

Ab Initio Polaron Wave Functions

Paul J. Robinson,¹ Joonho Lee,^{1,2} Ankit Mahajan,¹ and David R. Reichman^{1,*}

¹*Department of Chemistry, Columbia University, New York, New York 10027, United States*

²*Department of Chemistry and Chemical Biology,
Harvard University, Cambridge, Massachusetts 02138, United States*

In this work we demonstrate that accurate ground state wave functions may be constructed for polarons in a fully *ab initio* setting across the wide range of couplings associated with both the large and small polaron limits. We present a single general unitary transformation approach which encompasses an *ab initio* version of the Lee-Low-Pines theory at weak coupling and the coherent state Landau-Pekar framework at strong coupling while interpolating between these limits in general cases. We show that perturbation theory around these limits may be performed in a facile manner to assess the accuracy of the approach, as well as provide an independent route to the *ab initio* properties of polarons. We test these ideas on the case of LiF, where the electron-polaron is expected to be large and relatively weakly coupled, while the hole-polaron is expected to be a strongly coupled small polaron.

The interaction between charge carriers and phonons in solids controls a host important phenomena ranging from transport to superconductivity [1]. One of the most ubiquitous consequences of these interactions is the formation of polarons, quasiparticles composed of a charge carrier and a surrounding cloud of phonons [2, 3]. Polarons can have dramatically different properties from those of bare electrons and holes. For example, the quasi-particle band gap, carrier mass, and optical conductivity can all be strongly renormalized by polaron formation [4].

The theoretical description of polarons has evolved over the better part of the past 70 years, and has historically centered on a small number of canonical model Hamiltonians [4]. Such models represent idealized limits of the distinct physical situations in which polarons arise, stripped from the complications associated with the detailed microscopic electronic structure of the underlying solid. Some prominent examples include the Fröhlich model, the Holstein model and the Su-Schrieffer-Heeger (SSH) model. The Fröhlich model describes the interaction of a free electron with a continuous polarizable medium [5], and has provided important insights into large polaron formation in polar and ionic solids such as strontium titanate (SrTiO₃), silver chloride (AgCl), and lithium fluoride (LiF) [6]. The Holstein model describes electrons on a lattice interacting with non-dispersed local phonons with a uniform on-site electron-phonon interaction (EPI) [7, 8]. This model is appropriate for the study of small polarons in molecular crystals. The (electron-phonon) SSH model describes situations where nearest neighbor lattice displacements modulate the hopping of charge between lattice sites, and has been instrumental in describing excitations such as solitons in polyenes [9]. Each of these complementary models is rich in phenomenology, and the numerous studies of their solutions in different parameter regimes has provided a foundation

for the understanding of polarons in a multitude of settings. Real solids are expected to contain features of all of these models, as well as details excluded by them.

Recently, significant effort has been put forward to describe the properties of polarons starting from the *ab initio* Hamiltonian of a system that incorporates the full electron and phonon band structures, as well as the wave vector-dependent coupling of charge and phonons within these bands. Among these approaches, we highlight several which have a connection to the framework we describe in this paper. In Ref. [10] a Landau-Pekar-like variational ansatz was developed and employed to study electron and hole polarons in a variety of solids. The approach avoids the use of real space supercells by working directly within the first Brillouin zone, and provides a precise way to quantify which phonon modes contribute to specific polaronic properties. Ref. [11] approaches the same problem via the use of a simple canonical transformation on the *ab initio* Hamiltonian to efficiently recover polaron binding energies in specified limits. Connections between these two methods have been elucidated which demonstrate that these seemingly distinct theories effectively employ the same strong coupling ansatz, and thus are appropriate for strongly coupled, highly localized polarons such as those formed in the valence band of LiF [12]. More recently, a Green's function approach has been developed which provides a robust many-body framework for calculating polaron properties for *all* couplings [13, 14]. Although significantly more computationally involved than the simple approaches of Refs. [11, 12], this framework has the advantage of not only generality with respect to coupling strength, but also, in principle, the access to frequency-dependent quantities such as spectral functions as well as finite charge concentration effects.

Our goal in this work is to demonstrate that ground state wave functions may be constructed to accurately calculate *ab initio* polaronic properties for *arbitrary* coupling strengths in a simple and computationally efficient manner. Specifically, we develop a canonical transformation approach which is capable of capturing *ab initio*

* drr2103@columbia.edu

binding energies of polarons in both the large polaron, weak coupling limit, as well as small polaron, strong coupling limit. We then demonstrate that Rayleigh-Schrödinger (RS) perturbation theory may be successfully employed to capture these limits as well, with results that show remarkable consistency for polaron binding energies with the canonical transformation approach. These results demonstrate simple, accurate, and computationally efficient routes to the *ab initio* calculation of the properties of polarons in solids for arbitrary electron-phonon coupling strengths.

In the context of the Fröhlich model, the well-known theory of Lee, Low and Pines (LLP) is applicable for weakly coupled electron-phonon systems [17], while the adiabatic Landau-Pekar (LP) solution of the polaron problem is applicable for strongly coupled electron-phonon systems [4, 18, 19]. As discussed by Huybrechts [20, 21], both theories can be unified from the perspective of a single unitary transformation with variational parameters that account for the displacement of the origin of the phonon modes as well as the momentum transfer associated with charge-phonon scattering. Nagy and Markoš (NM) modified Huybrechts' ansatz by allowing the degree of momentum conservation to vary across the electronic band [22]. This modification rectifies several shortcomings of the original Huybrechts approach, and when applied to the Fröhlich model produces results for the binding energy over a very wide range of coupling parameters that are essentially as accurate with respect to absolute error as the approach of Ref. [13, 14] as shown in the Supplemental Material (SM). This approach forms the basis of our all-coupling *ab initio* wave function method, as outlined next.

Consider the general *ab initio* electron-phonon Hamiltonian, written to leading linear order in the EPI as

$$H = \sum_{iq} \epsilon_{iq} |i, q\rangle \langle i, q| + \sum_{q, \nu} \omega_{\nu q} b_{\nu q}^\dagger b_{\nu q} + \sum_{\substack{ij\nu \\ kq}} g_{ij\nu}^{ij}(k, q) |i, k+q\rangle \langle j, k| (b_{\nu q} + b_{\nu -q}^\dagger). \quad (1)$$

Here, the electronic bands, phonon frequencies and electron-phonon matrix elements are given by ϵ_{iq} , $\omega_{\nu q}$ and $g_{ij\nu}^{ij}(k, q)$, respectively. The phonon creation (annihilation) operator for momentum q is given by $b_{\nu q}^\dagger$ ($b_{\nu q}$). Since in this work we only consider the case of a single electron (hole) in the conduction (valence) bands, we work in the projector basis for electrons (holes), thus simplifying the application of the transformation defined below.

