Quasiparticles and phonon satellites in spectral functions of semiconductors and insulators: Cumulants applied to the full first-principles theory and the Fröhlich polaron

Jean Paul Nery[*] and Philip B. Allen[†]
Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794-3800, USA

Gabriel Antonius[‡]
Department of Physics, University of California at Berkeley, California 94720, USA
and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

Lucia Reining[§]
Laboratoire des Solides Irradiés, Ecole Polytechnique, CNRS, CEA/DSM and European Theoretical Spectroscopy Facility (ETSF), 91128 Palaiseau, France

Anna Miglio[‖] and Xavier Gonze[¶]
Institute of Condensed Matter and Nanoscience, European Theoretical Spectroscopy Facility (ETSF), Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium

(Received 19 October 2017; revised manuscript received 18 February 2018; published 22 March 2018)

The electron-phonon interaction causes thermal and zero-point motion shifts of electron quasiparticle (QP) energies $\epsilon_k(T)$. Other consequences of interactions, visible in angle-resolved photoemission spectroscopy (ARPES) experiments, are broadening of QP peaks and appearance of sidebands, contained in the electron spectral function $A(k,\omega) = -\Im m G_R(k,\omega)/\pi$, where G_R is the retarded Green's function. Electronic structure codes (e.g., using density-functional theory) are now available that compute the shifts and start to address broadening and sidebands. Here we consider MgO and LiF, and determine their nonadiabatic Migdal self-energy. The spectral function obtained from the Dyson equation makes errors in the weight and energy of the QP peak and the position and weight of the phonon-induced sidebands. Only one phonon satellite appears, with an unphysically large energy difference (larger than the highest phonon energy) with respect to the QP peak. By contrast, the spectral function from a cumulant treatment of the same self-energy is physically better, giving a quite accurate QP energy and several satellites approximately spaced by the LO phonon energy. In particular, the positions of the QP peak and first satellite agree closely with those found for the Fröhlich Hamiltonian by Mishchenko *et al.* [Phys. Rev. B **62**, 6317 (2000)] using diagrammatic Monte Carlo. We provide a detailed comparison between the first-principles MgO and LiF results and those of the Fröhlich Hamiltonian. Such an analysis applies widely to materials with infrared(IR)-active phonons.

DOI: 10.1103/PhysRevB.97.115145

I. INTRODUCTION

The notion of a single-particle (quasiparticle, or QP) spectrum ϵ_k for electrons (k is short for all needed quantum numbers—wave vector, band, spin—$\mathbf{k}$, n, σ) underlies much of solid state physics. Evidence for the existence of such quasiparticle spectra [1,2] relies on experiment. Optical experiments, combined with theoretical guidance, e.g., the "empirical pseudopotential method" [3], have been used for decades, and allowed extraction of accurate ϵ_k from reflectivity data for simple semiconductors. Excitonic effects cause deviation from an independent-particle interpretation, but theory can determine their consequences and help to extract one-electron properties. Angle-resolved photoemission spectroscopy (ARPES) provides a more direct QP spectrum [4,5]. The data can be approximately related to the rigorously defined one-particle spectral function, obtained from the retarded Green's function $G(k,\omega)$ as [2,6]

$$A(k,\omega) = -\frac{1}{\pi}\Im m G_R(k,\omega), \tag{1}$$

where $G_R(k,\omega)$ is the Fourier transform

$$\int dt \exp(i\omega t) G_R(k,t) \tag{2}$$

of the retarded Green's function

$$G_R(k,t) = -i\langle\{c_k(t), c_k^+(0)\}\rangle\theta(t), \tag{3}$$

and $\{a,b\}$ is an anticommutator.

[*]jeanpaul240@gmail.com
[†]philip.allen@stonybrook.edu
[‡]antonius@lbl.gov
[§]lucia.reining@polytechnique.fr
[‖]anna.miglio@uclouvain.be
[¶]xavier.gonze@uclouvain.be

 ©2018 American Physical Society

When the spectral function $A(k,\omega)$ exhibits a strong peak that correlates with the corresponding single-particle theory, this defines a QP energy. The total spectral weight $\int d\omega A(k,\omega)$ is 1, but the QP peak has reduced weight $Z_k < 1$. It is broadened, and accompanied by features at other energies. When clearly differing from a structureless background, these features are called satellites. This has been seen in many photoemission experiments, for example, Ref. [7] for polaron satellites, and Ref. [8] for plasmon satellites. Thus $A(k,\omega)$ contains more information than just the QP energy ϵ_k. Full interpretation is a challenge to theory; progress on plasmon [9,10] and polaron [11] satellites in real semiconductors is occurring.

On the basis of a one-band analysis for metals, Migdal [12] showed that the electron self-energy due to phonons has important low temperature dynamical effects, which can be accurately described by a lowest-order self-energy diagram Σ_M. Vertex corrections can be omitted because of the small ratio $\omega_{\rm ph}/E_F$ of phonon energies to the Fermi energy. Even though this argument that allows to neglect vertex corrections does not strictly apply in semiconductors, we will use the term "Migdal" to indicate the Migdal formula, Σ_M, given later as Eq.(6), where it is labeled $\Sigma^{\rm Fan}$, referring to earlier works on insulators by Fan in the fifties [13,14]. Using Dyson's equation, the corresponding Green's function is

$$G^D(k,\omega) = G_0(k,\omega) + G_0(k,\omega)\Sigma_M(k,\omega)G^D(k,\omega). \quad (4)$$

When inserted into Eq. (1), it gives a spectral function which we will call "Dyson-Migdal," $A^{\rm DM}$. Given the success of Migdal theory in metals and the expected success of low order perturbation theory for most electron-phonon problems, it is embarrassing to realize [15–17] that $A^{\rm DM}(k,\omega)$ often agrees poorly with measured $A(k,\omega)$. Typically, only one distinct satellite is found in the side-band spectrum, while in reality several satellites are possible [7], corresponding to emission of several phonons. Even worse, the QP peak is misplaced [18].

An alternative approach, also approximate, involves a cumulant treatment (see, e.g., Refs. [19–22] and pp. 410–415 of Ref. [2]). This will be denoted as $G^{\rm C}$ for the Green's function, and $A^{\rm C}$ for the spectral function. The cumulant was advocated for describing plasmon satellite effects in metals by Steiner *et al.* [8] and Hedin [23]. For electron-phonon effects, there is early work by Mahan [24] who used cumulant methods for the Urbach tail on optical $\sigma(\omega)$, and Dunn [15] for the polaron spectral function. Gunnarsson *et al.* [16] introduced the cumulant method for electron-phonon effects in metals. More recently, Verdi *et al.* [11] applied the cumulant method to doped semiconductors, where both plasmon and phonon effects occur. Here we apply the cumulant method to undoped semiconductors with electron-phonon renormalization but no dynamical electron-electron coupling.

To further motivate the interest to work with the cumulant approach instead of the Dyson-Migdal one, consider Fig. 1, which illustrates where the QP peak appears in various treatments of the Fröhlich polaron problem [19,25]. Lowest-order Rayleigh-Schrödinger perturbation theory puts the renormalized polaron energy below the conduction-band minimum by $\alpha\omega_{\rm LO}$, where α is the Fröhlich coupling constant. An accurate high-order treatment [26] gives a surprisingly similar answer, shown as blue squares on the graph. This is where the QP peak of the spectral function should be located. If the Migdal

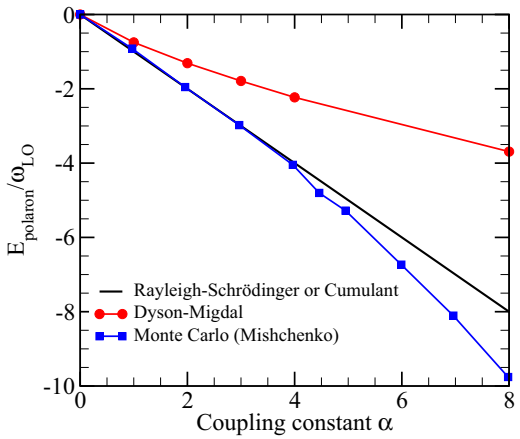

FIG. 1. Quasiparticle energy of the Fröhlich polaron, as a function of the Fröhlich coupling constant, in units of $\omega_{\rm LO}$: accurate results from Ref. [26] (blue squares), from the cumulant approach (which agrees with the lowest-order Rayleigh-Schrödinger theory) (black line), and from Dyson-Migdal spectral function [18,19] (red circles).

self-energy Σ_M is evaluated at $\omega = \epsilon_k$, the unperturbed energy, the corresponding energy shift $\Delta\epsilon_{k=0} = \Sigma_M(0,\epsilon_k)$ agrees with the Rayleigh-Schrödinger result $-\alpha\omega_{\rm LO}$. However, the actual QP peak of the spectral function $A^{\rm DM}$ is located at a fairly different energy, shown by the red circles on the graph [18,19]. The cumulant method used in the present paper puts the QP peak exactly on the Rayleigh-Schrödinger line, close to the correct polaron answer. The cumulant method also greatly improves the position of the first satellite peak, as will be shown later. It also gives additional peaks, corresponding to multiphonon excitations. However, they differ in location and strength from the other peaks in the accurate Fröhlich spectral function [26]. Another improvement given by the cumulant method is a much more reasonable value of the QP spectral weight Z, defined later in Eq. (11) and shown in Fig. 2.

Since the characteristic energy of plasmons is much bigger than that of phonons, the resolution needed to see plasmon satellites in photoemission experiments has been available for many years. The valence photoemission spectra of alkali metals, which exhibit multiple plasmon satellite structures caused by the Coulomb interaction, was modeled from first principles by Aryasetiawan *et al.* [9] using a cumulant treatment. More recently, Kheifets *et al.* [27] and Guzzo *et al.* [10] did ARPES for valence electrons in Si, and observed plasmon satellites. Their calculated spectrum based on a cumulant approach $G^{\rm C}$, agrees much better with their data than a theory based on the Dyson G^D. Similar observations were made for graphene, doped graphene, and graphite [28,29], and for several other materials [8,30]. The homogeneous electron gas has also been studied using the cumulant approach [31–33].

