Maximally localized exciton Wannier functions for solids

Jonah B. Haber [1,2,*] Diana Y. Qiu [3,4] Felipe H. da Jornada,[5] and Jeffrey B. Neaton [1,2,6,†]

[1]*Department of Physics, University of California Berkeley, Berkeley, California 94720, USA*
[2]*Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[3]*Department of Mechanical Engineering and Materials Science, Yale University, New Haven, Connecticut 06520, USA*
[4]*Energy Science Institute, Yale University, New Haven, Connecticut 06520, USA*
[5]*Department of Materials Science and Engineering, Stanford University, Stanford, California 94305, USA*
[6]*Kavli Energy Nanosciences Institute, Berkeley, California 94720, USA*

We introduce a maximally localized Wannier function representation of Bloch excitons, two-particle correlated electron-hole excitations, in crystalline solids, where the excitons are maximally localized with respect to an average electron-hole coordinate in real space. As a proof-of-concept, we illustrate this representation in the case of low-energy spin-singlet and -triplet excitons in cubic lithium fluoride, computed using the *ab initio* Bethe-Salpeter equation approach. We visualize the resulting maximally localized exciton Wannier functions (MLXWFs) in real space, detail the convergence of the exciton Wannier spreads, and demonstrate how Wannier-Fourier interpolation can be leveraged to obtain exciton energies and states at arbitrary exciton crystal momenta in the Brillouin zone. We further introduce an approach to treat the long-range dipolar coupling between singlet MLXWFs and discuss it in depth. The MLXWF representation sheds light on the fundamental nature of excitons and paves the way toward Wannier-based post-processing of excitonic properties, enabling the construction of *ab initio* exciton tight-binding models, efficient interpolation of the exciton-phonon vertex, the computation of Berry curvature associated with exciton bands, and beyond.

I. INTRODUCTION

Since their introduction 25 years ago [1], maximally localized Wannier functions (MLWFs) have had a transformative impact on our ability to compute and understand one-electron observables using density functional theory (DFT). Today MLWFs serve as a compact basis for linear scaling algorithms [2], allow for the computation of Berry phase (geometric) quantities [3–5] (e.g., electronic polarization [6]), and find application in efficient and accurate interpolation of linear response quantities (e.g., electron-phonon matrix elements) [7–9], and more [10].

In the MLWF scheme, periodic Bloch states are related to localized Wannier functions through a unitary transformation, one which simultaneously preserves the canonical commutation relations while also localizing the sum of the spreads of the states to the greatest extent possible [1,10,11]. While the scheme is most often applied to one-electron Bloch states, the MLWF procedure is general and can be applied to any lattice periodic function. Indeed this framework has been used to construct localized representations of lattice vibrations [12], the electromagnetic field in photonic crystals [13] and, recently, perturbations to the electronic wave function [14].

In principle, the scheme can formally be applied to multiparticle states that are Bloch periodic in some average coordinate [15–17]. One such excitation of broad theoretical and technological interest is the exciton, a two-particle correlated electron-hole state. To date, localized representations of the exciton have only been constructed in the isolated limit [18] and in one dimension [19]. Here, we introduce a general and practical extension of the MLWF scheme for excitons in a crystalline solid which may also serve as a blueprint for other multiparticle states.

Excitons are correlated electron-hole pairs which often dominate the low-energy optical response of semiconducting and insulating materials. Understanding these composite particles plays an increasingly important role in the design and development of next-generation of optoelectronic devices, especially those based on complex materials with strong light-matter interactions. Over the past two decades, *ab initio* many-body perturbation theory within the *GW* approximation, where *G* is the one-electron Green's function and *W* the screened Coulomb interaction, and the Bethe-Salpeter equation (BSE) approach [20–24] has rapidly emerged as a powerful and robust method for predicting excitonic properties for a wide range of increasingly complex materials including low-dimensional transition metal dichalcogenides [25], lead-halide perovskites [26–28], and organic crystals [29–32], in all cases yielding results in good agreement with experiment. Given the technological relevance and increasing maturity of computational architectures and algorithms [33], revisiting the MLWF scheme in the context of excitons is timely.

---
*Present address: Department of Materials Science and Engineering, Stanford University, Stanford, California 94305, USA; Corresponding author: jbhaber@stanford.edu
†Corresponding author: jbneaton@lbl.gov

In solid-state systems with translational symmetry, the exciton wave function can be written in Bloch periodic form with respect to an average electron-hole coordinate. A canonical example is the phenomenological Mott-Wannier model [34,35] of a weakly bound exciton; here, the exciton is described by a hydrogenlike wave function with the form

$$\Psi_{n\ell m,\mathbf{Q}}(\mathbf{R}, \mathbf{r}) = e^{i\mathbf{Q}\cdot\mathbf{R}}F_{n\ell m}(\mathbf{r}), \tag{1}$$

where $\mathbf{R}$ and $\mathbf{r}$ denote the average and relative coordinate of the electron-hole pair, respectively, while $F_{n\ell m}(\mathbf{r})$ is a hydrogenic wave function with quantum numbers $n\ell m$. Notably, $\Psi_{n\ell m,\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ obeys Bloch's theorem in $\mathbf{R}$—i.e., $\Psi_{\mathbf{Q},n\ell m}(\mathbf{R} + \bar{\mathbf{R}}, \mathbf{r}) = e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}}\Psi_{\mathbf{Q},n\ell m}(\mathbf{R}, \mathbf{r})$, where $\bar{\mathbf{R}}$ is a lattice vector—a general feature rigorously true of all descriptions of an exciton in a perfect crystal, e.g., excitons classified as Frenkel [36] or charge-transfer in crystals also obey Bloch's theorem in $\mathbf{R}$.

In this work we advance a new representation of the exciton in a periodic crystal, one which is maximally localized in an average electron-hole coordinate. This new representation is a natural but as-yet unexplored extension of one-electron Wannier functions to two-particle excitations and provides a rigorous particlelike picture of the exciton complimentary to the usual wavelike picture afforded by the Bloch representation of an exciton in a periodic solid.

The maximally localized exciton Wannier functions (MLXWFs) allow for post-processing of exciton related properties in analogy to the electronic case, for instance, the *ab initio* construction of exciton tight-binding models, the efficient interpolation of exciton eigenenergies [37,38] and exciton-phonon matrix elements [39–41] throughout the Brillouin zone, the computation of Berry curvature related properties at the excitonic level [42–45], and more. Our framework also deepens our understanding of the phenomenological Mott-Wannier [34,35] and Frenkel [36] descriptions of an exciton in periodic solids. It further provides a physical picture for the splitting of the transverse and longitudinal exciton branches of the exciton band structure [46,47] in terms of long-range dipole-dipole interactions between MLXWFs.

The remainder of this paper is organized as follows. In Sec. II we introduce exciton Wannier functions. In Sec. III we review the MLWF framework and adapt this formalism for excitons. In Sec. IV we focus on matrix elements of the exciton Hamiltonian in the MLXWF basis. We show singlet excitons are coupled over large distances through a dipolar interaction, and we draw an analogy to the theory of lattice vibrations in a hetero-polar crystal. In Sec. V we detail how to effectively partition the singlet exciton Hamiltonian into a short-range part, amenable to Wannier-based post-processing, and a long-range part which can be treated analytically. In Sec. VI we detail our Wannier-Fourier interpolation scheme for obtaining exciton band structures throughout the Brillouin zone. In Sec. VII we summarize the *GW*-BSE approach and computational details of our calculation. In Sec. VIII we apply our framework to cubic lithium fluoride, LiF. We tabulate the convergence of the Wannier spreads, visualize the MLXWFs in real-space, and show the results of our Wannier-Fourier interpolation for the lowest three exciton bands. We draw comparisons for both the singlet and triplet cases. In Sec. IX

we give a discussion detailing potential future applications of MLXWFs. We close in Sec. X with a summary of our work.

II. EXCITON WANNIER FUNCTION

An exciton is a composite quasiparticle consisting of a correlated electron-hole pair. The position of the electron, $\mathbf{r}_e$, and hole, $\mathbf{r}_h$, are correlated and encoded in the exciton wave function, $\Psi_{S\mathbf{Q}}(\mathbf{r}_e, \mathbf{r}_h)$, where S and $\mathbf{Q}$ denote the exciton's principle quantum number and crystal momentum, respectively. Physically, $\Psi_{S\mathbf{Q}}(\mathbf{r}_e, \mathbf{r}_h)$ is the probability amplitude to simultaneously find an electron and hole at $\mathbf{r}_e$ and $\mathbf{r}_h$, respectively.

An increasingly standard approach for computing exciton states and properties in solids is the *ab initio GW*-BSE method, mentioned above. In this approach, $\Psi_{S\mathbf{Q}}(\mathbf{r}_e, \mathbf{r}_h)$ is expressed as a coherent sum over noninteracting electron-hole product states, namely,

$$\Psi_{S\mathbf{Q}}(\mathbf{r}_e, \mathbf{r}_h) = \sum_{cv\mathbf{k}} A^{S\mathbf{Q}}_{cv\mathbf{k}} \psi_{c\mathbf{k}}(\mathbf{r}_e) \psi^{\star}_{v\mathbf{k}-\mathbf{Q}}(\mathbf{r}_h), \tag{2}$$

where $\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r})$ denotes a single-particle Bloch state with band index n and crystal momentum $\mathbf{k}$ (typically computed from Kohn-Sham DFT, Hartree Fock, etc.), while $A^{S\mathbf{Q}}_{cv\mathbf{k}}$ is the exciton expansion coefficient with subscript c (v) indexing conduction (valence) states. The same exciton wave function can be written in Bloch periodic form when re-expressed in the average $\mathbf{R} = (\mathbf{r}_e + \mathbf{r}_h)/2$ and relative $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_h$ coordinates. Explicitly,

$$\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{Q}}}}e^{i\mathbf{Q}\cdot\mathbf{R}}F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}), \tag{3}$$

where $F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ is given by

$$F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{k}}}}\sum_{cv\mathbf{k}} A^{S\mathbf{Q}}_{cv\mathbf{k}+\mathbf{Q}/2}e^{i\mathbf{k}\cdot\mathbf{r}}$$
$$\times u_{c\mathbf{k}+\mathbf{Q}/2}(\mathbf{R} + \mathbf{r}/2)u^{\star}_{v\mathbf{k}-\mathbf{Q}/2}(\mathbf{R} - \mathbf{r}/2). \tag{4}$$

Importantly, $F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ is cell-periodic in $\mathbf{R}$ but not in $\mathbf{r}$. [Note Eq. (1) is a specialized case of this general form.]

For present purposes, we assume that the exciton states are computed on regular $\mathbf{Q}$- and $\mathbf{k}$ meshes with $N_{\mathbf{Q}}$ and $N_{\mathbf{k}}$ points, respectively. Our exciton wave functions are normalized such that

$$\int |\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})|^2 d\mathbf{r}d\mathbf{R} = 1, \tag{5}$$

where the integrals in $\mathbf{R}$ and $\mathbf{r}$ are to be performed over supercells with volume $N_{\mathbf{Q}}V_{uc}$ and $N_{\mathbf{k}}V_{uc}$, respectively, with V_{uc} denoting the volume of the unit cell. From here on all integrals in $\mathbf{R}$ and $\mathbf{r}$ are understood to be taken over their respective supercell volumes unless otherwise stated. In Appendix A, we provide additional details on this change of coordinates and our normalization conventions. In the same Appendix we also discuss a generalization where the exciton is expressed in a weighted-average coordinate $\mathbf{R} = \alpha\mathbf{r}_e + \beta\mathbf{r}_h$, with $\alpha, \beta \geqslant 0$ and $\alpha + \beta = 1$. For the sake of conceptual simplicity, we will continue to work with the specialized case $\alpha = \beta = 1/2$ throughout the main text.

In this work, we define the exciton Wannier function as a Wannier transform of the Bloch exciton in the average electron-hole coordinate, explicitly,

$$W_{S\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{Q}}}} \sum_{\mathbf{Q}} e^{-i\mathbf{Q}\cdot\bar{\mathbf{R}}} \Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}), \quad (6)$$

where $\bar{\mathbf{R}}$ is a lattice vector. The exciton Wannier functions span the same functional space as the original excitonic states and by construction are orthonormal in indices $(S, \bar{\mathbf{R}})$ so that $\langle W_{S\bar{\mathbf{R}}}|W_{S'\bar{\mathbf{R}}'}\rangle = \delta_{SS'}\delta_{\bar{\mathbf{R}}\bar{\mathbf{R}}'}$. When translated by a lattice vector, they transform as $W_{S\bar{\mathbf{R}}}(\mathbf{R} - \bar{\mathbf{R}}', \mathbf{r}) = W_{S\bar{\mathbf{R}}+\bar{\mathbf{R}}'}(\mathbf{R}, \mathbf{r})$, a result which immediately follows from Bloch's theorem.

Further, the exciton Wannier functions are localized in the coordinate $\mathbf{R}$ and reside in the cell labeled by $\bar{\mathbf{R}}$. This is readily seen in the limit where the cell-periodic function is only weakly dependent on $\mathbf{Q}$ so that $F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}) \approx F_S(\mathbf{R}, \mathbf{r})$. Then, $W_{S\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r}) \approx F_S(\mathbf{R}, \mathbf{r})\sum_{\mathbf{Q}} e^{i\mathbf{Q}\cdot(\mathbf{R}-\bar{\mathbf{R}})}$. Since the only length scales appearing in the sum are the lattice parameters, $W_{S\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r})$ should decay rapidly in $\mathbf{R} - \bar{\mathbf{R}}$ beyond a few multiples of $\bar{\mathbf{R}}$. By contrast, the spread in the relative coordinate, $\mathbf{r}$, is related to the exciton radius, which, for weakly bound excitons, can be on the order of many unit cells. We expect there are many scenarios where the spread in the average coordinate is smaller than the spread in the relative coordinate.

III. MAXIMALLY LOCALIZED EXCITON WANNIER FUNCTIONS

In the MLWF scheme, extended one-electron Bloch states, $\psi_{n\mathbf{k}}(\mathbf{r}_e) = \langle \mathbf{r}_e|n\mathbf{k}\rangle$, are related to localized Wannier functions with orbital index m, $w_{m\bar{\mathbf{R}}}(\mathbf{r}_e) = \langle \mathbf{r}_e|m\bar{\mathbf{R}}\rangle$, through a unitary transformation

$$w_{m\bar{\mathbf{R}}}(\mathbf{r}_e) = \frac{1}{\sqrt{N_{\mathbf{k}}}} \sum_{n\mathbf{k}}^{N_{\mathcal{W}}} e^{-i\mathbf{k}\cdot\bar{\mathbf{R}}} U_{nm}(\mathbf{k})\psi_{n\mathbf{k}}(\mathbf{r}_e), \quad (7)$$

where $U_{nm}(\mathbf{k})$ is a unitary matrix which mixes some subset of states, $n \in \mathcal{W}$, at a given $\mathbf{k}$ point.

In developing this framework, Marzari and Vanderbilt [1] took advantage of the extra gauge freedom in $U_{nm}(\mathbf{k})$ to localize the sum of the spread of the Wannier functions to the greatest extent possible. For MLWFs, the sum of the spreads is defined as

$$\Omega^{\text{el}}[U] = \sum_m^{N_{\mathcal{W}}} \left[\langle m\bar{\mathbf{0}}|r_e^2|m\bar{\mathbf{0}}\rangle - \langle m\bar{\mathbf{0}}|\mathbf{r}_e|m\bar{\mathbf{0}}\rangle^2\right], \quad (8)$$

where $\mathbf{r}_e$ is understood as the position operator and the notation, $\Omega^{\text{el}}[U]$, indicates that the spread is a functional of the gauge, U.

Nowadays the minimization procedure is often performed using Wannier90, an open source, post-processing software for constructing MLWFs, compatible with many DFT codes [48,49]. Wannier90 minimizes Eq. (8) in reciprocal space and requires as a key input the overlaps of Bloch periodic states at neighboring $\mathbf{k}$ points, $M_{nm}^{(\mathbf{k},\mathbf{b})} = \langle u_{n\mathbf{k}}|u_{m\mathbf{k}+\mathbf{b}}\rangle_{\text{uc}}$, where the subscript "uc" indicates that the overlaps are to be computed in the unit cell. As emphasized by the developers of Wannier90, these inputs are entirely agnostic to the underlying electronic structure theory calculation [48,49].

TABLE I. Notation for, and analogy between, electron and exciton Wannier functions and related parameters. As defined in the text, $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_h$ and $\mathbf{R} = \alpha\mathbf{r}_e + \beta\mathbf{r}_h$, with $\alpha = \beta = 1/2$ being used throughout this work (see Appendix A).

	Electron	Exciton
Bloch state	$\psi_{n\mathbf{k}}(\mathbf{r}_e)$	$\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$
Wannier coordinate	$\mathbf{r}_e$	$\mathbf{R}$
Conjugate momentum	$\mathbf{k}$	$\mathbf{Q}$
Conjugate position	$\bar{\mathbf{R}}$	$\bar{\mathbf{R}}$
Wannier function	$w_{m\bar{\mathbf{R}}}(\mathbf{r}_e)$	$W_{M\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r})$
Rotation matrix	$U_{nm}(\mathbf{k})$	$U_{SM}(\mathbf{Q})$

A more subtle point is that MLWF procedure itself is agnostic to the type of quasiparticle excitation which one wishes to localize so long as the excitation can be written in Bloch form. Said another way, the procedure can be used to find localized representations of any lattice periodic excitation. Accordingly, in analogy with Eq. (7), exciton states, $\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}) = \langle \mathbf{R}\mathbf{r}|S\mathbf{Q}\rangle$, are also related to localized exciton Wannier functions, $W_{M\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r}) = \langle \mathbf{R}\mathbf{r}|M\bar{\mathbf{R}}\rangle$, through a unitary transformation, namely,

$$W_{M\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{Q}}}} \sum_{S\mathbf{Q}} e^{-i\mathbf{Q}\cdot\bar{\mathbf{R}}} U_{SM}(\mathbf{Q})\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}), \quad (9)$$

where M denotes the principal quantum number of the exciton Wannier function.

In further analogy with the one-electron case, we can leverage the gauge freedom and choose $U_{MS}(\mathbf{Q})$ to minimize the sum of spread of the exciton Wannier functions defined as

$$\Omega^{\text{Xct}}[U] = \sum_M^{N_{\mathcal{W}}} \left[\langle M\bar{\mathbf{0}}|R^2|M\bar{\mathbf{0}}\rangle - \langle M\bar{\mathbf{0}}|\mathbf{R}|M\bar{\mathbf{0}}\rangle^2\right], \quad (10)$$

with

$$\langle M\bar{\mathbf{0}}|R^2|M\bar{\mathbf{0}}\rangle = \int |W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r})|^2 R^2 d\mathbf{r}\,d\mathbf{R}. \quad (11)$$

As emphasized previously, the spread here is with respect to the average, and not the relative, coordinate.

We use the Wannier90 package to minimize Eq. (10). In practice this is done by passing overlap matrices, $M_{SS'}(\mathbf{Q}, \mathbf{B}) = \langle F_{S\mathbf{Q}}|F_{S'\mathbf{Q}+\mathbf{B}}\rangle$, to Wannier90. We compute these overlaps in the electron-hole basis as follows:

$$M_{SS'}(\mathbf{Q}, \mathbf{B}) = \sum_{cc'vv'\mathbf{k}} \left[A_{cv\mathbf{k}}^{S\mathbf{Q}}\right]^{\star} A_{c'v'\mathbf{k}+\mathbf{B}/2}^{S'\mathbf{Q}+\mathbf{B}}$$
$$\times \langle u_{c\mathbf{k}}|u_{c'\mathbf{k}+\mathbf{B}/2}\rangle_{\text{uc}} \langle u_{v'\mathbf{k}-\mathbf{Q}-\mathbf{B}/2}|u_{v\mathbf{k}-\mathbf{Q}}\rangle_{\text{uc}}. \quad (12)$$

Equation (12) constitutes a pivotal result of this work. Importantly, it allows us to compute overlaps between the cell-periodic part of exciton wave functions with quantities readily obtained from standard electronic software codes. A generalization of Eq. (12) for arbitrary $\mathbf{R} = \alpha\mathbf{r}_e + \beta\mathbf{r}_h$ is derived in Appendix B, see Eq. (B3) specifically.