We define a trial wave function

$$|\Psi_{tikh\nu qa_q}\rangle = U|0\rangle_{\text{ph}} \otimes |\psi\rangle_{\text{el}}. \quad (2)$$

The electronic portion of the wave function is given by the familiar configuration interaction (CI) singlets form

$$|\psi\rangle_{\text{el}} = \sum_{ik} t_{ik} |i, k\rangle, \quad (3)$$

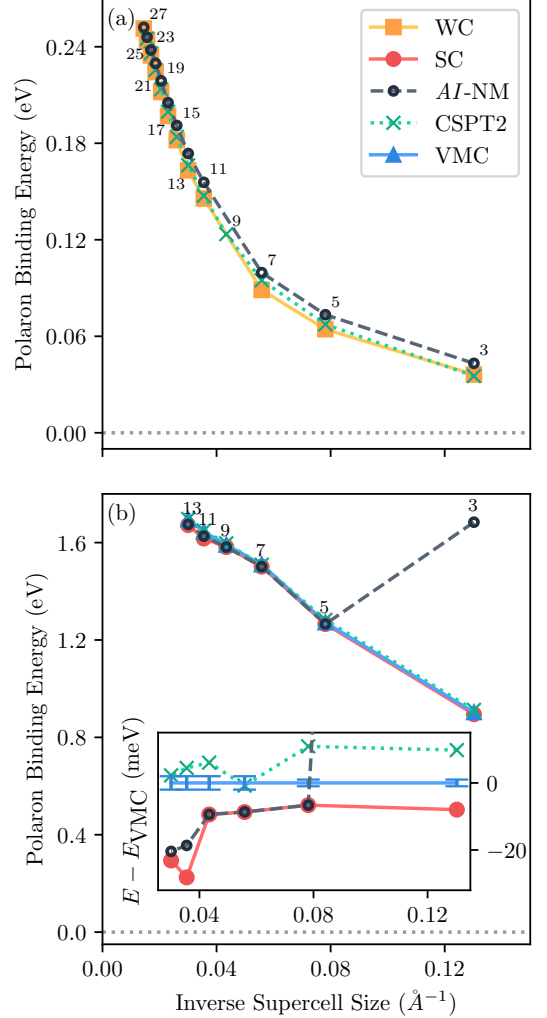


FIG. 1. LiF electron- and hole-polaron binding energies calculated via the WC ansatz (orange squares), the SC ansatz (red circles), the AI-NM ansatz (open black circles), CSPT2 (green Xs and dotted line), and the exact VMC approach of Ref. [15] (blue triangles). Each of these approaches are described in the main text and SM. (a) Plot of the convergence of the electron-polaron binding energy with respect to the inverse supercell size (as defined in Ref. [16]). The small numbers above the data indicate the size of the supercell, e.g. the size N of the N^3 k-grid. (b) Plot of the convergence of the hole-polaron binding energy with respect to inverse supercell size. The inset highlights the small differences between the methods.

where the label i denotes the band index and k the wave vector within the 1st Brillouin zone. $|0\rangle_{\text{ph}}$ denotes the phonon vacuum, and the unitary transformation $U = e^S$

System	WC (eV)	SC (eV)	<i>AI</i> -NM (eV)	CSPT2 (eV)	VMC (eV)
electron-polaron	0.35	0.24	0.33	0.35	–
hole-polaron	0.30	1.97	1.94	1.96	1.97

TABLE I. Polaron binding energies extrapolated to the thermodynamic limit for the electron and hole polarons in LiF via the WC, SC, *AI*-NM, CSPT2 methods (see text and SM), and a numerically exact VMC approach [15]. The binding energies were determined by linearly extrapolating the energies to the thermodynamic limit in the reciprocal supercell volume using only the 23^3 and 25^3 k-grids for the electron-polaron and only the 11^3 and 13^3 k-grids for the hole-polaron.

is defined as

$$S = \sum_{\substack{i\nu \\ kq}} B_{\nu q} |i, k + a_q q\rangle \langle i, k|, \quad (4)$$

$$B_{\nu q} \equiv h_{\nu q} b_{\nu q} - h_{\nu-q}^* b_{\nu-q}^\dagger. \quad (5)$$

In the following, we call this transformation the *AI*-NM (*ab initio* Nagy-Markoš) approach.

There are three sets of variational parameters to be determined for this ansatz, namely the CI coefficients $\{t_{ik}\}$, the coefficients that modify electron-phonon scattering momentum $\{a_k\}$, and the variational phonon mode displacements $\{h_{\nu k}\}$. Thus, when the wave vector grid is sizable, thousands of variational parameters are used to parameterize the ground state. Note that when the set of parameters $\{a_k\} = 1$, the transformation coincides with an *ab initio* version of the LLP transformation, while when $\{a_k\} = 0$, an *ab initio* version of the Landau-Pekar theory results. Importantly, we remark that in general the unitary rotation of the Hamiltonian cannot be carried out in explicit closed form except in particular limits. Specifically, as discussed in the SM, the limits of weak coupling, strong coupling, and arbitrarily strong (electron) momentum-independent charge-phonon coupling with a parabolic band structure all yield simple closed expressions for the unitary transformation of the Hamiltonian. Given the wide scope of these limits, the landscape of the energy functional will have many local minima, and we proceed by determining the transformation variationally, even without a rigorous minimum principle. A similar procedure has been employed in electronic structure theory with respect to the unitary Coupled Cluster method. In particular, our approach is equivalent to the UCC(2) method [23].