Electron-phonon interaction (EPI) effects are ubiquitous in solids [34], but were not seen in ARPES until improved resolution became available. In metals, temperature shifts of photoemission linewidths were seen for surface states in Cu [35] and Ag [36]. Phonon-induced "kinks" in the quasiparticle dispersion were seen by photoemission from surface states [37] of Be(0001) [38,39] and Mo(110) [40]. Bulk electron-phonon effects were seen by ARPES in superconducting and normal Pb

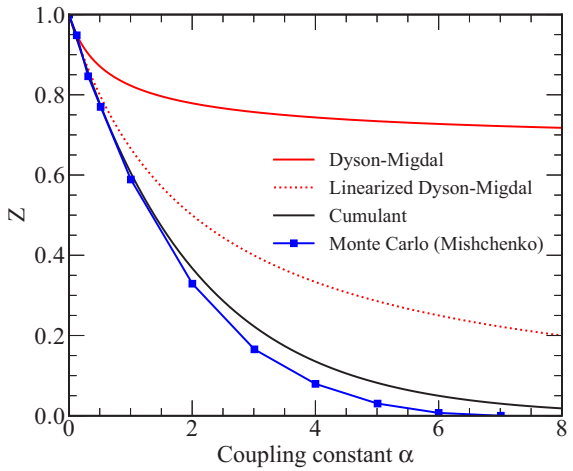

FIG. 2. Quasiparticle spectral weight $Z_{k=0}$ of the Fröhlich polaron, as a function of the Fröhlich coupling constant. The blue squares are accurate Monte Carlo results from Ref. [26]. The solid red line comes from the full Dyson-Migdal spectral function, Eq. (11). The dotted red line comes from the linearized approach, Eq. (14). The black line is the cumulant result using Eq. (26) with the retarded Green's function.

[41]. These experiments on metals did not resolve EPI-induced satellite structures. Recent ARPES experiments on monolayer metallic (and superconducting) FeSe, grown on both $SrTiO_3$ [42] and rutile TiO_2 [43], have seen replica bands displaced by the energy of a phonon of the substrate. ARPES studies of EPI effects on electrons near the band gaps of semiconductors are now available, with resolved EPI-induced satellites for electrons and holes doped into nonmetals, for bulk [7], surface [44,45], and interface [46] doped regions.

First-principle studies of the EPI effects on the electronic structure of insulators (zero-point motion as well as temperature dependence) have mostly used perturbation theory to second order in displacement (Fan and Debye-Waller terms) [47]. The focus has mostly been on the quasiparticle shift in simple solids, and occasionally on the quasiparticle broadening [34,48–67]. Polarons, by contrast, have been studied for a long time using high-order perturbative treatment of the singular Fröhlich first-order matrix element, with no second or higher-order matrix elements, and simplified model band structures [19,68]. When the Fröhlich coupling constant α exceeds $\approx$5, nonperturbative corrections start to be needed, and eventually small polarons may form [69] in real materials where the polaron radius in the continuum approach would be comparable to or smaller than the interatomic distance. Also, defects or higher than harmonic lattice displacement terms cannot be ignored in real materials. Only recently [60] it was noticed that first-principles studies of semiconductors had incorrectly ignored nonadiabatic aspects familiar in polaron literature for the quasiparticle energy shift (also sometimes referred to as renormalization) due to LO phonons in IR-active semiconductors. Perturbation theory had previously been simplified by omission of $\pm\omega_Q$ from denominators. Because of the $1/q$ divergence of the Fröhlich coupling, LO phonon contributions cause unphysical divergences (3D integral of $1/q^4$ at small q phonon wave vectors) if the $\pm\omega_Q$ pieces are omitted [60].

The literature on large polarons has focused on models such as the Fröhlich Hamiltonian, generally ignoring the existence of multiple phonon branches, nonparabolic electronic bands, Debye-Waller, and interband effects. We advocate unification of the separate skills of polaron and energy-band communities. This has started, with the above-mentioned understanding of the LO-phonon role in first-principles calculations [60,65,70], as well as the first-principles approach to the Fröhlich vertex developed by Sjakste *et al.* [71] and Verdi and Giustino [72].

Beyond the computation of quasiparticle shifts, first-principles studies of spectral functions, side bands and satellites have also appeared. Cannuccia and Marini [50,51] showed that optical data for diamond and polyethylene contain subgap EPI effects also visible in the single-electron theoretical spectral function $A(k,\omega)$. Spectral functions were computed by electron-phonon perturbation theory for the full band structure of C, BN, MgO and LiF by Antonius *et al.* [61]. A satellite is distinctly seen in the spectral function at the top of the valence band of LiF. However, these computations used G^D with a Migdal self-energy. Cumulant studies of EPI effects [15,24] have recently been revived by Verdi *et al.* [11] who discussed the doped TiO_2 data of Ref. [7], explaining the evolution from polaron to metallic-type EPI-renormalization as the doping level increases. The current paper reconsiders the results of Antonius *et al.* [61] using G^C, and shows, on the basis of the comparison with the Fröhlich Hamiltonian, how a cumulant calculation improves such results.

Cumulants [73], and their relatives, have been discussed a lot in older [9,15,16,21,23,74,75] as well as more recent [2,10,17,22,28,33,76–79] literature. The improvement they give is incomplete [22,79,80]. Despite the abundant literature, quantitative tests for EPI effects in non-metals treated with a cumulant approach are almost nonexistent. We are only aware of the above-mentioned model calculations for the Fröhlich Hamiltonian by Mahan [24] and Dunn [15], and first-principles study of TiO_2 by Verdi *et al.* [11].

The current paper deals with the zero-temperature spectral function $A(k,\omega,T=0)$, for the top of the valence band or bottom of the conduction band of two real insulators, MgO and LiF, arising from EPI effects. We first summarize the theoretical background for the (a) computation of self-energies and (b) spectral functions. Then we consider the Fröhlich Hamiltonian, and discuss results for several values of the coupling constant α. Subsequently, we calculate the self-energy and the spectral function of MgO and LiF, using *ab initio* density functional perturbation theory (DFPT) calculations with the code ABINIT [81,82] to determine phonons and their coupling with electrons. We consider, for the mentioned states, the Dyson-Migdal approach to the spectral function $[A^{\rm DM}(k,\omega)]$, as well as the cumulant approach $[A^{\rm C}(k,\omega)]$, and show that the cumulant results are physically more sensible. The DM approach gives qualitatively very different results, which calls for a reconsideration of the results given in Refs. [50,51,61].

II. SELF-ENERGY

The Hartree atomic unit system is used throughout ($\hbar = m_e = e = 1$). Starting from now, we will use the more explicit notations $\mathbf{k}$ for wave vectors and n for bands, instead of k. Spin is irrelevant in this article. The self-energy of an

unperturbed state, labeled by wave vector and band, includes two contributions at the lowest order of perturbation theory (quadratic in the strength of the EPI), namely, the Fan self-energy and the Debye-Waller self-energy [34]:

$$\Sigma(\mathbf{k}n,\omega) = \Sigma^{\mathrm{Fan}}(\mathbf{k}n,\omega) + \Sigma^{\mathrm{DW}}(\mathbf{k}n). \tag{5}$$

The Debye-Waller self-energy (see its expression in Refs. [34,56], a first-order effect from the second-order matrix elements) is static and real. On the contrary, the Fan self-energy is dynamical (but often treated adiabatically). In matrix notation, the Fan self-energy is given by $\Sigma = iG\Gamma D$, where G is the electron propagator, Γ is the vertex, and D is the phonon propagator. As often done [12], the vertex is approximated as $\Gamma = 1$. Approximating the electronic Green's function by its noninteracting counterpart, $G = G^{(0)}$, e.g., the Kohn-Sham Green's function without electron-phonon corrections, corresponds to a non-self-consistent treatment, and gives the standard result for the retarded Fan self-energy [34] :

$$\begin{aligned}\Sigma^{\mathrm{Fan}}&(\mathbf{k}n,\omega)\\ &= \frac{1}{N_{\mathbf{q}}}\sum_{\mathbf{q}j}^{\mathrm{BZ}}\sum_{n'}\left|\langle\mathbf{k}n|H^{(1)}_{\mathbf{q}j}|\mathbf{k}+\mathbf{q}n'\rangle\right|^2\\ &\times\left(\frac{n_{\mathbf{q}j}+1-f_{\mathbf{k}+\mathbf{q}n'}}{\omega-\varepsilon_{\mathbf{k}+\mathbf{q}n'}-\omega_{\mathbf{q}j}+i\eta}+\frac{n_{\mathbf{q}j}+f_{\mathbf{k}+\mathbf{q}n'}}{\omega-\varepsilon_{\mathbf{k}+\mathbf{q}n'}+\omega_{\mathbf{q}j}+i\eta}\right).\end{aligned} \tag{6}$$

In this expression, contributions from phonon modes with harmonic phonon energy $\omega_{\mathbf{q}j}$ and occupation number $n_{\mathbf{q}j}$ are summed for all phonon branches, labeled j, and wave vectors, labeled $\mathbf{q}$, in the entire Brillouin zone. The latter is discretized, with $N_{\mathbf{q}}$ indicating the number of wave vectors in the sum. The limit for infinite number of wave vectors is implied. Contributions from intermediate electronic states $|\mathbf{k}+\mathbf{q}n'\rangle$ with electron energy $\varepsilon_{\mathbf{k}+\mathbf{q}n'}$ (not renormalized by phonons) and occupation number $f_{\mathbf{k}+\mathbf{q}n'}$ are summed for all bands n' (valence and conduction). $H^{(1)}_{\mathbf{q}j}$ is the self-consistent change of potential due to the phonon labeled $\mathbf{q}j$ [34,83]. Equation (6) is also the Migdal result Σ_M.

We will work with semiconductors at zero temperature, in which case the phonon occupation numbers vanish, and the electron occupation numbers $f_{\mathbf{k}+\mathbf{q}n'}$ are either one, for the valence states, or zero, for the conduction states. At zero temperature, the phonon occupation numbers vanish. Since we will work with semiconductors, the electron occupation numbers $f_{\mathbf{k}+\mathbf{q}n'}$ are either one, for the valence states, or zero, for the conduction states. For semiconductors with IR-active phonons, the intraband ($n' = n$) contribution with small-q LO phonons (the Fröhlich problem) gives the most important dynamical features for the frequency range near the bare electronic energy ($\omega \approx \epsilon_{\mathbf{k}n}$) due to the combination of small-q diverging matrix element $\langle\mathbf{k}n|H^{(1)}_{\mathbf{q}j}|\mathbf{k}+\mathbf{q}n'\rangle \to C_n\delta_{nn'}/q$, see Ref. [60], and small denominator ($\omega \approx \epsilon_{\mathbf{k}n} \approx \epsilon_{\mathbf{k}+\mathbf{q}n}$) in Eq. (6).