In Table I, we distinguish the notation used for, and draw an analogy between, electron and exciton Wannier functions.

For use in the next section, it is convenient to note that Eq. (9) can be inverted to give

$$\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{Q}}}} \sum_{M\bar{\mathbf{R}}} e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}} U_{MS}^{\dagger}(\mathbf{Q}) W_{M\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r}). \quad (13)$$

Equation (13) can be derived from Eq. (9), with the use of the identity $\sum_{\bar{\mathbf{R}}} e^{i(\mathbf{Q}-\mathbf{Q}')\cdot\bar{\mathbf{R}}} = N_{\mathbf{Q}}\delta_{\mathbf{Q}\mathbf{Q}'}$.

IV. ANALYTIC PROPERTIES OF THE EXCITON HAMILTONIAN

An important property of one-electron MLWFs is that they form a compact basis for expanding the local one-electron Hamiltonian. In this case locality of the MLWFs and one-electron Hamiltonian guarantees that matrix elements in the MLWF basis—i.e., $\langle m\bar{\mathbf{0}}|H^{\text{el}}|n\bar{\mathbf{R}}\rangle$—decay rapidly with increasing $\bar{\mathbf{R}}$. The same is not immediately apparent for the exciton Hamiltonian which contains nonlocal interactions that couple the electron and hole degrees of freedom.

In the absence of spin-orbit coupling, we can partition the exciton Hamiltonian into a spin-singlet and spin-triplet sector. In this total spin basis, the exciton Hamiltonian reads

$$H^{\text{Xct}} = T - K^{\text{D}} + 2\delta_S K^{\text{X}}, \quad (14)$$

where T is a kinetic term related to the propagation of a free electron-hole pair while $-K^{\text{D}}$ is an attractive screened direct interaction and K^{X} a repulsive bare exchange interaction coupling the electron and hole [22]. δ_S is 1 for singlets and 0 for triplet excitons.

Below we show that the direct interaction is short-range so that triplet MLXWFs at $\bar{\mathbf{0}}$ and $\bar{\mathbf{R}}$ couple only when the spatial part of their MLXWF's overlap. By contrast, the exchange term is long-range and singlet MLXWFs couple even when there is minimal real space overlap. For large $\bar{\mathbf{R}}$, we perform a multipole expansion in this exchange term and show that at leading order MLXWFs are coupled through a dipole-dipole interaction which decays as $1/|\bar{\mathbf{R}}|^3$. In reciprocal space ($\mathbf{Q}$ space), this long-range coupling is nonanalytic in the $\mathbf{Q} \to 0$ limit and can drive a splitting of the longitudinal and transverse branches of the exciton band structure.

A. Exciton Hamiltonian in the MLXWF basis

The exciton Hamiltonian is invariant under translations by a lattice vector, $\bar{\mathbf{R}}$, explicitly

$$H^{\text{Xct}}(\mathbf{R}, \mathbf{r}) = H^{\text{Xct}}(\mathbf{R} + \bar{\mathbf{R}}, \mathbf{r}), \quad (15)$$

so that matrix elements of H^{Xct} in the Wannier basis obey the following identity:

$$\langle M\bar{\mathbf{R}}|H^{\text{Xct}}|N\bar{\mathbf{R}}'\rangle = \langle M\bar{\mathbf{0}}|H^{\text{Xct}}|N\bar{\mathbf{R}}' - \bar{\mathbf{R}}\rangle. \quad (16)$$

It follows that the matrix element of the exciton Hamiltonian between any two MLXWFs can be constructed from knowledge of

$$H_{MN}^{\text{Xct}}(\bar{\mathbf{R}}) \equiv \langle M\bar{\mathbf{0}}|H^{\text{Xct}}|N\bar{\mathbf{R}}\rangle. \quad (17)$$

We can relate $H_{MN}^{\text{Xct}}(\bar{\mathbf{R}})$ to the exciton Hamiltonian in the Bloch basis

$$H_{SS'}^{\text{Xct}}(\mathbf{Q}) \equiv \langle S\mathbf{Q}|H^{\text{Xct}}|S'\mathbf{Q}\rangle. \quad (18)$$

Going forward, we use subscripts "M, N" to label matrix elements in the MLXWF basis and "S, S'" to label matrix elements in the Bloch basis. In cases where orbital or band indices are omitted the result is independent of the basis set.

Using Eqs. (9) and (13) we find

$$H_{SS'}^{\text{Xct}}(\mathbf{Q}) = \sum_{\bar{\mathbf{R}}} e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}} U_{SM}(\mathbf{Q}) H_{MN}^{\text{Xct}}(\bar{\mathbf{R}}) U_{NS'}^{\dagger}(\mathbf{Q}) \quad (19)$$

and

$$H_{MN}^{\text{Xct}}(\bar{\mathbf{R}}) = \sum_{\mathbf{Q}} e^{-i\mathbf{Q}\cdot\bar{\mathbf{R}}} U_{MS}^{\dagger}(\mathbf{Q}) H_{SS'}^{\text{Xct}}(\mathbf{Q}) U_{S'N}(\mathbf{Q}), \quad (20)$$

where sums over repeated band indices are implied. By construction H^{Xct} is diagonal in the Bloch basis. Its elements are the exciton eigenenergies, $E_{S\mathbf{Q}}$.

Like H^{Xct}, the direct, K^{D}, and exchange, K^{X}, kernels are translationally invariant so it is sufficient to consider the matrix elements $K_{MN}^{\text{D}}(\bar{\mathbf{R}}) = \langle M\bar{\mathbf{0}}|K^{\text{D}}|N\bar{\mathbf{R}}\rangle$ and $K_{MN}^{\text{X}}(\bar{\mathbf{R}}) = \langle M\bar{\mathbf{0}}|K^{\text{X}}|N\bar{\mathbf{R}}\rangle$.

In the MLXWF basis, the commonly used static approximation [24] to the direct interaction kernel reads

$$K_{MN}^{\text{D}}(\bar{\mathbf{R}}) = \int W_{M\bar{\mathbf{0}}}^{\star}(\mathbf{R}, \mathbf{r})\varepsilon^{-1}(\mathbf{R}, \mathbf{r})\frac{e^2}{r}W_{N\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r})d\mathbf{R}d\mathbf{r}, \quad (21)$$

where e is the electronic charge and $\varepsilon^{-1}(\mathbf{R}, \mathbf{r})$ is the inverse dielectric function expressed in terms of the average, $\mathbf{R}$, and relative, $\mathbf{r}$, coordinates (see Appendix C 1 for derivation). Notably, the direct term involves interexciton overlaps of the electron-hole coordinates and only couples $W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r})$ to $W_{N\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r})$ when there is nonzero overlap in both $\mathbf{R}$ and $\mathbf{r}$. It immediately follows that $K_{MN}^{\text{D}}(\bar{\mathbf{R}})$ decays exponentially in $\bar{\mathbf{R}}$ when the MLXWFs are exponentially localized in $\mathbf{R}$.

Meanwhile the exchange interaction reads

$$K_{MN}^{\text{X}}(\bar{\mathbf{R}}) = \int W_{M\bar{\mathbf{0}}}^{\star}(\mathbf{R}, \mathbf{0})\frac{e^2}{|\mathbf{R} - \mathbf{R}'|}W_{N\bar{\mathbf{R}}}(\mathbf{R}', \mathbf{0})d\mathbf{R}d\mathbf{R}', \quad (22)$$

(see Appendix C 1 for derivation). The exchange interaction involves intraexciton overlaps of electron-hole pairs—i.e., the electron and hole within a single exciton must overlap which is why $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_h = \mathbf{0}$. Such overlaps are independent of distance between exciton Wannier functions and for this reason the exchange term couples MLXWFs which reside in cells $\bar{\mathbf{0}}$ and $\bar{\mathbf{R}}$ through the Coulomb interaction even when these MLXWFs are exponentially localized. We expect the exchange interaction to decay slowly with $\bar{\mathbf{R}}$.

To see how $K_{MN}^{\text{X}}(\bar{\mathbf{R}})$ decays at leading order in $\bar{\mathbf{R}}$ we perform a multipole expansion in the Coulomb interaction. This is justified in the limit where $\bar{\mathbf{R}}$ is much larger than the spatial extent of the MLXWF in the $\mathbf{R}$ coordinate. Retaining only the lowest-order dipole-dipole term, we arrive at the dipolar exchange kernel

$$K_{MN}^{\text{X,Dip}}(\bar{\mathbf{R}}) = e^2\left[\frac{\mathbf{P}_M^{\star}\cdot\mathbf{P}_N}{|\bar{\mathbf{R}}|^3} - 3\frac{(\mathbf{P}_M^{\star}\cdot\bar{\mathbf{R}})(\mathbf{P}_N\cdot\bar{\mathbf{R}})}{|\bar{\mathbf{R}}|^5}\right], \quad (23)$$

where $\mathbf{P}_M$ denotes the dipole moment associated with the exciton Wannier function M, namely,

$$\mathbf{P}_M = \int \mathbf{R} W_{M0}(\mathbf{R}, \mathbf{r} = 0) d\mathbf{R} \qquad (24)$$

(see Appendix C 2 a for derivation). We see that for $\mathbf{P}_M \neq \mathbf{0}$, $K_{MN}^{\mathrm{X}}(\bar{\mathbf{R}})$ decays like $1/|\bar{\mathbf{R}}|^3$ in the large $\bar{\mathbf{R}}$ limit.

For later use it is convenient to introduce the exciton dipole moment in the Bloch basis, namely,

$$\mathbf{P}_S = \int \mathbf{R} \Psi_{S0}(\mathbf{R}, \mathbf{r} = 0) d\mathbf{R}. \qquad (25)$$

We emphasize that $\mathbf{P}_S$ is, up to a multiplicative constant, equivalent to the usual exciton transition dipole matrix element used for studying optical processes. The connection is made explicitly in Eq. (62). The $\mathbf{r} = 0$ argument, often not explicitly expressed, ensures that $\mathbf{r}_e = \mathbf{r}_h$: the photoexcited state initially creates an electron and hole at the same position so as not to violate causality.

The exciton dipole matrix elements in the Bloch and Wannier basis are related through the following expressions:

$$\mathbf{P}_M = \frac{1}{\sqrt{N_{\mathbf{Q}}}} \sum_S U_{SM}(\mathbf{0}) \mathbf{P}_S,$$

$$\mathbf{P}_S = \sqrt{N_{\mathbf{Q}}} \sum_M U_{MS}^{\dagger}(\mathbf{0}) \mathbf{P}_M, \qquad (26)$$

provided $U_{SM}(\mathbf{Q})$ is analytic in the $\mathbf{Q} \to 0$ limit (see Appendix C 3 for derivation).

B. Dipolar coupling and nonanalyticity

In reciprocal space, the dipolar interaction can lead to a nonanalyticity in $\mathbf{Q}$ which in turn has important implications for the construction of exponentially localized MLXWFs. To see how this comes about, we Fourier transform $K_{MN}^{\mathrm{X,Dip}}(\bar{\mathbf{R}})$ to find

$$K_{MN}^{\mathrm{X,Dip}}(\mathbf{Q}) = \sum_{\bar{\mathbf{R}}} e^{i\mathbf{Q} \cdot \bar{\mathbf{R}}} K_{MN}^{\mathrm{X,Dip}}(\bar{\mathbf{R}})$$

$$= \frac{4\pi e^2}{V_{\mathrm{uc}}} \sum_{\mathbf{G}} \frac{[\mathbf{P}_M^{\star} \cdot (\mathbf{Q} + \mathbf{G})][\mathbf{P}_N \cdot (\mathbf{Q} + \mathbf{G})]}{|\mathbf{Q} + \mathbf{G}|^2}, \qquad (27)$$

where $\mathbf{G}$ denotes a reciprocal lattice vector, and it is understood that $\mathbf{Q}$ is restricted to the first Brillouin zone (see Appendix C 2 b for derivation). Equation (27) formally diverges for all $\mathbf{Q}$, which can be seen by noting that the summand remains finite for large $\mathbf{G}$. This divergence is a well known consequence of not removing the self-interaction term, $K_{MN}^{\mathrm{X,Dip}}(\bar{\mathbf{R}} = \mathbf{0})$, in Eq. (23), before Fourier transforming [50]. In Sec. V we will regulate this divergence. Presently we are interested in the nonanalytic structure of $K_{MN}^{\mathrm{X,Dip}}(\mathbf{Q})$ near the origin, which can be studied at this level despite the divergence. To isolate the small $\mathbf{Q}$ behavior, we define the nonanalytic kernel as

$$K_{MN}^{\mathrm{NA}}(\mathbf{Q}) \equiv \lim_{\mathbf{G}=0, \mathbf{Q} \to 0} K_{MN}^{\mathrm{X,Dip}}(\mathbf{Q}), \qquad (28)$$

where the notation indicates that $\mathbf{G}$ should be set to zero prior to taking the limit. We then find

$$K_{MN}^{\mathrm{NA}}(\mathbf{Q}) = \frac{4\pi e^2}{V_{\mathrm{uc}}} \frac{(\mathbf{P}_M^{\star} \cdot \mathbf{Q})(\mathbf{P}_N \cdot \mathbf{Q})}{|\mathbf{Q}|^2}. \qquad (29)$$

For general $\mathbf{P}_M$, $K_{MN}^{\mathrm{NA}}(\mathbf{Q})$ will approach different values as $\mathbf{Q} \to 0$ depending on the path taken in $\mathbf{Q}$ space, which is why we refer to this as the nonanalytic kernel. The nonanalyticity persists in Bloch basis where

$$K_{SS'}^{\mathrm{X,Dip}}(\mathbf{Q}) = \frac{4\pi e^2}{V_{\mathrm{uc}}} \sum_{MN} U_{SM}(\mathbf{Q}) U_{NS'}^{\dagger}(\mathbf{Q})$$

$$\times \sum_{\mathbf{G}} \frac{[\mathbf{P}_M^{\star} \cdot (\mathbf{Q} + \mathbf{G})][\mathbf{P}_N \cdot (\mathbf{Q} + \mathbf{G})]}{|\mathbf{Q} + \mathbf{G}|^2} \qquad (30)$$

and

$$K_{SS'}^{\mathrm{NA}}(\mathbf{Q}) = \frac{4\pi e^2}{V_{\mathrm{uc}} N_{\mathbf{Q}}} \frac{(\mathbf{P}_S^{\star} \cdot \mathbf{Q})(\mathbf{P}_{S'} \cdot \mathbf{Q})}{|\mathbf{Q}|^2}. \qquad (31)$$

In deriving Eq. (31) we have made use of Eq. (26) to write this expression in terms of $\mathbf{P}_S$. This nonanalyticity is well known [22,46,47,51] and was recently revisited in the context of first-principles BSE calculations of exciton dispersion [37,38], where it was calculated by directly taking the $\mathbf{Q} \to \mathbf{0}$ limit of $K_{SS'}^{\mathrm{X}}(\mathbf{Q})$. Our present derivation reinterprets this nonanalyticity as stemming from dipolar coupling between MLXWFs.

The nonanalytic kernel can lift the degeneracy between longitudinal (L) and transverse (T) exciton states, a phenomena known as LT splitting [22,46,47,51]. Further, in low-symmetry systems, as a direct result of this nonanalyticity, singlet exciton eigenvalues can exhibit rapid angular variation about $\mathbf{Q} = 0$, and may appear discontinuous when plotted along certain directions in reciprocal space.

Finally, we note that the multipole expansion presented in this section is especially applicable to cases where MLXWFs are centered at the origin of their home cell. In cases with multiple MLXWFs centered at different places in the home cell, it may be more appropriate to expand about the distances between Wannier centers, i.e., $\bar{\mathbf{d}} = \bar{\mathbf{R}} + \boldsymbol{\tau}_N - \boldsymbol{\tau}_M$, where $\boldsymbol{\tau}_N$ denotes the position of the Wannier center of the Nth MLXWF in the home cell. The results given in the section are easily generalized in this case but left to future work.

C. Analogy with phonons

The notation, and relation between, $K^{\mathrm{X,Dip}}$ and K^{NA} resembles that found in the dynamical theory of lattice vibrations where infrared (IR)-active phonons lead to similar dipolar coupling, LT splittings, and, in low-symmetry systems, a $\mathbf{Q}$-space nonanalyticity [50,52]. We summarize the situation for phonons below and then draw an analogy to excitons.

In many materials, polar ionic displacement patterns can induce an electric dipole moment. For long wavelength, IR-active modes, the phonon eigenvector is such that these dipole moments add constructively and we can associate a net microscopic dipole with each unit cell. Coupling between these dipoles drives a splitting of the longitudinal and transverse optical phonon branches and in low-symmetry systems can give rise to a nonanalyticity in the dynamical matrix [12,50,52].

With the aide of Eq. (13) we can arrive at a similar physical interpretation for excitons. Notably, Eq. (13) allows us to express the exciton wave function as a sum over MLXWFs. Each MLXWF resides in a cell labeled by $\bar{\mathbf{R}}$, and carries a dipole moment $\mathbf{P}_M$. For bright excitonic states ($\mathbf{P}_S \neq \mathbf{0}$), the Wannier rotation matrices are such that these dipole moments add constructively and we can associate a net microscopic dipole with each unit cell. Long-range interactions between these dipoles drive a splitting of the longitudinal and transverse exciton branches and in low-symmetry systems can give rise to a nonanalyticity in the exciton Hamiltonian.

More generally, we note that dipolar interactions, LT-splitting, and nonanalyticity are phenomena associated with polar excitations, IR active phonons, or bright excitons.

V. PARTITIONING THE EXCITON HAMILTONIAN IN THE WANNIERIZATION SUBSPACE

The long-range dipolar coupling in the singlet exciton Hamiltonian is critical to account for in post-processing applications like Wannier-Fourier interpolation, where $H_{MN}^{\mathrm{Xct}}(\bar{\mathbf{R}})$ must decay rapidly with increasing $\bar{\mathbf{R}}$ to obtain accurate interpolations. These long-range couplings constitute a nontrivial difference between MLXWFs and their one-electron counterparts.

A strategy for handling these interactions is to first remove the long-range coupling to arrive at a short-range (SR) Hamiltonian which we expect to be more suitable for use in Wannier-based post-processing applications.

