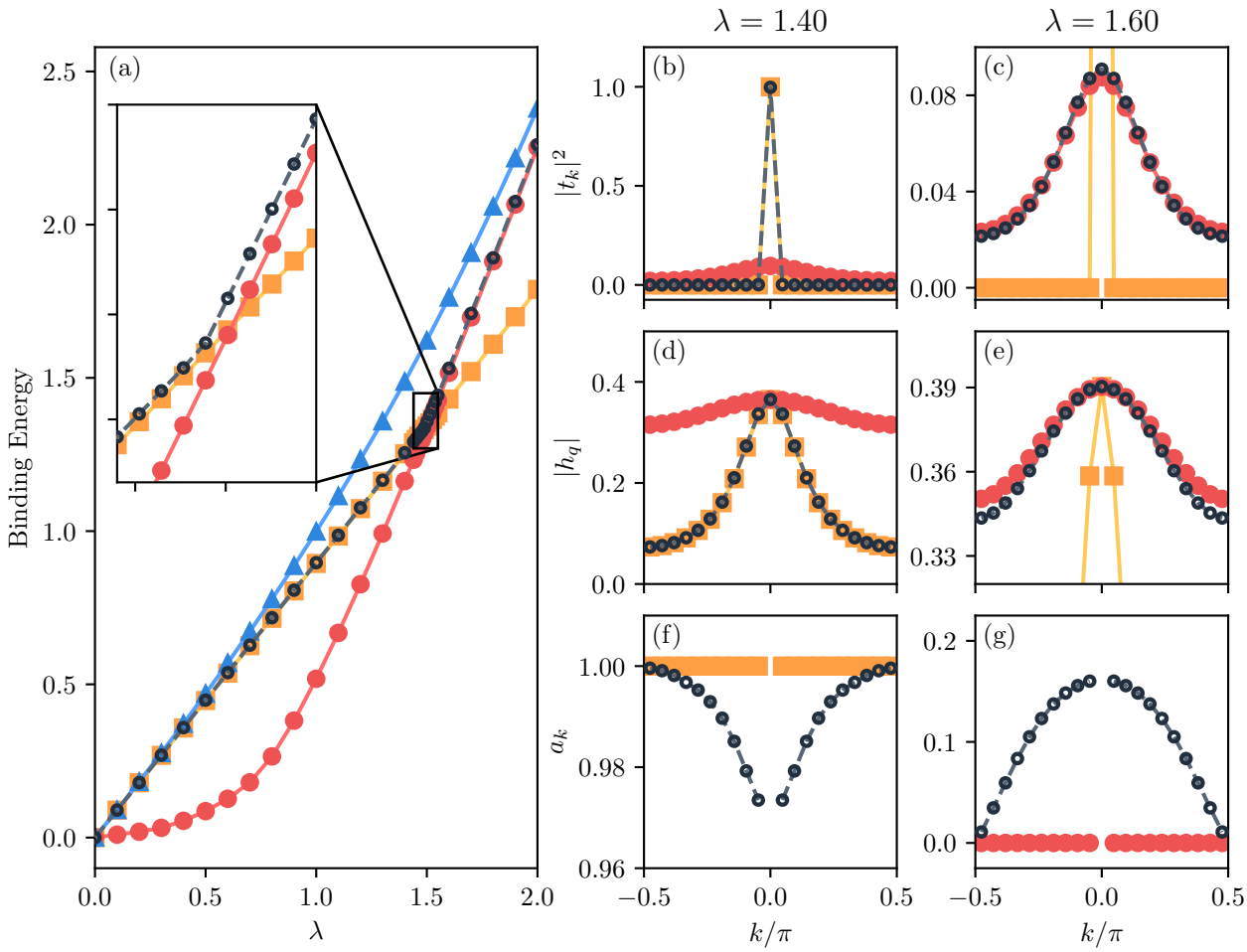

Figure 4.2: The 21-site 1D Holstein model with $\omega_0 = t_0 = 1$. We compare DMRG (blue triangles), weak coupling (orange squares), strong coupling (red circles) and the *AI*-NM ansatz (open black circles and dashed line). (a) Comparison of the polaron binding energy from each method across a wide range of dimensionless couplings λ. The zoomed inset further details the transition from weak to strong coupling and demonstrates that the *AI*-NM ansatz always improves on the better of the strong or weak coupling solution. (b), and (c): The electron densities in momentum space. (d), and (e): The magnitude of the phonon displacements $|h_q|$. (f), and (g): The deviations a_k from the WC and SC solutions respectively. (b), (d), and (f) show the variational quantities before the transition from weak to strong coupling at $\lambda = 1.4$ while (c), (e), and (g) display them after the transition at $\lambda = 1.6$.

Holstein model is of the form of Eq. 4.1 with the choices of $\epsilon_k = -2t_0\cos(k)$, and $g(k,q) = g/\sqrt{N}$. N is the number of sites, and t_0 is the electronic hopping frequency. We define the dimensionless coupling as $\lambda = g^2/2t_0\omega_0$. We compare the results from *AI*-NM against numerically exact density matrix renormalization group (DMRG) results.[129,130]

In Fig. 4.2(a), we show that the *AI*-NM ansatz generally tracks the better of the WC or SC solution and slightly improves upon each. The inset demonstrates that, like in the case of the Fröhlich model, there is a cusp in the energy as the system transitions from WC-like to SC-like. It is known, and reflected by the DMRG data, that the one-dimensional Holstein model exhibits a WC-SC crossover rather than a phase transition, but the slight kink in the energy is consistent with the original NM method and better than a true energy discontinuity.

In Fig. 4.2(b) (d), and (f) we more closely examine the variational parameters on either side of this transition. We demonstrate that the WC portion of the *AI*-NM solution is nearly identical to the WC-solution itself. In the original NM solution for the Fröhlich model NM solution is exactly the WC solution until $\alpha = 6$, so it may be the case that when WC is the better solution, *AI*-NM will always converge to it exactly in the thermodynamic limit. Conversely, in Fig. 4.2(c),(e), and (g) we find that *AI*-NM is visibly different from the SC solution and the energy improvement is larger.

Compared to the DMRG results, we do not fare as well as with the Fröhlich model, which may be attributable to the BCH truncation having a large effect on this system, or may indicate that the *AI*-NM is better suited to Fröhlich-like coupling. Either way, while the crossover in the Holstein model is not as dramatic as in the Fröhlich model, it demonstrates that the ansatz is capable of switching between the two limits correctly.

In applying *AI*-NM to the Holstein model, we encountered many possible solutions. This differs from the original NM approach where the nearly analytic form presented only one case of multiple solutions which hardly changed the energy. Here we locate the minimum by using starting guesses from the WC solution, the SC solution and a set of randomly chosen distributions.

4.3.2 Application to LiF

We now move to address the case of a fully *ab initio* system and apply *AI*-NM to LiF. We follow the procedure for generating ϵ_{ik}, $\omega_{\nu q}$ and $g^{ij}_{\nu}(k,q)$ discussed in Ref. 122. We begin by optimizing the geometry in Quantum Espresso[131,132] with a 12x12x12 k-grid using the the PBE exchange-correlation functional[133] with Optimized Norm-Conserving Vanderbilt Pseudopotentials,[134,135] and a 150 Ha energy cutoff. We find an optimized lattice constant of $a_0 = 4.035$ Å. We then apply density functional perturbation theory (DFPT) as implemented in the PHonon code within Quantum Espresso to find the phonon frequencies and electron-phonon couplings on an 11x11x11 grid. We next construct the maximally localized Wannier functions with Wannier90[136] and use EPW[137] to interpolate the parameters onto all of the k-grids applied in this chapter. For the hole-polaron we consider only the three highest energy valence bands while for the electron-polaron we consider only the lowest energy conduction band.

To better understand the *AI*-NM approach, we compare its against two other methods; one exact and one perturbative. Although the largest grids are eventually the ones necessary to extrapolate to the thermodynamic limit we can use smaller grids to compare our method against data from the exact neural-network quantum state (NNQS) approach of Ref. 138. For all system sizes we can apply perturbation theory on top of the SC result. This procedure is known as coherent-state Møller-Plesset perturbation theory (CSPT).[81]

It is well-understood that the hole-polaron is spatially localized within a single unit cell and is well-described by classical nuclear displacements.[45–47,121,122] In Fig. 4.3, we show that our ansatz appropriately recovers this SC limit.[1] In Fig. 4.3(a) we plot the convergence of the polaron ground state energy with respect to the k-mesh size. We find that by even a modestly sized 13x13x13

[1]We note that the 3x3x3 hole-polaron (which is not shown in Fig. 4.3) predicts an optimized binding energy of 1.68 eV, far below the numerically exact NNQS value of 0.90 eV from Ref. 138. Further examination of this result reveals large values of $h_{\nu q}$ with $\max(|h_{\nu q}|) = 1.15$ and $a_k \approx 1 \approx WC$. For comparison the optimized *AI*-SC and WC results from 5x5x5 have maximum magnitudes of 0.08 and 0.06 respectively. The regime of large a_k and large $h_{\nu q}$ is exactly where we expect the truncation of the BCH expansion to most adversely effect our results when we discard solutions via a uniform $|h_{\nu q}|$ cutoff. A more sophisticated error analysis of Eqs. J.7 and J.14 could produce a cutoff procedure for $h_{\nu q}$ which is a function of a_k. One expects such an approach to naturally allow $h_{\nu q}$ to take arbitrarily large values when $a_k = 0$ and the the transformation closes and restrict $h_{\nu q}$ to small values when a_k approaches zero.

supercell size	WC (eV)	SC (eV)	*AI*-NM (eV)	CSPT2 (eV)	NNQS (eV)
$3 \times 3 \times 3$	0.23	0.90	1.68	0.91	0.90
$5 \times 5 \times 5$	0.25	1.27	1.27	1.28	1.27
$9 \times 9 \times 9$	0.35	1.58	1.58	1.60	1.59
$13 \times 13 \times 13$	0.36	1.67	1.67	1.70	–

Table 4.1: Selected binding energies for the LiF hole-polaron at various system sizes via the WC, SC, and *AI*-NM approaches as well as CSPT2 and the exact NNQS approach.[138]

k-grid, the energy can evaluated in the thermodynamic limit via Makov-Payne extrapolation.[139] *AI*-NM has an extrapolated binding energy of 1.94 eV, which is consistent with previous findings.[46,121,123] It is nearly identical to the SC result of 1.97 eV. The difference can be attributed to an artifact of our two-point linear extrapolation. The WC result on the other hand begins plateauing to a binding energy of 0.30 eV reflecting its qualitatively incorrect nature for the hole-polaron. In Fig. 4.3(b) the electron densities ($|t_{ik}|^2$) are nearly identical between the SC and the *AI*-NM ansatzes except for a slightly sharper peak at the Gamma point. The WC result predicts a fully delocalized polaron state we would expect. The plot of displacements squared in Fig. 4.3(c) similarly shows nearly identical values between *AI*-NM and SC. The optimized a_k presented in Fig. 4.3(c) have vanishingly small values indicating that the *AI*-NM is effectively variational and in the SC limit.

Table 4.1 zooms in on the results of Fig. 4.3(a) and demonstrates just how closely the exact NNQS approach of Ref. 138 matches the SC, *AI*-NM, and CSPT2 results. The SC result is essentially exact for the LiF hole-polaron. Further, at larger system sizes, where the exact method of Ref. 138 is unavailable, CSPT2 remains in close agreement with the SC and *AI*-NM results.

On the opposite end of the WC-SC spectrum, we consider the LiF electron-polaron. Conventionally, it is considered a close analogue of the Fröhlich large polaron. However, it has long been suggested that the model's sensitivity to input parameters may render it inadequate to fully understand polarons in alkali halides,[110] making it a good target for *ab initio* approaches. Unlike the hole-polaron, only a single electronic band contributes to the electron-polaron. However the electron-polaron converges much more slowly with system size than the hole-polaron, and we compute k-grids up to 27x27x27 for our Makov-Payne extrapolation.