III. QUASIPARTICLES AND SPECTRAL FUNCTION

The second-order self-energy $\Sigma(\mathbf{k}n,\omega)$ is the basis of different approximations for the quasiparticle shift and the spectral function. In the Rayleigh-Schrödinger (RS) approximation, the new quasiparticle energy $E_{\mathbf{k}n}$ is shifted from its initial value $\epsilon_{\mathbf{k}n}$ by the real part of the self-energy evaluated at $\epsilon_{\mathbf{k}n}$:

$$E^{\mathrm{RS}}_{\mathbf{k}n} = \epsilon_{\mathbf{k}n} + \Re e\,\Sigma(\mathbf{k}n,\omega=\epsilon_{\mathbf{k}n}). \tag{7}$$

The spectral function $-\Im m\,G_R(\mathbf{k}n,\omega)/\pi$ has dynamical effects beyond Rayleigh-Schrödinger. In the Dyson-Migdal (DM) approach [34], the spectral function obtained from the self-energy is

$$\begin{aligned}A^{\mathrm{DM}}(\mathbf{k}n,\omega) &= -\frac{1}{\pi}\Im m\,G^{\mathrm{DM}}_R(\mathbf{k}n,\omega)\\ &= \frac{(-1/\pi)\Im m\,\Sigma(\mathbf{k}n,\omega)}{(\omega-\varepsilon_{\mathbf{k}n}-\Re e\,\Sigma(\mathbf{k}n,\omega))^2+(\Im m\,\Sigma(\mathbf{k}n,\omega))^2}.\end{aligned} \tag{8}$$

There is typically a QP peak at $\omega = E^{\mathrm{DM}}_{\mathbf{k}n}$, where $\omega - \epsilon_{\mathbf{k}n} + \Re e\,\Sigma(\mathbf{k}n,\omega) = 0$ [i.e., where the first term in the denominator of Eq. (8) vanishes]. This assumes small values of $\Im m\,\Sigma$ near $\omega = E^{\mathrm{DM}}_{\mathbf{k}n}$. If $\Im m\,\Sigma$ is not small at $E^{\mathrm{DM}}_{\mathbf{k}n}$, the QP peak can be strongly broadened. The value of $E^{\mathrm{DM}}_{\mathbf{k}n}$ is usually shifted from $E^{\mathrm{RS}}_{\mathbf{k}n}$ [Eq. (7)] by a non-negligible ω dependence of $\Re e\,\Sigma(\mathbf{k}n,\omega)$. An additional possible slight shift of the QP peak in Eq. (8) can occur if the ω dependence of $\Im m\,\Sigma$ is not negligible at $E^{\mathrm{DM}}_{\mathbf{k}n}$. Equation (8) usually also gives one broad satellite, above the quasiparticle peak (for the conduction states), or below it (for the valence states).

A QP part of the spectral function, in Dyson-Migdal theory, can be separated out by Taylor expanding Σ^{Fan} around $\omega = E^{\mathrm{DM}}_{\mathbf{k}n}$, keeping only the constant term in the imaginary part and both constant and linear terms in the real part. The answer is

$$A^{\mathrm{DM}}_{\mathrm{QP}}(\mathbf{k}n,\omega) = \frac{Z^{\mathrm{DM}}_{\mathbf{k}n}}{\pi}\frac{\Gamma^{\mathrm{DM}}_{\mathbf{k}n}}{\left(\omega-E^{\mathrm{DM}}_{\mathbf{k}n}\right)^2+\left(\Gamma^{\mathrm{DM}}_{\mathbf{k}n}\right)^2}, \tag{9}$$

where the QP energy is

$$E^{\mathrm{DM}}_{\mathbf{k}n} = \epsilon_{\mathbf{k}n} + \Re e\,\Sigma\left(\mathbf{k}n,\omega=E^{\mathrm{DM}}_{\mathbf{k}n}\right), \tag{10}$$

the quasiparticle weight Z^{DM} is

$$Z^{\mathrm{DM}}_{\mathbf{k}n} = \left(1-\Re e\frac{\partial\,\Sigma(\mathbf{k}n,\omega)}{\partial\omega}\bigg|_{\omega=E^{\mathrm{DM}}_{\mathbf{k}n}}\right)^{-1}, \tag{11}$$

and the QP broadening is

$$\Gamma^{\mathrm{DM}}_{\mathbf{k}n} = Z^{\mathrm{DM}}_{\mathbf{k}n}\left|\Im m\,\Sigma\left(\mathbf{k}n,\omega=E^{\mathrm{DM}}_{\mathbf{k}n}\right)\right|. \tag{12}$$

If $T = 0$ and the imaginary part of the self-energy vanishes at the QP energy, Eqs. (10) and (11) give the exact location and weight of a scaled Dirac delta peak $Z^{\mathrm{DM}}_{\mathbf{k}n}\delta(\omega - E^{\mathrm{DM}}_{\mathbf{k}n})$.

Often [2,21], Eq.(10) is linearized to find an approximate quasiparticle energy and the related weight from quantities defined at the bare eigenenergy $\epsilon_{\mathbf{k}n}$:

$$E^{\mathrm{DM0}}_{\mathbf{k}n} = \epsilon_{\mathbf{k}n} + Z^{\mathrm{DM0}}_{\mathbf{k}n}\Re e\,\Sigma(\mathbf{k}n,\omega=\epsilon_{\mathbf{k}n}), \tag{13}$$

with

$$Z^{\mathrm{DM0}}_{\mathbf{k}n} = \left(1-\Re e\frac{\partial\,\Sigma(\mathbf{k}n,\omega)}{\partial\omega}\bigg|_{\omega=\epsilon_{\mathbf{k}n}}\right)^{-1}. \tag{14}$$

The Lehmann representation of the spectral function is derived from Eqs. (1) and (3) using the exact eigenstates $|m\rangle$,

of energy E_m, of the full Hamiltonian,

$$A(\mathbf{k}n,\omega) = \sum_{m',m} \frac{e^{-\beta E_m}}{Z}[|\langle m'|c^+_{\mathbf{k}n}|m\rangle|^2\delta(\omega - E_{m'} + E_m) + |\langle m'|c_{\mathbf{k}n}|m\rangle|^2\delta(\omega + E_{m'} - E_m)], \quad (15)$$

where $Z = \Sigma_m \exp(-\beta E_m)$ is the partition function and β is $1/k_BT$. Only $T = 0$ is directly relevant to the computations of this paper, but the $T > 0$ properties are also important and interesting. From Eq. (15), it is clear that $A(\mathbf{k}n,\omega) \geqslant 0$ at all ω, and the integrated spectral weight $[\int d\omega A(\mathbf{k}n,\omega)]$ is 1. The QP part [Eq. (9)] has total weight $Z^D_{\mathbf{k}n}$, which must therefore be less than 1. The non-QP part $A - A_{\rm QP}$ has the rest of the spectral weight.

There is an interesting property of the first moment [79],

$$\int_{-\infty}^{\infty} d\omega\omega A(\mathbf{k}n,\omega) = \langle\{[c_{\mathbf{k}n},H],c^+_{\mathbf{k}n}\}\rangle. \quad (16)$$

For noninteracting electrons, the right-hand side of Eq. (16) is the band energy $\epsilon_{\mathbf{k}n}$. If the only interaction is with phonons, the right-hand side has an extra piece, $\langle\{[c_{\mathbf{k}n},H_{e\text{-}p}],c^+_{\mathbf{k}n}\}\rangle$. Terms in the electron-phonon interaction $H_{e\text{-}p}$ which have odd powers of lattice displacement do not contribute. The even powers do, however, and the total right-hand side of Eq. (16) is $\langle\mathbf{k}n|\langle H\rangle|\mathbf{k}n\rangle$, where $\langle H\rangle$ is the thermal average of the vibrating one-electron Hamiltonian. This is exactly a Debye-Waller shifted single-particle energy, both zero-point and thermal if $T > 0$. If only the leading second-order term in a vibrational Taylor expansion is kept, the answer is $\epsilon_{\mathbf{k}n} + \Sigma^{\rm DW}(\mathbf{k}n)$, with only the Debye-Waller part of Eq. (5). It is interesting that an observable (in principle) property can separate the Debye-Waller from the Fan effects, given that translational invariance forces a partial cancellation [47] and indicates an underlying connection between these terms.

Although the quasiparticle energy from the DM spectral function for the Fröhlich problem does not occur at the right place and the DM quasiparticle spectral weight is quite bad, as shown by Figs. 1 and 2, the integral of the DM spectral function is correctly 1, and its first moment is correctly $\epsilon_{\mathbf{k}n}$ for the Fröhlich problem, which has no two-phonon matrix elements and therefore no D-W term.

IV. CUMULANT

The cumulant expansion is an alternative to the usual Dyson diagrammatic perturbation theory. It derives from an exponential representation of the retarded Green's function $G(\mathbf{k}n,t)$ in the time domain:

$$G^{\rm C}(\mathbf{k}n,t) = G_0(\mathbf{k}n,t)e^{C(\mathbf{k}n,t)}. \quad (17)$$

To lowest order, it treats the Fan diagram, Fig. 3(a), exactly, and higher-order diagrams, such as Figs. 3(b) and 3(c), approximately.

Different methods are used to derive the cumulant $C(\mathbf{k}n,t)$ from the self-energy, including identification of the terms of the same power of the interaction in a diagrammatic expansion of the left-hand and right-hand sides of Eq. (17). In the case of one isolated electronic level (without dispersion), the cumulant approach gives the exact result using only a second-order self-energy [74]. Among others, Langreth [74], Hedin [23],

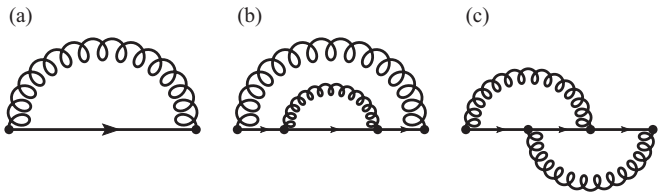

FIG. 3. Some of the lowest-order Feynman diagrams. (a) Usual Fan self-energy diagram. (b) Diagram that contributes to the self-consistency of the electron propagator. (c) Diagram that contains a vertex correction.

Gunnarsson *et al.* [16], Aryasetiawan *et al.* [9], Guzzo *et al.* [10], and Kas *et al.* [76] examined, developed, and applied the cumulant expansion.