Accordingly, we partition the exciton Hamiltonian as follows:

$$H^{\mathrm{Xct}}(\mathbf{Q}) = H^{\mathrm{SR}}(\mathbf{Q}) + 2\delta_S K^{\mathrm{LR}}(\mathbf{Q}), \tag{32}$$

where the long-range kernel is defined as

$$K^{\mathrm{LR}}(\mathbf{Q}) = K^{\mathrm{X,Dip}}(\mathbf{Q}) - \bar{K}^{\mathrm{X,Dip}}(\mathbf{0}) \tag{33}$$

and where expressions for $K^{\mathrm{X,Dip}}(\mathbf{Q})$ are given in the MLXWF and Bloch bases in Eqs. (27) and (30), respectively. The bar on $\bar{K}^{\mathrm{X,Dip}}(\mathbf{0})$ indicates that the $\mathbf{G} = \mathbf{0}$ component should be omitted when computing this term. This subtraction regulates the divergence in $K^{\mathrm{X,Dip}}(\mathbf{Q})$ so that $K^{\mathrm{LR}}(\mathbf{Q} \neq \mathbf{0})$ is formally convergent. We have dropped orbital/band indices in these expressions to indicate that the partitioning can be made in any basis set. However, for the sake of concreteness it is convenient to specialize to the MLXWF basis. Then Eq. (33) reads

$$K_{MN}^{\mathrm{LR}}(\mathbf{Q}) = \frac{4\pi e^2}{V_{\mathrm{uc}}} \sum_{\mathbf{G}} \frac{[\mathbf{P}_M^\star \cdot (\mathbf{Q} + \mathbf{G})][\mathbf{P}_N \cdot (\mathbf{Q} + \mathbf{G})]}{|\mathbf{Q} + \mathbf{G}|^2} - \frac{4\pi e^2}{V_{\mathrm{uc}}} \sum_{\mathbf{G} \neq \mathbf{0}} \frac{[\mathbf{P}_M^\star \cdot \mathbf{G}][\mathbf{P}_N \cdot \mathbf{G}]}{|\mathbf{G}|^2}. \tag{34}$$

As written, $K_{MN}^{\mathrm{LR}}(\mathbf{Q})$ is well defined for all $\mathbf{Q} \neq \mathbf{0}$ and approaches the nonanalytic kernel in the limit where $\mathbf{Q} \to \mathbf{0}$. Explicitly,

$$\lim_{\mathbf{Q} \to \mathbf{0}} K_{MN}^{\mathrm{LR}}(\mathbf{Q}) = K_{MN}^{\mathrm{NA}}(\mathbf{Q}). \tag{35}$$

This limit should be compared to that given in Eq. (28). Note that here there is no need to set $\mathbf{G} = \mathbf{0}$ before taking the limit.

At $\mathbf{Q} = \mathbf{0}$ Eq. (34) evaluates to

$$K_{MN}^{\mathrm{LR}}(\mathbf{0}) = \frac{4\pi e^2}{V_{\mathrm{uc}}} \frac{[\mathbf{P}_M^\star \cdot \mathbf{G}][\mathbf{P}_N \cdot \mathbf{G}]}{|\mathbf{G}|^2}\bigg|_{\mathbf{G}=\mathbf{0}}, \tag{36}$$

which is indeterminate. In the following section we show that a similar indeterminate term appears in $H_{MN}^{\mathrm{Xct}}(\mathbf{0})$ which to lowest order cancels with $K_{MN}^{\mathrm{LR}}(\mathbf{0})$. The cancellation is a direct consequence of our partitioning and definition of $K_{MN}^{\mathrm{LR}}(\mathbf{Q})$. The net result is that at lowest order $H_{MN}^{\mathrm{SR}}(\mathbf{0})$ is well defined despite the indeterminate contributions from both $H_{MN}^{\mathrm{Xct}}(\mathbf{0})$ and $K_{MN}^{\mathrm{LR}}(\mathbf{0})$.

A. Partitioning for Q = 0

The indeterminate behavior of $H^{\mathrm{Xct}}(\mathbf{0})$ originates from the exchange kernel. To see this in detail we Fourier transform the exchange interaction in Eq. (22) and set $\mathbf{Q} = \mathbf{0}$ to find

$$K_{MN}^{\mathrm{X}}(\mathbf{0}) = \sum_{\mathbf{G}} \langle M\bar{\mathbf{0}}|e^{i\mathbf{G}\cdot\mathbf{R}}|0\rangle \, v(\mathbf{G}) \, \langle 0|e^{-i\mathbf{G}\cdot\mathbf{R}'}|N\bar{\mathbf{0}}\rangle, \tag{37}$$

where

$$v(\mathbf{G}) = \frac{4\pi e^2}{V_{\mathrm{uc}}} \frac{1}{|\mathbf{G}|^2} \tag{38}$$

and

$$\langle 0|e^{-i\mathbf{G}\cdot\mathbf{R}}|N\bar{\mathbf{0}}\rangle = \int e^{-i\mathbf{G}\cdot\mathbf{R}} W_{N\bar{\mathbf{0}}}(\mathbf{R},\mathbf{0}) d\mathbf{R}. \tag{39}$$

Now, we isolate the head ($\mathbf{G} = \mathbf{0}$) contribution to the sum in Eq. (37) and expand the complex exponential to lowest order to find

$$K_{MN;\mathbf{G}=\mathbf{0}}^{\mathrm{X}}(\mathbf{0}) = \frac{4\pi e^2}{V_{\mathrm{uc}}} \frac{(\mathbf{P}_M^\star \cdot \mathbf{G})(\mathbf{P}_N \cdot \mathbf{G})}{|\mathbf{G}|^2}\bigg|_{\mathbf{G}=\mathbf{0}} + \cdots, \tag{40}$$

where we have used $\langle 0|e^{-i\mathbf{G}\cdot\mathbf{R}}|N\bar{\mathbf{0}}\rangle = -i\mathbf{G}\cdot\mathbf{P}_N + \cdots$. This limiting behavior should be compared with $K_{MN}^{\mathrm{LR}}(\mathbf{0})$, given in Eq. (36). We see that at $\mathbf{Q} = \mathbf{0}$ the long-range kernel exactly equals the lowest order term in the expansion of the head of the exchange kernel. For singlets, while both $H_{MN}^{\mathrm{Xct}}(\mathbf{0})$ and $K_{MN}^{\mathrm{LR}}(\mathbf{0})$ are indeterminate, at lowest order, $H_{MN}^{\mathrm{SR}}(\mathbf{0}) = H_{MN}^{\mathrm{Xct}}(\mathbf{0}) - 2\delta_S K_{MN}^{\mathrm{LR}}(\mathbf{0})$ is well defined; the indeterminate contributions cancel. Indeed it is this cancellation which motivated the definition of $K_{MN}^{\mathrm{LR}}(\mathbf{Q})$ given in Eq. (33).

When higher order terms in Eq. (40) are negligible, subtracting $2\delta_S K_{MN}^{\mathrm{LR}}(\mathbf{0})$ from $H_{MN}^{\mathrm{Xct}}(\mathbf{0})$ is equivalent to setting the head of the Coulomb interaction in the exchange term to zero. Explicitly,

$$H_{MN}^{\mathrm{SR}}(\mathbf{0}) = T_{MN}(\mathbf{0}) - K_{MN}^{\mathrm{D}}(\mathbf{0}) + 2\delta_S \bar{K}_{MN}^{\mathrm{X}}(\mathbf{0}), \tag{41}$$

where $\bar{K}_{MN}^{\mathrm{X}}(\mathbf{0})$ is equivalent to $K_{MN}^{\mathrm{X}}(\mathbf{0})$ but with the $v(\mathbf{G})$ replaced by the modified Coulomb interaction

$$\bar{v}(\mathbf{G}) = \frac{4\pi e^2}{V_{\mathrm{uc}}} \frac{1}{|\mathbf{G}|^2}(1 - \delta_{\mathbf{G0}}). \tag{42}$$

While we have only shown the cancellation between $K_{MN;\mathbf{G}=\mathbf{0}}^{\mathrm{X}}(\mathbf{0})$ and $K_{MN}^{\mathrm{LR}}(\mathbf{0})$ to first order, it is easy to see that including higher order long-range couplings in $K_{MN}^{\mathrm{LR}}(\mathbf{Q})$, e.g., dipole-quadrupole, quadrupole-quadrupole, etc., systematically leads to further cancellation between these two terms.

Further, the cancellation is independent of the basis set. Motivated by these results, we define the short-range Hamiltonian at $\mathbf{Q} = \mathbf{0}$ to be

$$H^{\mathrm{SR}}(\mathbf{0}) \equiv T(\mathbf{0}) - K^{\mathrm{D}}(\mathbf{0}) + 2\delta_S \bar{K}^{\mathrm{X}}(\mathbf{0}), \quad (43)$$

where we have dropped the orbital indices to indicate that the definition is basis set independent. Because $H^{\mathrm{Xct}}(\mathbf{0})$ is ill-defined, it is natural to define the exciton states at $\mathbf{Q} = \mathbf{0}$ as eigenstates of $H^{\mathrm{SR}}(\mathbf{0})$. Notably, our definition of the short-range Hamiltonian coincides with the usual *ab initio* BSE Hamiltonian used for studying optical properties where it can be shown that using the modified Coulomb interaction in Eq. (42) facilitates the inclusion of local field effects [53].

In practice, $H^{\mathrm{SR}}(\mathbf{0})$ is constructed and diagonalized in a basis of electron-hole product states to obtain exciton eigenstates, $\Psi_{S\mathbf{0}}(\mathbf{R}, \mathbf{r})$, and eigenvalues $E_{S\mathbf{0}}$ prior to Wannierization. Throughout this work whenever we reference $\mathbf{Q} = \mathbf{0}$ exciton states, e.g., as in Eq. (9) or Eq. (25), we have in mind the eigenstates of $H^{\mathrm{SR}}(\mathbf{0})$.

B. Partitioning for $\mathbf{Q} \neq \mathbf{0}$

Working in a basis of $\mathbf{Q} = \mathbf{0}$ exciton states, we can construct an exciton Hamiltonian for states near $\mathbf{Q} = \mathbf{0}$ as follows:

$$\lim_{\mathbf{Q} \to \mathbf{0}} H^{\mathrm{Xct}}_{SS'}(\mathbf{Q}) = H^{\mathrm{SR}}_{SS'}(\mathbf{0}) + 2\delta_S K^{\mathrm{NA}}_{SS'}(\mathbf{Q}), \quad (44)$$

where $H^{\mathrm{SR}}_{SS'}(\mathbf{0}) = E_{S\mathbf{0}}\delta_{SS'}$ and $K^{\mathrm{NA}}_{SS'}(\mathbf{Q})$ was defined in Eq. (31).

When we try to rotate Eq. (44) into the MLXWF basis using the exciton rotation matrices, $U_{MS}(\mathbf{Q})$, we encounter a challenge. Because Wannierization is always performed in a subspace, it is only possible to rotate the sector of $\lim_{\mathbf{Q} \to \mathbf{0}} H^{\mathrm{Xct}}_{SS'}(\mathbf{Q})$ for which $S, S' \in \mathcal{W}$. By contrast, the nonanalytic interaction mixes the entire space of bright ($\mathbf{P}_S \neq 0$) excitonic states. Inevitably, the long-range interaction couples MLXWFs with states outside the Wannierization subspace.

At lowest order we could simply neglect the coupling to states outside of the Wannierization window, however numerical results indicate that this tends to greatly exaggerate the effect of $K^{\mathrm{NA}}_{SS'}(\mathbf{Q})$, leading to nonphysical results. Our approach here is to first derive an effective Hamiltonian which downfolds the nonanalytic interaction into the Wannierization subspace and then rotate the Hamiltonian into the MLXWF basis.

Let $S, S' \in \mathcal{W}$ denote excitonic states in the Wannierization window (also referred to as the active space) and $R, R' \notin \mathcal{W}$ label excitonic states outside the Wannierization window (the passive space). Adopting a Lowdin partitioning strategy [54], we treat $H^{\mathrm{SR}}_{SS'}(\mathbf{0})$ as our unperturbed Hamiltonian, $K^{\mathrm{NA}}_{SS'}(\mathbf{Q})$ as a perturbation, and sum over all transitions involving states outside the Wannierization subspace, effectively integrating out the passive states. We arrive at the following expression:

$$\lim_{\mathbf{Q} \to \mathbf{0}} H^{\mathrm{Xct,eff}}_{SS'}(\mathbf{Q}, \omega)$$

$$= H^{\mathrm{SR}}_{SS'}(\mathbf{0}) + 2\delta_S \Bigg[K^{\mathrm{NA}}_{SS'}(\mathbf{Q}) + 2 \sum_{R \notin \mathcal{W}} \frac{K^{\mathrm{NA}}_{SR}(\mathbf{Q}) K^{\mathrm{NA}}_{RS'}(\mathbf{Q})}{\omega - E_{R\mathbf{0}}}$$

$$+ 2^2 \sum_{R, R' \notin \mathcal{W}} \frac{K^{\mathrm{NA}}_{SR}(\mathbf{Q}) K^{\mathrm{NA}}_{RR'}(\mathbf{Q}) K^{\mathrm{NA}}_{R'S'}(\mathbf{Q})}{(\omega - E_{R\mathbf{0}})(\omega - E_{R'\mathbf{0}})} + \cdots \Bigg], \quad (45)$$

where the superscript "eff" serves as a reminder that this effective Hamiltonian mixes only states in the Wannierization subspace. Our analysis implies that even after downfolding, the exciton Hamiltonian can still be partitioned into a short- and long-range part

$$\lim_{\mathbf{Q} \to \mathbf{0}} H^{\mathrm{Xct,eff}}_{SS'}(\mathbf{Q}, \omega) = H^{\mathrm{SR}}_{SS'}(\mathbf{0}) + 2\delta_S K^{\mathrm{NA,eff}}_{SS'}(\mathbf{Q}, \omega), \quad (46)$$

where $K^{\mathrm{NA,eff}}_{SS'}(\mathbf{Q}, \omega)$ denotes the bracketed term in Eq. (45) summed to infinite order. The price to pay for downfolding is that the nonanalytic part is now frequency dependent and Eq. (46) must be solved self-consistently to obtain exciton eigenvalues. When all terms in the infinite series in Eq. (45) are retained, eigenvalues of this effective Hamiltonian exactly coincide with those of the original Hamiltonian in Eq. (44) for $S \in \mathcal{W}$.

Fortunately, the separable structure of $K^{\mathrm{NA}}_{SS'}(\mathbf{Q})$ allows the series to be summed to infinite order to obtain the exact result,

$$K^{\mathrm{NA,eff}}_{SS'}(\mathbf{Q}, \omega) = \varepsilon^{-1}_{\mathrm{BG}}(\mathbf{Q}, \omega) K^{\mathrm{NA}}_{SS'}(\mathbf{Q}), \quad (47)$$

where $S, S' \in \mathcal{W}$ and $\varepsilon_{\mathrm{BG}}(\mathbf{Q}, \omega)$ is the "background" dielectric function and is given by

$$\varepsilon_{\mathrm{BG}}(\mathbf{Q}, \omega) = 1 - \frac{8\pi e^2}{V_{\mathrm{uc}} N_{\mathbf{Q}} |\mathbf{Q}|^2} \sum_{R \notin \mathcal{W}} \frac{(\mathbf{P}_R \cdot \mathbf{Q})(\mathbf{P}^\star_R \cdot \mathbf{Q})}{\omega - E_{R\mathbf{0}}}, \quad (48)$$

where $\mathbf{P}_R$ are dipole moments associated with Bloch excitons [see Eq. (25)]. Importantly, the sum on R in $\varepsilon_{\mathrm{BG}}(\mathbf{Q}, \omega)$ runs only over Bloch exciton states outside the Wannierization subspace (see Appendix D for derivation).

We see the effect of downfolding is for states in the passive space to screen the nonanalytic coupling between excitons in the active space. The result is closely related to the S approximation, first introduced in Refs. [55,56], where the exchange term in the BSE is screened to compensate for the fact that the BSE is solved in a finite subspace of electron-hole transitions. Similar downfolding techniques were recently used to study x-ray absorption spectra in liquid water [57].

Using Wannier rotation matrices at $\mathbf{Q} = \mathbf{0}$, we can rotate Eq. (46) into the Wannier basis to find

$$\lim_{\mathbf{Q} \to \mathbf{0}} H^{\mathrm{Xct,eff}}_{MN}(\mathbf{Q}, \omega) = H^{\mathrm{SR}}_{MN}(\mathbf{0}) + 2\delta_S K^{\mathrm{NA,eff}}_{MN}(\mathbf{Q}, \omega). \quad (49)$$

Let $E_{S\mathbf{Q}}$ and $C_{MS}(\mathbf{Q})$ denote the eigenvalues and eigenvectors of Eq. (49), respectively. By construction $E_{S\mathbf{Q}}$ exactly coincide with the eigenenergies obtain by diagonalizing Eq. (44) for $S \in \mathcal{W}$. Once $E_{S\mathbf{Q}}$ and $C_{MS}(\mathbf{Q})$ are known, it is possible to cast Eq. (49) as a static, albeit non-Hermitian, Hamiltonian, namely,

$$\lim_{\mathbf{Q} \to \mathbf{0}} H^{\mathrm{Xct,eff}}_{MN}(\mathbf{Q}) = H^{\mathrm{SR}}_{MN}(\mathbf{0}) + 2\delta_S K^{\mathrm{NA,eff}}_{MN}(\mathbf{Q}), \quad (50)$$

where

$$K^{\mathrm{NA,eff}}_{MN}(\mathbf{Q}) = \sum_{N'} K^{\mathrm{NA}}_{MN'}(\mathbf{Q}) \big[\varepsilon^{-1}_{\mathrm{BG}}(\mathbf{Q})\big]_{N'N}, \quad (51)$$

with

$$\big[\varepsilon^{-1}_{\mathrm{BG}}(\mathbf{Q})\big]_{N'N} = \sum_S C_{N'S}(\mathbf{Q}) \varepsilon^{-1}_{\mathrm{BG}}(\mathbf{Q}, E_{S\mathbf{Q}}) C^{-1}_{SN}(\mathbf{Q}) \quad (52)$$

and with $K^{\mathrm{NA}}_{MN}(\mathbf{Q})$ given explicitly in Eq. (31). Derivations of these results are given in Appendix D, where we further show

that Eq. (50) has the same eigenvalues and eigenvectors as Eq. (49).

To recap, we have developed a partitioning of the exciton Hamiltonian near $\mathbf{Q} = \mathbf{0}$ which (1) mixes only states in the Wannierization subspace, (2) reproduces the exciton eigenvalues near $\mathbf{Q} = \mathbf{0}$, and (3) decouples into a short-range and long-range (nonanalytic) part. Motivated by these results, we define an effective short-range Hamiltonian near $\mathbf{Q} = \mathbf{0}$ as

$$\lim_{\mathbf{Q} \to \mathbf{0}} H_{MN}^{\mathrm{SR,eff}}(\mathbf{Q}) = \lim_{\mathbf{Q} \to \mathbf{0}} H_{MN}^{\mathrm{Xct}}(\mathbf{Q}) - 2\delta_S K_{MN}^{\mathrm{NA,eff}}(\mathbf{Q}), \quad (53)$$

where $H_{MN}^{\mathrm{Xct}}(\mathbf{Q})$ is the original exciton Hamiltonian. We can generalize this definition to arbitrary $\mathbf{Q}$ as follows:

$$H_{MN}^{\mathrm{SR,eff}}(\mathbf{Q}) = H_{MN}^{\mathrm{Xct}}(\mathbf{Q}) - 2\delta_S K_{MN}^{\mathrm{LR,eff}}(\mathbf{Q}), \quad (54)$$

where

$$K_{MN}^{\mathrm{LR,eff}}(\mathbf{Q}) = \sum_{N'} K_{MN'}^{\mathrm{LR}}(\mathbf{Q})\big[\varepsilon_{\mathrm{BG}}^{-1}(\mathbf{Q})\big]_{N'N} \quad (55)$$

and with $[\varepsilon_{\mathrm{BG}}^{-1}(\mathbf{Q})]_{N'N}$ and $K_{MN}^{\mathrm{LR}}(\mathbf{Q})$ defined in Eqs. (52) and (34), respectively. From Eq. (35) it immediately follows that when $\mathbf{Q} \to \mathbf{0}$ Eq. (54) reduces to Eq. (53). In practice, $K_{MN}^{\mathrm{LR,eff}}(\mathbf{Q})$ is constructed as follows. First, we build and diagonalize Eq. (49) to obtain $E_{S\mathbf{Q}}$ and $C_{MS}(\mathbf{Q})$. With these eigenvalues and eigenvectors, we construct $[\varepsilon_{\mathrm{BG}}^{-1}(\mathbf{Q})]_{N'N}$ according to Eq. (52). The effective long-range kernel is then constructed according to Eqs. (55). The procedure is computationally inexpensive, involving only a few diagonalizations of a small $N_{\mathcal{W}} \times N_{\mathcal{W}}$ matrix at each $\mathbf{Q}$ point.