We determine the ground state energy variationally via minimization

$$E_g \leq E = \min_{t_{ik} h_{\nu q} a_q} \frac{\langle \Psi_{t_{ik} h_{\nu q} a_q} | H | \Psi_{t_{ik} h_{\nu q} a_q} \rangle}{\langle \Psi_{t_{ik} h_{\nu q} a_q} | \Psi_{t_{ik} h_{\nu q} a_q} \rangle}, \quad (6)$$

where E_g is the true ground state of the system. Through the second order in $h_{\nu q}$ the transformed energy E is given

as

$$E = \frac{1}{\sum_{ik} |t_{ik}|^2} \left\{ \sum_{iq} \left(\epsilon_{iq} - \frac{1}{2} \sum_{\nu k} |h_{\nu k}|^2 \mathcal{E}_{iqk} \right) |t_{iq}|^2 + \sum_{\nu q} \omega_{\nu q} |h_{\nu q}|^2 \sum_{iq} |t_{iq}|^2 - \sum_{\substack{ij\nu \\ kq}} (g_{\nu}^{ij}(k, q) h_{\nu-q} t_{ik+q}^* t_{jk+a_q q} + \text{c.c.}) \right\}, \quad (7)$$

$$\mathcal{E}_{iqk} \equiv 2\epsilon_{iq} - \epsilon_{iq-ak} - \epsilon_{iq+ak}, \quad (8)$$

We define all of our “off-grid” quantities via Fourier interpolation. This is similar in spirit to Wannier interpolation, a well-known approach to bridge low density k-mesh calculations and higher density ones [24]. This procedure not only provides a well-defined procedure to evaluate our energy for any set of variational parameters, but as is detailed in the SM, we can also evaluate the gradients of Eq. 7 analytically, allowing efficient and direct minimization for Eq. 6. A discussion of the procedure for determining the variational parameters may be found in the SM.

To illustrate the range and accuracy of our all-coupling *AI*-NM approach, we consider the case of LiF where the electron-polaron is expected to behave as a relatively large polaron while the hole-polaron is expected to form a small polaron. We work entirely within the first Brillouin zone and converge energies via a simple Makov-Payne extrapolation. 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. Details of the *ab initio* calculation and parameters are contained in the SM.

In addition to the *AI*-NM ansatz, one can restrict the form of Eq. 2 to the extreme limits of weak-coupling (WC), for which we set $a_k = 1$ for all k and strong-coupling (SC) for which we set $a_k = 0$ for all k . Note that for the SC case, the resulting transformation is one where coherent-state nuclear wave functions are employed as a basis. We also consider second-order RS perturbation theory in conjunction with coherent states (CSPT2) (see SM for details). This approach is the *ab initio* version of the coherent-state Møller-Plesset perturbation theory formulated for the Hubbard-Holstein model previously [25]. In the limits of very strong or very weak coupling, this method provided a second order energy correction

about the Landau-Pekar or non-interacting limits, respectively. As discussed in the SM, it is not possible to *a priori* determine whether the variational SC solution or the zero-displacement solution is the more appropriate reference for perturbation theory. For the case of the electron- and hole-polarons we uniformly apply the zero-displacement or variational displacement solutions as the reference states respectively. Lastly, in the case of the hole-polaron we employ the essentially exact VMC approach of Ref. [15] as a benchmark. As we will see, we find a remarkable consistency between all of these approaches for both the electron and hole polaron systems.

In Fig. 1 we show the grid size dependence of the electron and hole polaron binding energies for the different approaches mentioned above. Extrapolation with respect to increasingly larger wave vector grids yields a rather consistent picture of the binding energy for both the electron and hole polarons in LiF. These extrapolated values can be found in Table I. The following conclusions may be reached. With respect to the hole-polaron, a consistent binding energy close to 2 eV is found. We expect that this result is essentially exact within the linear electron-phonon coupling model and the level of *ab initio* theory used to construct the Hamiltonian. The largest uncertainty in the value of our binding energies likely arises due to the use of a crude extrapolation to the infinitely fine grid (thermodynamic) limit. The simple SC theory ($a_k = 0$ for all k) already appears sufficient to capture this result, with almost no change in the binding energy when second order perturbation theory is performed with these states, again suggesting that convergence has been achieved. The final value is very similar to the results for the hole-polaron found in Refs. [10–14]. This is not surprising, as the SC result is equivalent to an *ab initio* version of Landau-Pekar theory.

The case of the electron-polaron is more subtle. Here we do not have VMC results to compare with. The pure SC result gives values essentially identical to [10]. Again this is expected as the Landau-Pekar result. However the WC and *AI-NM* approach give results consistent with each other which are close to fifty percent larger than the result found for the electron-polaron in Ref. [10]. Remarkably, the same value is found from the CSPT2 approach, again suggesting some level of convergence and confidence in the value of the electron-polaron binding energy. It should be pointed out that this value is smaller than that found more recently in Refs. [13, 14]. However it is difficult to compare these values as [13, 14] include effects that arise from the second-order Debye-Waller term absent in our Hamiltonian (Eq.1) and our calculations.

It may appear that sole the utility of the full variational *AI-NM* approach resides in interpolating between the WC and SC limits in an *ab initio* fashion. However we note that the ansatz of Eq.2-5 is flexible enough to describe physics beyond these strict limits. To illustrate this we note that in the pure WC theory, the momentum density distribution of the polaron is a delta function in