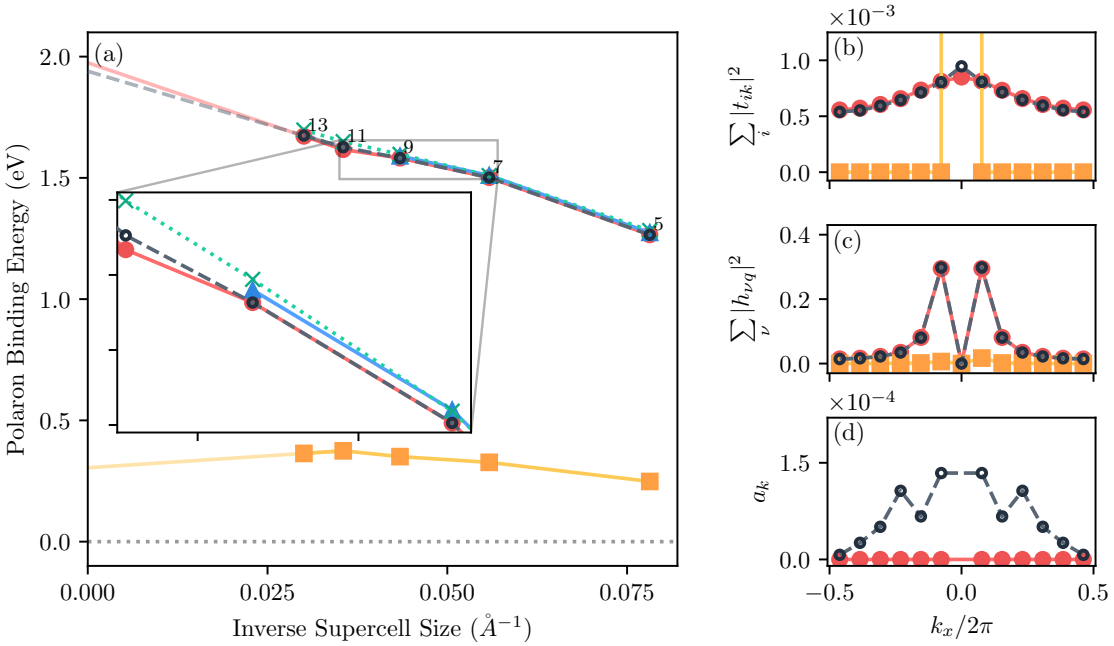

Figure 4.3: LiF hole-polaron properties via the WC ansatz (orange squares), the SC ansatz (red circles), the *AI*-NM ansatz (open black circles), CSPT2 (green Xs and dotted line), and the exact NNQS approach of Ref. 138 (blue triangles). (a) plot of the convergence of the polaron binding energy with respect to the inverse supercell size (as defined in Ref. 46). The small numbers above the data indicate the size of the supercell. The zoomed inset higlights the small differences between the methods. (b), (c), and (d) are slices of variational parameters along a single reciprocal lattice vector for the 13x13x13 k-mesh. (b) and (c) present the sums of the electron densities and displacements respectively and both nearly exactly match the SC result. (d) shows that a_k is almost exactly that of the SC limit.

Fig. 4.4(a) presents the energy convergence with respect so system size, and *AI*-NM predicts a polaron binding energy of 0.34 eV while WC predicts 0.35 eV.[2] SC predicts a significantly lower 0.25 eV binding energy, and does not even bind at all until N=17 which is similar to the result of Ref. 46. The discrepancy between *AI*-NM and WC can again be attributed to small artifacts in our extrapolation. Unlike in the hole-polaron case, *AI*-NM visibly departs from the WC solution at lower system sizes which can be understood by examining the variational parameters. The electron densities in Fig. 4.4(b) show an optimal *AI*-NM solution between the delocalized and localized WC and SC solutions. The *AI*-NM phonon displacements in Fig. 4.4(c) only slightly deviate from the WC displacements. Finally, Fig. 4.4(d) shows that the a_k are not simply one or zero, but a smooth looking function falling off from near one around the band edges to 0.84 around the Gamma point (the Gamma point itself does not have a well-defined a_k). This shape recalls the a_k profiles of Fig. 4.1(c) but is slightly non-monotonic on a scale imperceptible in the plot.

To better visualize the differences between the electron densities predicted by the methods, we plot a three-dimensional representation of the k-space electron density for *AI*-NM in Fig. 4.4(e), SC in Fig. 4.4(f) and WC in Fig. 4.4(g). As expected, WC predicts nearly a delta function in k-space corresponding to a fully delocalized polaron, SC predicts a very broad and spherical k-space density profile. *AI*-NM predicts a more complex electron density. There is a non-spherical central peak as well as local maxima along the reciprocal lattice vectors. This suggests that *AI*-NM will be able to provide further insights into the structure of the polarons in real-space.

In Table 4.1 we more closely examine the binding energies including small systems not shown in Fig. 4.4(a). For the two smallest system sizes we compare our results to the exact NNQS approach of Ref. 138. The exact binding energies predicted by that approach are 0.036 and 0.068 eV respectively which compares with our best *AI*-NM results of 0.040 and 0.079 eV. For those system sizes WC theory is nearly correct and produces binding energies of 0.034 and 0.069 eV.

[2]Because our parameters are interpolated from coarser to finer k-grids, there can be inaccuracies in that mapping. Here we find that for the largest k-grids, the energetic minimum shifts slightly off of the Gamma point. To minimize the effect of this shift in the energy, all energy differences are taken from the minimum of the smallest k-grid. However the CSPT2 results shown in Fig. 4.4(a) have an anomalously low energy difference between 19x19x19 and 21x21x21 because of this artifact.

supercell size	WC (eV)	SC (eV)	*AI*-NM (eV)	CSPT2 (eV)	NNQS (eV)
$3 \times 3 \times 3$	0.034	0.00	0.041	0.034	0.036
$5 \times 5 \times 5$	0.070	0.00	0.079	0.069	0.068
$15 \times 15 \times 15$	0.18	0.00	0.19	0.18	–
$23 \times 23 \times 23$	0.23	0.087	0.23	0.23	–

Table 4.2: Selected binding energies for the LiF electron-polaron at various system sizes via the WC, SC, and *AI*-NM approaches as well as the CSPT2 and the exact NNQS approach.[138]

While this makes it clear that our results are non-variational in this case, Fig. 4.4(a) demonstrates that the *AI*-NM approach tends towards the WC result, so the large deviations from variationality may be exacerbated by the intrinsic interpolation errors *AI*-NM introduces in small systems. For these system sizes, CSPT2 only deviates from the NNQS approach by 0.002 eV at a maximum suggesting that it is a good benchmark to compare against *AI*-NM for the larger systems. In the larger system sizes presented in Table 4.2 and Fig. 4.4(a), the WC approach, the *AI*-NM approach and CSPT2 are nearly identical to one another.

4.4 Conclusion

In this chapter we introduced the *AI*-NM ansatz, a flexible variational ansatz for polaron problems of any coupling strength. Like the original NM method, the *AI*-NM method joins the well-known SC and WC ansatzes by introducing an additional set of variational parameters into the unitary transform. *AI*-NM allows for the treatment arbitrary electron-phonon Hamiltonians, but because the generalization prevents the equations from closing exactly it is not variational except in some special cases. We highlighted the abilities of the original NM ansatz on the Fröhlich model, and showed that the *AI*-NM behaves similarly on the Holstein model. In both cases the ansatz always tracks and potentially improves the better of the WC or SC solution. When we applied the method to a fully *ab initio* model of LiF, we found that it was capable of accurately predicting the polaron binding energies for both the small hole-polaron and the large electron-polaron.

As we demonstrated with LiF, the *AI*-NM can be used to compute polaron binding energies as a stand-alone tool. However, future developments could apply *AI*-NM as a starting point for more accurate calculations. One clear improvement would be to include the higher-order displacements

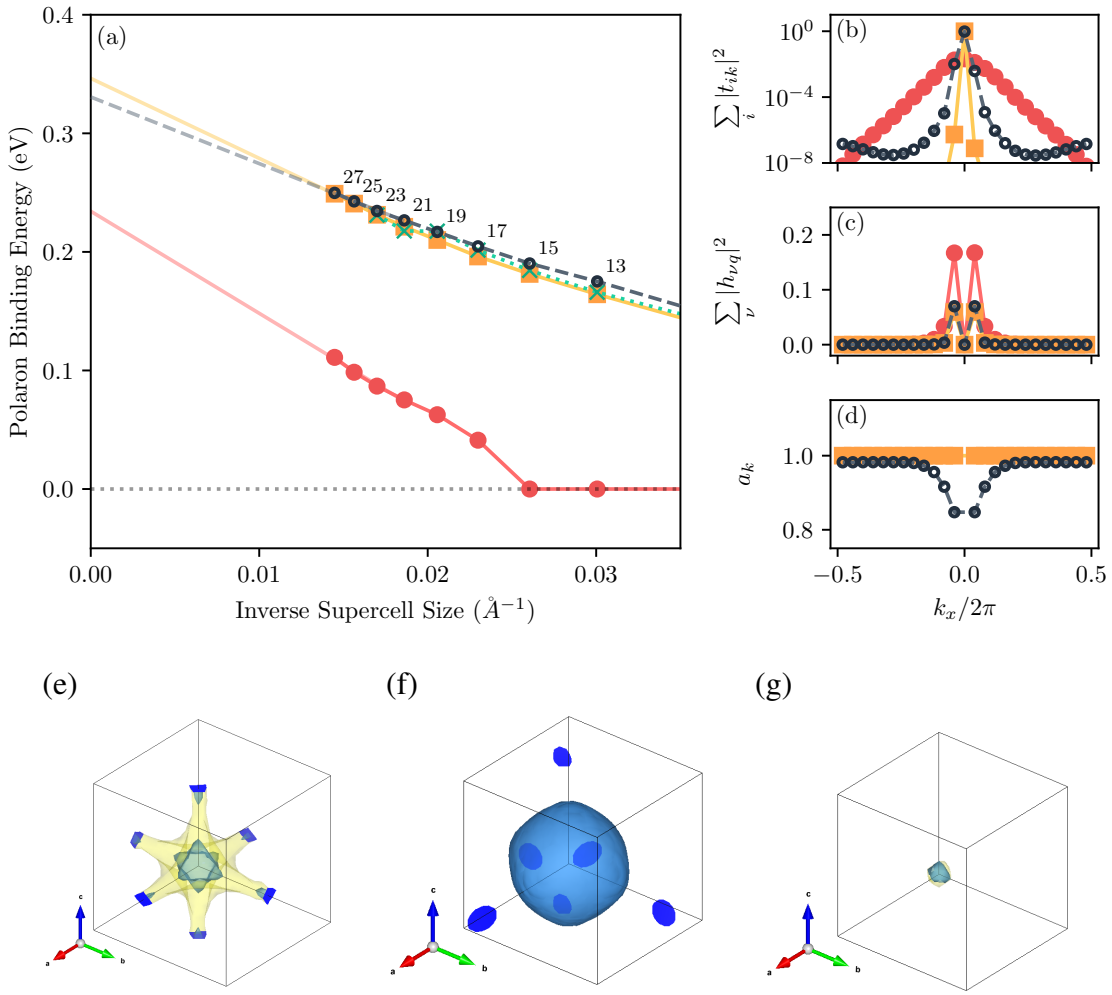

Figure 4.4: (a)-(d) have an identical layout and key to Fig. 4.3. (a) demonstrates that *AI*-NM converges towards WC in the thermodynamic limit. (b), (c), and (d) are slices of the variational parameters from a 25x25x25 k-mesh. (b) shows that *AI*-NM produces visibly different electron densities from WC and SC. (c) demonstrates that the displacements are similar between the *AI*-NM and WC results. (d) demonstrates that the optimized a_k deviate far from the WC limits. (e)-(g) Electron density isosurface plots of the LiF electron-polaron calculated via the *AI*-NM (e), the strong-coupling ansatz (f), and the weak-coupling ansatz (g). All isosurfaces are from 21x21x21 k-grids. The blue isosurface corresponds to an isovalue of 10^{-7} while the yellow isosurface corresponds to an isovalue of 10^{-9}.

either explicitly or pertrubatively. Also, we did not consider the Debye-Waller terms which are second-order in the EPI, and are included in AHC theory and the methods of Refs. 46 and 122. In a different vein, *AI*-NM is an ideal starting point for perturbation theory because it is typically close in energy to the exact solution. Similarly, using the *AI*-NM transformed Hamiltonian, one could construct a low-order Green's function theory and generate spectral functions with the second-order cumulant expansion.