We emphasize that all results presented here are for bulk materials where the $\mathbf{Q} \to \mathbf{0}$ limit of the Coulomb interaction is given by $v(\mathbf{Q}) = 4\pi e^2/|\mathbf{Q}|^2$. For confined and low dimensional systems, the $\mathbf{Q} \to \mathbf{0}$ limit of $v(\mathbf{Q})$ changes [58] and expressions in this section should be modified accordingly.

We have further neglected the effects of spin-orbit coupling. When included, the exciton Hamiltonian can no longer be partitioned into a singlet and triplet sector, instead excitons should be labeled by their total angular momentum. In general, all excitons receive an exchange contribution and the partitioning described in this section should be applied to all states. In practice, the formulas of this section can be used as they stand provided the definition of $\mathbf{P}_S$ in Eq. (25) is appropriately generalized for spinors.

VI. WANNIER-FOURIER INTERPOLATION

We now turn to the interpolation of exciton band structures, making use of the results derived in the previous section. We start with triplet excitons, where there is no exchange coupling and hence no dipolar interactions which simplifies matters. Here, the interpolation procedure is entirely analogous to what is done at the one-electron level [7,10]. We then discuss what modifications must be made to carry out the procedure for singlets.

A. Triplets

Given the exciton energies, $E_{S\mathbf{Q}_c}$, from, e.g., an *ab initio* BSE calculation, and exciton Wannier rotation matrices, $U_{MS}(\mathbf{Q}_c)$, on a coarse uniform grid of $\mathbf{Q}_c$ points, we first

Wannier-Fourier transform to obtain hopping matrix elements

$$H_{MN}^{\mathrm{Xct}}(\bar{\mathbf{R}}) = \sum_{S\mathbf{Q}_c} e^{-i\mathbf{Q}_c \cdot \bar{\mathbf{R}}} U_{MS}^{\dagger}(\mathbf{Q}_c) E_{S\mathbf{Q}_c} U_{SN}(\mathbf{Q}_c), \quad (56)$$

and subsequently Fourier transform to an arbitrary $\mathbf{Q}$ point to find

$$H_{MN}^{\mathrm{Xct}}(\mathbf{Q}) = \sum_{\bar{\mathbf{R}}} e^{i\mathbf{Q} \cdot \bar{\mathbf{R}}} H_{MN}^{\mathrm{Xct}}(\bar{\mathbf{R}}). \quad (57)$$

Triplet exciton eigenenergies at $\mathbf{Q}$ are obtained by diagonalizing $H_{MN}^{\mathrm{Xct}}(\mathbf{Q})$. The success of this scheme hinges critically on the rapid decay of $H_{MN}^{\mathrm{Xct}}(\bar{\mathbf{R}})$ with increasing $\bar{\mathbf{R}}$, which, for triplets, is guaranteed when $W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r})$ is well localized in $\mathbf{R}$.

B. Singlets

For singlet excitons, even when $W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r})$ is well-localized in $\mathbf{R}$, the matrix element $H_{MN}^{\mathrm{Xct}}(\bar{\mathbf{R}})$ will still decay as $1/|\bar{\mathbf{R}}|^3$ as discussed in Sec. IV A, which in many cases will be too gradual to perform Wannier-Fourier interpolation efficiently.

To deal with this issue, we remove the long-range interaction before Wannier-Fourier interpolating, and add it back after the interpolation is complete. This strategy is inspired by a widely used technique for Fourier interpolating phonon eigenfrequencies in many semiconductors and insulators [59]. We outline our four-step procedure below.

(1) Subtract the effective long-range interaction, defined in Eq. (55), to obtain the effective short range Hamiltonian on a coarse grid:

$$H_{MN}^{\mathrm{SR,eff}}(\mathbf{Q}_c) = H_{MN}^{\mathrm{Xct}}(\mathbf{Q}_c) - 2\delta_S K_{MN}^{\mathrm{LR,eff}}(\mathbf{Q}_c). \quad (58)$$

(2) Fourier interpolate to obtain $H^{\mathrm{SR,eff}}$ at an arbitrary $\mathbf{Q}$ point:

$$\begin{aligned} H_{MN}^{\mathrm{SR,eff}}(\bar{\mathbf{R}}) &= \sum_{\mathbf{Q}_c} e^{-i\mathbf{Q}_c \cdot \bar{\mathbf{R}}} H_{MN}^{\mathrm{SR,eff}}(\mathbf{Q}_c), \\ H_{MN}^{\mathrm{SR,eff}}(\mathbf{Q}) &= \sum_{\bar{\mathbf{R}}} e^{i\mathbf{Q} \cdot \bar{\mathbf{R}}} H_{MN}^{\mathrm{SR,eff}}(\bar{\mathbf{R}}). \end{aligned} \quad (59)$$

(3) Add back the effective long-range interaction to obtain H^{Xct} at finite $\mathbf{Q}$:

$$H_{MN}^{\mathrm{Xct}}(\mathbf{Q}) = H_{MN}^{\mathrm{SR,eff}}(\mathbf{Q}) + 2\delta_S K_{MN}^{\mathrm{LR,eff}}(\mathbf{Q}). \quad (60)$$

(4) Diagonalize $H_{MN}^{\mathrm{Xct}}(\mathbf{Q})$ to obtain singlet exciton eigenvalues at arbitrary $\mathbf{Q}$ points.

We emphasize here that it is critical to use the effective long-range interaction, defined in Eq. (55), rather than the bare long-range interaction, defined in Eq. (34). The former properly accounts for long-range coupling to states outside the Wannierization window while the latter does not.

VII. IMPLEMENTATION AND COMPUTATIONAL DETAILS

As previously emphasized, the Wannierization procedure is independent of computational method used to compute excitonic properties. Popular methods include time-dependent density functional theory [53], equation-of-motion coupled

cluster [60], and the *ab initio GW*-Bethe-Salpeter equation (BSE) approach [24], to name a few. Here we adopt the latter method and summarize the approach below.

A. Implementation

For computational tractability, the BSE is cast as an eigenvalue equation in a basis of electron-hole product states $\langle \mathbf{r}_e \mathbf{r}_h | cv\mathbf{k}\mathbf{Q} \rangle \equiv \psi_{c\mathbf{k}}(\mathbf{r}_e)\psi_{v\mathbf{k}-\mathbf{Q}}^{\star}(\mathbf{r}_h)$ and reads

$$(\varepsilon_{c\mathbf{k}} - \varepsilon_{v\mathbf{k}-\mathbf{Q}})A_{cv\mathbf{k}}^{S\mathbf{Q}}$$
$$+ \sum_{c'v'\mathbf{k}'} \langle cv\mathbf{k}\mathbf{Q}|K^{\mathrm{eh}}|c'v'\mathbf{k}'\mathbf{Q}\rangle A_{c'v'\mathbf{k}'}^{S\mathbf{Q}} = E_{S\mathbf{Q}}A_{cv\mathbf{k}}^{S\mathbf{Q}}, \quad (61)$$

where $\varepsilon_{n\mathbf{k}}$ are *GW* quasiparticle energies and $K^{\mathrm{eh}} = -K^{\mathrm{D}} + 2\delta_S K^{\mathrm{X}}$ is the electron-hole interaction kernel. Upon solving the BSE, one obtains exciton eigenenergies, $E_{S\mathbf{Q}}$, and exciton wave functions in terms of the expansion coefficients, $A_{cv\mathbf{k}}^{S\mathbf{Q}}$, see Eq. (2).

At $\mathbf{Q} = \mathbf{0}$, the exchange term is ill-defined because the Coulomb interaction is divergent in the long-wavelength limit. As discussed in Sec. V A we build $K^{\mathrm{X}}(\mathbf{Q} = \mathbf{0})$ with the modified Coulomb interaction, $\bar{v}(\mathbf{G}) = 4\pi e^2/|\mathbf{G}|^2(1 - \delta_{\mathbf{G0}})$, which is equivalent to the usual Coulomb interaction except the $\mathbf{G} = \mathbf{0}$ component has been set to zero. For all other $\mathbf{Q}$, the usual Coulomb interaction, $v(\mathbf{Q} + \mathbf{G}) = 4\pi e^2/|\mathbf{Q} + \mathbf{G}|^2$, is used to construct $K^{\mathrm{X}}(\mathbf{Q})$. Modifying the Coulomb interaction in this way regulates the nonanalyticity contained in $K^{\mathrm{X}}(\mathbf{Q})$ and ensures that the exciton coefficients are well defined at $\mathbf{Q} = \mathbf{0}$. While this is not the only regulation scheme, it is convenient and consistent with our definition of the long-range kernel in Eq. (33).

The exciton dipole matrix elements, in the Bloch basis are computed from exciton expansion coefficients and single-particle Bloch states as follows:

$$\mathbf{P}_S = \int \mathbf{R}\Psi_{S0}(\mathbf{R}, \mathbf{r} = \mathbf{0})d\mathbf{R}$$
$$= \sqrt{\frac{N_\mathbf{Q}}{N_\mathbf{k}}} \sum_{cv\mathbf{k}} A_{cv\mathbf{k}}^{S0} \langle u_{v\mathbf{k}}|\mathbf{R}|u_{c\mathbf{k}}\rangle_{\mathrm{uc}}, \quad (62)$$

where $A_{cv\mathbf{k}}^{S0}$ should be calculated from the modified $\mathbf{Q} = \mathbf{0}$ kernel as discussed in the previous paragraph. Note $\mathbf{P}_S$ is equivalent to the usual exciton transition dipole matrix elements used in optical studies [24]. From Eqs. (26) and (62) it immediately follows that

$$\mathbf{P}_M = \frac{1}{\sqrt{N_\mathbf{k}}} \sum_S U_{SM}(\mathbf{0}) \sum_{cv\mathbf{k}} A_{cv\mathbf{k}}^{S0} \langle u_{v\mathbf{k}}|\mathbf{R}|u_{c\mathbf{k}}\rangle_{\mathrm{uc}}. \quad (63)$$

With $\mathbf{P}_M$, we construct the $K_{MN}^{\mathrm{LR,eff}}(\mathbf{Q}, \omega)$ as prescribed in Eq. (55) and detailed in the bottom right-hand corner of Fig. 1. In practice Ewald's technique is used to efficiently evaluate Eq. (55). Additional details can be found in Appendix E.

B. Computational details

As a first application, we apply the above formalism to Wannierize the low-lying excitons in LiF, a prototypical wide band-gap insulator with weak dielectric screening and correspondingly strong electron-hole interactions. We note that LiF
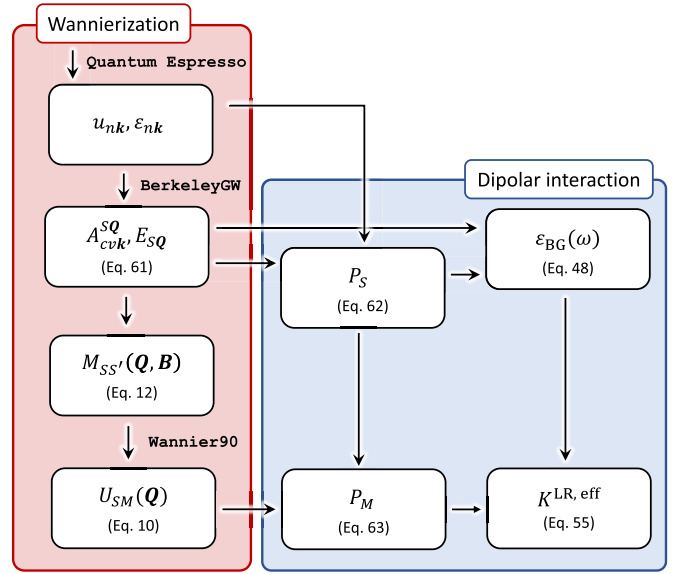

FIG. 1. Workflows for obtaining the unitary matrix $U_{SM}(\mathbf{Q})$ that leads to MLXWFs (red section) and computing the dipolar coupling between MLXWFs (blue section). Software used in this work is indicated.

was one of the first systems studied within the *ab initio GW* plus BSE approach [24] while the closely related compound LiCl appeared in the original work on MLWFs [1].

The starting point for all calculations is a ground state density functional theory (DFT) calculation to obtain Kohn-Sham energies and eigenstates [61,62]. We use a planewave basis set, the local density approximation (LDA), and norm-conserving pseudopotentials taken from `pseudo-dojo` [63,64]. To converge the ground state density we use an 80 Ry planewave cutoff and $8 \times 8 \times 8$ **k** mesh. All DFT calculations are performed with `Quantum Espresso` [65].

We perform BSE calculations for LiF atop the Kohn-Sham eigenvalues and eigenstates. The electron-hole kernel, K^{eh}, is expanded in one conduction band and three valence bands. For the direct kernel we assume static screening, and compute the static dielectric matrix within the random phase approximation using a sum-over-bands approach [66,67]. In computing the susceptibility we include 195 unoccupied bands and use a 20 Ry planewave cutoff. All BSE calculations are performed with `BerkeleyGW` [33].

Our workflow requires exciton eigenenergies, $E_{S\mathbf{Q}}$, and eigenstates, $\Psi_{S\mathbf{Q}}$, on a regular $\mathbf{Q}$ mesh. In principle the **k** and $\mathbf{Q}$ meshes on which we solve the BSE are decoupled, however in practice it is convenient to take them to be commensurate so that a single set of wave functions can be used to compute overlaps appearing in Eq. (12). From Eq. (12) we see that for a fixed $\mathbf{Q}$ mesh we require the following exciton expansion coefficients $A_{cv\mathbf{k}+\mathbf{B}/2}^{S\mathbf{Q}}$ and wave function overlaps $\langle u_{v'\mathbf{k}-\mathbf{Q}-\mathbf{B}/2}|u_{v\mathbf{k}-\mathbf{Q}}\rangle_{\mathrm{uc}}$. Because of the $\mathbf{B}/2$ shifts, for a fixed $\mathbf{Q}$ mesh we require a **k** mesh which is both commensurate and twice as dense, along each reciprocal lattice vector direction, as the $\mathbf{Q}$ mesh. To give an explicit example, Wannierizing the exciton on a $5 \times 5 \times 5$ $\mathbf{Q}$ mesh requires solving the BSE at 125 $\mathbf{Q}$-pts with the electron-hole kernel constructed on a

$10 \times 10 \times 10$ **k** mesh. Throughout this work, we use a Γ centered **Q** mesh and half-shifted **k** mesh.

The factor $\mathbf{B}/2$ is linked to our choice to Wannierize in $\mathbf{R} = (\mathbf{r}_e + \mathbf{r}_h)/2$. As discussed Appendix B it is more generally possible to Wannierize in $\mathbf{R} = \alpha \mathbf{r}_e + \beta \mathbf{r}_h$, where $\alpha + \beta = 1$ and $\alpha, \beta > 0$. Notably for $\alpha = 1, \beta = 0$ or $\alpha = 0, \beta = 1$, it is possible to take the two meshes perfectly commensurate. For illustrative purposes, in the main text we have taken $\alpha = 1/2, \beta = 1/2$ which represents a good compromise between conceptual simplicity and computational tractability. We have numerically confirmed that for LiF the choices $\alpha = 0, \beta = 1$ and $\alpha = 1, \beta = 0$ give MLXWFs with the same orbital shapes and centers as those discussed below.

Finally, we note that the total number of exciton bands is fixed by the basis size, in this case $N_{\text{Xct}} = N_{\mathbf{k}} \times N_c \times N_v$, where N_c and N_v are the number of conduction and valence states used to expand the electron-hole kernel. For the example above we have $N_{\text{Xct}} = 3000$. In this work, no symmetry or interpolation scheme is used when building the electron-hole kernel.

VIII. RESULTS

A. Exciton energies and dispersion

LiF takes up a rock-salt crystal structure and we use a lattice constant of 4.026 Å for all calculations. It is a direct gap material with a LDA Kohn-Sham gap of $E_g^{\text{LDA}} = 8.87$ eV, consistent with previously reported values between 8.3–8.8 eV for the LDA gap [68–71]. We attribute this range to slight variability in the lattice parameters, the inclusion or noninclusion of the 1s state in the Li pseudopotential, and a mixture of Gaussian versus planewave basis sets used. To account for quasiparticle corrections we rigidly shift the conduction bands so that the fundamental gap is $E_{\text{gap}}^{\text{QP}} = 14.3$ eV and stretch the valence bands by 15%, as was done in prior work [72,73].

Our BSE calculation of LiF shows that the lowest energy triplet (T) exciton, at Γ ($\mathbf{Q} = \mathbf{0}$), is threefold degenerate. The additional repulsive exchange interaction in the singlet kernel lifts the degeneracy, leading to a twofold degenerate transverse (S_t), and nondegenerate longitudinal (S_l) exciton, as expected. We obtain the following energies for these states: $E_T = 12.00$ eV, $E_{S_t} = 12.36$ eV and $E_{S_l} = 13.08$ eV, all of which are largely consistent with prior work. Our exciton binding energy, $E_{\text{B}} = E_{\text{gap}}^{\text{QP}} - E_{S_t} = 1.94$ eV, is slightly larger than what has been reported in the older literature (1.5–1.8 eV) [24,70,72] and slightly less than a recent report of a 2.05 eV [74]. Similarly, the LT-splitting $\Delta_{\text{LT}} = E_{S_l} - E_{S_t} = 0.72$ eV is somewhat larger than 0.5 eV previously reported [24]. For present purposes our results are in satisfactory agreement with prior work.

In Fig. 4 we show the BSE-computed exciton dispersion (blue curve) for the low energy triplet (right panel) bands. We find that the excitons at Γ are threefold degenerate. The three lowest exciton bands remain entangled throughout the Brillouin zone but are well separated from higher lying bands (see Fig. 4). In the left panel of the same figure, we show the exciton dispersion for the singlet states. We can clearly see that the threefold degeneracy at Γ is lifted by the exchange interaction, with the longitudinal branch higher in

TABLE II. Minimized spread Ω for singlet and triplet excitons in LiF and its decomposition into invariant, Ω_{I}, off-diagonal, Ω_{OD}, and diagonal Ω_{D} parts. The spread of the individual Wannier functions are reported in the final column. All values are reported in Å^2.

Q mesh	Ω	Ω_{I}	Ω_{OD}	Ω_{D}	WF spread
			Singlets		
$3 \times 3 \times 3$	4.799	4.738	0.061	0.0	1.600
$4 \times 4 \times 4$	5.881	5.823	0.058	0.0	1.960
$5 \times 5 \times 5$	6.503	6.470	0.060	0.0	2.176
			Triplets		
$3 \times 3 \times 3$	3.980	3.957	0.023	0.0	1.327
$4 \times 4 \times 4$	4.494	4.476	0.018	0.0	1.498
$5 \times 5 \times 5$	4.797	4.778	0.017	0.0	1.599

energy than the two degenerate transverse branches [75]. The eigenvalues approach a well defined value in the $\mathbf{Q} \to 0$ limit, independent of the path taken the $\mathbf{Q}$ space. The three lowest singlet bands are not entirely isolated from higher lying singlet states, however the disentanglement is fairly minimal occurring only near the L points at the BZ edge. Our calculations are in very good agreement with prior *ab initio* calculations of the exciton dispersion in LiF along high-symmetry paths [73].