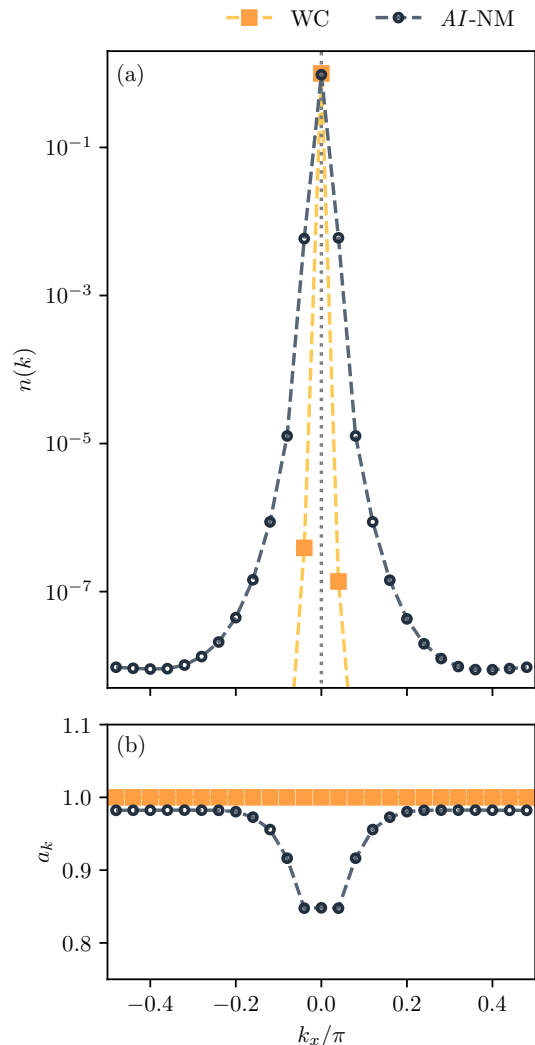


FIG. 2. Detailed view of the electron-polaron results for the 25^3 k-grid. (a) Comparison of a slice of the momentum density, $n(k) = \langle a_k a_k^\dagger \rangle$, for WC and *AI-NM*. The dotted line indicates the delta function of an infinitely localized $n(k)$ in the thermodynamic limit. (b) The value of the momentum-conservation-modulating variational parameter a_k for the *AI-NM* and WC methods demonstrating that non-trivial a_k structure can be present in large *ab initio* systems.

k-space, and thus the large polaron is unrealistically delocalized. Within the full approach of Eq.2-5, however, solutions are found that partially localize the polaron even though the binding energy is nearly identical to the WC value. This is illustrated in Fig.2. Localization occurs because low-lying states in the landscape of variational parameters can have non-uniform a_k values between 0 and 1. We note that a full reconstruction of the *observable* momentum density distribution ($n(k) = \sum_i n_i(k)$ with the sum taken over bands labels) requires undoing the

unitary transformation which in fully *ab initio* problems necessitates an expansion in the $h_{\nu q}$ parameters which is rather involved. Thus in Fig.2 only the leading-order term $n(k) \sim \sum_i |t_{ik}|^2$ is plotted, where the localization effect (in real-space) is already evident.

In conclusion, we have presented several routes to the *ab initio* calculation of the properties polarons based on ground state wave functions that encompass the full range of realistic coupling strengths in solids. The first approach is based an all-coupling unitary transformation which naturally interpolates between known extreme limits of the problem. Using the example of LiF, we demonstrate that this method can faithfully capture properties of both the large electron polaron and the small hole polaron, without predetermined knowledge of the physical properties of each system. We then demonstrate that perturbation theory around the strong coupling limit provides a distinct but harmonious way to compute these properties. Both approaches have the advantage of sim-

plicity and appear to be highly accurate in a realistic *ab initio* setting. Although not discussed here, these approaches may also form the basis of a facile means to compute real frequency spectral information as well as the properties of exciton-polarons [26–30]. These topics, as well as further calculations on a wide range of systems will be taken up in future work.

ACKNOWLEDGEMENTS

P.J.R. acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. DGE-2036197. D.R.R. and A.M. are partially supported by NSF CHE-2245592. We would like to thank Marco Bernardi and Yao Luo for helpful assistance during the early stages of this work.

-
- [1] N. W. Ashcroft and M. N. David, *Solid state physics* (Holt, Rinehart and Winston, New York, 1976).
 - [2] J. Devreese and F. Peeters., eds., *Polarons and excitons in polar semiconductors and ionic crystals*, NATO ASI series. Series B, Physics v. 108 (Plenum Press, New York, 1984).
 - [3] C. Franchini, M. Reticioli, M. Setvin, and U. Diebold, Polarons in materials, *Nat. Rev. Mater.* **6**, 560 (2021).
 - [4] G. D. Mahan, *Many-Particle Physics*, 3rd ed. (Springer US, Boston, MA, 2000).
 - [5] H. Fröhlich, Electrons in lattice fields, *Adv. Phys.* **3**, 325 (1954).
 - [6] T. D. Schultz, Slow electrons in polar crystals: Self-energy, mass, and mobility, *Phys. Rev.* **116**, 526 (1959).
 - [7] T. Holstein, Studies of polaron motion: Part i. the molecular-crystal model, *Ann. Phys.* **8**, 325 (1959).
 - [8] T. Holstein, Studies of polaron motion: Part ii. the “small” polaron, *Ann. Phys.* **8**, 343 (1959).
 - [9] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Solitons in Polyacetylene, *Phys. Rev. Lett.* **42**, 1698 (1979).
 - [10] W. H. Sio, C. Verdi, S. Poncé, and F. Giustino, Polarons from First Principles, without Supercells, *Phys. Rev. Lett.* **122**, 246403 (2019).
 - [11] 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.* **5**, 063805 (2021).
 - [12] 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* **105**, 155132 (2022).
 - [13] 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.* **129**, 076402 (2022).
 - [14] 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* **106**, 075119 (2022).
 - [15] A. Mahajan, P. J. Robinson, J. Lee, and D. R. Reichman, Structure and dynamics of electron-phonon coupled systems using neural quantum states, to be submitted; see parallel arXiv submission (2024).
 - [16] W. H. Sio, C. Verdi, S. Poncé, and F. Giustino, *Ab initio* theory of polarons: Formalism and applications, *Phys. Rev. B* **99**, 235139 (2019).
 - [17] T. D. Lee, F. E. Low, and D. Pines, The motion of slow electrons in a polar crystal, *Phys. Rev.* **90**, 297 (1953).
 - [18] S. I. Pekar, *Untersuchungen über die Elektronentheorie der Kristalle* (De Gruyter, 1954).
 - [19] L. D. Landau, Electron motion in crystal lattices, *Phys. Z. Sowjet.* , 664 (1933).
 - [20] W. J. Huybrechts, Note on the ground-state energy of the feynman polaron model, *J. Phys. C: Solid State Phys.* **9**, L211 (1976).
 - [21] W. J. Huybrechts, Internal excited state of the optical polaron, *J. Phys. C: Solid State Phys.* **10**, 3761 (1977).
 - [22] P. Nagy and P. Markoš, On the nature of the self-trapping of different types of polarons, *Phys. Status Solidi B* **151**, 571 (1989).
 - [23] R. J. Bartlett, S. A. Kucharski, and J. Noga, Alternative coupled-cluster ansätze ii. the unitary coupled-cluster method, *Chem. Phys. Lett.* **155**, 133 (1989).
 - [24] J. R. Yates, X. Wang, D. Vanderbilt, and I. Souza, Spectral and fermi surface properties from wannier interpolation, *Phys. Rev. B* **75**, 195121 (2007).
 - [25] J. Lee, S. Zhang, and D. R. Reichman, Constrained-path auxiliary-field quantum monte carlo for coupled electrons and phonons, *Phys. Rev. B* **103**, 115123 (2021).
 - [26] J. Pollmann and H. Büttner, Effective hamiltonians and bindings energies of wannier excitons in polar semiconductors, *Phys. Rev. B* **16**, 4480 (1977).
 - [27] A. M. Alvertis, J. B. Haber, Z. Li, C. J. N. Coveney, S. G. Louie, M. R. Filip, and J. B. Neaton, Phonon screening and dissociation of excitons at finite temperatures from first principles (2023), arXiv:2312.03841 [cond-mat.mtrl-sci].