In a recent paper, Kas *et al.* [76] considered a cumulant expansion of the retarded Green's function. See Appendix B for a discussion of the alternative time-ordered Green's function approach. Their equations, applied to the EPI self-energy (including Fan and Debye-Waller contributions), are

$$A^{\rm C}(\mathbf{k}n,\omega) = -\frac{1}{\pi}\Im m G^{\rm C}(\mathbf{k}n,\omega), \quad (18)$$

$$G^{\rm C}(\mathbf{k}n,\omega) = \int_{-\infty}^{\infty} e^{i\omega t} G^{\rm C}(\mathbf{k}n,t)dt, \quad (19)$$

$$G^{\rm C}(\mathbf{k}n,t) = -i\theta(t)e^{-i(\varepsilon_{\mathbf{k}n}+\Sigma^{\rm DW}_{\mathbf{k}n})t}e^{C(\mathbf{k}n,t)}, \quad (20)$$

$$C(\mathbf{k}n,t) = \int_{-\infty}^{\infty} \beta(\mathbf{k}n,\omega)\frac{e^{-i\omega t} + i\omega t - 1}{\omega^2}d\omega, \quad (21)$$

$$\beta(\mathbf{k}n,\omega) = \frac{1}{\pi}|\Im m\, \Sigma^{\rm Fan}(\mathbf{k}n,\omega + \varepsilon_{\mathbf{k}n})|. \quad (22)$$

Equation (20) shows that the static Debye-Waller self-energy $\Sigma^{\rm DW}_{\mathbf{k}n}$ shifts (in frequency) the whole spectral function with respect to a quasiparticle situated at $\varepsilon_{\mathbf{k}n}$.

Equation (21) also shows that the cumulant and its first time derivative vanish at $t = 0$, from which one deduces that the integral of the spectral function $A^{\rm C}(\mathbf{k}n,\omega)$ over all frequencies is 1, while its first moment is equal to $\varepsilon_{\mathbf{k}n} + \Sigma^{\rm DW}_{\mathbf{k}n}$. These properties agree perfectly with the exact results of the previous section. Because the phonon density of states (DOS) from acoustic modes vanishes quadratically at zero frequency, for the top of the valence band and the bottom of the conduction band, $\beta(\mathbf{k}n,\omega)$ vanishes quadratically around $\omega = 0$. Then following the method of Ref. [9], three separate effects can be attributed to the three pieces $(e^{-i\omega t} + i\omega t - 1)$ of Eq. (21). Specifically, $e^{-i\omega t}$ generates the satellites, $+i\omega t$ shifts the quasiparticle peak, and -1 generates the renormalization of the quasiparticle weight. The latter two effects make use of Kramers-Kronig relations (satisfied by the retarded self-energy),

$$\Re e\,\Sigma^{\rm Fan}(\mathbf{k}n,\varepsilon_{\mathbf{k}n}) = -P\int_{-\infty}^{\infty}\frac{\beta(\mathbf{k}n,\omega)}{\omega}d\omega, \quad (23)$$

$$\Re e\frac{\partial\,\Sigma^{\rm Fan}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}} = -\Re e\int_{-\infty}^{\infty}\frac{\beta(\mathbf{k}n,\omega)}{(\omega+i\delta)^2}d\omega, \quad (24)$$

where P denotes the principal part of the integral. The first of these gives the quasiparticle peak shift,

$$E_{\mathbf{k}n} = \epsilon_{\mathbf{k}n} + \Re e\Sigma^{\mathrm{Fan}}(\mathbf{k}n,\omega=\epsilon_{\mathbf{k}n}) + \Re e\Sigma^{\mathrm{DW}}(\mathbf{k}n) = \epsilon_{\mathbf{k}n} + \Re e\Sigma(\mathbf{k}n,\omega=\epsilon_{\mathbf{k}n}). \quad (25)$$

This is identical to the shift $E^{\mathrm{RS}}_{\mathbf{k}n}$ from Rayleigh-Schrödinger perturbation theory. The second yields the quasiparticle weight,

$$Z_{\mathbf{k}n} = \exp\left(\Re e\frac{\partial\Sigma^{\mathrm{Fan}}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}}\right). \quad (26)$$

In Eq. (17), or equivalently Eq. (20), the exponential can be Taylor-expanded; the Green's function and spectral function in the time domain are the sum of the product of the independent-electron contribution multiplied by powers of the cumulant, the latter being a linear functional of the imaginary part of the self-energy. In the frequency domain, this creates a satellite series [9,22], coming from repeated convolution in frequency of the undressed particle spectral function (a Dirac delta) by the "satellite spectral function" A^{S} (see Eqs. (8), (13), and (14) of Ref. [9] or Eqs. (3), (4), and (6) of the supplemental materials of Ref. [11]). Supposing A^{S} to be also a Dirac delta (as in the Lundqvist model Hamiltonian), one obtains for the spectral function a Poisson distribution of Dirac function satellites, each being spaced by the characteristic bosonic satellite energy. More generally, supposing the imaginary part of the self-energy to be nonzero only for positive frequencies beyond a given threshold, the threshold for each satellite contribution will be determined by the self-energy threshold multiplied by the order of the satellite. In the next section, such an effect can be seen in the case of the Fröhlich Hamiltonian.

When the imaginary part of the self-energy has contributions at both negative and positive frequencies, this simple picture is lost, unless the cumulant is clearly dominated by one of these. In the case of the first-principles EPI for wide-gap insulators, the self-energy has indeed both negative and positive contributions. However, only one of these contributions will dominate for wide-gap insulators, as will be seen later.

V. FRÖHLICH HAMILTONIAN SELF-ENERGY

The relationship between the Fröhlich coupling and first-principles calculations has been established in Refs. [71,72], in the general case of several electronic bands, several phonon branches, as well as anisotropic Born effective charge tensor and dielectric tensor. Here, we consider the simple original Fröhlich Hamiltonian, corresponding to the following hypotheses: (1) only one isotropic electronic parabolic band (we will first treat the conduction band), with a minimum at $\mathbf{k}=0$. We choose $\varepsilon_{\mathbf{k}=0,c}=0$, and use a parabolic dispersion governed by the effective mass m^*. (2) Only one LO-phonon branch with constant phonon frequency ω_{LO}. (3) Isotropic Born effective charge Z^*, isotropic electronic (optical) dielectric constant ϵ_∞, and thus, isotropic low-frequency dielectric constant ϵ_0 [84].

Only intraband terms $n'=n$ are present, thus, the general $g_{nn'j}(\mathbf{k},\mathbf{q}) = \langle\mathbf{k}n|H^{(1)}_{\mathbf{q}j}|\mathbf{k}+\mathbf{q}n'\rangle$ reduces to $g_{nn'j}(\mathbf{k},\mathbf{q}) = g_{\mathbf{q}}\delta_{nn'}$ for $j=LO$, with

$$g_{\mathbf{q}} = \frac{i}{q}\left[\frac{4\pi}{\Omega_0}\frac{\omega_{\mathrm{LO}}}{2}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_0}\right)\right]^{1/2}, \quad (27)$$

where Ω_0 is the volume of the primitive cell.

In computing the self-energy, the constant Debye-Waller shift is neglected. This might seem a strong approximation. However, hypothesis (2) implies the neglect of the Fan term from the acoustic modes as well, and it is known that the acoustic mode Fan contributions and the Debye-Waller contributions cancel each other in the vanishing-q limit [47]. Furthermore, the LO-phonon Fan term dominates, due to the integrable divergence mentioned at the end of Sec. II.

The expression for the Fröhlich self-energy is easily found, see, e.g., Ref. [19], but is presented here as well, for convenience and comparison with the first-principles results. The zero-temperature formula for the Fan self-energy of the (Fröhlich) electron state at the bottom of the conduction band ($\mathbf{k}=0$) comes from Eq. (6):

$$\Sigma^{\mathrm{Fr}}_e(\mathbf{k}=0,\omega) = \frac{1}{N_{\mathbf{q}}}\sum_{\mathbf{q}}\frac{|g_{\mathbf{q}}|^2}{\omega-\epsilon_{\mathbf{q}}-\omega_{\mathrm{LO}}+i\eta}. \quad (28)$$

The intermediate electron energy $\epsilon_{\mathbf{q}}$ becomes $q^2/2m^*$ in the parabolic band with effective mass approximation. Using the Debye sphere for the Brillouin zone, the equation becomes

$$\Sigma^{\mathrm{Fr}}_e(0,\omega) = \int_0^{q_D} dq\,\frac{\Omega_0}{(2\pi)^3}\frac{4\pi q^2|g_q|^2}{\omega-\frac{q^2}{2m^*}-\omega_{\mathrm{LO}}+i\eta}. \quad (29)$$

Provided the electronic energy on the boundary of the Debye sphere is much bigger than ω_{LO}, the upper limit q_D can be safely extended to infinity [19]. For $q_D\to\infty$, this gives

$$\Re e\,\Sigma^{\mathrm{Fr}}_e(\omega) = -\frac{\alpha\omega_{\mathrm{LO}}}{\sqrt{1-\omega/\omega_{\mathrm{LO}}}}\theta(\omega_{\mathrm{LO}}-\omega), \quad (30)$$

$$\Im m\,\Sigma^{\mathrm{Fr}}_e(\omega) = -\frac{\alpha\omega_{\mathrm{LO}}}{\sqrt{\omega/\omega_{\mathrm{LO}}-1}}\theta(\omega-\omega_{\mathrm{LO}}), \quad (31)$$

where the Fröhlich coupling constant α is

$$\alpha = \left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_0}\right)\left(\frac{m^*}{2\omega_{\mathrm{LO}}}\right)^{1/2}. \quad (32)$$

Above ω_{LO}, the self-energy is purely imaginary, while below ω_{LO}, it is purely real. Both are negative and diverge like an inverse square root of the frequency around ω_{LO}.

For the valence band, with the eigenenergy of the top of the valence band now taken as zero, the corresponding retarded self-energy is

$$\Re e\,\Sigma^{\mathrm{Fr}}_h(\omega) = +\frac{\alpha\omega_{\mathrm{LO}}}{\sqrt{1+\omega/\omega_{\mathrm{LO}}}}\theta(\omega+\omega_{\mathrm{LO}}), \quad (33)$$

$$\Im m\,\Sigma^{\mathrm{Fr}}_h(\omega) = -\frac{\alpha\omega_{\mathrm{LO}}}{\sqrt{-\omega/\omega_{\mathrm{LO}}-1}}\theta(-\omega-\omega_{\mathrm{LO}}). \quad (34)$$

For a given material with well-defined dielectric constant and LO frequency, the coupling constant α from Eq. (32) has different values for electron and hole polarons, due to differing effective masses.