Efficient and scalable computational methods for calculating properties of polarons from first-principles will enable deeper insights into previously inexcessible systems. There are still major questions about the lead-halide perovskites, for example, because their large scale and numerous complexities have made it difficult to fully model how EPIs effect their properties. The *AI*-NM variational ansatz provides a promising avenue of exploration which may be able to bring *ab initio* polaron calculations to the scale of next-generation materials.

References

[1] P. J. Robinson, I. S. Dunn, and D. R. Reichman, "Cumulant methods for electron-phonon problems. i. perturbative expansions," *Phys. Rev. B*, vol. 105, p. 224 304, 22 Jun. 2022.

[2] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, "Theory of Superconductivity," *Phys. Rev.*, vol. 108, no. 5, pp. 1175–1204, Dec. 1957.

[3] L. Wang, D. Beljonne, L. Chen, and Q. Shi, "Mixed quantum-classical simulations of charge transport in organic materials: Numerical benchmark of the Su-Schrieffer-Heeger model," *J. Chem. Phys.*, vol. 134, no. 24, p. 244 116, Jun. 2011.

[4] G. D. Mahan, *Many-Particle Physics*, 3rd. Boston, MA: Springer US, 2000, ISBN: 978-1-4419-3339-3.

[5] A. S. Mishchenko, N. Nagaosa, G. De Filippis, A. de Candia, and V. Cataudella, "Mobility of Holstein Polaron at Finite Temperature: An Unbiased Approach," *Phys. Rev. Lett.*, vol. 114, no. 14, p. 146 401, Apr. 2015.

[6] L. Song and Q. Shi, "Calculation of correlated initial state in the hierarchical equations of motion method using an imaginary time path integral approach," *J. Chem. Phys.*, vol. 143, no. 19, p. 194 106, Nov. 2015.

[7] F. C. Spano, "The Spectral Signatures of Frenkel Polarons in H- and J-Aggregates," *Acc. Chem. Res.*, vol. 43, no. 3, pp. 429–439, Mar. 2010.

[8] A. Oleson *et al.*, "Perylene Diimide-Based Hj- and hJ-Aggregates: The Prospect of Exciton Band Shape Engineering in Organic Materials," *J. Phys. Chem. C*, vol. 123, no. 33, pp. 20 567–20 578, Aug. 2019.

[9] T. Holstein, "Studies of polaron motion: Part i. the molecular-crystal model," *Ann. Phys.*, vol. 8, no. 3, pp. 325–342, 1959.

[10] T. Holstein, "Studies of polaron motion: Part ii. the "small" polaron," *Ann. Phys.*, vol. 8, no. 3, pp. 343–389, 1959.

[11] H. Fröhlich, "Electrons in lattice fields," *Adv. Phys.*, vol. 3, no. 11, pp. 325–361, Jul. 1954.

[12] W. P. Su, J. R. Schrieffer, and A. J. Heeger, "Solitons in Polyacetylene," *Phys. Rev. Lett.*, vol. 42, no. 25, pp. 1698–1701, Jun. 1979.

[13] S. A. Trugman, "General many-body systems," in *Applications of Statistical and Field Theory Methods to Condensed Matter*, D. Baeriswyl, A. R. Bishop, and J. Carmelo, Eds. Boston, MA: Springer US, 1990, pp. 253–263, ISBN: 978-1-4684-5763-6.

[14] J Bonča, S. A. Trugman, and I Batistić, "Holstein polaron," *Phys. Rev. B*, vol. 60, no. 3, pp. 1633–1642, Jul. 1999.

[15] L.-C. Ku, S. A. Trugman, and J. Bonča, "Dimensionality effects on the Holstein polaron," *Phys. Rev. B*, vol. 65, no. 17, p. 174 306, Apr. 2002.

[16] J. Ranninger and U. Thibblin, "Two-site polaron problem: Electronic and vibrational properties," *Phys. Rev. B*, vol. 45, no. 14, pp. 7730–7738, Apr. 1992.

[17] F. Marsiglio, "The spectral function of a one-dimensional Holstein polaron," *Phys. Lett. A*, vol. 180, no. 3, pp. 280–284, Sep. 1993.

[18] E. V. L. de Mello and J. Ranninger, "Dynamical properties of small polarons," *Phys. Rev. B*, vol. 55, no. 22, pp. 14 872–14 885, Jun. 1997.

[19] H. Fehske, J. Loos, and G. Wellein, "Spectral properties of the 2D Holstein polaron," *Zeitschrift für Phys. B Condens. Matter*, vol. 104, no. 4, pp. 619–627, Dec. 1997.

[20] A. S. Alexandrov, V. V. Kabanov, and D. K. Ray, "From electron to small polaron: An exact cluster solution," *Phys. Rev. B*, vol. 49, no. 14, pp. 9915–9923, Apr. 1994.

[21] M. Hohenadler, M. Aichhorn, and W. von der Linden, "Spectral function of electron-phonon models by cluster perturbation theory," *Phys. Rev. B*, vol. 68, no. 18, p. 184 304, Nov. 2003.

[22] G. De Filippis, V. Cataudella, V. Marigliano Ramaglia, and C. A. Perroni, "Static and dynamic polaron features in a coherent-state basis," *Phys. Rev. B*, vol. 72, no. 1, p. 014 307, Jul. 2005.

[23] M. Berciu, "Green's Function of a Dressed Particle," *Phys. Rev. Lett.*, vol. 97, no. 3, p. 036 402, Jul. 2006.

[24] G. L. Goodvin, M. Berciu, and G. A. Sawatzky, "Green's function of the Holstein polaron," *Phys. Rev. B*, vol. 74, no. 24, p. 245 104, Dec. 2006.

[25] A. S. Mishchenko, N. V. Prokof'ev, A. Sakamoto, and B. V. Svistunov, "Diagrammatic quantum Monte Carlo study of the Fröhlich polaron," *Phys. Rev. B*, vol. 62, no. 10, pp. 6317–6336, Sep. 2000.

[26] J. Jaklič and P. Prelovšek, "Finite-temperature properties of doped antiferromagnets," *Adv. Phys.*, vol. 49, no. 1, pp. 1–92, Jan. 2000.

[27] P. Prelovšek and J. Bonča, "Ground state and finite temperature lanczos methods," in *Strongly Correlated Systems: Numerical Methods*, A. Avella and F. Mancini, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 1–30, ISBN: 978-3-642-35106-8.

[28] J. Bonča, S. A. Trugman, and M. Berciu, "Spectral function of the Holstein polaron at finite temperature," *Phys. Rev. B*, vol. 100, no. 9, p. 094 307, Sep. 2019.

[29] A. S. Mishchenko, L. Pollet, N. V. Prokof'ev, A. Kumar, D. L. Maslov, and N. Nagaosa, "Polaron mobility in the "beyond quasiparticles" regime," *Phys. Rev. Lett.*, vol. 123, p. 076 601, 7 Aug. 2019.

[30] D. Jansen, J. Bonča, and F. Heidrich-Meisner, "Finite-temperature density-matrix renormalization group method for electron-phonon systems: Thermodynamics and holstein-polaron spectral functions," *Phys. Rev. B*, vol. 102, p. 165 155, 16 Oct. 2020.

[31] M. R. Carbone, D. R. Reichman, and J. Sous, "Numerically exact generalized green's function cluster expansions for electron-phonon problems," *Phys. Rev. B*, vol. 104, p. 035 106, 3 Jul. 2021.

[32] Y. Tanimura and R. Kubo, "Time Evolution of a Quantum System in Contact with a Nearly Gaussian-Markoffian Noise Bath," *J. Phys. Soc. Japan*, vol. 58, no. 1, pp. 101–114, Jan. 1989.

[33] A. Ishizaki and Y. Tanimura, "Quantum dynamics of system strongly coupled to low-temperature colored noise bath: Reduced hierarchy equations approach," *J. Phys. Soc. Japan*, vol. 74, no. 12, pp. 3131–3134, 2005.

[34] Y. Tanimura, "Stochastic Liouville, Langevin, Fokker-Planck, and Master Equation Approaches to Quantum Dissipative Systems," *J. Phys. Soc. Japan*, vol. 75, no. 8, p. 082 001, Aug. 2006.

[35] A. Ishizaki and G. R. Fleming, "Unified treatment of quantum coherent and incoherent hopping dynamics in electronic energy transfer: Reduced hierarchy equation approach," *J. Chem. Phys.*, vol. 130, no. 23, p. 234 111, Jun. 2009.

[36] Q. Shi, L. Chen, G. Nan, R.-X. Xu, and Y. Yan, "Efficient hierarchical Liouville space propagator to quantum dissipative dynamics," *J. Chem. Phys.*, vol. 130, no. 8, p. 084 105, Feb. 2009.

[37] H. Liu, L. Zhu, S. Bai, and Q. Shi, "Reduced quantum dynamics with arbitrary bath spectral densities: Hierarchical equations of motion based on several different bath decomposition schemes," *J. Chem. Phys.*, vol. 140, no. 13, p. 134 106, Apr. 2014.

[38] L. Chen, Y. Zhao, and Y. Tanimura, "Dynamics of a One-Dimensional Holstein Polaron with the Hierarchical Equations of Motion Approach," *J. Phys. Chem. Lett.*, vol. 6, no. 15, pp. 3110–3115, 2015.

[39] I. S. Dunn, R. Tempelaar, and D. R. Reichman, "Removing instabilities in the hierarchical equations of motion: Exact and approximate projection approaches," *J. Chem. Phys.*, vol. 150, no. 18, p. 184 109, May 2019.

[40] V. Janković and N. Vukmirović, "Spectral and thermodynamic properties of the holstein polaron: Hierarchical equations of motion approach," *Phys. Rev. B*, vol. 105, p. 054 311, 5 Feb. 2022.

[41] M. Lüders *et al.*, "Ab initio theory of superconductivity. I. Density functional formalism and approximate functionals," *Phys. Rev. B*, vol. 72, no. 2, p. 024 545, Jul. 2005.

[42] F. Giustino, "Electron-phonon interactions from first principles," *Rev. Mod. Phys.*, vol. 89, p. 015 003, 1 Feb. 2017.

[43] C. Verdi, F. Caruso, and F. Giustino, "Origin of the crossover from polarons to Fermi liquids in transition metal oxides," *Nat. Commun.*, vol. 8, no. 1, p. 15 769, Aug. 2017.

[44] J. P. Nery, P. B. Allen, G. Antonius, L. Reining, A. Miglio, and X. Gonze, "Quasiparticles and phonon satellites in spectral functions of semiconductors and insulators: Cumulants applied to the full first-principles theory and the Fröhlich polaron," *Phys. Rev. B*, vol. 97, no. 11, p. 115 145, Mar. 2018.

[45] W. H. Sio, C. Verdi, S. Poncé, and F. Giustino, "Polarons from First Principles, without Supercells," *Phys. Rev. Lett.*, vol. 122, no. 24, p. 246 403, Jun. 2019.

[46] W. H. Sio, C. Verdi, S. Poncé, and F. Giustino, "Ab initio theory of polarons: Formalism and applications," *Phys. Rev. B*, vol. 99, no. 23, p. 235 139, 2019.

[47] J.-J. Zhou and M. Bernardi, "Predicting charge transport in the presence of polarons: The beyond-quasiparticle regime in $SrTiO_3$," *Phys. Rev. Research*, vol. 1, p. 033 138, 3 Dec. 2019.

[48] B. K. Chang, J.-J. Zhou, N.-E. Lee, and M. Bernardi, "Intermediate polaronic charge transport in organic crystals from a many-body first-principles approach," *npj Computational Materials*, vol. 8, no. 1, p. 63, 2022.

[49] D. Dunn, "Electron-Phonon Interactions in an Insulator," *Can. J. Phys.*, vol. 53, no. 4, pp. 321–337, Feb. 1975.

[50] G. Mahan, "Phonon-broadened optical spectra: Urbach's rule," *Phys. Rev.*, vol. 145, no. 2, p. 602, 1966.

[51] O. Gunnarsson, V. Meden, and K. Schönhammer, "Corrections to Migdal's theorem for spectral functions: A cumulant treatment of the time-dependent Green's function," *Phys. Rev. B*, vol. 50, no. 15, pp. 10 462–10 473, 1994.

[52] S. Barišić, J. Labbé, and J. Friedel, "Tight Binding and Transition-Metal Superconductivity," *Phys. Rev. Lett.*, vol. 25, no. 14, pp. 919–922, Oct. 1970.