- [28] Z. Dai, C. Lian, J. Lafuente-Bartolome, and F. Giustino, Excitonic polarons and self-trapped excitons from first-principles exciton-phonon couplings, *Phys. Rev. Lett.* **132**, 036902 (2024).
- [29] Z. Dai, C. Lian, J. Lafuente-Bartolome, and F. Giustino, Theory of excitonic polarons: From models to first-principles calculations, *Phys. Rev. B* **109**, 045202 (2024).
- [30] P. J. Robinson, J. Lee, and D. R. Reichman, Unpublished.
- [31] 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* **62**, 6317 (2000).
- [32] B. Gerlach and H. Löwen, Analytical properties of polaron systems or: Do polaronic phase transitions exist or not?, *Rev. Mod. Phys.* **63**, 63 (1991).
- [33] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, Quantum espresso: a modular and open-source software project for quantum simulations of materials, *J. Phys. Condens. Matter* **21**, 395502 (2009).
- [34] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, Advanced capabilities for materials modelling with quantum espresso, *J. Phys. Condens. Matter* **29**, 465901 (2017).
- [35] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [36] D. R. Hamann, Optimized norm-conserving vanderbilt pseudopotentials, *Phys. Rev. B* **88**, 085117 (2013).
- [37] M. Schlipf and F. Gygi, Optimization algorithm for the generation of oncv pseudopotentials, *Comput. Phys. Commun.* **196**, 36 (2015).
- [38] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, An updated version of wannier90: A tool for obtaining maximally-localised wannier functions, *Comput. Phys. Commun.* **185**, 2309 (2014).
- [39] S. Poncé, E. Margine, C. Verdi, and F. Giustino, Epw: Electron-phonon coupling, transport and superconducting properties using maximally localized wannier functions, *Comput. Phys. Commun.* **209**, 116 (2016).
- [40] 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 $f(\{h_{\nu q}\}) = \sum_{\nu q} \exp(c(2|h_{\nu q}|^2 - h_{\text{cut}}^2)/h_{\text{cut}}^2) - \exp(c)$ if $|h_{\nu q}| > h_{\text{cut}}$ and zero otherwise. c is the strength of the forcing term.
- [41] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, *Nat. Methods* **17**, 261 (2020).

SUPPLEMENTAL MATERIAL

1. Fröhlich Model

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. [22] for a derivation and description. While the main method of this work is intended for application in *ab initio* systems, this section presents information about the general properties of the NM ansatz.

In Ref. [22] the authors considered a generalized Fröhlich model which allows for several types of polarons to be considered. The Hamiltonian we consider is

$$H = \sum_k \epsilon_k |k\rangle \langle k| + \omega_0 \sum_q b_q^\dagger b_q + \sum_{kq} g(q) |k+q\rangle \langle k| (b_q + b_{-q}^\dagger). \quad (\text{SM.1})$$

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

Nagy and Markoš employ a unitary transformation defined in Eq. 6 of Ref. [22] with the projection of the density written in the form $\sum_k |k + a_q q\rangle \langle k|$, where a_q is a variational parameter defined for all q to be between zero and one. Nagy and Markoš additionally consider an ansatz for the electronic wave function of the Gaussian form (an assumption which we relax in our *ab initio* version), with the k-space wave function given by $\sqrt{\frac{8}{V}} \left(\frac{\pi}{\lambda}\right)^{3/4} e^{-\frac{k^2}{2\lambda}}$. Using their transformation and this polaronic wave function one may write the energy as an integral equation dependent on λ .

The results of the NM ansatz for the Fröhlich model are summarized in Fig. SM.1. As mentioned previously, the Fröhlich model is challenging for variational approaches, and most incorrectly [32] predict a phase transition when transitioning from WC to SC in the form of a discontinuity 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]. Compared to the numerically exact results of Ref. [31], the NM ansatz never errs by more than $\approx 14\%$ over the full range of couplings. The Green's function approach of Ref. [13] has a slightly lower maximum error of $\approx 10\%$; however, it consistently overshoots the

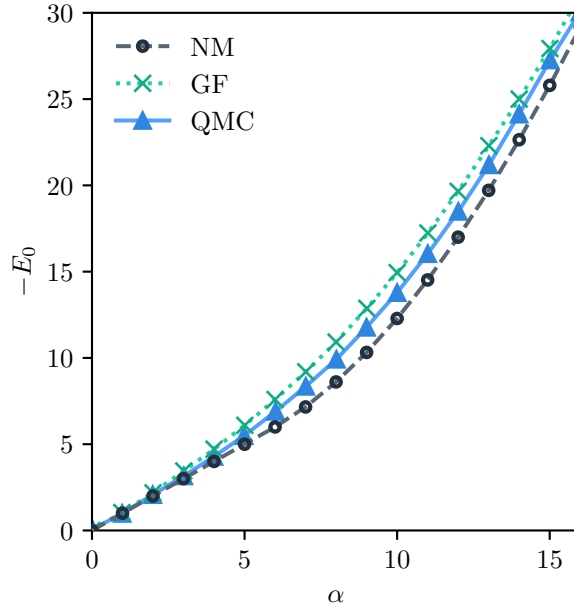


FIG. SM.1. Exact ground state energy of the Fröhlich model from Ref. [31] (blue triangles), the Green's function approach of Ref. [13] (green X with dotted line), and the NM ansatz (open black circles with dashed lines).

exact energy (the approach is not variational) and does not do as well as the NM ansatz or the pure WC result for very 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.