In Secs. VII and VIII dealing with first-principles calculations, we will maintain a small finite broadening factor δ [85], of order $0.12\omega_{\mathrm{LO}}$, for numerical reasons. Thus the

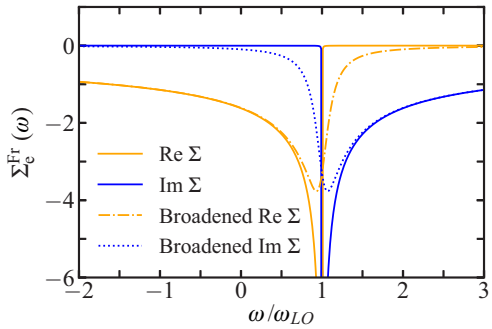

FIG. 4. Fröhlich Hamiltonian self-energy. Real part in orange, imaginary part in blue. The functions with a negligible broadening, $\delta = 0.001\omega_{\rm LO}$, are represented by a continuous lines, while functions with a broadening $\delta = 0.12\omega_{\rm LO}$, similar to the one used in first-principles calculations, are represented by dashed lines.

self-energy functions, Eqs.(30) and (34), will not retain their inverse-square-root shape. Equations (30) and (31) and their broadened versions are shown in Fig. 4.

The self-energies in Eqs. (30)–(34) were derived in the lowest order perturbation theory. They include only the Fan diagram, without vertex corrections. Calculations of self-energies at higher orders have been performed for the Fröhlich Hamiltonian, see, e.g., Refs. [15,19]. For values of α in the range considered in the present paper, those higher-order corrections to the quasiparticle shift are small, as argued by Migdal [12]. Calculations using these formulas for $\Sigma^{\rm Fan}$ and the corresponding spectral functions $A^{\rm DM}$ and $A^{\rm C}$, are discussed in the next section, and plotted in Figs. 5–8.

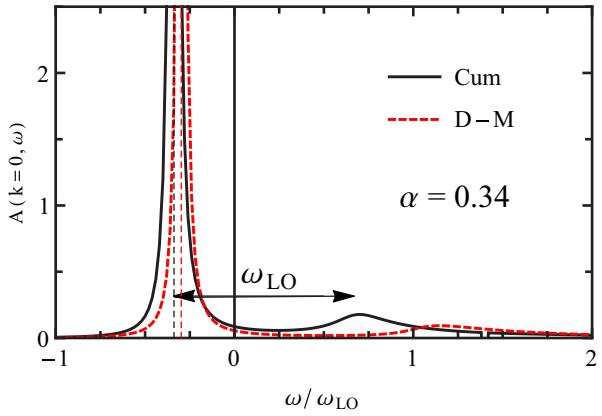

FIG. 5. Fröhlich Hamiltonian spectral function using the cumulant approach $G^{\rm C}$ (solid, black) and the Dyson-Migdal approach $G^{\rm DM}$ (dashed, red) for $\alpha = 0.34$. The position of the quasiparticle peak slightly differs between the two. The cumulant version deviates from the Fröhlich value $-\alpha\omega_{\rm LO}$ only because a nonzero broadening $\delta \approx 0.12\omega_{\rm LO}$ [85] is used in numerical evaluation of Eq. (29), for consistency with later calculations. In the DM case, the onset of the phonon-emission "satellite" is higher by $\omega_{\rm LO}$ than the bare band energy $\omega = \varepsilon^{(0)}_{\mathbf{k}=0,c} = 0$ [18]. By contrast, it is higher by $\omega_{\rm LO}$ than the quasiparticle peak in the cumulant method, corresponding to states that combine the dressed quasiparticle with one LO phonon.

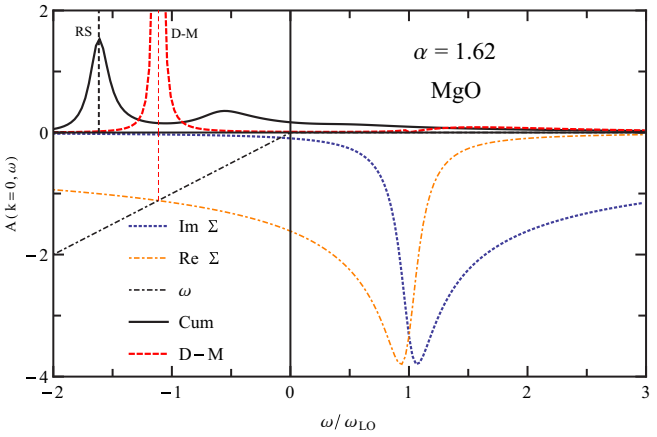

FIG. 6. The lower part shows the DM self-energy (in units of $\omega_{\rm LO}$) for a Fröhlich electron with $\alpha = 1.62$ (typical of the conduction-band minimum of MgO), using Eq. (30) except broadened with $\delta = 0.12\omega_{\rm LO}$ in Eq. (28). The position of the DM QP peak is at the crossing between the real part and the line $\Re e\Sigma = \omega$. The upper part shows both the resulting DM spectral function and the cumulant version from $G^{\rm C}$. The satellite setting in at $\omega/\omega_{\rm LO} = 1$ in the DM case is barely visible in this picture.

VI. FRÖHLICH HAMILTONIAN QUASIPARTICLE ENERGY AND SPECTRAL FUNCTION

The spectral function $-\Im m G_R(\mathbf{k}n,\omega)/\pi$ has "dynamical effects" beyond Rayleigh-Schrödinger. In Fig. 5, we show the spectral function for $\alpha = 0.34$ (a small value, typical of many semiconductors, e.g., electrons at the conduction-band minimum of GaN), for the Fröhlich Hamiltonian using two approximations. The black curve corresponds to the cumulant expansion Eqs. (18)–(22), while the dash-red curve was obtained with the DM approach Eq. (8). For consistency with later first-principles calculations, we used a small broadening factor $\delta \approx 0.12\omega_{\rm LO}$ for the self-energy, which causes a small artificial shift and broadening of the quasiparticle peak.

There are two effects. (i) In the cumulant case, the quasiparticle peak agrees with the value $-\alpha\omega_{\rm LO}$ predicted by Rayleigh-Schrödinger perturbation theory. (ii) A phonon-emission side band appears, with one clearly visible satellite, at different energies in the DM and cumulant cases. The separation between the quasiparticle and satellite is about $\omega_{\rm LO}$ in the cumulant expansion, but slightly higher than $(1+\alpha)\omega_{\rm LO}$ in the DM approach. Physically, this satellite ought to start at the quasiparticle energy $E_{\rm QP}$ plus $\omega_{\rm LO}$. The DM shift of $E_{\rm QP}$ away from $-\alpha\omega_{\rm LO}$ does not agree with the results obtained for many years by the polaron community [19,68].

Increasing the value of α to values typical for the valence and conduction-band extrema of MgO and LiF, 1.62, 4.01, and 8.00 (see later), gives the three next Figs. 6, 7, and 8. Note that the position $E_{\rm QP}$ of the quasiparticle peak differs more and more between the two, with the cumulant version staying at the RS answer $-\alpha\omega_{\rm LO}$, as expected. $A^{\rm DM}$ becomes increasingly unphysical for stronger couplings: the side band has one broad satellite, setting in at $\omega_{\rm LO}$, with a maximum at a frequency that increases with α. By contrast, $A^{\rm C}$ has the satellite onset at $E_{\rm QP} + \hbar\omega_{\rm LO}$. Several satellites are clearly visible in $A^{\rm C}$ of Fig. 7, spaced approximately by $\omega_{\rm LO}$. The side

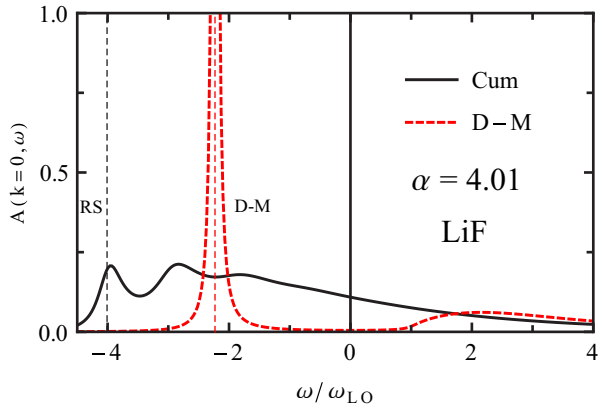

FIG. 7. Fröhlich Hamiltonian spectral function using the cumulant G^{C} (in black) and the Dyson-Migdal approach G^{DM} (dashed, in red) for $\alpha = 4.01$ (typical of the conduction-band minimum of LiF), from the Migdal self-energy broadened by $\delta = 0.12\omega_{\mathrm{LO}}$.

bands become broader and less well-defined as α increases, with a long tail extending to higher energies. The numerical value of the broadening $\delta \approx 0.12\omega_{\mathrm{LO}}$, albeit small, has an impact on the threshold at which it becomes impossible to distinguish the satellites from the overall smooth behavior. This broadening factor is multiplied by the order of the satellite in the repeated convolution of the undressed particle spectral function mentioned in Sec. IV.

These cumulant results globally agree with the previous cumulant-based study by Dunn [15], for $\alpha = 2$, 4, and 6. He worked, however, at finite temperature, and also included the next order of perturbation theory in his calculations of the self-energy. Higher orders of perturbation theory tend to sharpen features of the spectral function. In the case of a model core-electron spectrum, for which the exact solution is known [74], the next order of perturbation theory improves significantly the position of the peak, and sharpens it with respect to a first-order self-consistent treatment. On the other hand, a first-order non-self-consistent treatment also gives a sharper plasmon satellite

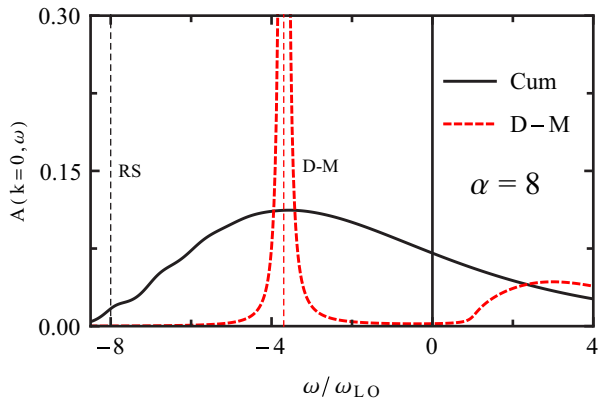

FIG. 8. Fröhlich Hamiltonian spectral function using the cumulant G^{C} (in black) and the Dyson-Migdal approach G^{DM} (dashed, in red) for $\alpha = 8$, from the Migdal self-energy broadened by $\delta = 0.12\omega_{\mathrm{LO}}$.

than the first-order self-consistent treatment, albeit located at nearly the same too low energy [75].