[53] J. Sous, "Peierls Bipolarons and Localization in Solid-State and Molecular Systems," Ph.D. dissertation, The University of British Columbia, 2018.

[54] J. Sous, M. Chakraborty, R. V. Krems, and M. Berciu, "Light Bipolarons Stabilized by Peierls Electron-Phonon Coupling," *Phys. Rev. Lett.*, vol. 121, no. 24, p. 247 001, Dec. 2018.

[55] J. T. Devreese and A. S. Alexandrov, "Fröhlich polaron and bipolaron: recent developments," *Reports Prog. Phys.*, vol. 72, no. 6, p. 066 501, Jun. 2009.

[56] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems*. Mineola, NY: Dover Publications, Inc., 2003, ISBN: 978-0486428277.

[57] G. D. Mahan, "Mobility of Polarons," *Phys. Rev.*, vol. 142, no. 2, pp. 366–371, Feb. 1966.

[58] A. Damascelli, Z. Hussain, and Z.-X. Shen, "Angle-resolved photoemission studies of the cuprate superconductors," *Rev. Mod. Phys.*, vol. 75, no. 2, pp. 473–541, Apr. 2003.

[59] D. S. Inosov, "Angle-Resolved Photoelectron Spectroscopy Studies of the Many-Body Effects in the Electronic Structure of High-Tc Cuprates," Ph.D. dissertation, Technischen Universität Dresden, Jul. 2008.

[60] R. Brout and F. Englert, "Linked-clusters expansion in quantum statistics: Petit canonical ensemble," *Phys. Rev.*, vol. 120, pp. 1519–1527, 5 Dec. 1960.

[61] G. Horwitz, R. Brout, and F. Englert, "Zero-temperature properties of the many-fermion system," *Phys. Rev.*, vol. 130, pp. 409–419, 1 Apr. 1963.

[62] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*. Elsevier, 1992.

[63] T. J. Zuehlsdorff, A. Montoya-Castillo, J. A. Napoli, T. E. Markland, and C. M. Isborn, "Optical spectra in the condensed phase: Capturing anharmonic and vibronic features using dynamic and static approaches," *J. Chem. Phys.*, vol. 151, no. 7, p. 074 111, 2019.

[64] A. Anda, L. De Vico, T. Hansen, and D. Abramavičius, "Absorption and fluorescence lineshape theory for polynomial potentials," *J. Chem. Theory Comput.*, vol. 12, no. 12, pp. 5979–5989, 2016, PMID: 27759961.

[65] J. McClain *et al.*, "Spectral functions of the uniform electron gas via coupled-cluster theory and comparison to the GW and related approximations," *Phys. Rev. B*, vol. 93, no. 23, p. 235 139, Jun. 2016.

[66] J. Strümpfer and K. Schulten, "Open Quantum Dynamics Calculations with the Hierarchy Equations of Motion on Parallel Computers," *J. Chem. Theory Comput.*, vol. 8, no. 8, pp. 2808–2816, Aug. 2012.

[67] T. Berkelbach, *Pyrho: A python package for reduced density matrix techniques*, 2017.

[68] P. J. Robinson, I. S. Dunn, and D. R. Reichman, "Cumulant methods for electron-phonon problems. ii. the self-consistent cumulant expansion," *Phys. Rev. B*, vol. 105, p. 224 305, 22 Jun. 2022.

[69] A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*. Englewood Cliffs: Prentice-Hall, 1963.

[70] L. Hedin, "New method for calculating the one-particle green's function with application to the electron-gas problem," *Phys. Rev.*, vol. 139, A796–A823, 3A Aug. 1965.

[71] B. I. Lundqvist, "Single-particle spectrum of the degenerate electron gas," *Phys. Kondens. Mater.*, vol. 6, pp. 193–205, 1967.

[72] B. I. Lundqvist, "Single-particle spectrum of the degenerate electron gas," *Phys. Kondens. Mater.*, vol. 6, pp. 206–217, 1967.

[73] B. I. Lundqvist, "Single-particle spectrum of the degenerate electron gas," *Phys. Kondens. Mater.*, vol. 7, pp. 117–123, 1968.

[74] M. S. Hybertsen and S. G. Louie, "Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies," *Phys. Rev. B*, vol. 34, pp. 5390–5413, 8 Oct. 1986.

[75] E. L. Shirley, "Self-consistent gw and higher-order calculations of electron states in metals," *Phys. Rev. B*, vol. 54, pp. 7758–7764, 11 Sep. 1996.

[76] U. von Barth and B. Holm, "Self-consistent GW_0 results for the electron gas: Fixed screened potential W_0 within the random-phase approximation," *Phys. Rev. B*, vol. 54, pp. 8411–8419, 12 Sep. 1996.

[77] N. Prokof'ev and B. Svistunov, "Fermi-polaron problem: Diagrammatic monte carlo method for divergent sign-alternating series," *Phys. Rev. B*, vol. 77, 020408(R), 2 Jan. 2008.

[78] L. Hedin, "On correlation effects in electron spectroscopies and the GW approximation," *J. Phys. Condens. Matter*, vol. 11, no. 42, R489–R528, Oct. 1999.

[79] C. Almbladh and L. Hedin, "Beyond the one-electron model: Many-electron effects in atoms, molecules, and solids," *Handbook on Synchrotron Radiation*, vol. 1b, pp. 607–904, 1983.

[80] B. Pandey and P. B. Littlewood, "Going beyond the cumulant approximation: Power series correction to the single-particle green's function in the holstein system," *Phys. Rev. Lett.*, vol. 129, p. 136 401, 13 Sep. 2022.

[81] J. Bonča and S. A. Trugman, "Dynamic properties of a polaron coupled to dispersive optical phonons," *Phys. Rev. B*, vol. 103, p. 054 304, 5 Feb. 2021.

[82] J. J. Kas, J. J. Rehr, and L. Reining, "Cumulant expansion of the retarded one-electron green function," *Phys. Rev. B*, vol. 90, p. 085 112, 8 Aug. 2014.

[83] P. Mitrić, V. Janković, N. Vukmirović, and D. Tanasković, "Spectral functions of the holstein polaron: Exact and approximate solutions," *Phys. Rev. Lett.*, vol. 129, p. 096 401, 9 Aug. 2022.

[84] M. Schlipf, S. Poncé, and F. Giustino, "Carrier lifetimes and polaronic mass enhancement in the hybrid halide perovskite $CH_3NH_3PbI_3$ from multiphonon fröhlich coupling," *Phys. Rev. Lett.*, vol. 121, p. 086 402, 8 Aug. 2018.

[85] F. Brown-Altvater *et al.*, "Band gap renormalization, carrier mobilities, and the electron-phonon self-energy in crystalline naphthalene," *Phys. Rev. B*, vol. 101, p. 165 102, 16 Apr. 2020.

[86] B. Holm and F. Aryasetiawan, "Self-consistent cumulant expansion for the electron gas," *Phys. Rev. B*, vol. 56, pp. 12 825–12 831, 20 Nov. 1997.

[87] D. J. J. Marchand and M. Berciu, "Effect of dispersive optical phonons on the behavior of a holstein polaron," *Phys. Rev. B*, vol. 88, 060301(R), 6 Aug. 2013.

[88] D. C. Leslie, "Developments in the theory of turbulence(book)," *Oxford, Clarendon Press*, 1973.

[89] W. D. McComb, *The Physics of Fluid Turbulence*. New York: Oxford University Press, 1990.

[90] S. Engelsberg and J. R. Schrieffer, "Coupled electron-phonon system," *Phys. Rev.*, vol. 131, pp. 993–1008, 3 Aug. 1963.

[91] R. Zwanzig, *Nonequilibrium Statistical Mechanics*. Oxford university press, 2001.

[92] E. Vanden Eijnden, "Studying random differential equations as a tool for turbulent diffusion," *Phys. Rev. E*, vol. 58, R5229–R5232, 5 Nov. 1998.

[93] P. R. Kramer, A. J. Majda, and E. Vanden-Eijnden, "Closure approximations for passive scalar turbulence: A comparative study on an exactly solvable model with complex features," *J. Stat. Phys.*, vol. 111, no. 3, pp. 565–679, 2003.

[94] M. Tokuyama and H. Mori, "Statistical-Mechanical Theory of Random Frequency Modulations and Generalized Brownian Motions*)," *Prog. Theor. Exp. Phys.*, vol. 55, no. 2, pp. 411–429, Feb. 1976.

[95] B. Yoon, J. M. Deutch, and J. H. Freed, "A comparison of generalized cumulant and projection operator methods in spin-relaxation theory," *J. Chem. Phys.*, vol. 62, no. 12, pp. 4687–4696, 1975.

[96] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*. Oxford: Oxford University Press, 2002.

[97] E Vanden Eijnden, "Contribution to the statistical theory of turbulence: Application to anomalous transport in plasmas," Ph.D. dissertation, Universite Libre de Bruxelles, 1997.

[98] S. Olver, "Numerical approximation of highly oscillatory integrals," Ph.D. dissertation, University of Cambridge Cambridge, 2008.

[99] S. Olver, "Fast, numerically stable computation of oscillatory integrals with stationary points," *BIT Numer. Math.*, vol. 50, no. 1, pp. 149–171, 2010.

[100] B. Holm and U. von Barth, "Fully self-consistent GW self-energy of the electron gas," *Phys. Rev. B*, vol. 57, pp. 2108–2117, 4 Jan. 1998.

[101] B. Holm and F. Aryasetiawan, "Total energy from the galitskii-migdal formula using realistic spectral functions," *Phys. Rev. B*, vol. 62, pp. 4858–4865, 8 Aug. 2000.

[102] A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory*. New York: McGraw-Hill, Inc., 1989.

[103] H. Y. Fan, "Temperature dependence of the energy gap in semiconductors," *Phys. Rev.*, vol. 82, pp. 900–905, 6 Jun. 1951.

[104] A. Migdal, "Interaction between electrons and lattice vibrations in a normal metal," *Sov. Phys. JETP*, vol. 7, no. 6, pp. 996–1001, 1958.

[105] F. Caruso, C. Verdi, and F. Giustino, "Many-body calculations of plasmon and phonon satellites in angle-resolved photoelectron spectra using the cumulant expansion approach," in *Handbook of Materials Modeling: Methods: Theory and Modeling*, W. Andreoni and S. Yip, Eds. Cham: Springer International Publishing, 2020, pp. 341–365, ISBN: 978-3-319-44677-6.

[106] B. Pandey and P. B. Littlewood, "Going beyond the cumulant approximation ii:power series correction to single particle green's function in 1d holstein chain," 2023, arXiv:2303.03503.

[107] P. Mitrixć, V. Janković, N. Vukmirović, and D. Tanasković, "Cumulant expansion in the holstein model: Spectral functions and mobility," *Phys. Rev. B*, vol. 107, p. 125 165, 12 Mar. 2023.

[108] F. Giustino, M. L. Cohen, and S. G. Louie, "Small phonon contribution to the photoemission kink in the copper oxide superconductors," *Nature*, vol. 452, no. 7190, pp. 975–978, 2008.

[109] M. J. Schilcher *et al.*, "The significance of polarons and dynamic disorder in halide perovskites," *ACS Energy Letters*, vol. 6, no. 6, pp. 2162–2173, 2021.

[110] T. D. Schultz, "Slow electrons in polar crystals: Self-energy, mass, and mobility," *Phys. Rev.*, vol. 116, pp. 526–543, 3 Nov. 1959.

[111] T. D. Lee, F. E. Low, and D. Pines, "The motion of slow electrons in a polar crystal," *Phys. Rev.*, vol. 90, pp. 297–302, 2 Apr. 1953.

[112] S. I. Pekar, *Untersuchungen über die Elektronentheorie der Kristalle*. De Gruyter, 1954.

[113] L. D. Landau, "Electron motion in crystal lattices," *Phys. Z. Sowjet.*, p. 664, 3 1933.

[114] W. J. Huybrechts, "Note on the ground-state energy of the feynman polaron model," *Journal of Physics C: Solid State Physics*, vol. 9, no. 8, p. L211, Apr. 1976.

[115] R. P. Feynman, R. W. Hellwarth, C. K. Iddings, and P. M. Platzman, "Mobility of slow electrons in a polar crystal," *Phys. Rev.*, vol. 127, pp. 1004–1017, 4 Aug. 1962.