2. Ab Initio Calculations

We follow the procedure for generating ϵ_{ik} , $\omega_{\nu q}$ and $g_{\nu}^{ij}(k, q)$ for LiF discussed in Ref. [14]. We begin by optimizing the geometry of the system using Quantum Espresso [33, 34] employing a 12x12x12 k-grid using the the PBE exchange-correlation functional [35] with Optimized Norm-Conserving Vanderbilt Pseudopotentials [36, 37], 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 [38] and use EPW [39] to interpolate the parameters onto all of the k-grids applied in this work.

3. Closure of All-Coupling Ansatz

While the unitary transformation defined by Eq. 4 has appealing properties, when combined with the *ab initio* Hamiltonian in Eq. 1 it does not close in general. We will instead work with a truncated version of the transformation which discards all terms greater than quadratic in h_{nq} . Given this truncation we examine the magnitudes of h_{nq} after optimization and discard solutions with large h_{nq} [40].

The transformed Hamiltonian can be expressed using the Baker–Campbell–Hausdorff (BCH) formula

$$e^{-S} O e^S = O + [O, S] + \frac{1}{2!} [[O, S], S] + \dots \quad (\text{SM.2})$$

The phonon creation/annihilation operators are readily transformed and become

$$e^{-S} b_{\nu q} e^S = b_{\nu q} - h_{\nu q}^* \sum_{ik} |i, k\rangle \langle i, k - a_q q|, \quad (\text{SM.3})$$

$$e^{-S} b_{\nu q}^\dagger e^S = b_{\nu q}^\dagger - h_{\nu q} \sum_{ik} |i, k - a_q q\rangle \langle i, k|. \quad (\text{SM.4})$$

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

$$H_e = \sum_{ik} \epsilon_{ik} |i, k\rangle \langle i, k| \quad (\text{SM.5})$$

$$e^{-S} H_e e^S = \sum_{i,k} \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\substack{q^{(1)} \dots q^{(n)} \\ \nu^{(1)} \dots \nu^{(n)}}} \left[\prod_{m=1}^n B_{\nu^{(m)} q^{(m)}} \right] C_{iq^{(1)} \dots q^{(n)}}^{(n)}(k) \left| i, k \right\rangle \left\langle i, k - \sum_{m=1}^n a_{q^{(m)}} q^{(m)} \right| \quad (\text{SM.6})$$

Here, the coefficients are given by the recursion relation

$$C_i^{(0)}(k) = \epsilon_{ik}, \quad (\text{SM.7})$$

$$C_{iq^{(1)} \dots q^{(n)}}^{(n)}(k) = C_{iq^{(1)} \dots q^{(n-1)}}^{(n-1)}(k) - C_{iq^{(1)} \dots q^{(n-1)}}^{(n-1)}(k - a_{q^{(n)}} q^{(n)}). \quad (\text{SM.8})$$

Evaluating Eq. SM.6 in it's entirety is impractical and we will use only up to second order. However, the structure of Eq. SM.8 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 [22]. As an example, the three non-vanishing coefficients for the Fröhlich model are

$$C_k^{(0)} = \frac{k^2}{2m} \quad (\text{SM.9})$$

$$C_{k,q}^{(1)} = -\frac{a_q q (a_q q - 2k)}{2m} \quad (\text{SM.10})$$

$$C_{k,q_1,q_2}^{(2)} = \frac{a_{q_1} q_1 a_{q_2} q_2}{m}. \quad (\text{SM.11})$$

It is clear that all higher order terms will vanish because $C^{(2)}$ has no k dependence, so further application of Eq. SM.8 will produce zero.

We now consider the electron-phonon interaction term

$$H_{\text{e-ph}} = \sum_{\substack{ij\nu \\ kq}} g_{\nu}^{ij}(k, q) |i, k + q\rangle \langle j, k| (b_{\nu q} + b_{\nu -q}^{\dagger}). \quad (\text{SM.12})$$

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

$$e^{-S} \tilde{\rho}_q^{ij} e^S = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\substack{k \\ q^{(1)} \dots q^{(n)} \\ \nu^{(1)} \dots \nu^{(n)}}} \left[\prod_{m=1}^n B_{\nu^{(m)} q^{(m)}} \right] D_{ij\nu q q^{(1)} \dots q^{(n)}}^{(n)}(k) \left| ik + q + \sum_{m=1}^n a_{q^{(m)}} q^{(m)} \right\rangle \left\langle ik \right|. \quad (\text{SM.13})$$

Similarly to Eq. SM.8, the coefficients are given by the recursion relation

$$D_{ij\nu q}^{(0)}(k) = g_{\nu}^{ij}(k, q), \quad (\text{SM.14})$$

$$D_{ij\nu q q^{(1)} \dots q^{(n)}}^{(n)}(k) = D_{ij\nu q q^{(1)} \dots q^{(n-1)}}^{(n-1)}(k + a_{q^{(n)}} q^{(n)}) - D_{ij\nu q q^{(1)} \dots q^{(n-1)}}^{(n-1)}(k). \quad (\text{SM.15})$$

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. SM.13. We also rearrange the first-order term so that the electron-phonon coupling elements are easily evaluated

$$e^{-S} \tilde{\rho}_q^{ij} e^S \approx \tilde{\rho}_q^{ij} + \sum_{k\nu'q'} B_{\nu'q'} g_{\nu}^{ij}(k, q) (|i, k + q\rangle \langle j, k - a_{q'} q'| - |i, k + q + a_{q'} q'\rangle \langle j, k|). \quad (\text{SM.16})$$

As in Eq. SM.8, if $g_{\nu}^{ij}(k, q)$ is a polynomial in k with degree m then all terms of order $n > m$ in Eq. SM.13 will vanish and the transformation closes. In Ref. [22] 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, as for the band term, the first- and all higher-order terms vanishes.

Overall, there are three cases where the transformation closes and thus the method is rigorously variational. These are the trivial weak-coupling limit ($g(k, q) \rightarrow 0$), the strong coupling ($a_k = 0$) limit, and, less trivially, cases which include the Fröhlich model where the electron-phonon coupling is independent of the electronic-momentum and the electronic bands are polynomials up to second-degree.