The first satellite shape and position, in our cumulant calculations, resemble reasonably well those of the diagrammatic Monte Carlo (MC) calculations of Mischenko *et al.* [26], apparently the best reference results available at present. However, the MC results do not show the second and third satellites and instead develop a satellite in the range from 3.5 to $4.0\omega_{\mathrm{LO}}$ if α is larger than one. For values of α larger than 4, another satellite appears in the range 8.0–$9.0\omega_{\mathrm{LO}}$. As MC results for optical $\sigma(\omega)$ compare well with other approaches [86], we believe that the cumulant approach for $A(k,\omega)$ has some errors for ω beyond the first satellite. The physical reason for the disappearance of the second and third multiphonon peaks, and the appearance of other peaks, has been discussed in Ref. [87]. The new peaks may be related to the so-called "relaxed excited states," not treated by the lowest-order cumulant approach, that dominate the spectral function in the energy range beyond the first phonon threshhold.

Let us analyze the behavior of the quasiparticle peak in the DM case in more detail, in the case without any numerical broadening. The quasiparticle energy is found from

$$E_{\mathrm{QP}}^{\mathrm{DM}} = \Re e\, \Sigma\left(E_{\mathrm{QP}}^{\mathrm{DM}}\right). \tag{35}$$

Using Eq. (30) for electron states gives a cubic equation,

$$\alpha^2 = \left(\frac{E_{\mathrm{QP}}^{\mathrm{DM}}}{\omega_{\mathrm{LO}}}\right)^2 - \left(\frac{E_{\mathrm{QP}}^{\mathrm{DM}}}{\omega_{\mathrm{LO}}}\right)^3 \quad (E_{\mathrm{QP}}^{D} < 0). \tag{36}$$

For all α, this has one negative real root; $E_{\mathrm{QP}}^{\mathrm{DM}} \sim -\alpha\omega_{\mathrm{LO}}$ for small α and $E_{\mathrm{QP}}^{\mathrm{DM}} \sim -\alpha^{2/3}\omega_{\mathrm{LO}}$ for large α. This agrees with Fig. 1 (red circles). For small α, to second order, the Dyson-Migdal QP energy is

$$E_{\mathrm{QP}}^{\mathrm{DM}} \approx -(\alpha - \alpha^2/2)\omega_{\mathrm{LO}} \tag{37}$$

The DM quasiparticle energy (35) corresponds to the intersection of the $\Re e\, \Sigma(\omega)$ function with the straight ω line, as shown in Fig. 6. As shown in Fig. 1, the second-order Rayleigh-Schrödinger answer $E_{\mathrm{QP}}^{\mathrm{RS}} = -\alpha\omega_{\mathrm{LO}}$, although not perfect, is better than the Dyson-Migdal answer (with $\Sigma_M = G_0 D$, not the self-consistent $\Sigma_M = GD$, see Ref. [18]).

The Fröhlich spectral function $A = -\Im m\, G_R(\omega)/\pi$, in DM approximation, has two parts. The quasiparticle part is $Z^{\mathrm{DM}}\delta(\omega - E_{\mathrm{QP}}^{\mathrm{DM}})$, where $Z^{\mathrm{DM}} = 1/(1 - d\Sigma/d\omega)$ is evaluated at $\omega = E_{\mathrm{QP}}^{\mathrm{DM}}$. To lowest order, the quasiparticle weight $Z^{\mathrm{DM}} \approx 1 - \alpha/2$. At large α, Z^{DM} tends asymptotically to 2/3, but the linearized weight (14) is $Z^{\mathrm{DMlin}} \approx 1/(1 + \alpha/2)$ for all values of α, which tends to zero asymptotically. The DM spectral function (without numerical broadening) is

$$A(x) = \frac{Z^{\mathrm{DM}}}{\omega_{\mathrm{LO}}}\delta(x - x_{\mathrm{QP}}) + \frac{1}{\pi\omega_{\mathrm{LO}}}\frac{\alpha\sqrt{x-1}}{x^2(x-1) + \alpha^2}\theta(x-1), \tag{38}$$

where $x = \omega/\omega_{\mathrm{LO}}$ and $x_{\mathrm{QP}} = E_{\mathrm{QP}}^{\mathrm{DM}}/\omega_{\mathrm{LO}}$. The DM side band always starts at $\omega = \omega_{\mathrm{LO}}$, rather than at the intuitively correct value of $E_{\mathrm{QP}} + \omega_{\mathrm{LO}}$. The Monte Carlo spectral functions [26] show sidebands starting close to the intuitive energy.

TABLE I. Computed basic characteristics of MgO and LiF. See text for the different symbols.

	Unit	MgO	LiF
a_0	(Å)	3.01	2.88
Ω_0	(Å^3)	19.2	16.9
ε_g (DFT-GGA)	(eV)	4.49	8.54
$\omega_{\rm LO}/\omega_{\rm TO}$	(eV)	0.0844/0.0454	0.0828/0.0466
$\epsilon_\infty/\epsilon_0$		3.23/11.14	2.04/6.44
m_e^*		0.340	0.873
m_{hh}^* (Γ-X/Γ-L)		2.164/3.822	3.622/11.955
m_{lh}^* (Γ-X/Γ-L)		0.387/0.335	1.346/0.887
α_e		1.624	4.009
α_{hh} (Γ-X/Γ-L)		4.101/5.450	8.165/14.834
α_{lh} (Γ-X/Γ-L)		1.734/1.610	4.977/4.040

VII. FULL MIGDAL SELF-ENERGY FROM FIRST-PRINCIPLES

We present now first-principles results for the full self-energies (real and imaginary parts, from all phonon modes, including interband and Debye-Waller effects) of MgO and LiF, at the valence-band maximum (VBM) and conduction-band minimum (CBM). These results will be used in the next section to find DM and cumulant spectral functions. This section also gives the related first-principles parameters to be fed into the Fröhlich model. For MgO and LiF band extrema, the Fröhlich coupling α ranges from about 1.5 to 15. We also tabulate the magnitudes of the separate Debye-Waller and Fan terms, as well as their contributions from unoccupied and occupied states.

Technical details of the first-principles calculations are in Appendix A. The most delicate issue concerns the sampling of phonon wave vectors in the Brillouin zone, and the numerical broadening factor needed to treat Eq. (6). To obtain well-converged self-energies, the eigenenergy differences between sampled wave vectors as well as the numerical broadening factor must be significantly smaller than the LO phonon frequency. This is especially important at the unperturbed quasiparticle eigenenergy, where the real and imaginary parts and their derivatives govern the asymptotics of the cumulant, and hence the quasiparticle peak characteristics. We choose a broadening of 0.01 eV, approximately $\omega_{\rm LO}/8$ (see Table I), and wave-vector grids up to $96 \times 96 \times 96$ points for MgO and $48 \times 48 \times 48$ points for LiF. This is considerably better than in Ref. [61] for the same materials (diamond and BN were also studied in that work). In Ref. [61], the broadening factors ranged between 0.1 eV and 0.4 eV, and phonon wave-vector grids had at most $32 \times 32 \times 32$ points.

MgO and LiF both crystallize in the (cubic) rocksalt structure, with one formula unit per primitive cell. The density functional theory (DFT-GGA) Kohn-Sham electronic structure of both materials can be found elsewhere [61], and will not be reproduced here. In both materials, the CBM is not degenerate. It is parabolic in a large region around Γ, so we expect the effective mass parabolic approximation to be adequate. The VBM is triply degenerate at Γ. One light-hole band rapidly separates from two heavy-hole bands away from Γ, the latter being degenerate along the Γ-X and Γ-L directions. The deviation with respect to parabolicity is faster than for the conduction band. The DFT eigenenergy is set to zero at the relevant band extremum. In both materials, there are three acoustic and three optic phonon branches. At Γ, the LO branch is separated from the doubly-degenerate TO branches.

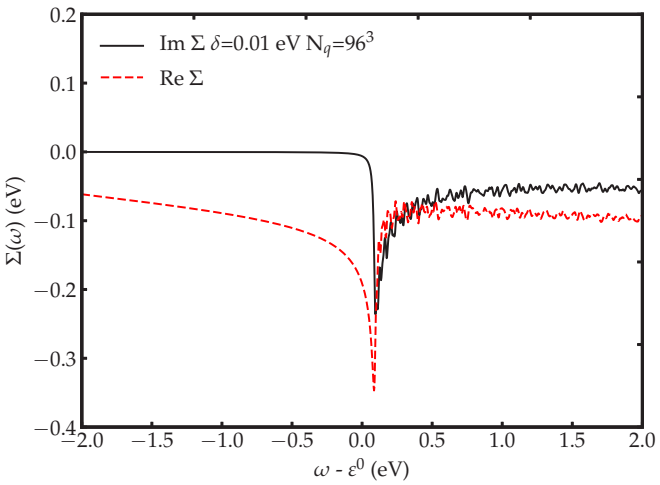

FIG. 9. The MgO conduction-band minimum retarded self-energy with $\delta = 0.01$ eV and a 96^3 q grid. Full black line: imaginary part; dotted red line: real part. The reference energy ε^0 is the unrenormalized conduction-band energy minimum. This figure is surprisingly similar to Fig. 6, considering that it includes all phonon and interband effects rather than just the analytic Fröhlich result.

Table I presents the computed geometric, electronic, dielectric, and dynamical properties of MgO and LiF that determine the corresponding Fröhlich parameter α, also reported in this table. The primitive cell parameter is $a_0 = a/\sqrt{2}$, where a is the size of the conventional cube. Different effective masses are mentioned for the valence bands, corresponding to the heavy holes (hh) and to the light holes (lh), and also to different directions of the nonspherical electronic structure [88]. Note that the $\omega_{\rm LO}$ and $\omega_{\rm TO}$ of the two materials are rather similar, while their dielectric properties and effective masses differ significantly. The Fröhlich parameter α provides a rough estimate of the phonon-induced zero-point renormalization of the quasiparticle energy. For the conduction-band minimum, the estimated shifts ($-\alpha_e\omega_{\rm LO}$) are -0.137 eV for MgO and -0.332 eV for LiF. For valence bands, we do not attempt to integrate over all effective mass directions and hole types, but simply provide the corresponding α values deduced from Eq. (32). The values presented in Table I are in reasonable agreement with those recently computed in Ref. [70] for the same materials. However, in Ref. [70], the Fröhlich polaron binding energy is defined as half the value from the usual theoretical approach (that we adopt), because the authors cutoff the q integral at π over the polaron radius instead of infinity. We find on the contrary that the Fröhlich values underestimate first-principles values, as will appear later.