[116] P. Nagy and P. Markoš, "On the nature of the self-trapping of different types of polarons," *Phys. Status Solidi B*, vol. 151, no. 2, pp. 571–580, 1989.

[117] W. J. Huybrechts, "Internal excited state of the optical polaron," *J. Phys. C: Solid State Phys.*, vol. 10, no. 19, p. 3761, Oct. 1977.

[118] L. Lin, P. B. Littlewood, and A. Edelman, "Analysis of fröhlich bipolarons," 2021, arXiv:2111.12250.

[119] P. B. Allen and V Heine, "Theory of the temperature dependence of electronic band structures," *Journal of Physics C: Solid State Physics*, vol. 9, no. 12, p. 2305, Jun. 1976.

[120] P. B. Allen and M. Cardona, "Theory of the temperature dependence of the direct gap of germanium," *Phys. Rev. B*, vol. 23, pp. 1495–1505, 4 Feb. 1981.

[121] J. Lafuente-Bartolome, C. Lian, W. H. Sio, I. G. Gurtubay, A. Eiguren, and F. Giustino, "Unified approach to polarons and phonon-induced band structure renormalization," *Phys. Rev. Lett.*, vol. 129, p. 076 402, 7 Aug. 2022.

[122] J. Lafuente-Bartolome, C. Lian, W. H. Sio, I. G. Gurtubay, A. Eiguren, and F. Giustino, "Ab initio self-consistent many-body theory of polarons at all couplings," *Phys. Rev. B*, vol. 106, p. 075 119, 7 Aug. 2022.

[123] N.-E. Lee, H.-Y. Chen, J.-J. Zhou, and M. Bernardi, "Facile ab initio approach for self-localized polarons from canonical transformations," *Phys. Rev. Mater.*, vol. 5, p. 063 805, 6 Jun. 2021.

[124] Y. Luo, B. K. Chang, and M. Bernardi, "Comparison of the canonical transformation and energy functional formalisms for ab initio calculations of self-localized polarons," *Phys. Rev. B*, vol. 105, p. 155 132, 15 Apr. 2022.

[125] I. Lang and Y. A. Firsov, "Kinetic theory of semiconductors with low mobility," *Sov. Phys. JETP*, vol. 16, no. 5, p. 1301, 1963.

[126] S. Engelsberg and J. R. Schrieffer, "Coupled electron-phonon system," *Phys. Rev.*, vol. 131, pp. 993–1008, 3 Aug. 1963.

[127] B. Gerlach and H. Löwen, "Analytical properties of polaron systems or: Do polaronic phase transitions exist or not?" *Rev. Mod. Phys.*, vol. 63, pp. 63–90, 1 Jan. 1991.

[128] B. S. Shastry and B. Sutherland, "Twisted boundary conditions and effective mass in heisenberg-ising and hubbard rings," *Phys. Rev. Lett.*, vol. 65, pp. 243–246, 2 Jul. 1990.

[129] E. Jeckelmann and S. R. White, "Density-matrix renormalization-group study of the polaron problem in the holstein model," *Phys. Rev. B*, vol. 57, pp. 6376–6385, 11 Mar. 1998.

[130] M. Fishman, S. R. White, and E. M. Stoudenmire, "The ITensor Software Library for Tensor Network Calculations," *SciPost Phys. Codebases*, p. 4, 2022.

[131] P. Giannozzi *et al.*, "Quantum espresso: A modular and open-source software project for quantum simulations of materials," *Journal of Physics: Condensed Matter*, vol. 21, no. 39, p. 395 502, Sep. 2009.

[132] P Giannozzi *et al.*, "Advanced capabilities for materials modelling with quantum espresso," *Journal of Physics: Condensed Matter*, vol. 29, no. 46, p. 465 901, Oct. 2017.

[133] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.*, vol. 77, pp. 3865–3868, 18 Oct. 1996.

[134] D. R. Hamann, "Optimized norm-conserving vanderbilt pseudopotentials," *Phys. Rev. B*, vol. 88, p. 085 117, 8 Aug. 2013.

[135] M. Schlipf and F. Gygi, "Optimization algorithm for the generation of oncv pseudopotentials," *Computer Physics Communications*, vol. 196, pp. 36–44, 2015.

[136] A. A. Mostofi *et al.*, "An updated version of wannier90: A tool for obtaining maximally-localised wannier functions," *Computer Physics Communications*, vol. 185, no. 8, pp. 2309–2310, 2014.

[137] S. Poncé, E. Margine, C. Verdi, and F. Giustino, "Epw: Electron–phonon coupling, transport and superconducting properties using maximally localized wannier functions," *Computer Physics Communications*, vol. 209, pp. 116–133, 2016.

[138] A. Mahajan, P. J. Robinson, J. Lee, and D. R. Reichman, "Neural network quantum states for electron-phonon coupled systems," *Unpublished*, 2023.

[139] G. Makov and M. C. Payne, "Periodic boundary conditions in ab initio calculations," *Phys. Rev. B*, vol. 51, pp. 4014–4022, 7 Feb. 1995.

[140] Y. Yan, T. Xing, and Q. Shi, "A new method to improve the numerical stability of the hierarchical equations of motion for discrete harmonic oscillator modes," *The Journal of Chemical Physics*, vol. 153, no. 20, p. 204 109, 2020.

[141] R. P. Feynman, "An operator calculus having applications in quantum electrodynamics," *Phys. Rev.*, vol. 84, pp. 108–128, 1 Oct. 1951.

[142] P. Linz, "Linear multistep methods for volterra integro-differential equations," *J. ACM*, vol. 16, no. 2, 295–301, Apr. 1969.

[143] P. Virtanen *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

Appendix A: Moments for the Holstein model

Plugging in the specific form of the Holstein interaction, performing the time integrals, and removing several of the internal momentum sums via conservation of momentum, we find the second moment is given by

$$M_2\left(k,t\right) = -\frac{g^2}{N} e^{-i\varepsilon_k t} \sum_q \Big[\left(N_0+1\right) A_q^- + N_0 A_q^+ \Big], \tag{A.1}$$

$$A_q^{\pm} \equiv \frac{\pm it\left(\omega_0 \pm \left(\varepsilon_k - \varepsilon_q\right)\right) - e^{\pm it\left(\omega_0 \pm \left(\varepsilon_k - \varepsilon_q\right)\right)} + 1}{\left(\omega_0 \pm \left(\varepsilon_k - \varepsilon_q\right)\right)^2}, \tag{A.2}$$

and the fourth moment is given by

$$
\begin{aligned}
M_4(k,t) = \frac{g^4}{N^2} e^{-i\varepsilon_k t} \sum_{q_1,q_2} \Bigg[& (N_0+1)^2 T_1\left(f_1^+, f_2^+; t\right) \\
& + N_0 (N_0+1) T_1\left(f_1^-, f_2^+; t\right) \\
& + N_0 (N_0+1) T_1\left(f_1^+, f_2^-; t\right) \\
& + N_0^2 T_1\left(f_1^-, f_2^-; t\right) \\
& + (N_0+1)^2 T_2\left(f_1^+, f_2^+, f\left(q_1+q_2, 0, 2\omega_0\right); t\right) \\
& + N_0 (N_0+1) T_2\left(f_1^-, f_2^+, f\left(q_1+q_2, 0, 0\right); t\right) \\
& + N_0 (N_0+1) T_2\left(f_1^+, f_2^-, f\left(q_1+q_2, 0, 0\right); t\right) \\
& + N_0^2 T_2\left(f_1^-, f_2^-, f\left(q_1+q_2, 0, -2\omega_0\right); t\right) \\
& + (N_0+1)^2 T_3\left(f_1^+, f\left(q_1+q_2, 0, 2\omega_0\right); t\right) \\
& + N_0 (N_0+1) T_3\left(f_1^-, f\left(q_1+q_2, 0, 0\right); t\right) \\
& + N_0 (N_0+1) T_3\left(f_1^+, f\left(q_1+q_2, 0, 0\right); t\right) \\
& + N_0^2 T_3\left(f_1^-, f\left(q_1+q_2, 0, -2\omega_0\right); t\right) \Bigg],
\end{aligned}
\tag{A.3}
$$

where

$$
\begin{aligned}
T_1(a,b;t) = \frac{1}{b} \Bigg[& \frac{a\frac{t^2}{2} - t - h(a,t)}{a^2} \\
& + \frac{1}{b}\left(\frac{h(a,t) - h(b,t)}{a-b} - \frac{t + h(a,t)}{a}\right)\Bigg],
\end{aligned}
\tag{A.4}
$$

$$
\begin{aligned}
T_2(a,b,c;t) = \frac{1}{b} \Bigg[& \frac{1}{c}\left(\frac{t + h(a,t)}{a} - \frac{h(a,t) - h(c,t)}{a-c}\right) \\
& - \frac{1}{c-b}\left(\frac{h(a,t) - h(b,t)}{a-b} - \frac{h(a,t) - h(c,t)}{a-c}\right)\Bigg],
\end{aligned}
\tag{A.5}
$$

$$T_3(a,b;t) = \frac{1}{a}\left[\frac{1}{b}\left(\frac{t+h(a,t)}{a} - \frac{h(a,t)-h(b,t)}{a-b}\right)\right.$$
$$\left. + \frac{1}{b-a}\left(\frac{e^{-at}t+h(a,t)}{a} + \frac{h(a,t)-h(b,t)}{a-b}\right)\right], \tag{A.6}$$

$$f(a,b,c) = i(\varepsilon_{k+a} - \varepsilon_{k+b} + c), \tag{A.7}$$

$$f_i^{\pm} = f(q_i, 0, \pm\omega_0), \tag{A.8}$$

$$h(x,t) = \frac{e^{-xt}-1}{x}. \tag{A.9}$$

For both M_2 and M_4, singular terms within the momentum sums are evaluated in a limiting sense using L'Hopital's rule.

Appendix B: Hierarchical equations of motion

For an exact benchmark of $\mathcal{G}(k,t)$ we will use the Hierarchical Equations of Motion (HEOM) approach. First popularized for solving vibronic models with continuous bath spectral densities, [32–36] HEOM has recently been adapted to solve discrete bath models such as the Holstein and SSH models. [37–39] While we have recently shown that the finite truncation of HEOM can lead to long-time instability in such models, [39] for the present application the converged short and intermediate time behavior is sufficient to provide benchmarks for $\mathcal{G}(t)$ and $A(\omega)$. Two recent versions of HEOM have provided practical routes to circumventing instabilities.[40, 140]

To compute $\mathcal{G}(k,t)$ with HEOM we rewrite

$$\begin{aligned}\mathcal{G}(k,t) = -i\Theta(t)\,\mathrm{Tr}_S\Big[a_k \mathrm{Tr}_B\Big[e^{-iHt} \\ \times \left(e^{-\beta H_b} \otimes a_k^\dagger \rho_{vac}\right) e^{iHt}\Big]\Big],\end{aligned} \tag{B.1}$$

where

$$\rho_{vac} = |0\rangle\langle 0| = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{B.2}$$

is the pure-state electronic density matrix representing the zero-electron vacuum, written in a basis of zero-electron and one-electron states. The S and B subscripts denote partial traces over the electron and phonon subspaces, respectively. One-electron states are described in the site basis. In

this basis,

$$a_k^\dagger = \frac{1}{\sqrt{N}} \begin{bmatrix} 0 & 0 & \dots & 0 \\ e^{-ik} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ e^{-ik(N-1)} & 0 & 0 & 0 \end{bmatrix}, \tag{B.3}$$

$$a_k = \frac{1}{\sqrt{N}} \begin{bmatrix} 0 & e^{ik} & \dots & e^{ik(N-1)} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{B.4}$$

Thus, to calculate $\mathcal{G}(k,t)$ we initialize a hierarchy of auxiliary density matrices, each of dimension $(N+1)\times(N+1)$. All matrices $\rho_{m_{1\pm},\dots,m_{N\pm}}(t=0)$ are set to zero except for

$$\rho_{0,..,0}(t=0) = a_k^\dagger \rho_{vac}\ . \tag{B.5}$$

Then, we propagate in time using the discrete-bath HEOM [37–39]

$$\begin{aligned} \frac{d}{dt}\rho_{m_{1\pm},\dots,m_{N\pm}}(t) &= -i\mathcal{L}\rho_{m_{1\pm},\dots,m_{N\pm}}(t) \\ &\quad - i\sum_{n=1}^{N}\omega_0\left(m_{n-}-m_{n+}\right)\rho_{m_{1\pm},\dots,m_{N\pm}}(t) \\ &\quad + \sum_{n=1}^{N}\Big[\Phi_n\Big(\rho_{m_{1\pm},\dots,m_{n+}+1,\dots,m_{N\pm}}(t) \\ &\quad + \rho_{m_{1\pm},\dots,m_{n-}+1,\dots,m_{N\pm}}(t)\Big) \\ &\quad + m_{n+}\Theta_{n+}\rho_{m_{1\pm},\dots,m_{n+}-1,\dots,m_{N\pm}}(t) \\ &\quad + m_{n-}\Theta_{n-}\rho_{m_{1\pm},\dots,m_{n-}-1,\dots,m_{N\pm}}(t)\Big], \end{aligned} \tag{B.6}$$

where

$$\mathcal{L} = [\hat{H}_e, ...], \tag{B.7}$$

$$\Phi_n = [\hat{V}_n, ...] \tag{B.8}$$

$$\hat{V}_n = a_n^\dagger a_n, \tag{B.9}$$

and

$$\Theta_{n\pm} = -\frac{(g\omega_0)^2}{2}\left([\hat{V}_n, ...]\coth\left(\frac{\beta\omega_0}{2}\right) \mp \{\hat{V}_n, ...\}\right). \tag{B.10}$$

Finally, we compute the Green's function as

$$\mathcal{G}(k,t) = -i\Theta(t)\,Tr\left[a_k \rho_{0,..,0}(t)\right]. \tag{B.11}$$

Converging with respect to the hierarchy depth L, we obtain the exact $\mathcal{G}(k,t)$ for the Holstein model.