B. Wannierization of excitons in LiF

Based on the exciton dispersion, we restrict our analysis to the subspace $\mathcal{W}$ which contains the lowest three exciton bands, $N_{\mathcal{W}} = 3$. For triplet states we do not provide a starting guess for the Wannierization procedure. For singlets, which are minimally entangled, we find that it is helpful to provide `Wannier90` with some initial guess for the exciton Wannier functions. Here we use the already Wannierized triplet states as our initial guess. In Appendix B we detail how these overlaps are constructed. We utilize the standard disentanglement scheme [11] implemented in `Wannier90` to select the optimal three-band subspace for constructing MLXWFs.

In Table II we report on the convergence of the Wannierization procedure for both singlet and triplet excitons with increasing **Q** mesh. We decompose the total spread, Ω, into its invariant, Ω_{I}, off-diagonal, Ω_{OD}, and diagonal, Ω_{D}, contributions (see Ref. [1] for details on the decomposition). As discussed previously, the underlying **k** mesh is chosen twice as dense as the **Q** mesh. Despite this additional variability, we still observe clear convergence trends. The convergence of the total spread, Ω, is relatively slow, stemming primarily from the gauge invariant part of the spread, Ω_{I}, which is fixed after the disentanglement procedure is complete. By contrast, the gauge dependent part of the spread, $\Omega_{\text{D}} + \Omega_{\text{OD}}$, the part which is actually minimized after the subspace is fixed, converges rapidly with increasing **Q** mesh, changing by less than 0.002 Å^2 between the final two steps. This behavior is analogous, and the level of convergence in Ω similar, to what was reported at one-electron level in the original paper by Marzari and Vanderbilt [1] in simple semiconducting systems. We further find that Ω_{D} is strictly zero, an indication that all Wannier functions possess an inversion center.

In the final column of Table II, we report the spread of the individual exciton Wannier functions, i.e., $\langle M\bar{\mathbf{0}}|R^2|M\bar{\mathbf{0}}\rangle -$
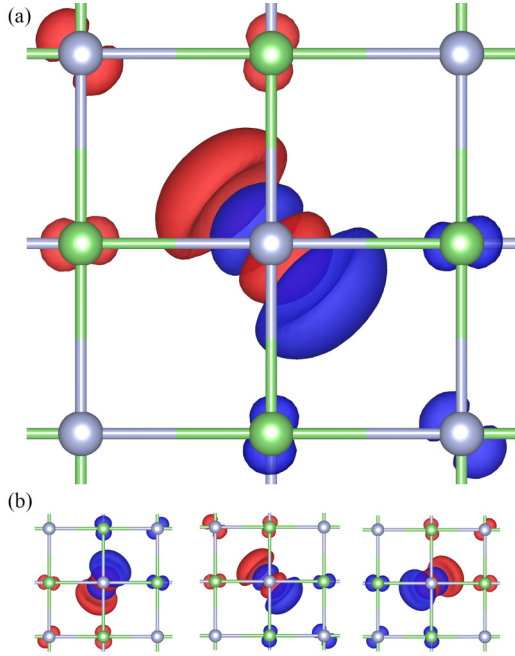
FIG. 2. Triplet MLXWF for LiF, plotted as a function of the average coordinate $\mathbf{R}$ with the relative coordinate $\mathbf{r} = \mathbf{0}$—i.e., $W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{0})$ vs $\mathbf{R}$. Green and gray spheres denote Li and F atoms respectively. The Wannier function is entirely real so that blue and red lobes denote regions of positive and negative probability amplitude, respectively. In panel (a) we show an enlarged version of the MLXWF for $M = 2$. In panel (b) we show the three MLXWFs ($M = 1$–3).

TABLE III. Magnitude of MLXWF dipole moments (in Å), $|\mathbf{P}_M|$ and the background dielectric constant evaluated at the longitudinal and transverse exciton energies.

$\mathbf{k}$ mesh	N	$|\mathbf{P}_M|$	$\varepsilon_{\mathrm{BG}}(E_{S_l}, \mathbf{Q} \to \mathbf{0})$	$\varepsilon_{\mathrm{BG}}(E_{S_t}, \mathbf{Q} \to \mathbf{0})$
$3 \times 3 \times 3$	645	0.260	2.06	1.58
$4 \times 4 \times 4$	1533	0.261	2.06	1.59
$5 \times 5 \times 5$	2997	0.261	2.04	1.58

In analogy to the one-electron case, our numerical results indicate that the Wannierized excitons can be made entirely real through multiplication by a complex phase. For triplets we find ImW/Re$W \approx 0.0001$ while for singlets ImW/Re$W \approx 0.01$. Further, we see the MLXWFs transform as odd functions under inversion symmetry and have a nonzero dipole moment. For singlet excitons, this dipole moment gives rise to long-range dipolar interactions between MLXWFs [see Eq. (23)] and drives the splitting of the longitudinal and transverse exciton branches as previously discussed in Sec. IV B.

D. MLXWF dipole moments and LT splitting

The MLXWF dipole moments can in principle be computed through real-space integration of the MLXWF depicted in Fig. 2, following Eq. (24). However, here it is more convenient to compute them via Eq. (63) using $A_{cv\mathbf{k}}^{S\mathbf{0}}$ and $U_{SM}(\mathbf{0})$, quantities we already have in hand at the end of our BSE and Wannierization calculation, respectively. In Table III we report the magnitude of the MLXWF dipoles.

The MLXWF dipole moments are the main ingredients for constructing the dipolar interaction between MLXWFs. As discussed in Sec. V B, it is important to further screen this interaction to correctly account of coupling to excitons outside the Wannierization subspace.

In Table III we report the values of background dielectric function evaluated at the longitudinal and transverse exciton energies computed with different $\mathbf{Q}$ meshes. As the $\mathbf{Q}$ mesh increases, so does the $\mathbf{k}$ mesh, and consequently the number of solutions to the BSE, N_{Xct}. In constructing the effective dielectric function, see Eq. (48), we must sum over all states outside Wannierization window $N = N_{\mathrm{Xct}} - N_{\mathcal{W}}$. Presently $N_{\mathcal{W}}$ is held fixed at 3. Despite the increase in the number of states, $\varepsilon(E_{S_l}, \mathbf{Q} \to \mathbf{0}) \approx 2.0$ and $\varepsilon(E_{S_t}, \mathbf{Q} \to \mathbf{0}) \approx 1.6$ across all $\mathbf{Q}$ meshes, pointing to the stability of the procedure.

$\langle M\bar{\mathbf{0}}|\mathbf{R}|M\bar{\mathbf{0}}\rangle^2$. We find all three Wannier functions have the same spread, stemming from the cubic symmetry of LiF, and in Table II we report a single value in this column. For a given $\mathbf{Q}$ mesh, we find the singlet MLXWFs have larger spreads relative to the triplets. The result is expected as singlets experience an additional exchange force which further delocalizes these states relative to their triplet counterparts.

C. Visualizing the MLXWFs

To examine the nature of the MLXWF, in Fig. 2, we plot isosurfaces of the MLXWF in $\mathbf{R}$ with $\mathbf{r}$ held fixed at $\mathbf{0}$. Explicitly, we plot

$$W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r} = \mathbf{0}) = \frac{1}{\sqrt{N_{\mathbf{Q}}}} \sum_{S\mathbf{Q}} U_{SM}(\mathbf{Q}) \Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r} = \mathbf{0}) \quad (64)$$

versus $\mathbf{R}$. While it is possible more generally to restrict $\mathbf{r} = \mathbf{c}$, where $\mathbf{c}$ is some constant vector, our choice $\mathbf{r} = \mathbf{0}$ is motivated by the fact that it is precisely this component which appears in the exchange integral and MLXWF dipole moments defined in Eqs. (22) and (24), respectively. We observe that the MLXWF is well localized in $\mathbf{R}$ in the cell $\bar{\mathbf{R}} = \bar{\mathbf{0}}$. This should be contrasted with $\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r} = \mathbf{0})$, which is delocalized in $\mathbf{R}$ across the entire crystal. In Fig. 2(a) we show an enlarged version of the MLXWF. We find that the MLXWF are centered on the fluorine site. The MLXWF has nontrivial nodal structure and in future work it will be interesting to rationalize its shape in terms of linear combinations of electron-hole product states. In Fig. 2(b) we plot the isosurfaces for the other two MLXWFs.

E. Exciton band structure interpolation

As a first application, we show how MLXWFs can be used to interpolate exciton energies throughout the Brillouin zone. Our goal here is to illustrate that the MLXWFs functions can be used as a basis for constructing *ab initio* tight-binding models which faithfully reproduces the exciton dispersion and contains the same physics as the usual Bloch description. Additional applications are expounded upon in the following section.

In Fig. 3 we plot the magnitude of the BSE Hamiltonian in the Wannier basis, $H_{MN}^{\mathrm{Xct}}(\bar{\mathbf{R}})$, as a function of $|\bar{\mathbf{R}}|$ in a manner analogous to what is done in Ref. [7] for H^{DFT}. Physically these matrix elements can be interpreted as hopping matrix
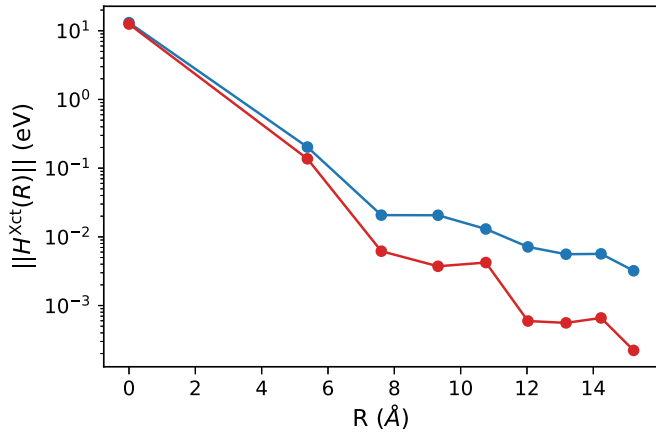
FIG. 3. Decay of the singlet (blue curve) and triplet (red curve) BSE Hamiltonian in the Wannier representation $H_{MN}^{Xct}(\bar{\mathbf{R}})$ as a function of $\bar{\mathbf{R}}$ for LiF. The data points correspond to the largest value of the matrix element at a given $|\bar{\mathbf{R}}|$—i.e., $||H(\bar{\mathbf{R}})|| = \max_{MN,|\mathbf{R}|=R} |H_{MN}^{Xct}(\mathbf{R})|$.

elements in the MLXWF basis. As expected, we observe that the triplet matrix elements decay exponentially in $|\bar{\mathbf{R}}|$ while the singlet matrix elements appear to decay more slowly. Theoretically, we expect $H_{MN}^{Xct}(\bar{\mathbf{R}})$ to decay as $1/|\bar{\mathbf{R}}|^3$ for singlets. Our numerical results highlight the importance of removing the long-range dipolar coupling before interpolating.

In Fig. 4 we show the results of our Wannier-Fourier interpolated exciton dispersion for LiF. In blue, we plot a linear interpolation of the singlet and triplet exciton eigenenergies obtained through explicitly diagonalizing the BSE Hamiltonian at 72 **Q** points along a high symmetry path. In red we overlay our Wannier-Fourier interpolated exciton dispersion which we computed starting from a $5 \times 5 \times 5$ coarse **Q** mesh. We find our Wannier-Fourier interpolation scheme (see Sec. VI) gives excellent agreement for both the singlet and triplet exciton dispersions. In the inset of the left panel, we show how Wannier-Fourier interpolation preforms without

correcting for long-range interactions. In this case, we find that the dispersion agrees relatively well away from Γ, but completely fails to capture the LT-splitting in the $\mathbf{Q} \to \mathbf{0}$ limit, highlighting the importance of analytically handling the long-range term.

IX. DISCUSSION

We expect that MLXWFs could find wide application much as MLWFs have for one-electron states. By design we have formulated our framework to mirror as closely as possible the one-electron case. In doing so we are hopeful that with minimal effort many of the same Wannier-based post-processing tools available for electrons may be readily adapted for the excitonic case. In this work, we have already shown that the internal routines in `Wannier90` can, with minimal modification, be used to plot MLXWFs and interpolate the exciton band structure. We expect the same to be true for other quantities. For instance, using MLXWFs we can interpolate the exciton-phonon vertex [39–41] in a manner analogous to how MLWFs are used to interpolate the electron-phonon vertex [7]. Another example is the use of MLXWFs to rotate into a smooth gauge which will facilitate the computation of Berry-phase related properties, like the anomalous hall effect [3], at the excitonic level.

In adapting these workflows to the excitonic case, it will be important to carefully treat the long-range dipolar coupling between singlet MLXWFs. As previously discussed these long-range couplings are inherently tied to the polarization carried by the exciton, and have no analog at the single electron-level (though there is an analogy with phonons, see Sec. IV C). These long-range interactions are responsible for many interesting properties of excitons. For instance, we have already shown they give rise to splitting of the longitudinal and transverse exciton branches in LiF. In lower symmetry systems these dipolar interactions will lead to nonanalytic behavior in the exciton eigenvalues. In the context of exciton-phonon coupling we anticipate a long-range polar coupling
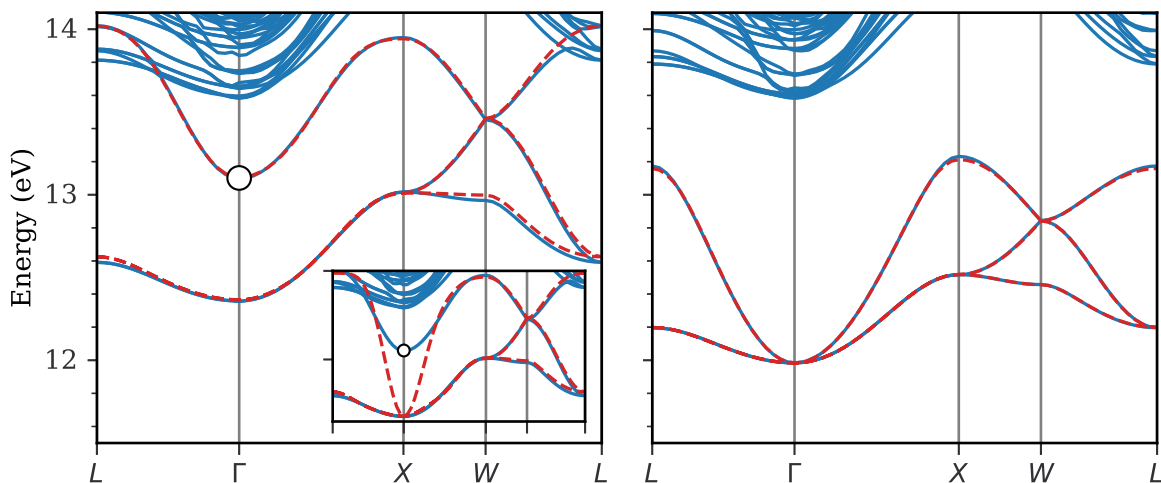

FIG. 4. Singlet (left panel) and triplet (right panel) exciton dispersion for LiF. Linear interpolation of exciton eigenergeies explicitly explicitly obtained by diagonalizing the BSE at 72 **Q** points along the high-symmetry path (blue curve). Wannier-Fourier interpolated exciton dispersion starting from a Γ-centered $5 \times 5 \times 5$ **Q** mesh (dotted red curve). The inset of the left panel shows how Wannier-Fourier interpolation performs without isolating the long-range contribution.

between bright excitons and polar phonons. While long-range Fröhlich [76] and piezoelectric [77] electron-phonon coupling have been studied from first-principles, to the best of our knowledge, the polar exciton-lattice coupling mentioned above has yet to be explored. By contrast, this long-range coupling has already been linked to the anomalous enhancement of the exciton spin-hall effect relative to the case where electron-hole interactions are neglected [42–44]. Our formalism provides a way to compute this enhancement from first-principles.

Beyond Wannier-based post-processing applications we expect the MLXWFs will provide valuable insight into the fundamental nature of the exciton itself. In much the same way that MLWFs provide a bridge between Bloch states and tight-binding orbitals, MLXWFs provide a bridge between the Mott-Wannier [35] and Frenkel [36] starting points for describing an exciton. More specifically, just as the shape of MLWFs can often be rationalized in terms of linear combinations of atomic orbitals, the real space orbital structure of MLXWFs can be rationalized in terms of linear combinations of electron-hole product orbitals. Such analysis could help to better understand the most dominant electron-hole transitions which compose excitons in complex materials. MLXWFs could also serve as a practical tool for quantifying the degree of electron-hole charge-transfer and compliment existing *ab initio* methods [31].

MLXWFs also provide an optimal orbital basis for describing the low-energy physics of excitons near the optical band-edge in much the same way that MLWFs provide a efficient basis for describing quasiparticle states near the fundamental band edge, or Fermi surface for metallic systems. Indeed with the quantities computed in this work it is possible to write down an effective tight-binding Hamiltonian for low-lying excitons, namely,

$$
\begin{aligned}
H^{\text{Xct}} = & \sum_{\substack{MN \\ \mathbf{RR}'}} \langle M\mathbf{0}|H^{\text{SR}}|N\mathbf{D}\rangle\, c^{\dagger}_{M\mathbf{R}} c_{N\mathbf{R}'} \\
& + \sum_{\substack{MN \\ \mathbf{RR}'}} 2\delta_S \langle M\mathbf{0}|K^{\text{LR,eff}}|N\mathbf{D}\rangle\, c^{\dagger}_{M\mathbf{R}} c_{N\mathbf{R}'}, \quad (65)
\end{aligned}
$$

where $\mathbf{D} = \mathbf{R}' - \mathbf{R}$. The first term is readily understood as the usual hopping term while the second term is the dipolar long-range interaction which couples exciton orbitals over long distances. We note that this is not the only partitioning which can be made. For certain applications it may be useful to further isolate higher-order dipole-quadrupole and quadrupole-quadrupole couplings as has recently been done for the electron-phonon vertex [77]. However as shown in the main text all $\mathbf{Q}$-space nonanalyticity is contained in the lowest order dipolar term and for many applications we expect Eq. (65) may serve as a suitable effective Hamiltonian.

One such application of particular interest is the study of exciton transport. For tightly bound excitons, the two terms in Eq. (65) lead to two different transport mechanisms. The first term gives rise to short-range Dexter-like [78] transport while the second term gives rise to long-range Förster-like transport [79,80]. The MLXWF computed in this work can be used as a diabatic basis for studying exciton dynamics [19], shedding light on exciton dynamics.

Finally, in this work, we have generalized the one-particle Wannierization procedure to handle a specific two-particle excitation (the exciton). In principle this procedure may readily be adapted to handle other two-particle excitations, for example the bipolaron, as well as excitations involving more than two particles, for instance, trions and biexcitons, and beyond.

X. CONCLUSIONS

In this work, we have developed a procedure for constructing maximally localized exciton Wannier functions (MLXWFs). We have benchmarked our work on LiF, and demonstrated that the MLXWFs converge smoothly with increasingly dense $\mathbf{Q}$ meshes and can be used for post-processing applications, like Wannier-Fourier interpolation of the exciton dispersion. We expect this work will serve as a starting point for many other post-processing applications as discussed in Sec. IX. Our framework helps to connect the classic Mott-Wannier and tight-binding pictures of excitons in a first principles context further bridging the gap between condensed matter and quantum chemistry methods for computing neutral excitations.