The self-energy $\Sigma_1 + i\Sigma_2$ for the CBM of MgO, in a 2-eV window around the bare electronic energy ϵ^0, is in Fig. 9. Fröhlich-type real and imaginary peaks, both negative, occur at $\omega_{\rm LO} = 0.0844$ eV, just as in Fig. 6. Despite a very fine 96^3 q-point grid sampling, numerical noise is still visible for the small 0.01-eV broadening of the denominators (see Appendix A).

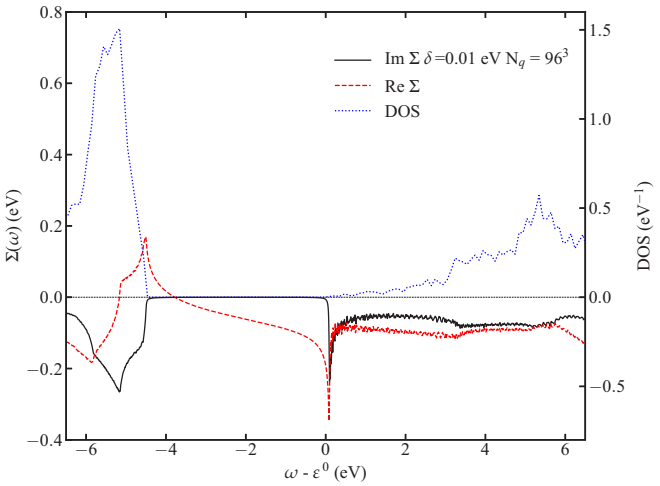

FIG. 10. Retarded self-energy for the bottom of the conduction band of MgO in a wider range of energy than in Fig. 9: imaginary part in black, real part in dashed red. The electronic DOS is also shown (dotted blue) for comparison.

The Debye-Waller contribution, and the Fan contributions from bands other than the lowest conduction shift Σ_1 downwards in Fig. 9 compared to Fig. 6, and give it a slight linear slope in the vicinity of the conduction-band minimum. Figure 10 shows the same data as Fig. 9 in a wider energy window, also with the electronic density of states (DOS). Additional structures are present, in the valence-band region (below −4 eV), with the same van Hove singularities as the electronic DOS. Small structures in the conduction-band region are seen as well.

Similarly, Σ_1 and Σ_2 for the VBM of MgO are shown in Fig. 11. With the same sampling and broadening as the CBM, the noise can hardly be seen. The curvature of the hole band is less pronounced for the VBM than for the CBM, making the numerical work less difficult. Close to $\omega - \epsilon^0 = 0$, the

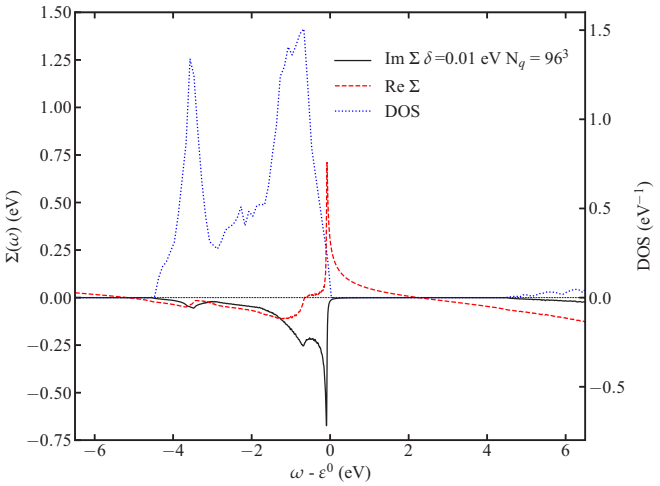

FIG. 11. Retarded self-energy for the top of the valence band of MgO: imaginary part in black, real part in dashed red. The electronic DOS is also shown (dotted blue), for comparison. The VBM eigenenergy is the reference energy ε^0, which explains the horizontal shift of the DOS with respect to Fig. 10.

TABLE II. MgO and LiF first-principles self-energy (eV) and frequency derivative (dimensionless) at $\omega = \varepsilon_{\mathbf{k}n}$, and their components, for the conduction-band minimum and valence-band maximum. The Debye-Waller self-energy is static (frequency-independent) and real. The quasiparticle weights, from linearized DM and retarded cumulant approaches are also given, as well as their occupied and unoccupied band factors. For the real part of the self-energy, results are reported with two different phonon wave-vector grids (96^3 and 48^3), while for the imaginary part, the derivatives, and the quasiparticle weights, only the results obtained with the 48^3 grid are reported. Both Σ_2 and $\partial\Sigma_2/\partial\omega$ vanish; their nonzero values arise from artificial broadening by $\delta = 0.12\omega_{\mathrm{LO}}$. See Appendix B for explicit expressions of $\Sigma_{\mathrm{oc}}^{\mathrm{Fan}}$ and $\Sigma_{\mathrm{un}}^{\mathrm{Fan}}$.

	MgO CBM	MgO VBM	LiF CBM	LiF VBM
96^3 grid				
Σ_1	−0.191	0.302	−0.370	0.723
Σ_1^{DW}	−0.056	4.271	0.078	6.785
$\Sigma_{1,\mathrm{un}}^{\mathrm{Fan}}$	−0.371	−4.336	−0.524	−6.911
$\Sigma_{1,\mathrm{oc}}^{\mathrm{Fan}}$	0.235	0.367	0.077	0.850
48^3 grid				
Σ_1	−0.175	0.285	−0.342	0.695
Σ_1^{DW}	−0.054	4.263	0.078	6.772
$\Sigma_{1,\mathrm{un}}^{\mathrm{Fan}}$	−0.354	−4.327	−0.497	−6.898
$\Sigma_{1,\mathrm{oc}}^{\mathrm{Fan}}$	0.233	0.349	0.077	0.821
Σ_2	−0.005	0.016	−0.014	0.053
$\Sigma_{2,\mathrm{un}}^{\mathrm{Fan}}$	−0.005	0.000	−0.014	0.000
$\Sigma_{2,\mathrm{oc}}^{\mathrm{Fan}}$	−0.000	0.016	0.000	0.053
$\partial\Sigma_1/\partial\omega$	−0.455	−1.594	−1.353	−4.780
$\partial\Sigma_{1,\mathrm{un}}^{\mathrm{Fan}}/\partial\omega$	−0.446	−0.007	−1.345	−0.012
$\partial\Sigma_{1,\mathrm{oc}}^{\mathrm{Fan}}/\partial\omega$	−0.009	−1.587	−0.008	−4.768
$\partial\Sigma_2/\partial\omega$	−0.058	−0.252	−0.206	−1.248
$\partial\Sigma_{2,\mathrm{un}}^{\mathrm{Fan}}/\partial\omega$	−0.058	0.000	−0.206	0.000
$\partial\Sigma_{2,\mathrm{oc}}^{\mathrm{Fan}}/\partial\omega$	0.000	−0.252	0.000	−1.248
$Z_{\mathbf{k}n}^{Dlin}$	0.687	0.386	0.425	0.173
$Z_{\mathbf{k}n}$	0.634	0.204	0.258	0.008
$Z_{\mathbf{k}n}^{\mathrm{un}}$	0.640	0.993	0.260	0.988
$Z_{\mathbf{k}n}^{\mathrm{oc}}$	0.991	0.205	0.992	0.008

structure of the self-energy is close to the Fröhlich self-energy, with the appropriate sign change for a hole polaron. Additional valence-band characteristic features are seen, clearly related to the electronic DOS. By contrast, the imaginary part of the self-energy in the conduction region is very small, and the real part is nearly structureless.

Unlike the electronic dispersion, the role of the phonon dispersion is apparently minor. For the phonon frequencies to have an impact on the self-energy, the difference between $\omega \approx \epsilon_{\mathbf{k}n}$ and the electronic eigenenergies $\epsilon_{\mathbf{k}+\mathbf{q}n'}$ must be comparable to phonon frequencies [see Eqs. (B4) and (B5)]. This happens only in a small Brillouin zone region around Γ, in which the phonon branches are practically constant.

The characteristics of the self-energy, evaluated at the bare eigenenergy $\varepsilon_{\mathbf{k}n}$, are reported in Table II, including the decomposition into Debye-Waller and Fan (and unoccu-

TABLE III. MgO self-energy (eV) and frequency derivative (adimensional) at $\omega = \varepsilon_{\mathbf{k}n}$, for the conduction-band minimum, for different wave-vector samplings and broadening factors δ.

Wave-vector grid	δ(eV)	Σ	$\partial\Sigma/\partial\omega$
20^3	0.01	$-0.1391 - i\,0.0008$	$-0.1603 - i\,0.0089$
20^3	0.02	$-0.1390 - i\,0.0032$	$-0.1589 - i\,0.0177$
20^3	0.05	$-0.1381 - i\,0.0078$	$-0.1498 - i\,0.0421$
20^3	0.10	$-0.1352 - i\,0.0147$	$-0.1232 - i\,0.0712$
20^3	0.01	$-0.1391 - i\,0.0016$	$-0.1604 - i\,0.0089$
32^3	0.01	$-0.1602 - i\,0.0031$	$-0.3086 - i\,0.0294$
48^3	0.01	$-0.1746 - i\,0.0050$	$-0.4975 - i\,0.0619$
96^3	0.01	$-0.1912 - i\,0.0062$	$-0.6264 - i\,0.0968$

pied/occupied) components. Also, the frequency derivative of the self-energy and components are given. A $48 \times 48 \times 48$ q-wave-vector grid has been used by default for this table, except for the real part of the self-energies and their decomposition, which is also given using the more converged $96 \times 96 \times 96$ q-wave-vector grid.

The convergence of the Allen-Heine-Cardona zero-point renormalization (ZPR) with respect to the wave-vector sampling has been thoroughly analyzed in Sec. IV B 2 of Ref. [60]. In particular, for IR-active materials treated in the nonadiabatic approximation, at the band structure extrema, a $1/N_q$ behavior is obtained, where N_q is the linear density of q points of the three-dimensional sampling. As shown in Appendix A, Table III, such a trend matches well numerical results. Thus the 48^3 and 96^3 grids Σ_1 results can be extrapolated to infinity, giving for the CBM and the VBM MgO, respectively, a ZPR (or polaron binding energy) of -207 and 319 meV, and for the CBM and VBM of LiF, respectively, a ZPR of -398 and 751 meV. The total band gap ZPR for MgO is 526 meV, while for LiF it is 1149 meV.