Appendix C: K-phonon approximation

For analyzing finite-size effects in an inexpensive, approximate way, we will also compute $\mathcal{G}(k,t)$ via numerical diagonalization of the Hamiltonian within a truncated basis. Toward this end we introduce the momentum-space basis kets

$$|\nu_0,\ldots,\nu_N\rangle_0, \tag{C.1}$$

and

$$|k,\nu_0,\ldots,\nu_N\rangle_1, \tag{C.2}$$

which represent states with zero and one electron, respectively. The electronic quantum number k indicates the momentum of the electron. The vibrational quantum numbers ν_i denote the number of vibrational quanta in each normal mode, such that

$$\begin{aligned} &b_q^\dagger|0,\ldots,\nu_q,\ldots,0\rangle_0 = \\ &\quad\sqrt{\nu_q+1}|0,\ldots,\nu_q+1,\ldots,0\rangle_0, \end{aligned} \tag{C.3}$$

$$\begin{aligned} &b_q^\dagger|k,0,\ldots,\nu_q,\ldots,0\rangle_1 = \\ &\quad\sqrt{\nu_q+1}|k,0,\ldots,\nu_q+1,\ldots,0\rangle_1. \end{aligned} \tag{C.4}$$

We work within a truncated K phonon basis such that

$$\sum_{q=1}^{N}\nu_q \le K. \tag{C.5}$$

Using this basis to represent the Hamiltonian, we can then compute the matrix exponential necessary to determine $\mathcal{G}(k,t)$ by numerically diagonalizing the Hamiltonian. We will refer to this

approach as the "$K-$phonon approximation." In the text only $K = 1$ results are shown.

Appendix D: Derivation of SC-CE

Closely following Dunn,[49] we start with the Green's function in the form Eq. 2.9, with the goal to evaluate the trace of the exponential to all orders. Since there is a single electron on the lattice, we can easily determine all of the matrix elements by expanding the exponent as

$$\begin{aligned}\mathcal{G}_k(t) = \mathcal{G}_k^{(0)}(t)\Bigg\langle \sum_{n=0}^{\infty} T\Bigg[\frac{(-i)^n}{n!}\sum_{\{q_j\}_{j=1}^n}\int_0^t dt_1\cdots\int_0^t dt_n \\ \times g_{q_n,k}e^{-it_n\epsilon_k}e^{it_n\epsilon_{k+q_n}}A_{q_n}(t_n) \\ \times g_{q_{n-1},k+q_n}e^{-it_{n-1}\epsilon_{k+q_n}}e^{it_{n-1}\epsilon_{k+q_n+q_{n-1}}}A_{q_{n-1}}(t_{n-1}) \\ \cdots\times g_{q_1,k+q_2+\cdots+q_n}e^{-it_n\epsilon_{k+q_2+\cdots+q_n}}e^{it_n\epsilon_k}A_{q_1}(t_1)\Bigg]\Bigg\rangle.\end{aligned} \tag{D.1}$$

Eq. D.1 makes apparent that each application of V has the effect of shifting all subsequent k-dependent terms by q_i. Here, $A_q(t) = b_q e^{-i\omega_q t} + b_{-q}^\dagger e^{-i\omega_q t}$, and $\sum_i q_i = 0$. Note that we have kept notation general beyond the Holstein model such that the electron-phonon coupling constant depends on momentum. This relationship can be represented with the differential operator $e^{q\cdot\frac{d}{dk}}$. We define a vertex operator $\Gamma_{qk}(t) = g_{qk}e^{-it\epsilon_k}e^{q\cdot\frac{d}{dk}}e^{it\epsilon_k}$ which allows us to rewrite the electron Green's function as an exponential

$$\mathcal{G}_k(t) = \mathcal{G}_k^{(0)}(t)\left\langle Te^{-i\sum_q\int_0^t d\tau\Gamma_{qk}(\tau)A_q(\tau)}\right\rangle. \tag{D.2}$$

The remaining trace is only over the phonons. Assuming that the phonons are purely harmonic, we can apply the standard result $\langle e^B\rangle = e^{\frac{1}{2}\langle B^2\rangle}$, which is valid for an operator B which is a linear combination of b and $b^\dagger$. This expectation value is readily evaluated as $\langle TA_q(t)A_{q'}(t')\rangle = iD_q^{(0)}(t-t')\delta_{q,-q'}$. This final trace further constrains the values that q_i can take, so that only phys-

ical electron-phonon vertices are possible. This completes the first portion of the derivation, and we are left with an explicit form for S

$$S = -i \sum_q \int_0^t d\tau \int_0^\tau d\tau' D_q^{(0)}(\tau - \tau')\Gamma_{-qk}(\tau)\Gamma_{qk}(\tau'). \tag{D.3}$$

To generate a self-consistent expression for S, we apply the Feynman operator ordering theorem.[141] The theorem states that given a functional of time dependent operators $F[\hat{A}(\tau), \hat{B}(\tau), \dots]$ where $\tau \in [0, t]$ and with a unitary operator of the form $\hat{U}(\tau') = T \exp\left(\int_0^{\tau'} d\tau \hat{P}(\tau)\right)$ then

$$\begin{aligned} &\hat{U}(t)F[\hat{A}(\tau), \hat{B}(\tau), \dots] = \\ &T\left[\exp\left(\int_0^t d\tau \hat{P}(\tau)\right) F[\hat{U}(\tau)\hat{A}(\tau)\hat{U}^{-1}(\tau), \dots]\right]. \end{aligned} \tag{D.4}$$

Considering that we have expressed $\mathcal{G}_k(t)$ as a functional of differential operators, we choose $\hat{P} = -i\left(\epsilon_k + \frac{d}{d\tau}\phi_k(\tau)\right)$. The scalar function $\phi_k(t)$ is only required to vanish at $t = 0$ and the self-consistent approach presented in the main text corresponds to the choice $\phi_k(t) = -iC_k(t)$.

The transformed vertex functional can be simplified to

$$\begin{aligned} \bar{\Gamma}_{qk}(\tau) &\equiv \hat{U}(\tau)\Gamma_{qk}(\tau)\hat{U}^{-1}(\tau) \\ &= g_{qk} e^{-i\phi_k(\tau)} e^{q \cdot \frac{d}{dk}} e^{+i\phi_k(\tau)}, \end{aligned} \tag{D.5}$$

and the transformed S is then given by

$$\begin{aligned} \bar{S} = &-i \int_0^t d\tau \left(\epsilon_k + \frac{d}{d\tau}\phi_k(\tau)\right) \\ &- i \sum_q \int_0^t d\tau \int_0^\tau d\tau' D_q^{(0)}(\tau - \tau')\bar{\Gamma}_{-qk}(\tau)\bar{\Gamma}_{qk}(\tau'). \end{aligned} \tag{D.6}$$

Just as in the standard cumulant expansion, the cumulants of a given order can be computed by

evaluating time-ordered powers of S

$$C_k(t) = i\phi_k(t) + \sum_{n=1}^{\infty} \frac{1}{n!} T[\bar{S}^n]_c, \tag{D.7}$$

where the notation $[...]_c$ denotes the cumulant of an operator, and $T[\bar{S}]_c = T[\bar{S}]$. Subsequent terms are given by

$$T[\bar{S}^n]_c = T[\bar{S}^n] - \sum_{m=1}^{n-1} \frac{(n-1)!}{m!(n-m-1)!} T[\bar{S}^m] T[\bar{S}^{n-m}]_c. \tag{D.8}$$

Evaluating Eq. D.7 with $\phi_k(t) = C_k(t)$ to leading order produces Eq. 2.12.

Appendix E: SC-CE Recursive Expression

Here we illustrate how the lowest-order SC-CE generates the exact second-order CE and approximations to all higher order CEs. Considering only the lowest order in Eqs. D.7 and D.8, we can write Eq. 2.13 for the SC-CE as a self-consistent equation for the Green's function

$$\begin{aligned}\mathcal{G}_k(t) = \mathcal{G}_k^{(0)}(t) \exp\Bigg(i \sum_q \int_0^t d\sigma \int_0^\sigma d\tau |g_{qk}|^2 \\ \times D_q^{(0)}(\sigma - \tau) \frac{\mathcal{G}_{k-q}(\sigma)\mathcal{G}_k(\tau)}{\mathcal{G}_{k-q}(\tau)\mathcal{G}_k(\sigma)} \Bigg).\end{aligned} \tag{E.1}$$

In the weak coupling limit, we find an iterative solution starting with the bare Green's function. The first iteration is straightforward and replicates exactly the second cumulant

$$\mathcal{G}_k^{(1)}(t) = \mathcal{G}_k^{(0)}(t) \exp(C_2(k,t)). \tag{E.2}$$

While the integrals for subsequent iterations may be analytically evaluated, it is more instructive to consider the further iterations as functions of $C_2(k,t)$.

The second and all subsequent iterations now have exponential resummations of the second-order cumulant itself. Explicitly,

$$\begin{aligned}\mathcal{G}_k^{(2)}(t) = &\mathcal{G}_k^{(0)}(t) \exp\Bigg(- i \sum_q \int_0^t d\sigma \int_0^\sigma d\tau |g_{qk}|^2 \\ &\times D_q^{(0)}(\sigma - \tau) \mathcal{G}_{k-q}^{(0)}(\sigma - \tau) \mathcal{G}_k^{(0)}(\tau - \sigma) \\ &\times e^{C_2(k,\tau) + C_2(k-q,\sigma) - C_2(k,\sigma) - C_2(k-q,\tau)} \Bigg).\end{aligned} \tag{E.3}$$

The n^{th} cumulant is defined as being proportional to the n^{th} power of the coupling constant, so by

expanding the exponent we find an approximation for all higher order cumulants

$$\mathcal{G}_k^{(2)}(t) = \mathcal{G}_k^{(0)}(t) \exp\Bigg(\sum_{n=0}^{\infty} \frac{-i}{n!} \sum_q \int_0^t d\sigma \int_0^\sigma d\tau \tag{E.4}$$
$$\times |g_{qk}|^2 D_q^{(0)}(\sigma-\tau)\mathcal{G}_{k-q}^{(0)}(\sigma-\tau)\mathcal{G}_k^{(0)}(\tau-\sigma)$$
$$\times \left(C_2(k,\tau) + C_2(k-q,\sigma) - C_2(k,\sigma) - C_2(k-q,\tau)\right)^n \Bigg).$$

We can simplify this expression by introducing $C_m^{(n)}(k,t)$, which represents the approximation to the n^{th} iteration of the m^{th} cumulant. This results in

$$\mathcal{G}_k^{(2)}(t) = \mathcal{G}_k^{(0)}(t) \exp\left(\sum_{n=1}^{\infty} C_{2n}^{(2)}(k,t) \right). \tag{E.5}$$

Importantly, because the second iteration produces cumulants of all orders, further iterations require expanding the exponent into moments.