4. Numerical Evaluation and Optimization of AI-NM Ansatz

In order to evaluate Eq. 7 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

$$\eta_{k+a_q q} = \sum_r \eta_r e^{-ia_q q \cdot r} e^{-ik \cdot r} = \mathcal{F} [\eta_r e^{-ia_q q \cdot r}; r, k], \quad (\text{SM.17})$$

$$\eta_r = \mathcal{F}^{-1}[\eta_k]. \quad (\text{SM.18})$$

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 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. 6 [41]. 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. 6 which we denote as E_0 . The full gradient is then constructed via standard calculus.

Starting with the displacements

$$\frac{\partial E_0}{\partial h_{\nu p}^*} = -\frac{1}{2} h_{\nu p} \sum_{iq} \mathcal{E}_{iqp} |t_{iq}|^2 + \omega_{\nu p} h_{\nu p} \sum_{ik} |t_{ik}|^2 - \sum_{ijk} g_{\nu}^{ij}(k, p) t_{ik+p(1-a_p)}^* t_{jk} \quad (\text{SM.19})$$

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

$$\mathcal{E}_{ikp} = 4 \sum_r \epsilon_{ir} \sin^2 \left(\frac{k a_k \cdot r}{2} \right) e^{-iq \cdot r}, \quad (\text{SM.20})$$

For the CI amplitudes

$$\begin{aligned} \frac{\partial E_0}{\partial t_{ip}^*} = & \left(\epsilon_{ip} - \frac{1}{2} \sum_{\nu k} |h_{\nu k}|^2 \mathcal{E}_{ipk} \right) t_{ip} \\ & - \sum_{\substack{j\nu \\ q}} \left(g_{\nu}^{ij}(p-q, q) h_{\nu-q} t_{jp-q(1-a_q)} + \frac{1}{N} \sum_{kr} g_{\nu}^{ij}(k, p) h_{\nu-q}^* t_{ik+q} e^{i(k-p+a_q q) \cdot r} \right). \end{aligned} \quad (\text{SM.21})$$

The gradient with respect to a_k is given by

$$\begin{aligned} \frac{\partial E_0}{\partial a_p} = & - \sum_{q\nu} |t_{iq}|^2 (|h_{\nu p}|^2 + |h_{\nu-p}|^2) \sum_r \epsilon_{ir}(p \cdot r) \sin(a_p p \cdot r) e^{-iq \cdot r} \\ & + 2\Re \left(i \sum_{\substack{ij\nu \\ k}} g_{\nu}^{ij}(k, p) h_{\nu-q} t_{ik+p}^* \sum_r t_{jr}(p \cdot r) e^{-ia_p p \cdot r} e^{-ik \cdot r} \right) \\ & - 2\Re \left(i \sum_{\substack{ij\nu \\ k}} g_{\nu}^{ij}(k, -p) h_{\nu q} t_{ik-p}^* \sum_r t_{jr}(p \cdot r) e^{ia_p p \cdot r} e^{-ik \cdot r} \right). \end{aligned} \quad (\text{SM.22})$$

The application of analytic gradients allows us to efficiently run the L-BFGS-B routine; however, the optimization itself is non-trivial and presents several challenges. We found that there were multiple competing solutions, so the solver could, for example, optimize to the SC state even when there was a lower minimum available. Because of this difficulty, we ran the minimization procedures from a number of starting points including some randomly generated. Additionally, we interpolated the lowest energy solutions from smaller k-mesh sizes to larger k-meshes in order to speed up convergence. We also note a numerical error that arises in small systems because the computed values of $g_{\nu}^{ij}(k, q)$ are not precisely Hermitian. This non-physicality results in a slight disagreement between the evaluated energies and the gradient minimization because Hermiticity of the electron-phonon coupling is analytically assumed in our ansatz. Finally, the optimization via the approach outlined above with the minimizer we have utilized often terminates where notable gradient structure is still present. This may be a consequence of our interpolation procedures, but further investigation is needed to improve the optimization, ensuring the lowest lying solutions are obtained.

5. CSPT2

CSPT2 is a version of MP2 which applies coherent state displacements to the electron-phonon Hamiltonian in order to generate the reference Hamiltonian for perturbation theory [25]. Beginning with Eq. 1 we apply the coherent state transformation to displace each phonon operator by a complex-scalar displacement $\phi_{\nu\mathbf{q}}$. Unlike in previous sections, we consider both the variationally minimized non-zero $\phi_{\nu\mathbf{q}}$ and the zero-displacement, $\phi_{\nu\mathbf{q}} = 0$, solutions. This additional consideration is important because it is not possible to know which solution is the best representation of the true wavefunction *a priori*, so we use the perturbative energy to determine the most appropriate reference *a posteriori*.

Beginning with Eq. 1 and applying the coherent state transformation, we define the zeroth-order Hamiltonian as

$$\hat{H}_0 = \sum_{\mathbf{k}m} \epsilon_{m\mathbf{k}} a_{m\mathbf{k}}^\dagger a_{m\mathbf{k}} + \sum_{\mathbf{k}\mathbf{q}mn\nu} g_\nu^{nm}(\mathbf{k}, \mathbf{q}) a_{n\mathbf{k}+\mathbf{q}}^\dagger a_{m\mathbf{k}} (\phi_{\nu\mathbf{q}} + \phi_{\nu-\mathbf{q}}^*) + \sum_{\mathbf{q}\nu} \omega_{\nu\mathbf{q}} |\phi_{\nu\mathbf{q}}|^2 + \sum_{\mathbf{q}\nu} \omega_{\nu\mathbf{q}} b_{\nu\mathbf{q}}^\dagger b_{\nu\mathbf{q}} \quad (\text{SM.23})$$

The ground state of Eq. SM.23 is the same as the optimized SC state,

$$|\Psi_0^{(0)}\rangle = \sum_{\mathbf{k}m} t_{m\mathbf{k}} |m\mathbf{k}\rangle |0\rangle. \quad (\text{SM.24})$$

The excited states of Eq. SM.23 can be labeled by electronic state index α and the phonon occupation string $\{n_{\nu\mathbf{q}}\}$. The corresponding electronic states can be obtained by diagonalizing the Fock operator,

$$\hat{F} = \sum_{\mathbf{k}m} \epsilon_{m\mathbf{k}} a_{m\mathbf{k}}^\dagger a_{m\mathbf{k}} + \sum_{\mathbf{k}\mathbf{q}mn\nu} g_\nu^{nm}(\mathbf{k}, \mathbf{q}) a_{n\mathbf{k}+\mathbf{q}}^\dagger a_{m\mathbf{k}} (\phi_{\nu\mathbf{q}} + \phi_{\nu-\mathbf{q}}^*). \quad (\text{SM.25})$$

We label eigenvalues of this Fock operator by ϵ_α with the corresponding eigenvectors $\mathbf{t}_\alpha$.