ACKNOWLEDGMENTS

APPENDIX A: DEFINITIONS AND CONVENTIONS

1. One-electron states

We use the following standard definitions and conventions for the single-particle Bloch states

$$
\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{k}}}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) \quad (A1)
$$

and

$$
u_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V_{\text{uc}}}} \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (A2)
$$

where $u_{n\mathbf{k}}(\mathbf{r})$ is the cell-periodic part of the Bloch state, $\mathbf{G}$ a reciprocal lattice vector, V_{uc} the volume of the unit cell and $N_{\mathbf{k}}$ the number of $\mathbf{k}$ points in the regular mesh the on which the wave functions are computed (equivalently the number of unit cell in the periodic Born–von Karman supercell). Taking the Fourier components $c_{n\mathbf{k}}(\mathbf{G})$ to be orthonormal in n for a fixed

k, i.e.,

$$\sum_{\mathbf{G}} c_{n\mathbf{k}}^{\star}(\mathbf{G})c_{n'\mathbf{k}}(\mathbf{G}) = \delta_{nn'}, \quad \text{(A3)}$$

it immediately follows that $u_{n\mathbf{k}}(\mathbf{r})$ and $\psi_{n\mathbf{k}}(\mathbf{r})$ are orthonormal in the unit cell and crystal supercell, respectively, that is

$$\int_{\text{uc}} d\mathbf{r} u_{n\mathbf{k}}^{\star}(\mathbf{r})u_{n'\mathbf{k}}(\mathbf{r}) = \delta_{nn'} \quad \text{(A4)}$$

and

$$\int_{V_{\mathbf{k}}} d\mathbf{r} \psi_{n\mathbf{k}}^{\star}(\mathbf{r})\psi_{n'\mathbf{k}'}(\mathbf{r}) = \delta_{\mathbf{k}\mathbf{k}'}\delta_{nn'}, \quad \text{(A5)}$$

where the notation $\int_{V_{\mathbf{k}}}$ indicated that the integral is to be performed over $V_{\mathbf{k}} = V_{\text{uc}}N_{\mathbf{k}}$.

2. Exciton States

a. Electron-hole coordinates

The exciton wave function expressed in electron-hole coordinates takes the form

$$\Psi_{S\mathbf{Q}}(\mathbf{r}_e, \mathbf{r}_h) = \sum_{cv\mathbf{k}} A_{cv\mathbf{k}}^{S\mathbf{Q}} \psi_{c\mathbf{k}}(\mathbf{r}_e)\psi_{v\mathbf{k}-\mathbf{Q}}^{\star}(\mathbf{r}_h), \quad \text{(A6)}$$

where $A_{cv\mathbf{k}}^{S\mathbf{Q}}$ are the exciton expansion coefficients. For a fixed **Q**, we take the exciton expansion coefficients to be orthonormal in the band index S, explicitly

$$\sum_{cv\mathbf{k}} A_{cv\mathbf{k}}^{S'\mathbf{Q}\star} A_{cv\mathbf{k}}^{S\mathbf{Q}} = \delta_{SS'}. \quad \text{(A7)}$$

It follows that the exciton wave function as defined in Eq. (A6) is orthonormal in the crystal supercell.

b. Weighted-average and relative coordinates

For the purpose of this work it is convenient re-express the exciton in average **R** and relative **r** coordinates. In the main text, we defined these coordinates as $\mathbf{R} = (\mathbf{r}_e + \mathbf{r}_h)/2$ and $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_h$. Here we consider the more general transformation for **R** given below,

$$\mathbf{R} = \alpha\mathbf{r}_e + \beta\mathbf{r}_e, \quad \text{(A8)}$$

where $\alpha + \beta = 1$ and $\alpha, \beta \geqslant 0$. We now refer to **R** as the weighted-average. In this coordinate system, the exciton wave function can be written in explicit Bloch periodic form,

$$\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{Q}}}} e^{i\mathbf{Q}\cdot\mathbf{R}} F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}), \quad \text{(A9)}$$

where $F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ is the cell-periodic part of the exciton wave function and is given by

$$F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$$
$$= \frac{1}{\sqrt{N_{\mathbf{k}}}} \sum_{cv\mathbf{k}} A_{cv\mathbf{k}+\alpha\mathbf{Q}}^{S\mathbf{Q}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{c\mathbf{k}+\alpha\mathbf{Q}}(\mathbf{R}+\beta\mathbf{r})u_{v\mathbf{k}-\beta\mathbf{Q}}^{\star}(\mathbf{R}-\alpha\mathbf{r}),$$
$$\text{(A10)}$$

where $N_{\mathbf{Q}}$ and $N_{\mathbf{k}}$ are the number of mesh points used in the **Q** and **k** meshes, respectively.

Note by taking $\alpha = \beta = 1/2$, $F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ reduces to the specific form reported in Eq. (4) of the main text. Generally, the functional form of $F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ hinges on ones choice of weighted-average coordinate, i.e., the choice of α, β. This dependence is implicit in the **R** coordinate. To keep the notation light, we will not explicitly label $F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ with an α, β index.

c. Normalization

Below we demonstrate that the exciton wave function as defined above is orthonormal in the **k**/**Q** crystal supercells. By this we mean

$$\int_{V_{\mathbf{Q}}} d\mathbf{R} \int_{V_{\mathbf{k}}} d\mathbf{r} \Psi_{S\mathbf{Q}}^{\star}(\mathbf{R}, \mathbf{r})\Psi_{S'\mathbf{Q}'}(\mathbf{R}, \mathbf{r}) = \delta_{\mathbf{Q}\mathbf{Q}'}\delta_{SS'}, \quad \text{(A11)}$$

where the notation indicates that the integrals over the weighted-average and relative coordinate should be performed over a crystal supercell with volume $V_{\mathbf{Q}} = N_{\mathbf{Q}}V_{\text{uc}}$ and $V_{\mathbf{k}} = N_{\mathbf{k}}V_{\text{uc}}$, respectively. We expect there will be many cases where the **Q** mesh needed for Wannierization will be less dense than the **k** mesh on which the BSE must be diagonalized to obtain converged excitonic states. For this reason we delineate between the two meshes at all stages in the following derivations.

To demonstrate orthonormality it is convenient to further partition $F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ as follows:

$$F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}) = \sum_{cv\mathbf{k}} A_{cv\mathbf{k}+\alpha\mathbf{Q}}^{S\mathbf{Q}} \chi_{cv\mathbf{k}\mathbf{Q}}(\mathbf{R}, \mathbf{r}), \quad \text{(A12)}$$

where

$$\chi_{cv\mathbf{k}\mathbf{Q}}(\mathbf{R}, \mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{k}}}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{c\mathbf{k}+\alpha\mathbf{Q}}(\mathbf{R}+\beta\mathbf{r})u_{v\mathbf{k}-\beta\mathbf{Q}}^{\star}(\mathbf{R}-\alpha\mathbf{r}),$$
$$\text{(A13)}$$

and $\chi_{cv\mathbf{k}\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ is cell-periodic in **R**, explicitly $\chi_{cv\mathbf{k}\mathbf{Q}}(\mathbf{R}+\bar{\mathbf{R}}, \mathbf{r}) = \chi_{cv\mathbf{k}\mathbf{Q}}(\mathbf{R}, \mathbf{r})$.

On the way to demonstrating orthonormality of the exciton wave functions, it is helpful to prove the following result:

$$\int_{\text{uc}} d\mathbf{R} \int_{V_{\mathbf{k}}} d\mathbf{r} \chi_{cv\mathbf{k}\mathbf{Q}}^{\star}(\mathbf{R}, \mathbf{r})\chi_{c'v'\mathbf{k}'\mathbf{Q}'}(\mathbf{R}, \mathbf{r}) = \langle u_{c\mathbf{k}+\alpha\mathbf{Q}}|u_{c'\mathbf{k}+\alpha\mathbf{Q}'}\rangle_{\text{uc}} \langle u_{v'\mathbf{k}-\beta\mathbf{Q}'}|u_{v\mathbf{k}-\beta\mathbf{Q}}\rangle_{\text{uc}} \delta_{\mathbf{k}\mathbf{k}'}, \quad \text{(A14)}$$

which will be used in Appendix B. We make use of the Fourier expansion in Eq. (A2) to express

$$\chi_{cv\mathbf{k}\mathbf{Q}}(\mathbf{R}, \mathbf{r}) = \frac{1}{\sqrt{V_{\mathbf{k}}}} \frac{1}{\sqrt{V_{\text{uc}}}} e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}_c, \mathbf{G}_v} c_{c\mathbf{k}+\alpha\mathbf{Q}}(\mathbf{G}_c)c_{v\mathbf{k}-\beta\mathbf{Q}}^{\star}(\mathbf{G}_v)e^{i\mathbf{G}_c\cdot(\mathbf{R}+\beta\mathbf{r})}e^{-i\mathbf{G}_v\cdot(\mathbf{R}-\alpha\mathbf{r})}. \quad \text{(A15)}$$

Then

$$\int_{\text{uc}} d\mathbf{R} \int_{V_\mathbf{k}} d\mathbf{r}\, \chi^\star_{cv\mathbf{k}\mathbf{Q}}(\mathbf{R}, \mathbf{r}) \chi_{c'v'\mathbf{k}'\mathbf{Q}'}(\mathbf{R}, \mathbf{r}) = \sum_{\mathbf{G}_c, \mathbf{G}_v, \mathbf{G}_{c'}, \mathbf{G}_{v'}} c^\star_{c\mathbf{k}+\alpha\mathbf{Q}}(\mathbf{G}_c) c_{v\mathbf{k}-\beta\mathbf{Q}}(\mathbf{G}_v) c_{c'\mathbf{k}'+\alpha\mathbf{Q}'}(\mathbf{G}_{c'}) c^\star_{v'\mathbf{k}'-\beta\mathbf{Q}'}(\mathbf{G}_{v'})$$

$$\times \frac{1}{V_{\text{uc}}} \int_{\text{uc}} d\mathbf{R} e^{i(-\mathbf{G}_c + \mathbf{G}_v + \mathbf{G}_{c'} - \mathbf{G}_{v'})\cdot\mathbf{R}} \frac{1}{V_\mathbf{k}} \int_{V_\mathbf{k}} d\mathbf{r} e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} e^{i\beta(-\mathbf{G}_c+\mathbf{G}_{c'})\cdot\mathbf{r}} e^{i\alpha(-\mathbf{G}_v+\mathbf{G}_{v'})\cdot\mathbf{r}}$$

$$= \sum_{\mathbf{G}_c, \mathbf{G}_v, \mathbf{G}_{c'}, \mathbf{G}_{v'}} c^\star_{c\mathbf{k}+\alpha\mathbf{Q}}(\mathbf{G}_c) c_{v\mathbf{k}-\beta\mathbf{Q}}(\mathbf{G}_v) c_{c'\mathbf{k}'+\alpha\mathbf{Q}'}(\mathbf{G}_{c'}) c^\star_{v'\mathbf{k}'-\beta\mathbf{Q}'}(\mathbf{G}_{v'}) \delta_{\mathbf{G}_c \mathbf{G}_{c'}} \delta_{\mathbf{G}_v \mathbf{G}_{v'}} \delta_{\mathbf{k}\mathbf{k}'}$$

$$= \langle u_{c\mathbf{k}+\alpha\mathbf{Q}} | u_{c'\mathbf{k}+\alpha\mathbf{Q}'} \rangle_{\text{uc}} \langle u_{v'\mathbf{k}-\beta\mathbf{Q}'} | u_{v\mathbf{k}-\beta\mathbf{Q}} \rangle_{\text{uc}} \delta_{\mathbf{k}\mathbf{k}'}. \tag{A16}$$

From this result and orthonormality of the cell-periodic states [see Eq. (A4)], it immediately follows that

$$\langle \chi_{cv\mathbf{k}\mathbf{Q}} | \chi_{c'v'\mathbf{k}'\mathbf{Q}} \rangle = \delta_{cc'} \delta_{vv'} \delta_{\mathbf{k}\mathbf{k}'}. \tag{A17}$$

We can show that the exciton wave function is normalized to 1 in the crystal supercell

$$\langle \Psi_{S\mathbf{Q}} | \Psi_{S'\mathbf{Q}'} \rangle = \frac{1}{N_\mathbf{Q}} \int_{V_\mathbf{Q}} d\mathbf{R} e^{-i(\mathbf{Q}-\mathbf{Q}')\cdot\mathbf{R}} \int_{V_\mathbf{k}} d\mathbf{r} F^\star_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}) F_{S'\mathbf{Q}'}(\mathbf{R}, \mathbf{r})$$

$$= \delta_{\mathbf{Q}\mathbf{Q}'} \int_{\text{uc}} d\mathbf{R} \int_{V_\mathbf{k}} d\mathbf{r} F^\star_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}) F_{S'\mathbf{Q}}(\mathbf{R}, \mathbf{r})$$

$$= \delta_{\mathbf{Q}\mathbf{Q}'} \sum_{cv\mathbf{k}, c'v'\mathbf{k}'} A^{S\mathbf{Q}\star}_{cv\mathbf{k}+\alpha\mathbf{Q}} A^{S'\mathbf{Q}}_{c'v'\mathbf{k}'+\alpha\mathbf{Q}} \left[\int_{\text{uc}} d\mathbf{R} \int_{V_\mathbf{k}} d\mathbf{r} \chi^\star_{cv\mathbf{k}\mathbf{Q}}(\mathbf{R}, \mathbf{r}) \chi_{c'v'\mathbf{k}'\mathbf{Q}}(\mathbf{R}, \mathbf{r}) \right]$$

$$= \delta_{\mathbf{Q}\mathbf{Q}'} \sum_{cvc'v'\mathbf{k}} A^{S\mathbf{Q}\star}_{cv\mathbf{k}+\alpha\mathbf{Q}} A^{S'\mathbf{Q}}_{c'v'\mathbf{k}'+\alpha\mathbf{Q}} \delta_{cc'} \delta_{vv'} \delta_{\mathbf{k}\mathbf{k}'}$$

$$= \delta_{\mathbf{Q}\mathbf{Q}'} \delta_{SS'},$$

where we have used the cell-periodicity of $F_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r})$ in $\mathbf{R}$ to reduce the integral over $V_\mathbf{Q}$ to an integral in the unit cell and in passing from line 3 to line 4, we have used Eq. (A17).

For later use it will also be helpful to note that

$$\int_{V_\mathbf{Q}} \Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{0}) d\mathbf{R} = \int_{\text{uc}} F_{S\mathbf{0}}(\mathbf{R}, \mathbf{0}) d\mathbf{R} = \frac{1}{\sqrt{N_\mathbf{k}}} \sum_{cv\mathbf{k}} A^{S\mathbf{0}}_{cv\mathbf{k}} \int_{\text{uc}} u_{c\mathbf{k}}(\mathbf{R}) u^\star_{v\mathbf{k}}(\mathbf{R}) d\mathbf{R} = 0. \tag{A18}$$

APPENDIX B: DERIVATION OF EXCITON OVERLAPS

Because the exciton wave function has a Bloch periodic structure when expressed in $\mathbf{R} + \alpha\mathbf{r}_e + \beta\mathbf{r}_h$, it is possible to Wannierize in $\mathbf{R} = \alpha\mathbf{r}_e + \beta\mathbf{r}_h$ for any $\alpha, \beta \geqslant 0$ pair which sum to 1. In other words one can minimize the sum of the spreads of the exciton Wannier function, i.e.,

$$\Omega^{\text{Xct}}[U; \alpha, \beta] = \sum_{M}^{J} [\langle M\bar{\mathbf{0}} | R^2 | M\bar{\mathbf{0}} \rangle - \langle M\bar{\mathbf{0}} | \mathbf{R} | M\bar{\mathbf{0}} \rangle^2] \tag{B1}$$

for any $\mathbf{R} = \alpha\mathbf{r}_e + \beta\mathbf{r}_h$. In the main text we choose $\alpha = \beta = 1/2$ for its conceptual simplicity and computational tractability. Below we derive the relevant overlaps which must be passed to `Wannier90` to Wannierize for arbitrary $\mathbf{R} + \alpha\mathbf{r}_e + \beta\mathbf{r}_h$.

Here it is helpful to explicitly label quantities by their (α, β) pair. To Wannierize in the coordinate $\mathbf{R} = \alpha\mathbf{r}_e + \beta\mathbf{r}_h$ we must pass to `Wannier90` overlaps of the form

$$M^{(\alpha,\beta)}_{SS'}(\mathbf{Q}, \mathbf{B}) = \langle F^{(\alpha,\beta)}_{S\mathbf{Q}} | F^{(\alpha,\beta)}_{S'\mathbf{Q}+\mathbf{B}} \rangle. \tag{B2}$$

Starting from Eq. (A12) we find

$$M^{(\alpha,\beta)}_{SS'}(\mathbf{Q}, \mathbf{B}) = \sum_{cv\mathbf{k}, c'v'\mathbf{k}'} A^{S\mathbf{Q}\star}_{cv\mathbf{k}+\alpha\mathbf{Q}} A^{S'\mathbf{Q}+\mathbf{B}}_{c'v'\mathbf{k}'+\alpha\mathbf{Q}+\alpha\mathbf{B}} \left[\int_{\text{uc}} d\mathbf{R} \int_{V_\mathbf{k}} d\mathbf{r} \chi^{(\alpha,\beta)\star}_{cv\mathbf{k}\mathbf{Q}}(\mathbf{R}, \mathbf{r}) \chi^{(\alpha,\beta)}_{c'v'\mathbf{k}'\mathbf{Q}+\mathbf{B}}(\mathbf{R}, \mathbf{r}) \right]$$

$$= \sum_{cvc'v'\mathbf{k}} A^{S\mathbf{Q}\star}_{cv\mathbf{k}+\alpha\mathbf{Q}} A^{S'\mathbf{Q}+\mathbf{B}}_{c'v'\mathbf{k}'+\alpha\mathbf{Q}+\alpha\mathbf{B}} \langle u_{c\mathbf{k}+\alpha\mathbf{Q}} | u_{c'\mathbf{k}+\alpha\mathbf{Q}+\alpha\mathbf{B}} \rangle_{\text{uc}} \langle u_{v'\mathbf{k}-\beta\mathbf{Q}-\beta\mathbf{B}} | u_{v\mathbf{k}-\beta\mathbf{Q}} \rangle_{\text{uc}} \delta_{\mathbf{k}\mathbf{k}'}$$

$$= \sum_{cvc'v'\mathbf{k}} A^{S\mathbf{Q}\star}_{cv\mathbf{k}} A^{S'\mathbf{Q}+\mathbf{B}}_{c'v'\mathbf{k}+\alpha\mathbf{B}} \langle u_{c\mathbf{k}} | u_{c'\mathbf{k}+\alpha\mathbf{B}} \rangle_{\text{uc}} \langle u_{v'\mathbf{k}-\mathbf{Q}-\beta\mathbf{B}} | u_{v\mathbf{k}-\mathbf{Q}} \rangle_{\text{uc}}. \tag{B3}$$

In going from line 1 to line 2, we have made use of Eq. (A16) and in the last line we have made a change of variables, $\mathbf{k} \rightarrow \mathbf{k} - \alpha \mathbf{Q}$. To recover Eq. (12) in the main text, take $\alpha = \beta = 1/2$.

It is important to understand that while different choices of α, β will change the functional which is minimized, it may happen that the $U_{SM}(\mathbf{Q})$ which minimize $\Omega_{\alpha\beta}^{\text{Xct}}[U; \alpha, \beta]$ are independent of α, β. For LiF we have numerically verified that the choices $\alpha = 1, \beta = 0, \alpha = 1/2, \beta = 1/2$, and $\alpha = 0, \beta = 1$ all give the same Wannier rotation matrices. Only the spread $\Omega[U; \alpha, \beta]$ differs with choice of α and β.