After the second iteration, we can generalize to an arbitrary iteration:

$$\mathcal{G}_{\mathbf{k}}^{(n+1)}(t) = \mathcal{G}_{\mathbf{k}}^{(0)}(t) \exp\Bigg(-i \sum_{\mathbf{q}} \int_0^t d\sigma \int_0^\sigma d\tau |g_{qk}|^2$$
$$\times D_{\mathbf{q}}^0(\sigma-\tau)\mathcal{G}_{\mathbf{k}}^{(0)}(\tau-\sigma)\mathcal{G}_{\mathbf{k}-\mathbf{q}}^{(0)}(\sigma-\tau) e^{F^{(n)}(\mathbf{k},\mathbf{q},\sigma,\tau)} \Bigg). \tag{E.6}$$

Here, we introduce the notation:

$$F^{(n)}(\mathbf{k},\mathbf{q},\sigma,\tau) \equiv F_2^{(n)}(\mathbf{k},\mathbf{q},\sigma,\tau) + F_4^{(n)}(\mathbf{k},\mathbf{q},\sigma,\tau) + \ldots, \tag{E.7}$$
$$F_i^{(n)}(\mathbf{k},\mathbf{q},\sigma,\tau) \equiv C_i^{(n)}(\mathbf{k}-\mathbf{q},\sigma) - C_i^{(n)}(\mathbf{k}-\mathbf{q},\tau)$$
$$- C_i^{(n)}(\mathbf{k},\sigma) + C_i^{(n)}(\mathbf{k},\tau). \tag{E.8}$$

The procedure for constructing the approximation to the n^{th} iteration of the m^{th} cumulant is given

by the moment expansion

$$\sum_{m=0}^{\infty} \lambda^{2m} W_{2m}^{(n)}(\mathbf{k}, \mathbf{q}, \sigma, \tau) = e^{\sum_{m=1}^{\infty} \lambda^{2m} F_{2m}^{(n)}(\mathbf{k}, \mathbf{q}, \sigma, \tau)}, \tag{E.9}$$

where the first few moments are given by:

$$W_0^{(n)}(\mathbf{k}, \mathbf{q}, \sigma, \tau) = 1 \tag{E.10}$$

$$W_2^{(n)}(\mathbf{k}, \mathbf{q}, \sigma, \tau) = F_2^{(n)}(\mathbf{k}, \mathbf{q}, \sigma, \tau) \tag{E.11}$$

$$\begin{aligned} W_4^{(n)}(\mathbf{k}, \mathbf{q}, \sigma, \tau) &= F_4^{(n)}(\mathbf{k}, \mathbf{q}, \sigma, \tau) \\ &\quad + \frac{1}{2}\left(F_2^{(n)}(\mathbf{k}, \mathbf{q}, \sigma, \tau)\right)^2. \end{aligned} \tag{E.12}$$

Because each iteration introduces another power of g^2 to all of the approximated cumulants, and the n^{th} order cumulant is defined as being proportional to g^n, the n^{th} iteration cannot alter any of the approximated cumulants of order lower than n. Because of this, and the fact that the approximate cumulant for a given order is an integral over lower order approximations, the iterative method is actually recursive. Dropping the superscripts indicating iteration, the general form for the SC-CE approximation to the n^{th} cumulant is

$$\begin{aligned} C_{2n}^{\text{SC-CE}}(\mathbf{k}, t) = -i \sum_{\mathbf{q}} \int_0^t d\sigma \int_0^\sigma d\tau |g_{qk}|^2 D_{\mathbf{q}}(\sigma - \tau) \\ \times \mathcal{G}_{\mathbf{k}}^{(0)}(\tau - \sigma) \mathcal{G}_{\mathbf{k}-\mathbf{q}}^{(0)}(\sigma - \tau) W_{2n-2}(\mathbf{k}, \mathbf{q}, \sigma, \tau). \end{aligned} \tag{E.13}$$

Appendix F: Numerical VIDE Solution

Eq. 2.13 is not convenient for numerical calculations. Instead, we define, $y_k(t) \equiv e^{C_k(t)}$ and solve directly for $y_k(t)$,

$$\begin{aligned}\frac{dy_k(t)}{dt} = &- i \sum_q \int_0^t d\tau |g_{qk}|^2 D_q^{(0)}(t-\tau) \\ &\times e^{i(\epsilon_k - \epsilon_{k-q})(t-\tau)} \frac{y_k(\tau) y_{k-q}(t)}{y_{k-q}(\tau)}.\end{aligned} \tag{F.1}$$

Numerical solutions to equation F.1 were carried out with a multi-step predictor corrector method as derived by Linz.[142] His method is general for first order Volterra Integro-Differential Equations (VIDEs) of the form

$$y'(x) = F\left(x, y(x), \int_0^x dt K(x,t,y(t))\right). \tag{F.2}$$

The form of Eq. F.1 implies that F can be written purely as a function of $\int_0^x dt K(x,t,y(t))$. We used an Adams-Bashforth second-order predictor step and a third-order Adams-Moulton corrector step. Following Linz, for the starting procedure of the Holstein model we applied the self-consistent Simpson's method iterated five times.[142]

Appendix G: PS Derivation

We begin with the temporal Dyson equation

$$0 = i\frac{d\mathcal{G}_k(t)}{dt} - \epsilon_k \mathcal{G}_k(t) - \int_0^t d\tau \Sigma_k(t-\tau)\,\mathcal{G}_k(\tau)\,. \tag{G.1}$$

The approach is leading order in the self-consistency, so we consider only the first-order self-energy

$$\Sigma_k^{(1)}(t) = \frac{ig^2}{N}\sum_q \mathcal{D}_q^{(0)}(t)\,\mathcal{G}_{k-q}^{(0)}(t)\,, \tag{G.2}$$

$$\mathcal{G}_k^{(0)}(t) = -ie^{-i\epsilon_k t}. \tag{G.3}$$

We then replace the bare Green's function with the interacting Green's function to find the SC-M self-energy. We then shift our summation variable in the self-energy.

$$\Sigma_k(t-\tau) = \frac{ig^2}{N}\sum_q \mathcal{D}_{k-q}^{(0)}(t-\tau)\,\mathcal{G}_q(t-\tau)\,. \tag{G.4}$$

We then consider the interacting Green's function to be a product of the bare Green's function and some other function representing a power series of corrections

$$\mathcal{G}_k(t) \equiv \mathcal{G}_k^{(0)}(t)\,\mathcal{P}_k(t)\,. \tag{G.5}$$

The time derivative is easily taken as

$$i\frac{d\mathcal{G}_k(t)}{dt} = i\mathcal{G}_k^{(0)}(t)\frac{d\mathcal{P}_k(t)}{dt} + \epsilon_k \mathcal{G}_k(t)\,. \tag{G.6}$$

One can also define a non-self-consistent self-energy, $\Sigma^{(0)}_{kq}(t)$, so that we can rewrite the PS self-energy as

$$\Sigma_k(t-\tau) = \frac{-g^2}{N}\sum_q \Sigma^{(0)}_{kq}(t-\tau)\,\mathcal{P}_q(t-\tau)\,, \tag{G.7}$$

$$\Sigma^{(0)}_{kq}(t) = -i\mathcal{D}^{(0)}_{k-q}(t)\,\mathcal{G}^{(0)}_q(t)\,. \tag{G.8}$$

Combining Eqs. G.1, G.6, and G.7 we find

$$\frac{d\mathcal{P}_k(t)}{dt} = \frac{ig^2}{N}\sum_q \int_0^t d\tau e^{i\epsilon_k(t-\tau)}\Sigma^{(0)}_{kq}(t-\tau)\,\mathcal{P}_q(t-\tau)\,\mathcal{P}_k(\tau)\,. \tag{G.9}$$

The novelty of the Power Series method is introduced by making a further approximation on top of the Migdal approximation exclusively for $q = k$

$$\mathcal{P}_k(\tau)\,\mathcal{P}_k(t_2-\tau) \to \mathcal{P}_k(t_2)\,. \tag{G.10}$$

This results in the final set of equations which are identical to those in Ref. 80 and 106.

$$\begin{aligned}\frac{d\mathcal{P}_k(t)}{dt} =& \frac{ig^2}{N}\sum_{q\neq k}\int_0^t d\tau e^{i\epsilon_k(t-\tau)}\Sigma^{(0)}_{kq}(t-\tau)\,\mathcal{P}_q(t-\tau)\,\mathcal{P}_k(\tau) \\ &+\frac{ig^2}{N}\int_0^t d\tau e^{i\epsilon_k(t-\tau)}\Sigma^{(0)}_{kk}(t-\tau)\,\mathcal{P}_k(t)\,.\end{aligned} \tag{G.11}$$

To see the exact correspondence with equation 11 in Ref. 106 one simply shifts the summation in A11 via $q' = k - q$ and does a u-substitution in the integration to $\tau' = t - \tau$

$$\begin{aligned}\frac{d\mathcal{P}_k(t)}{dt} =& -\frac{ig^2}{N}\sum_{q\neq 0}\int_0^t d\tau e^{i\epsilon_k\tau}\tilde{\Sigma}^{(0)}_{kq}(\tau)\mathcal{P}_k(\tau)\,\mathcal{P}_{k-q}(t-\tau) \\ &-\frac{ig^2}{N}\int_0^t d\tau e^{i\epsilon_k\tau}\tilde{\Sigma}^{(0)}_{k0}(\tau)\mathcal{P}_k(t)\,,\end{aligned} \tag{G.12}$$

$$\tilde{\Sigma}^{(0)}_{kq}(t) \equiv \Sigma^{(0)}_{k,k-q}(t)\,. \tag{G.13}$$

Appendix H: Recasting the PS into a Memory Kernel

Starting with Eq. G.11, we move back to formulating the method directly in terms of Green's functions.

$$\frac{1}{\mathcal{G}_k^{(0)}(t)}\left(\frac{d\mathcal{G}_k(t)}{dt}+i\epsilon_k\mathcal{G}_k(t)\right)=\frac{g^2}{N}\sum_{q\neq k}\int_0^t d\tau\frac{\mathcal{G}_k^{(0)}(\tau)}{\mathcal{G}_k^{(0)}(t)}\mathcal{D}_{k-q}^{(0)}(t-\tau)\,\mathcal{G}_q^{(0)}(t-\tau)\,\mathcal{P}_q(t-\tau)\,\mathcal{P}_k(\tau) \tag{H.1}$$
$$+\frac{g^2}{N}\int_0^t d\tau\frac{\mathcal{G}_k^{(0)}(\tau)}{\mathcal{G}_k^{(0)}(t)}\mathcal{D}_0^{(0)}(t-\tau)\,\mathcal{G}_k^{(0)}(t-\tau)\,\mathcal{P}_k(t).$$

Moving closer to the Dyson form we rearrange some terms, and recombining the PS terms with the associated bare Green's functions, we can find the following expression which is nearly in the form we want.

$$0=i\frac{d\mathcal{G}_k(t)}{dt}-\epsilon_k\mathcal{G}_k(t)-\frac{ig^2}{N}\sum_{q\neq k}\int_0^t d\tau\mathcal{D}_{k-q}^{(0)}(t-\tau)\,\mathcal{G}_q(t-\tau)\,\mathcal{G}_k(\tau) \tag{H.2}$$
$$-\frac{g^2}{N}\int_0^t d\tau\mathcal{D}_0^{(0)}(t-\tau)\,\mathcal{G}_k(t)\,\frac{\mathcal{G}_k(\tau)}{\mathcal{G}_k(\tau)}.$$

For the second term we have taken advantage of the fact that $\mathcal{G}_k^{(0)}(t-\tau)=-i\frac{\mathcal{G}_k^{(0)}(t)}{\mathcal{G}_k^{(0)}(\tau)}$ Collecting the terms, into a memory kernel.

$$0=i\frac{d\mathcal{G}_k(t)}{dt}-\epsilon_k\mathcal{G}_k(t)-\int_0^t d\tau\left[\frac{g^2}{N}\sum_q\begin{cases}i\mathcal{D}_{k-q}^{(0)}(t-\tau)\,\mathcal{G}_q(t-\tau) & q\neq k\\ \mathcal{D}_0^{(0)}(t-\tau)\,\frac{\mathcal{G}_k(t)}{\mathcal{G}_k(\tau)} & q=k\end{cases}\right]\mathcal{G}_k(\tau). \tag{H.3}$$

The quantity in square brackets is exactly the memory kernel of Eq. 3.8.