In general, the exact ground state can be written using the basis of these states,

$$|\Psi_0\rangle = \sum_{\alpha, \{n_{\nu\mathbf{q}}\}} C_{\alpha, \{n_{\nu\mathbf{q}}\}} |\psi_\alpha\rangle |\{n_{\nu\mathbf{q}}\}\rangle. \quad (\text{SM.26})$$

In CSPT2, we obtain $C_{\alpha, \{n_{\nu\mathbf{q}}\}}$ by perturbative expansion. Our perturbing Hamiltonian follows

$$\hat{V} = \sum_{\mathbf{k}\mathbf{q}nm\nu} g_\nu^{nm}(\mathbf{k}, \mathbf{q}) a_{n\mathbf{k}+\mathbf{q}}^\dagger a_{m\mathbf{k}} (b_{\nu\mathbf{q}} + b_{\nu-\mathbf{q}}^\dagger) + \sum_{\mathbf{q}\nu} \omega_{\nu\mathbf{q}} (b_{\nu\mathbf{q}}^\dagger \phi_{\nu\mathbf{q}} + \phi_{\nu\mathbf{q}}^* b_{\nu\mathbf{q}}). \quad (\text{SM.27})$$

It is immediately clear that the first-order energy contribution, $E^{(1)}$, is zero. The first-order wave function is determined by (observing only one phonon states survive)

$$C_{\alpha, 1_{\nu\mathbf{q}}}^{(1)} = \frac{\langle \Psi_{\alpha, 1_{\nu\mathbf{q}}} | \hat{V} | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_\alpha^{(0)}} = \frac{\sum_{\mathbf{k}nm} g_\nu^{nm}(\mathbf{k}, \mathbf{q}) (t_{\alpha, n\mathbf{k}+\mathbf{q}})^* t_{0, m\mathbf{k}} + \delta_{\alpha, 0} \omega_{\nu\mathbf{q}} \phi_{\nu\mathbf{q}}}{\epsilon_0 - \epsilon_\alpha - \omega_{\nu\mathbf{q}}}. \quad (\text{SM.28})$$

Using this, the second-order energy contribution is

$$E^{(2)} = \sum_{\alpha\nu\mathbf{q}} \frac{|\langle \Psi_{\alpha, 1_{\nu\mathbf{q}}} | \hat{V} | \Psi_0^{(0)} \rangle|^2}{E_0^{(0)} - E_\alpha^{(0)}} = \sum_{\alpha\nu\mathbf{q}} \frac{|\sum_{\mathbf{k}nm} g_\nu^{nm}(\mathbf{k}, \mathbf{q}) (t_{\alpha, n\mathbf{k}+\mathbf{q}})^* t_{0, m\mathbf{k}} + \delta_{\alpha, 0} \omega_{\nu\mathbf{q}} \phi_{\nu\mathbf{q}}|^2}{\epsilon_0 - \epsilon_\alpha - \omega_{\nu\mathbf{q}}}. \quad (\text{SM.29})$$

To generate the energies presented in the main text the following protocol is used. First we evaluate the energy expectation value of Eq. SM.25 with both $\phi_{\nu\mathbf{q}} = 0$ and a $\phi_{\nu\mathbf{q}}$ variationally minimized from a small random displacement.

We can also compute other properties than energy order-by-order. In particular, we focus on evaluating the expectation value for

$$\rho_n(\mathbf{k}) = a_{n\mathbf{k}}^\dagger a_{n\mathbf{k}}. \quad (\text{SM.30})$$

The first order correction is

$$\langle \Psi_0^{(1)} | \rho_n(\mathbf{k}) | \Psi_0^{(0)} \rangle + h.c. = 0. \quad (\text{SM.31})$$

The second order correction is

$$(\langle \Psi_0^{(2)} | \rho_n(\mathbf{k}) | \Psi_0^{(0)} \rangle + h.c.) + \langle \Psi_0^{(1)} | \rho_n(\mathbf{q}) | \Psi_0^{(1)} \rangle. \quad (\text{SM.32})$$

We only compute the last term for computational efficiency reasons. This term is given as

$$\begin{aligned} \langle \Psi_0^{(1)} | \rho_n(\mathbf{k}) | \Psi_0^{(1)} \rangle &= \sum_{\alpha\beta\nu\mathbf{q}'\lambda\mathbf{q}''} (C_{\alpha, 1_{\nu\mathbf{q}'}}^{(1)})^* \langle \Psi_{\alpha, 1_{\nu\mathbf{q}'}} | \rho_n(\mathbf{q}) | \Psi_{\beta, 1_{\lambda\mathbf{q}''}} \rangle C_{\beta, 1_{\lambda\mathbf{q}''}}^{(1)} \\ &= \sum_{\alpha\beta\nu\mathbf{q}'} (C_{\alpha, 1_{\nu\mathbf{q}'}}^{(1)})^* (t_{\alpha, n\mathbf{k}})^* t_{\beta, n\mathbf{k}} C_{\beta, 1_{\nu\mathbf{q}'}}^{(1)}. \end{aligned} \quad (\text{SM.33})$$

Calculation of this quantity can be efficiently performed by forming the following intermediate

$$D_{n_{\mathbf{k}}, \nu_{\mathbf{q}}} = \sum_{\alpha} C_{\alpha, 1\nu_{\mathbf{q}}}^{(1)} t_{\alpha, n_{\mathbf{k}}}. \quad (\text{SM.34})$$

With this,

$$\langle \Psi_0^{(1)} | \rho_n(\mathbf{q}) | \Psi_0^{(1)} \rangle = \sum_{\nu_{\mathbf{q}}} (D_{n_{\mathbf{k}}, \nu_{\mathbf{q}}})^* D_{n_{\mathbf{k}}, \nu_{\mathbf{q}}}. \quad (\text{SM.35})$$