For singlet excitons we find it is also necessary to provide Wannier90 with a starting guess. Here we use already Wannierized triplet states as a starting guess for Wannierization of the singlet excitons. To do so, we pass Wannier90 the following overlaps:

$$A_{SM}^{(\mathbf{Q})} = \langle \Psi_{S\mathbf{Q}} | W_{M\bar{\mathbf{0}}} \rangle = \sum_{cv\mathbf{k}T} A_{cv\mathbf{k}}^{S\mathbf{Q}\star} A_{cv\mathbf{k}}^{T\mathbf{Q}} U_{TM}(\mathbf{Q}). \tag{B4}$$

APPENDIX C: DERIVATION OF DIRECT AND EXCHANGE TERMS IN THE WANNIER BASIS

1. Electron-hole kernel in the MLXWF basis

To derive the direct and exchange terms we work in electron-hole coordinates and subsequently transform to average/relative coordinates. Both the exchange and direct electron-hole kernel take the form

$$K_{MN}(\bar{\mathbf{R}}) = \int W_{M\bar{\mathbf{0}}}^{\star}(\mathbf{r}_e, \mathbf{r}_h) K(\mathbf{r}_e, \mathbf{r}_h, \mathbf{r}_e', \mathbf{r}_h') W_{N\bar{\mathbf{R}}}(\mathbf{r}_e', \mathbf{r}_h') d[\mathbf{r}], \tag{C1}$$

where $d[\mathbf{r}] = d\mathbf{r}_e d\mathbf{r}_h d\mathbf{r}_e' d\mathbf{r}_h'$ and the integrals are over the supercells associated with the electron and hole $\mathbf{k}$-point meshes. In this real space electron-hole basis, the screened direct term takes the form

$$K^{\text{D}}(\mathbf{r}_e, \mathbf{r}_h; \mathbf{r}_e, \mathbf{r}_h') = \varepsilon^{-1}(\mathbf{r}_e, \mathbf{r}_h) v(\mathbf{r}_e, \mathbf{r}_h) \delta(\mathbf{r}_e, \mathbf{r}_e') \delta(\mathbf{r}_h, \mathbf{r}_h'), \tag{C2}$$

while the exchange operator takes the form

$$K^{\text{X}}(\mathbf{r}_e, \mathbf{r}_h; \mathbf{r}_e', \mathbf{r}_h') = v(\mathbf{r}_e, \mathbf{r}_h) \delta(\mathbf{r}_e, \mathbf{r}_h) \delta(\mathbf{r}_e', \mathbf{r}_h'). \tag{C3}$$

Substituting these expressions in we find

$$K_{MN}^{\text{D}}(\bar{\mathbf{R}})$$
$$= \int W_{M\bar{\mathbf{0}}}^{\star}(\mathbf{r}_e, \mathbf{r}_h) \varepsilon^{-1}(\mathbf{r}_e, \mathbf{r}_h) v(\mathbf{r}_e, \mathbf{r}_h) W_{N\bar{\mathbf{R}}}(\mathbf{r}_e, \mathbf{r}_h) d\mathbf{r}_e d\mathbf{r}_h \tag{C4}$$

and

$$K_{MN}^{\text{X}}(\bar{\mathbf{R}}) = \int W_{M\bar{\mathbf{0}}}^{\star}(\mathbf{r}_e, \mathbf{r}_e) v(\mathbf{r}_e, \mathbf{r}_h) W_{N\bar{\mathbf{R}}}(\mathbf{r}_h, \mathbf{r}_h) d\mathbf{r}_e d\mathbf{r}_h. \tag{C5}$$

When re-expressed in the weighted-average $\mathbf{R} = \alpha \mathbf{r}_e + \beta \mathbf{r}_h$ and relative $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_h$ coordinate $W(\mathbf{r}_e, \mathbf{r}_h) \rightarrow W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r})$ while $W(\mathbf{r}_e, \mathbf{r}_e) \rightarrow W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{0})$. With these substitutions, and adopting the normalization scheme described in Appendix A,

we find

$$K_{MN}^{\text{D}}(\bar{\mathbf{R}}) = \int_{V_{\mathbf{Q}}} d\mathbf{R} \int_{V_{\mathbf{k}}} d\mathbf{r} \, W_{M\mathbf{0}}^{\star}(\mathbf{R}, \mathbf{r}) \varepsilon^{-1}(\mathbf{R}, \mathbf{r}) \frac{e^2}{r} W_{N\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r}) \tag{C6}$$

and

$$K_{MN}^{\text{X}}(\bar{\mathbf{R}}) = \int_{V_{\mathbf{Q}}} d\mathbf{R} \int_{V_{\mathbf{Q}}} d\mathbf{R}' \, W_{M\mathbf{0}}^{\star}(\mathbf{R}, \mathbf{0}) \frac{e^2}{|\mathbf{R} - \mathbf{R}'|} W_{N\bar{\mathbf{R}}}(\mathbf{R}', \mathbf{0}), \tag{C7}$$

which are equivalent to Eqs. (21) and (22) reported in the main text.

2. Leading-order expansion of the exchange interaction

Here we provide details on the lowest order expansion of $K_{MN}^{\text{X}}(\bar{\mathbf{R}})$ in $\bar{\mathbf{R}}$.

a. Real space

We start from the exchange interaction in the MLXWF basis

$$K_{MN}^{\text{X}}(\bar{\mathbf{R}}) = \int W_{M\mathbf{0}}^{\star}(\mathbf{R}, \mathbf{0}) \frac{e^2}{|\mathbf{R} - \mathbf{R}'|} W_{N\bar{\mathbf{R}}}(\mathbf{R}', \mathbf{0}) d\mathbf{R} d\mathbf{R}'$$
$$= \int W_{M\mathbf{0}}^{\star}(\mathbf{R}, \mathbf{0}) \frac{e^2}{|\mathbf{R} - \mathbf{R}' - \bar{\mathbf{R}}|} W_{N\bar{\mathbf{0}}}(\mathbf{R}', \mathbf{0}) d\mathbf{R} d\mathbf{R}', \tag{C8}$$

where we have made a change of integration $\mathbf{R}' \rightarrow \mathbf{R}' + \bar{\mathbf{R}}$ and used the identity $W_{M\bar{\mathbf{R}}}(\mathbf{R} + \bar{\mathbf{R}}, \mathbf{r}) = W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r})$ to arrive at the second line.

To derive the leading order contribution to this integral we make a multipole expansion in Coulomb interaction. The lowest order term in this expansion is

$$\frac{e^2}{|\mathbf{R} - \mathbf{R}' - \bar{\mathbf{R}}|} = e^2 \left[\frac{\mathbf{R} \cdot \mathbf{R}'}{|\bar{\mathbf{R}}|^3} - 3 \frac{(\mathbf{R} \cdot \bar{\mathbf{R}})(\mathbf{R}' \cdot \bar{\mathbf{R}})}{|\bar{\mathbf{R}}|^5} \right] + \cdots \tag{C9}$$

Plugging this first-order expansion into Eq. (C8), we arrive at the dipolar kernel [see Eq. (23) in the main text]

$$K_{MN}^{\text{X,Dip}}(\bar{\mathbf{R}}) = e^2 \left[\frac{\mathbf{P}_M^{\star} \cdot \mathbf{P}_N}{|\bar{\mathbf{R}}|^3} - 3 \frac{(\mathbf{P}_M^{\star} \cdot \bar{\mathbf{R}})(\mathbf{P}_N \cdot \bar{\mathbf{R}})}{|\bar{\mathbf{R}}|^5} \right], \tag{C10}$$

where

$$\mathbf{P}_M = \int_{V_{\mathbf{Q}}} \mathbf{R} W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{0}) d\mathbf{R}. \tag{C11}$$

b. Reciprocal space

In Eq. (27) of the main text we give the Fourier transform of the dipolar interaction. To derive this result we start with the Fourier expansion of the Coulomb interaction

$$\frac{e^2}{|\mathbf{R} - \mathbf{R}' - \bar{\mathbf{R}}|} = \frac{4\pi e^2}{V_{\mathbf{q}}} \sum_{\mathbf{q}\mathbf{G}} e^{+i(\mathbf{q}+\mathbf{G})\cdot\mathbf{R}}$$
$$\times \frac{1}{|\mathbf{q} + \mathbf{G}|^2} e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{R}'} e^{-i\mathbf{q}\cdot\bar{\mathbf{R}}}, \tag{C12}$$

where $V_{\mathbf{q}} = N_{\mathbf{q}} V_{\text{uc}}$ and we have used $e^{i\mathbf{G}\cdot\bar{\mathbf{R}}} = 1$.

Multiplying both sides by $e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}}$ and summing over $\bar{\mathbf{R}}$, we find

$$\sum_{\bar{\mathbf{R}}} \frac{e^2}{|\mathbf{R} - \mathbf{R}' - \bar{\mathbf{R}}|} e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}} = \frac{4\pi e^2}{V_{uc}} \sum_{\mathbf{G}} e^{+i(\mathbf{Q}+\mathbf{G})\cdot\mathbf{R}} \frac{1}{|\mathbf{Q}+\mathbf{G}|^2} e^{-i(\mathbf{Q}+\mathbf{G})\cdot\mathbf{R}'}, \tag{C13}$$

where we have used $\sum_{\bar{\mathbf{R}}} e^{i(\mathbf{Q}-\mathbf{q})\cdot\bar{\mathbf{R}}} = N_{\mathbf{Q}}\delta_{\mathbf{Qq}}$.

Plugging Eq. (C13) into Eq. (C8) we obtain

$$\sum_{\bar{\mathbf{R}}}\left[\int W_{M\bar{\mathbf{0}}}^{\star}(\mathbf{R},\mathbf{0})\frac{e^2}{|\mathbf{R}-\mathbf{R}'-\bar{\mathbf{R}}|}W_{N\bar{\mathbf{0}}}(\mathbf{R}',\mathbf{0})d\mathbf{R}d\mathbf{R}'\right]e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}}$$

$$= \frac{4\pi e^2}{V_{uc}}\sum_{\mathbf{G}}\left[\int d\mathbf{R}W_{M\bar{\mathbf{0}}}^{\star}(\mathbf{R},\mathbf{0})e^{i(\mathbf{Q}+\mathbf{G})\cdot\mathbf{R}}\right]\frac{1}{|\mathbf{Q}+\mathbf{G}|^2}\left[\int d\mathbf{R}'e^{-i(\mathbf{Q}+\mathbf{G})\cdot\mathbf{R}'}W_{N\bar{\mathbf{0}}}(\mathbf{R}',\mathbf{0})\right]. \tag{C14}$$

To arrive at the dipolar contribution we expand both sides to lowest order in $\mathbf{R}$ and $\mathbf{R}'$. We find that

$$e^2\sum_{\bar{\mathbf{R}}}\left[\frac{\mathbf{P}_M^{\star}\cdot\mathbf{P}_N}{|\bar{\mathbf{R}}|^3} - \frac{(\mathbf{P}_M^{\star}\cdot\bar{\mathbf{R}})(\mathbf{P}_N\cdot\bar{\mathbf{R}})}{|\bar{\mathbf{R}}|^5}\right]e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}} = \frac{4\pi e^2}{V_{uc}}\sum_{\mathbf{G}}\frac{[\mathbf{P}_M^{\star}\cdot(\mathbf{Q}+\mathbf{G})][\mathbf{P}_N\cdot(\mathbf{Q}+\mathbf{G})]}{|\mathbf{Q}+\mathbf{G}|^2}, \tag{C15}$$

where MLXWF dipole moments, $\mathbf{P}_M$, are defined in Eq. (C11).

3. Exciton transition dipole matrix elements and their relations

The MLXWF transition dipole matrix elements naturally appear when making the multipole expansion of the exchange interaction. Here we discuss properties of these dipole matrix elements and their relation to the usual dipole matrix elements used to study optical properties.

In our multipole expansion we found that only the MLXWF dipoles at $\bar{\mathbf{R}} = \bar{\mathbf{0}}$ were required. Nevertheless in the discussion below it is helpful to consider the MLXWF dipole matrix element at nonzero $\bar{\mathbf{R}}$, defined as

$$\mathbf{P}_{M\bar{\mathbf{R}}} = \int_{V_{\mathbf{Q}}} \mathbf{R}W_{M\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r}=\mathbf{0})d\mathbf{R}. \tag{C16}$$

We also introduce the Bloch exciton dipole matrix element at nonzero $\mathbf{Q}$, defined as

$$\mathbf{P}_{S\mathbf{Q}} = \int_{V_{\mathbf{Q}}} \mathbf{R}\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r}=\mathbf{0})d\mathbf{R}. \tag{C17}$$

Below we will prove

$$\mathbf{P}_{S\mathbf{Q}} = \mathbf{P}_{S\mathbf{0}}\delta_{\mathbf{Q0}}, \tag{C18}$$

$$\mathbf{P}_{M\bar{\mathbf{R}}} = \mathbf{P}_{M\bar{\mathbf{0}}}. \tag{C19}$$

These results can be derived by considering how the Bloch and Wannier exciton dipole moments transform under translation in the weighted-average coordinate.

For the Bloch exciton dipole moment we have

$$\mathbf{P}_{S\mathbf{Q}} = \int_{V_{\mathbf{Q}}} \mathbf{R}\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{0})\,d\mathbf{R}$$

$$= \int_{V_{\mathbf{Q}}} (\mathbf{R}+\bar{\mathbf{R}})\Psi_{S\mathbf{Q}}(\mathbf{R}+\bar{\mathbf{R}}, \mathbf{0})\,d\mathbf{R}$$

$$= \int_{V_{\mathbf{Q}}} (\mathbf{R}+\bar{\mathbf{R}})\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{0})e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}}\,d\mathbf{R}$$

$$= e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}}\int_{V_{\mathbf{Q}}} \mathbf{R}\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{0})d\mathbf{R} + \bar{\mathbf{R}}e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}}\int_{V_{\mathbf{Q}}} \Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{0})\,d\mathbf{R}$$

$$= e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}}\int_{V_{\mathbf{Q}}} \mathbf{R}\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{0})d\mathbf{R} + \mathbf{0}$$

$$= e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}}\mathbf{P}_{S\mathbf{Q}}, \tag{C20}$$

where we have shifted the integration variable by a lattice constant, applied Bloch's theorem, and used Eq. (A18) to arrive at the final line. To complete the argument, note that the only way for $\mathbf{P}_{S\mathbf{Q}} = e^{i\mathbf{Q}\cdot\bar{\mathbf{R}}}\mathbf{P}_{S\mathbf{Q}}$ to hold for arbitrary $\bar{\mathbf{R}}$ is for $\mathbf{P}_{S\mathbf{Q}} = \mathbf{P}_{S\mathbf{0}}\delta_{\mathbf{Q0}}$ which proves Eq. (C18).

Along similar lines, we have

$$\mathbf{P}_{M\bar{\mathbf{R}}} = \int_{V_{\mathbf{Q}}} \mathbf{R}W_{M\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{0})d\mathbf{R}$$

$$= \int_{V_{\mathbf{Q}}} (\mathbf{R}+\bar{\mathbf{R}})W_{M\bar{\mathbf{R}}}(\mathbf{R}+\bar{\mathbf{R}}, \mathbf{0})d\mathbf{R}$$

$$= \int_{V_{\mathbf{Q}}} \mathbf{R}W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{0})d\mathbf{R} + \bar{\mathbf{R}}\int_{V_{\mathbf{Q}}} W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{0})d\mathbf{R}$$

$$= \int_{V_{\mathbf{Q}}} \mathbf{R}W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{0})d\mathbf{R} + \mathbf{0}$$

$$= \mathbf{P}_{M\bar{\mathbf{0}}}, \tag{C21}$$

where we have used the relation $W_{M\bar{\mathbf{R}}}(\mathbf{R}+\bar{\mathbf{R}}, \mathbf{r}) = W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r})$ and $\int W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r}=\mathbf{0})d\mathbf{R} = 0$, this last relation follows from Eq. (A18).

With these results we can show that the Bloch and MLXWF transition dipole matrix elements are related through the Wannier rotation matrices at $\mathbf{Q} = \mathbf{0}$. We start from

$$\Psi_{S\mathbf{0}}(\mathbf{R}, \mathbf{r}=\mathbf{0}) = \frac{1}{\sqrt{N_{\mathbf{Q}}}}\sum_{M\bar{\mathbf{R}}}U_{MS}^{\dagger}(\mathbf{0})W_{M\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r}=\mathbf{0}), \tag{C22}$$

which follows from Eq. (13), and integrate both sides against $\mathbf{R}$ to find

$$\mathbf{P}_{S\mathbf{0}} = \frac{1}{\sqrt{N_{\mathbf{Q}}}}\sum_{M\bar{\mathbf{R}}}U_{MS}^{\dagger}(\mathbf{0})\mathbf{P}_{M\bar{\mathbf{R}}} = \sqrt{N_{\mathbf{Q}}}\sum_{M}U_{MS}^{\dagger}(\mathbf{0})\mathbf{P}_{M\bar{\mathbf{0}}}, \tag{C23}$$

where we have used $\mathbf{P}_{M\bar{\mathbf{R}}} = \mathbf{P}_{M\bar{\mathbf{0}}}$ [see Eq. (C19)] and the trivial relation $\sum_{\bar{\mathbf{R}}} 1 = N_{\mathbf{Q}}$. Equation (C23) can be directly inverted to express $\mathbf{P}_{M\bar{\mathbf{0}}}$ in terms of $\mathbf{P}_{S0}$. Alternatively, we can derive the inverse relation starting from

$$W_{M\bar{\mathbf{0}}}(\mathbf{R}, \mathbf{r} = \mathbf{0}) = \frac{1}{\sqrt{N_{\mathbf{Q}}}} \sum_{S\mathbf{Q}} U_{SM}(\mathbf{Q})\Psi_{S\mathbf{Q}}(\mathbf{R}, \mathbf{r} = \mathbf{0}), \quad (C24)$$

which follows from Eq. (9) in the main text. Again we integrate both sides with respect to $\mathbf{R}$ to find

$$\mathbf{P}_{M\bar{\mathbf{0}}} = \frac{1}{\sqrt{N_{\mathbf{Q}}}} \sum_{S\mathbf{Q}} U_{SM}(\mathbf{Q})\mathbf{P}_{S\mathbf{Q}} = \frac{1}{\sqrt{N_{\mathbf{Q}}}} \sum_{S} U_{SM}(\mathbf{0})\mathbf{P}_{S0}, \quad (C25)$$

where we have used $\mathbf{P}_{S\mathbf{Q}} = \mathbf{P}_{S0}\delta_{\mathbf{Q}\mathbf{0}}$ [see Eq. (C18)]. Equations (C23) and (C25) coincide with Eq. (26) in the main text.

4. Exciton position matrix element

The exciton transition dipole matrix elements should be distinguished from the expectation value of the position operator. The latter, in the MLXWF basis, is defined as

$$\langle \mathbf{R} \rangle_{M\bar{\mathbf{R}}} = \int_{V_{\mathbf{Q}}} d\mathbf{R} \int_{V_{\mathbf{k}}} d\mathbf{r} W_{M\bar{\mathbf{R}}}^{\star}(\mathbf{R}, \mathbf{r})\mathbf{R}W_{M\bar{\mathbf{R}}}(\mathbf{R}, \mathbf{r}). \quad (C26)$$

Importantly, the position operator matrix elements of MLXWFs centered on different cells are related by the addition of a lattice vector $\bar{\mathbf{R}}$. Explicitly,

$$\langle \mathbf{R} \rangle_{M\bar{\mathbf{R}}} = \langle \mathbf{R} \rangle_{M} + \bar{\mathbf{R}}. \quad (C27)$$

This relation is analogous to what is seen at the one-electron level and plays an important role in many post-processing applications where a consistent choice of $\bar{\mathbf{R}}$ is crucial to obtain correct results, e.g., the Berry phase theory of polarization. Equation (C27) should be contrasted with Eq. (C21), notably the dipole matrix elements are invariant under translation.