Appendix I: Validation of PS Implementation

Like in Ref. 68, we implement the PS method in the integro-differential equation form using a general code for evaluating equations of the form

$$y'(t) = F\left(t, y(t), \int_0^t d\tau K[t, \tau, y(\tau)]\right). \tag{I.1}$$

The PS formulation is the best way to numerically simulate the method because it removes the trivial oscillations of the bare Green's function. So the exact equation we implement (which follows directly from Eq. G.11) is

$$\frac{d\mathcal{P}_k(t)}{dt} = \frac{-ig^2}{N}\int_0^t d\tau \sum_q e^{i(\epsilon_k-\epsilon_q)(t-\tau)} \begin{cases} \mathcal{D}^{(0)}_{k-q}(t-\tau)\,\mathcal{P}_q(t-\tau)\,\mathcal{P}_k(\tau) & k \neq q \\ \mathcal{D}^{(0)}_0(t-\tau)\,\mathcal{P}_k(t) & k = q \end{cases}. \tag{I.2}$$

Appendix J: Closure of All-Coupling Ansatz

While the unitary transformation defined by Eq. 4.5 has appealing properties, when combined with the *ab initio* Hamiltonian in Eq. 4.2 it does not close.

We will instead work with a truncated version of the transformation which discards all terms greater than quadratic in h_{nq} While this approximation does change the Hamiltonian, we can examine the magnitudes of h_{nq} after optimization and discard solutions with large h_{nq}. [1]

We now work out the transformed Hamiltonian using the standard Baker–Campbell–Hausdorff (BCH) formula

$$e^{-S}Oe^{S} = O + [O,S] + \frac{1}{2!}[[O,S],S] + \dots. \tag{J.1}$$

The phonon creation/annihilation operators are readily transformed and become

$$e^{-S}b_{\nu q}e^{S} = b_{\nu q} - h^{*}_{\nu q}\sum_{ik}|i,k\rangle\langle i,k-a_q q|, \tag{J.2}$$

$$e^{-S}b^{\dagger}_{\nu q}e^{S} = b^{\dagger}_{\nu q} - h_{\nu q}\sum_{ik}|i,k-a_q q\rangle\langle i,k|. \tag{J.3}$$

[1] To preference against solutions which fall outside our acceptable bounds for $h_{\nu q}$ we set a cutoff h_{cut} and then apply a forcing function to our energy of the form

$$f(\{h_{\nu q}\}) = \sum_{\nu q}\begin{cases}\exp\left(c\frac{2|h_{\nu q}|^2 - h^2_{\text{cut}}}{h^2_{\text{cut}}}\right) - \exp(c) & |h_{\nu q}| > h_{\text{cut}} \\ 0 & |h_{\nu q}| \le h_{\text{cut}}\end{cases}.$$

c is the strength of the forcing term.

For the band term, we can write out the full transformation analytically.

$$H_e = \sum_{ik} \epsilon_{ik} |i,k\rangle\langle i,k| \tag{J.4}$$

$$e^{-S} H_e e^{S} = \sum_{i,k} \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\substack{q^{(1)}\ldots q^{(n)} \\ \nu^{(1)}\ldots\nu^{(n)}}} \left[\prod_{m=1}^{n} B_{\nu^{(m)} q^{(m)}} \right] C^{(n)}_{iq^{(1)}\ldots q^{(n)}}(k) \left| ik \right\rangle \left\langle ik - \sum_{m=1}^{n} a_{q^{(m)}} q^{(m)} \right| \tag{J.5}$$

Here, the coefficients are given by the recursion relation

$$C^{(0)}_{i}(k) = \epsilon_{ik}, \tag{J.6}$$

$$C^{(n)}_{iq^{(1)}\ldots q^{(n)}}(k) = C^{(n-1)}_{iq^{(1)}\ldots q^{(n-1)}}(k) - C^{(n-1)}_{iq^{(1)}\ldots q^{(n-1)}}(k - a_{q^{(n)}} q^{(n)}). \tag{J.7}$$

Evaluating Eq. J.5 in it's entirety is impractical and we will use only up to second order. However, the structure of Eq. J.7 demonstrates that if ϵ_{ik} is an m^{th}-order polynomial, then all $C^{(n>m)} = 0$.

Considering the case of the Fröhlich model, which has parabolic bands, we now understand why the band term for the original ansatz closed exactly on that model.[116] As an example, the three non-vanishing coefficients for the Fröhlich model are

$$C^{(0)}_{k} = \frac{k^2}{2m} \tag{J.8}$$

$$C^{(1)}_{k,q} = -\frac{a_q q \left(a_q q - 2k\right)}{2m} \tag{J.9}$$

$$C^{(2)}_{k,q_1,q_2} = \frac{a_{q_1} q_1 a_{q_2} q_2}{m}. \tag{J.10}$$

It is clear that all higher order terms will vanish because $C^{(2)}$ has no k dependence, so further application of Eq. J.7 will produce zero.

We now consider the electron-phonon interaction term,

$$H_{\text{e-ph}} = \sum_{\substack{ijv \\ kq}} g^{ij}_{\nu}(k,q) |i, k+q\rangle\langle j,k| (b_{\nu q} + b^{\dagger}_{\nu -q}). \tag{J.11}$$

The phonon operators are easily transformed so we need only consider transforming a single momentum sum over the density $\tilde{\rho}_q^{ij} = \sum_k g_\nu^{ij}(k,q)|ik+q\rangle\langle jk|$. Like before we can write a recursive expression for the full transformation

$$e^{-S}\tilde{\rho}_q^{ij}e^S = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{\substack{k\\ q^{(1)}\ldots q^{(n)}\\ \nu^{(1)}\ldots\nu^{(n)}}}\left[\prod_{m=1}^{n}B_{\nu^{(m)}q^{(m)}}\right]D^{(n)}_{ij\nu qq^{(1)}\ldots q^{(n)}}(k)\left|ik+q+\sum_{m=1}^{n}a_{q^{(m)}}q^{(m)}\right\rangle\left\langle ik\right|. \quad \text{(J.12)}$$

Similarly to Eq. J.7, the coefficients are given by the recursion relation

$$D^{(0)}_{ij\nu q}(k) = g_\nu^{ij}(k,q), \quad \text{(J.13)}$$

$$D^{(n)}_{ij\nu qq^{(1)}\ldots q^{(n)}}(k) = D^{(n-1)}_{ij\nu qq^{(1)}\ldots q^{(n-1)}}\left(k+a_{q^{(n)}}q^{(n)}\right) - D^{(n-1)}_{ij\nu qq^{(1)}\ldots q^{(n-1)}}(k). \quad \text{(J.14)}$$

Because we only consider terms which do not vanish upon application of the phonon vacuum and are overall second order or lower in the phonon displacement, our energy expression only requires the zeroth- and first-order terms in Eq. J.12. We also rearrange the first-order term so that the electron-phonon coupling elements are easily evaluated

$$e^{-S}\tilde{\rho}_q^{ij}e^S \approx \tilde{\rho}_q^{ij} + \sum_{k\nu'q'}B_{\nu'q'}g_\nu^{ij}(k,q)\left(|i,k+q\rangle\langle j,k-a_{q'}q'| - |i,k+q+a_{q'}q'\rangle\langle j,k|\right). \quad \text{(J.15)}$$

Like in Eq. J.7, if $g_\nu^{ij}(k,q)$ is a polynomial in k with degree m then all terms of order $n > m$ in Eq. J.12 will vanish and the transformation closes. In Ref. 116 the interaction term closed without approximation for the Fröhlich model, which can now be understood because in that model the coupling is not k-dependent and only depends on the momentum of the emitted/absorbed phonon. It is important to note that this term also closes in the strong coupling limit where like in the band term the first- and all higher-order terms vanishes.

In full, there are three cases where our method is rigorously variational. There is the trivial weak-coupling limit ($g(k,q) = 0$), and the strong coupling ($a_k = 0$) limit. Finally, there is the case which includes the Fröhlich model where the electron-phonon coupling is independent of the

electronic-momentum and the electronic bands polynomials up to second-degree.

Appendix K: Numerical Evaluation of All-Coupling Ansatz

In order to evaluate Eq. 4.8 we must be more precise about what it means to have a continuous-valued variational parameter in the index of another parameter. It would seem at first that this type of parameter would only make sense in the thermodynamic limit when the parameters become functions with continuous domains.

Nonetheless, we can define the method for finite-sized systems in an unambiguous manner via the Fourier transform. For some parameter η which is defined in the discrete reciprocal lattice and two vectors k and q which are on the lattice, we define

$$\eta_{k+a_q q} = \sum_r \eta_r e^{-ia_q q\cdot r} e^{-ik\cdot r} = \mathcal{F}\left[\eta_r e^{-ia_q q\cdot r}; r, k\right], \tag{K.1}$$

$$\eta_r = \mathcal{F}^{-1}[\eta_k]. \tag{K.2}$$

The trade-off for using the a_k parameters is that now the ansatz is in a mixed momentum-position representation and requires Fourier transformations at each step. The Fourier interpolation allows us to have a well-defined energy expression and also then prescribes a straightforward procedure to compute the gradients.

We note that for our optimization it is numerically more convenient for all of our parameters to have domains of $(-\infty, \infty)$, so we define $a_k = S(z_k)$ where $S(x) = (1 + \exp(-x))^{-1}$ is the Sigmoid function.

We optimize the variational parameters via L-BFGS-B minimization of Eq. 4.7.[143] In the equations below, we display the gradients of the energy with respect to the variational parameters. We only show the gradients of the numerator of Eq. 4.7 which we denote as E_0. The full gradient is then constructed via the standard rules of calculus.

Starting with the displacements

$$\frac{\partial E_0}{\partial h^*_{\nu p}} = -\frac{1}{2} h_{\nu p} \sum_{iq} \mathcal{E}_{iqp} |t_{iq}|^2 + \omega_{\nu p} h_{\nu p} \sum_{ik} |t_{ik}|^2 - \sum_{ijk} g^{ij}_{\nu}(k,p) t^*_{ik+p(1-a_p)} t_{jk} \tag{K.3}$$

In the Fourier representation of the problem we can show that our modified band energy can be evaluated as

$$\mathcal{E}_{iqk} = 4 \sum_r \epsilon_{ir} \sin^2 \left(\frac{k a_k \cdot r}{2} \right) e^{-iq\cdot r}, \tag{K.4}$$

For the CI amplitudes

$$\begin{aligned} \frac{\partial E_0}{\partial t^*_{ip}} = &\left(\epsilon_{ip} - \frac{1}{2} \sum_{\nu k} |h_{\nu k}|^2 \mathcal{E}_{ipk} \right) t_{ip} \\ &- \sum_{\substack{j\nu \\ q}} \left(g^{ij}_{\nu}(p-q,q) h_{\nu-q} t_{jp-q(1-a_q)} + \frac{1}{N} \sum_{kr} g^{ij}_{\nu}(k,q)^* h^*_{\nu-q} t_{ik+q} e^{i(k-p+a_q q)\cdot r} \right). \end{aligned} \tag{K.5}$$

The gradient with respect to a_k is given by

$$\begin{aligned} \frac{\partial E_0}{\partial a_p} = &- \sum_{qi\nu} |t_{iq}|^2 \left(|h_{\nu p}|^2 + |h_{\nu -p}|^2 \right) \sum_r \epsilon_{ir} (p\cdot r) \sin(a_p p \cdot r) e^{-iq\cdot r} \\ &+ 2\Re \left(i \sum_{\substack{ij\nu \\ k}} g^{ij}_{\nu}(k,p) h_{\nu-q} t^*_{ik+p} \sum_r t_{jr} (p\cdot r) e^{-ia_p p\cdot r} e^{-ik\cdot r} \right) \\ &- 2\Re \left(i \sum_{\substack{ij\nu \\ k}} g^{ij}_{\nu}(k,-p) h_{\nu q} t^*_{ik-p} \sum_r t_{jr} (p\cdot r) e^{ia_p p\cdot r} e^{-ik\cdot r} \right). \end{aligned} \tag{K.6}$$