APPENDIX D: DERIVATION OF THE EFFECTIVE NONANALYTIC KERNEL

a. Downfolded nonanalytic kernel

In this section we derive the downfolded nonanalytic kernel [see Eq. (47)]. Our starting point is the infinite series in Eq. (45),

$$K_{SS'}^{\text{NA,eff}}(\mathbf{Q}) = K_{SS'}^{\text{NA}}(\mathbf{Q}) + 2 \sum_{R \notin \mathcal{W}} \frac{K_{SR}^{\text{NA}}(\mathbf{Q})K_{RS'}^{\text{NA}}(\mathbf{Q})}{\omega - E_{R0}}$$

$$+ 2^2 \sum_{R,R' \notin \mathcal{W}} \frac{K_{SR}^{\text{NA}}(\mathbf{Q})K_{RR'}^{\text{NA}}(\mathbf{Q})K_{R'S'}^{\text{NA}}(\mathbf{Q})}{(\omega - E_{R0})(\omega - E_{R'0})} + \cdots, \quad (D1)$$

where $K_{SS'}^{\text{NA}}(\mathbf{Q})$, given in Eq. (31), is reported here for clarity:

$$K_{SS'}^{\text{NA}}(\mathbf{Q}) = \frac{4\pi e^2}{V_{\text{uc}}N_{\mathbf{Q}}} \frac{[\mathbf{P}_S^{\star} \cdot \mathbf{Q}][\mathbf{P}_{S'} \cdot \mathbf{Q}]}{|\mathbf{Q}|^2}. \quad (D2)$$

The separable form of $K_{SS'}^{\text{NA}}(\mathbf{Q})$, allows Eq. (D1) to be summed to infinite order. We find

$$K_{SS'}^{\text{NA,eff}}(\mathbf{Q}) = K_{SS'}^{\text{NA}}(\mathbf{Q})[1 + 2v(\mathbf{Q})\chi^{\text{BG}}(\mathbf{Q}) + [2v(\mathbf{Q})\chi^{\text{BG}}(\mathbf{Q})]^2 + \cdots], \quad (D3)$$

where the common ratio,

$$2v(\mathbf{Q})\chi_{\text{BG}}(\mathbf{Q}) = 2\frac{4\pi e^2}{V_{\text{uc}}N_{\mathbf{Q}}|\mathbf{Q}|^2} \sum_{R \notin \mathcal{W}} \frac{[\mathbf{P}_R \cdot \mathbf{Q}][\mathbf{P}_R^{\star} \cdot \mathbf{Q}]}{\omega - E_{R0}}, \quad (D4)$$

can physically be interpreted as the product of the Coulomb interaction, $v(\mathbf{Q})$, and background susceptibility, $\chi_{\text{BG}}(\mathbf{Q})$, in the $\mathbf{Q} \to \mathbf{0}$ limit. Equation (D3) sums to

$$K_{SS'}^{\text{NA,eff}}(\mathbf{Q}) = \varepsilon_{\text{BG}}^{-1}(\mathbf{Q}, \omega)K_{SS'}^{\text{NA}}(\mathbf{Q}), \quad (D5)$$

where

$$\varepsilon_{\text{BG}}(\mathbf{Q}, \omega) = 1 - 2v(\mathbf{Q})\chi_{\text{BG}}(\mathbf{Q}, \omega). \quad (D6)$$

We see that states in the passive space dynamically screen the nonanalytic coupling between active space states.

b. Static nonanalytic kernel

In this section we show how to cast the frequency-dependent effective Hamiltonian [see Eq. (49)] as a static non-Hermitian Hamiltonian [see Eqs. (50)–(52)] in the MLXWF basis. Our starting point is Eq. (49), reproduced here for clarity:

$$\lim_{\mathbf{Q} \to \mathbf{0}} H_{MN}^{\text{Xct,eff}}(\mathbf{Q}, \omega) = H_{MN}^{\text{SR}}(\mathbf{0}) + 2\delta_S K_{MN}^{\text{NA,eff}}(\mathbf{Q}, \omega), \quad (D7)$$

where

$$K_{MN}^{\text{NA,eff}}(\mathbf{Q}, \omega) = \varepsilon_{\text{BG}}^{-1}(\mathbf{Q}, \omega)K_{MN}^{\text{NA}}(\mathbf{Q}), \quad (D8)$$

with $K_{MN}^{\text{NA}}(\mathbf{Q})$ defined in Eq. (29) in the main text. Let $E_{S\mathbf{Q}}$ and $C_{MS}(\mathbf{Q})$ denote the eigensolutions of Eq. (D7) so that

$$\sum_N \lim_{\mathbf{Q} \to \mathbf{0}} H_{MN}^{\text{Xct,eff}}(\mathbf{Q}, E_{S\mathbf{Q}})C_{NS}(\mathbf{Q}) = E_{S\mathbf{Q}}C_{MS}(\mathbf{Q}). \quad (D9)$$

Define

$$\lim_{\mathbf{Q} \to \mathbf{0}} H_{MN}^{\text{Xct,eff}}(\mathbf{Q})$$

$$= \sum_{N'} \sum_{S'} \lim_{\mathbf{Q} \to \mathbf{0}} H_{MN'}^{\text{Xct,eff}}(\mathbf{Q}, E_{S'\mathbf{Q}})C_{N'S'}(\mathbf{Q})C_{S'N}^{-1}(\mathbf{Q}). \quad (D10)$$

Note that by construction $\lim_{\mathbf{Q} \to \mathbf{0}} H_{MN}^{\text{Xct,eff}}(\mathbf{Q})$ has the same eigenvalues and eigenvectors as Eq. (D9).

Plugging Eqs. (D7) and (D8) into Eq. (D10) we find

$$\lim_{\mathbf{Q} \to \mathbf{0}} H_{MN}^{\text{Xct,eff}}(\mathbf{Q}) = H_{MN}^{\text{SR}}(\mathbf{0}) + 2\delta_S K_{MN}^{\text{NA,eff}}(\mathbf{Q}), \quad (D11)$$

where

$$K_{MN}^{\text{NA,eff}}(\mathbf{Q}) = \sum_{N'} K_{MN'}^{\text{NA}}(\mathbf{Q})[\varepsilon_{\text{BG}}^{-1}(\mathbf{Q})]_{N'N} \quad (D12)$$

and

$$[\varepsilon_{\text{BG}}^{-1}(\mathbf{Q})]_{N'N} = \sum_S C_{N'S}(\mathbf{Q})\varepsilon_{\text{BG}}^{-1}(\mathbf{Q}, E_{S\mathbf{Q}})C_{SN}^{-1}(\mathbf{Q}). \quad (D13)$$

Equations (D11)–(D13) are equivalent to those given in the main text [see Eqs. (50)–(52)].

Note, because $C_{SN}(\mathbf{Q})$ are obtained by diagonalizing Eq. (D7) they are unique only up to a phase. For example, if $C_{NS}(\mathbf{Q})$ satisfies Eq. (D7), then so will $e^{i\phi_S(\mathbf{Q})}C_{NS}(\mathbf{Q})$. In Eq. (D13) the arbitrary phases associated with $C_{N'S}(\mathbf{Q})$ and $C_{SN}^{-1}(\mathbf{Q})$ cancel so that $[\varepsilon_{\text{BG}}^{-1}(\mathbf{Q})]_{N'N}$ is well defined.

APPENDIX E: EWALD SUMMATION

The reciprocal space dipolar kernel, $K^{\mathrm{X,Dip}}(\mathbf{Q})$, defined in Eq. (27) as

$$K_{MN}^{\mathrm{X,Dip}}(\mathbf{Q}) = \frac{4\pi e^2}{V_{\mathrm{uc}}} \sum_{\mathbf{G}} \frac{[\mathbf{P}_M^{\star} \cdot (\mathbf{Q} + \mathbf{G})][\mathbf{P}_N \cdot (\mathbf{Q} + \mathbf{G})]}{|\mathbf{Q} + \mathbf{G}|^2},$$

(E1)

formally diverges. This divergence stems from the self-interaction ($\bar{\mathbf{R}} = \mathbf{0}$) term in the real space dipolar kernel, $K^{\mathrm{X,Dip}}(\bar{\mathbf{R}})$ [see Eq. (23)]. By contrast, the long-range kernel, defined in Eq. (34) as

$$K_{MN}^{\mathrm{LR}}(\mathbf{Q}) = \frac{4\pi e^2}{V_{\mathrm{uc}}} \sum_{\mathbf{G}} \frac{[\mathbf{P}_M^{\star} \cdot (\mathbf{Q} + \mathbf{G})][\mathbf{P}_N \cdot (\mathbf{Q} + \mathbf{G})]}{|\mathbf{Q} + \mathbf{G}|^2}$$
$$- \frac{4\pi e^2}{V_{\mathrm{uc}}} \sum_{\mathbf{G} \neq \mathbf{0}} \frac{[\mathbf{P}_M^{\star} \cdot \mathbf{G}][\mathbf{P}_N \cdot \mathbf{G}]}{|\mathbf{G}|^2},$$

(E2)

is convergent. Still, the convergence with respect to the $\mathbf{G}$ sum is slow, a consequence of subtracting two divergent quantities from one another. To efficiently evaluate $K_{MN}^{\mathrm{LR}}(\mathbf{Q})$, we make use of Ewald's technique [50]. We partition $K_{MN}^{\mathrm{X,Dip}}(\mathbf{Q})$ into a rapidly converging reciprocal and real-space sum [see, for instance, Eq. (73) of Ref. [52] for additional details on this partitioning].

For a sufficiently large range separation parameter, Λ, the real-space contribution to the Ewald sum can be made negligible. Retaining only the reciprocal space contribution we arrive at the following approximation to $K_{MN}^{\mathrm{LR}}(\mathbf{Q})$:

$$K_{MN}^{\mathrm{LR}}(\mathbf{Q})$$
$$= \frac{4\pi e^2}{V_{\mathrm{uc}}} \left[\sum_{\mathbf{G}} \frac{[\mathbf{P}_M^{\star} \cdot (\mathbf{Q} + \mathbf{G})][\mathbf{P}_N \cdot (\mathbf{Q} + \mathbf{G})]}{|\mathbf{Q} + \mathbf{G}|^2} e^{-(\mathbf{Q}+\mathbf{G})^2/(4\Lambda^2)} \right.$$
$$\left. - \sum_{\mathbf{G} \neq \mathbf{0}} \frac{[\mathbf{P}_M^{\star} \cdot \mathbf{G}][\mathbf{P}_N \cdot \mathbf{G}]}{|\mathbf{G}|^2} e^{-(\mathbf{G})^2/(4\Lambda^2)} \right].$$

(E3)

Note that for $\Lambda \to \infty$ Eq. (E3) reduces to Eq. (E2) and is exact. For sufficiently large Λ, Eq. (E3) gives a very good approximation to Eq. (E2) and converges much faster, so that only a few $\mathbf{G}$-shells need to be included. For LiF we take Λ to be one inverse Bohr.

[1] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[2] A. J. Williamson, R. Q. Hood, and J. C. Grossman, Phys. Rev. Lett. **87**, 246406 (2001).

[3] X. Wang, J. R. Yates, I. Souza, and D. Vanderbilt, Phys. Rev. B **74**, 195118 (2006).

[4] J. Ibañez-Azpiroz, S. S. Tsirkin, and I. Souza, Phys. Rev. B **97**, 245143 (2018).

[5] S. S. Tsirkin, P. A. Puente, and I. Souza, Phys. Rev. B **97**, 035158 (2018).

[6] R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).

[7] F. Giustino, M. L. Cohen, and S. G. Louie, Phys. Rev. B **76**, 165108 (2007).

[8] J. Noffsinger, F. Giustino, B. D. Malone, C.-H. Park, S. G. Louie, and M. L. Cohen, Comput. Phys. Commun. **181**, 2140 (2010).

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

[10] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Rev. Mod. Phys. **84**, 1419 (2012).

[11] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2001).

[12] K. M. Rabe and U. V. Waghmare, Phys. Rev. B **52**, 13236 (1995).

[13] A. García-Martín, D. Hermann, F. Hagmann, K. Busch, and P. Wölfle, Nanotechnology **14**, 177 (2003).

[14] J.-M. Lihm and C.-H. Park, Phys. Rev. X **11**, 041053 (2021).

[15] W. Kohn, Phys. Rev. **133**, A171 (1964).

[16] E. Koch and S. Goedecker, Solid State Commun. **119**, 105 (2001).

[17] I. Souza, T. Wilkens, and R. M. Martin, Phys. Rev. B **62**, 1666 (2000).

[18] H. Tamura, J. Phys. Chem. A **120**, 9341 (2016).

[19] Z. Jin and J. E. Subotnik, J. Chem. Phys. **146**, 244110 (2017).

[20] L. Hedin, Phys. Rev. **139**, A796 (1965).

[21] L. Hedin and S. Lundqvist, in *Solid State Physics*, Vol. 23, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic Press, San Diego, CA, 1970), pp. 1–181.

[22] G. Strinati, Riv. Nuovo Cimento (1978-1999) **11**, 1 (1988).

[23] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[24] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).

[25] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Phys. Rev. Lett. **111**, 216805 (2013).

[26] X. Zhu, H. Su, R. A. Marcus, and M. E. Michel-Beyerle, J. Phys. Chem. Lett. **5**, 3061 (2014).

[27] A. Molina-Sánchez, ACS Appl. Energy Mater. **1**, 6361 (2018).

[28] R. I. Biega, M. R. Filip, L. Leppert, and J. B. Neaton, J. Phys. Chem. Lett. **12**, 2057 (2021).

[29] M. L. Tiago, J. E. Northrup, and S. G. Louie, Phys. Rev. B **67**, 115212 (2003).

[30] S. Sharifzadeh, A. Biller, L. Kronik, and J. B. Neaton, Phys. Rev. B **85**, 125307 (2012).

[31] S. Sharifzadeh, P. Darancet, L. Kronik, and J. B. Neaton, J. Phys. Chem. Lett. **4**, 2197 (2013).

[32] T. Rangel, K. Berland, S. Sharifzadeh, F. Brown-Altvater, K. Lee, P. Hyldgaard, L. Kronik, and J. B. Neaton, Phys. Rev. B **93**, 115206 (2016).

[33] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, Comput. Phys. Commun. **183**, 1269 (2012).

[34] G. H. Wannier, Phys. Rev. **52**, 191 (1937).

[35] N. F. Mott and M. J. Littleton, Trans. Faraday Soc. **34**, 485 (1938).

[36] J. Frenkel, Phys. Rev. **37**, 17 (1931).

[37] D. Y. Qiu, T. Cao, and S. G. Louie, Phys. Rev. Lett. **115**, 176801 (2015).

[38] D. Y. Qiu, G. Cohen, D. Novichkova, and S. Refaely-Abramson, Nano Lett. **21**, 7644 (2021).

[39] Y. Toyozawa, Prog. Theor. Phys. **20**, 53 (1958).

[40] G. Antonius and S. G. Louie, Phys. Rev. B **105**, 085111 (2022).

[41] H.-Y. Chen, D. Sangalli, and M. Bernardi, Phys. Rev. Lett. **125**, 107401 (2020).

[42] V. K. Kozin, V. A. Shabashov, A. V. Kavokin, and I. A. Shelykh, Phys. Rev. Lett. **126**, 036801 (2021).

[43] M. Onga, Y. Zhang, T. Ideue, and Y. Iwasa, Nat. Mater. **16**, 1193 (2017).

[44] S.-I. Kuga, S. Murakami, and N. Nagaosa, Phys. Rev. B **78**, 205201 (2008).

[45] W. Yao and Q. Niu, Phys. Rev. Lett. **101**, 106401 (2008).

[46] W. R. Heller and A. Marcus, Phys. Rev. **84**, 809 (1951).

[47] J. J. Hopfield and D. G. Thomas, J. Phys. Chem. Solids **12**, 276 (1960).

[48] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, Comput. Phys. Commun. **185**, 2309 (2014).

[49] G. Pizzi, V. Vitale, R. Arita, S. Blügel, F. Freimuth, G. Géranton, M. Gibertini, D. Gresch, C. Johnson, T. Koretsune, J. Ibañez-Azpiroz, H. Lee, J.-M. Lihm, D. Marchand, A. Marrazzo, Y. Mokrousov, J. I. Mustafa, Y. Nohara, Y. Nomura, L. Paulatto *et al.*, J. Phys.: Condens. Matter **32**, 165902 (2020).

[50] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon Press, Oxford, UK, 1956).

[51] R. S. Knox, *Theory of Excitons* (Academic Press, New York, NY, 1963).

[52] X. Gonze and C. Lee, Phys. Rev. B **55**, 10355 (1997).

[53] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).

[54] P. Löwdin, J. Chem. Phys. **19**, 1396 (1951).

[55] L. X. Benedict, Phys. Rev. B **66**, 193105 (2002).

[56] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Phys. Rev. B **103**, 045117 (2021).

[57] F. Tang, Z. Li, C. Zhang, S. G. Louie, R. Car, D. Y. Qiu, and X. Wu, Proc. Natl. Acad. Sci. U.S.A. **119**, e2201258119 (2022).

[58] S. Ismail-Beigi, Phys. Rev. B **73**, 233103 (2006).

[59] X. Gonze, J.-C. Charlier, D. C. Allan, and M. P. Teter, Phys. Rev. B **50**, 13035 (1994).

[60] X. Wang and T. C. Berkelbach, J. Chem. Theory Comput. **16**, 3095 (2020).

[61] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[62] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[63] D. R. Hamann, Phys. Rev. B **88**, 085117 (2013).

[64] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G.-M. Rignanese, Comput. Phys. Commun. **226**, 39 (2018).

[65] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo *et al.*, J. Phys.: Condens. Matter **29**, 465901 (2017).

[66] S. L. Adler, Phys. Rev. **126**, 413 (1962).

[67] N. Wiser, Phys. Rev. **129**, 62 (1963).

[68] E. L. Shirley, Phys. Rev. B **58**, 9579 (1998).

[69] E. L. Shirley, L. J. Terminello, J. E. Klepeis, and F. J. Himpsel, Phys. Rev. B **53**, 10296 (1996).

[70] N.-P. Wang, M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **67**, 115111 (2003).

[71] T. Deilmann, Phys. Rev. B **98**, 165105 (2018).

[72] L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. Lett. **80**, 4514 (1998).

[73] M. Gatti and F. Sottile, Phys. Rev. B **88**, 155113 (2013).

[74] J. Sun, J. Yang, and C. A. Ullrich, Phys. Rev. Res. **2**, 013091 (2020).

[75] Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. **22**, 833 (1967).

[76] C. Verdi and F. Giustino, Phys. Rev. Lett. **115**, 176401 (2015).

[77] V. A. Jhalani, J.-J. Zhou, J. Park, C. E. Dreyer, and M. Bernardi, Phys. Rev. Lett. **125**, 136602 (2020).

[78] D. L. Dexter, J. Chem. Phys. **21**, 836 (1953).

[79] T. Förster, Ann. Phys. **437**, 55 (1948).

[80] T. Förster, Discuss. Faraday Soc. **27**, 7 (1959).