The Fröhlich estimated CBM shifts ($-\alpha_e\omega_{\mathrm{LO}}$), i.e., -137 meV for MgO and -332 meV for LiF, are in qualitative agreement with first-principles results, but underestimate their absolute value by about 50–70 meV. The first-principles ZPR of the CBM in these materials is thus largely dominated by the Fröhlich part. The analysis of the VBM shift is more complex due to the band warping, and will not be given here. Still, the range of α for holes mentioned in Table I and the zero-point renormalisation for the VBM in Table II are quite consistent. A similar dominance of the Fröhlich part of the electron-phonon interaction in other IR-active materials with large LO-TO splitting is expected, and would be consistent with the widespread use of the Fröhlich Hamiltonian for the interpretation of many experimental results.

Our self-energy values compare favorably with those of Table I, column $\Sigma^{\mathrm{dyn}}(\varepsilon^0)$ of Antonius *et al.* [61]. Remember, however, that in the latter study, the broadening factor δ was much larger and the sampling of phonon wave vectors much coarser than in the present study, see Appendix A. Actually, the quantities reported in the column $\Sigma^{\mathrm{stat}}(\varepsilon^0)$ of Table I of Ref. [61] should diverge for vanishing broadening factor and perfect Brillouin zone sampling, for the IR-active materials BN, MgO, and LiF, as shown in Ref. [60]. The similarity of $\Sigma^{\mathrm{stat}}(\varepsilon^0)$ and $\Sigma^{\mathrm{dyn}}(\varepsilon^0)$ is thus an artifact, simply due to the similarity of the chosen broadening factor value ($\delta = 0.1$ eV) with the LO phonon frequency in these materials (see, e.g., Table I).

The decomposition of Σ_1 into its Debye-Waller and Fan components highlights the dramatic cancellation between the Debye-Waller component and the unoccupied band Fan components for the VBM of the two materials. As a consequence, the occupied band Fan component has the same magnitude as the total zero-point renormalization value. By contrast, the CBM zero-point renormalization comes from contributions with different signs, without noticeable cancellation. Thus it is surprising that the total shift is given quite accurately by the Fröhlich part. The sum rule for acoustic modes, presented in Ref. [47], contributes to the final dominance of the Fröhlich estimation.

In Table II, we also report the quasiparticle weights, from linearized DM and retarded cumulant approaches, that are directly obtained from the derivative of the real part of the self-energy with respect to the frequency at the bare electronic energy, see Eqs. (14) and (26). The weights can be decomposed in their hole and electron factors, as discussed in Appendix B.

$Z^{\mathrm{DM0}}_{\mathbf{k}n}$ and $Z_{\mathbf{k}n}$ spectral weights differ the most in the VBM case; for LiF, the ratio exceeds one order of magnitude. It is slightly less than two for the MgO VBM. Taking into account the results from large polaron studies of the Fröhlich Hamiltonian [clearly favoring the Rayleigh or cumulant shifts, Eqs. (7) and (25)], the values from Table I of Ref. [61], column $\Sigma^{\mathrm{dyn}}(\varepsilon^0)$ have also to be preferred over the values in columns $Z\,\Sigma^{\mathrm{dyn}}(\varepsilon^0)$ or $\Sigma^{\mathrm{dyn}}(\varepsilon)$. Similarly, the values published in Ref. [60], Table VII, column "ZPR Nonadiabatic" correspond to the preferred expressions (7) and (25).

For completeness, Table II also gives the imaginary parts of the self-energy and their derivatives, which are, respectively, linked physically to the broadening of the quasiparticle peak and its asymmetry. Note that these values are actually artificial effects of numerical broadening. Our computations only include band extrema at zero temperature, for which the imaginary parts of Eqs. (B4) and (B5) vanish exactly if there is no artificial broadening. Nonzero values of $\Im m\,\Sigma$ at the QP energy thus indicate the magnitude of the broadening parameter.

VIII. FIRST-PRINCIPLES SPECTRAL FUNCTIONS

Figures 12 to 15 present DM and retarded cumulant spectral functions, for the CBM and VBM of MgO and LiF, and also show the Fröhlich spectral function obtained with estimated α (Table II) for the CBM case. The LiF VBM spectral function A^{DM} from G^{DM} was previously given in Fig. 1 of Ref. [61], but our numerical treatment of G^{DM} is significantly improved. The Brillouin zone has been sampled by a $96 \times 96 \times 96$ grid for the CBM of MgO, and a $48 \times 48 \times 48$ grid for the other cases. In all cases, a 0.01 eV $\approx 0.12\omega_{\mathrm{LO}}$ broadening of the self-energy has been used.

We only use the retarded cumulant approach due to the discussion in the previous section. It also appears to be the preferred method to obtain the quasiparticle spectral functions for insulators as well as metals.

For the MgO CBM, with Fröhlich $\alpha = 1.62$, Fig. 12, the position of the quasiparticle peak in the first-principles DM case is higher than from the first-principles cumulant case.

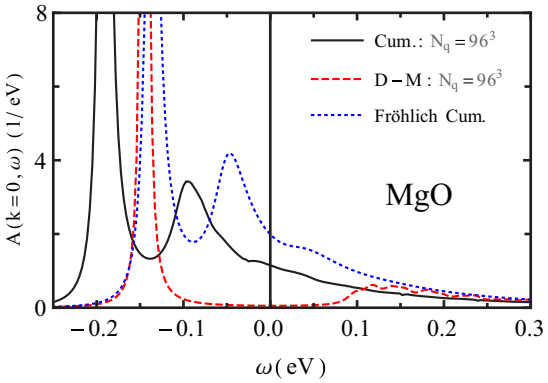

FIG. 12. Cumulant (black) and Dyson-Migdal (dashed red) spectral functions for the conduction-band minimum of MgO. The Fröhlich spectral function (dotted blue) with $\alpha = 1.62$ is also shown for comparison. The good correspondance between the Fröhlich and Dyson-Migdal peaks is fortuitous, see text.

Also, the Fröhlich peak position is closer to the first-principles DM position than with the first-principles cumulant one. This agreement between the DM position and the Fröhlich peak position is accidental: the Fröhlich constant is too small to reproduce the band gap shift from first-principles (cf. the above mentioned 50–70 meV underestimation), while the DM shift is also too small, but this is due to the incorrect underestimation highlighted in Fig. 1. The position of the satellite follows the same pattern as observed for the Fröhlich Hamiltonian; the DM satellite is separated from the quasiparticle peak by much more than the LO phonon frequency, while the distance between the satellite and the quasiparticle peak in the cumulant case is close to the LO phonon frequency value. Hence we conclude that for the CBM of MgO, the spectral function shape is dominated by the LO phonon. First-principles and Fröhlich Hamiltonian approaches yield very similar shapes, although the Fröhlich approach underestimates the QP energy shift. This is an important result of the present work. The same conclusion will be obtained for the other band extrema, for both MgO and LiF.

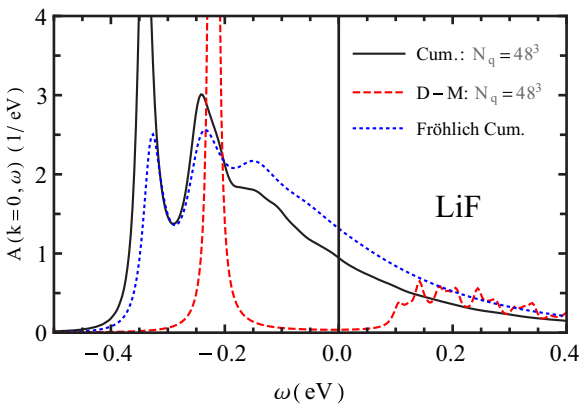

FIG. 13. Cumulant (black) and Dyson-Migdal (dashed red) spectral functions for the conduction-band minimum of LiF. The Fröhlich spectral function (dotted blue) with $\alpha = 4.01$ is also shown for comparison.

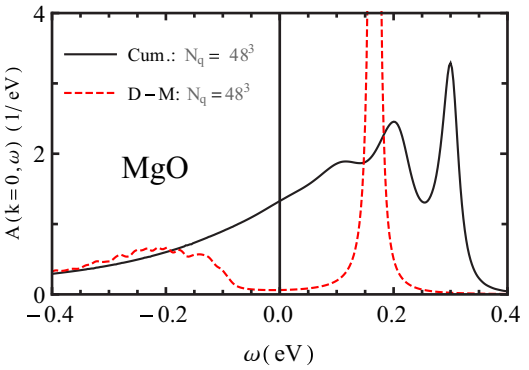

FIG. 14. Cumulant (black) and Dyson-Migdal (dashed red) spectral functions for the valence-band maximum of MgO.

In the LiF CBM case, with Fröhlich $\alpha = 4.01$ (Fig. 13), the position of the quasiparticle peak in the DM case is much higher than in the cumulant case. The Fröhlich-only cumulant peak position is close to the full-band cumulant one. Because the value of α is larger than for MgO, LiF has a larger second satellite before the smoothing of the spectral function. The same observations as for the MgO CBM, concerning the shape and position of the peaks, can also be made.

Despite the MgO valence band being threefold degenerate at Γ, the MgO VBM case, Fig. 14, is actually very similar to the LiF CBM case, with positive energy shifts seen instead of negative energy shifts.

Finally, in the case of LiF VBM, with the Fröhlich α being at most 14.8 for the heavy-hole effective mass, Fig. 15, the cumulant spectral function has become a broad peak (similar to Fig. 8) without any quasiparticle peak or satellite structure, unlike in the DM case. The lack of structure in the cumulant spectral function results from the large value of the Fröhlich α, directly linked to the large hole effective mass, i.e., the rather flat LiF valence bands [89].

IX. SUMMARY AND PERSPECTIVES

In this work, we compute from first principles the electron-phonon self-energy for the electronic states at the band extrema of MgO and LiF, at zero temperature, from which we obtain the spectral function using both Dyson-Migdal and the cumulant

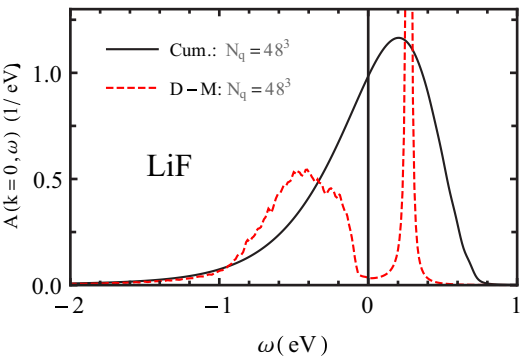

FIG. 15. Cumulant (black) and Dyson-Migdal (dashed red) spectral functions for the valence-band maximum of LiF.

methodologies. We also analyze the self-energy and spectral function of the Fröhlich Hamiltonian using the same methodologies.

For the imaginary part of the self-energy, the characteristic inverse square root behavior of the Fan self-energy found in the Fröhlich Hamiltonian, starting at the LO phonon frequency threshold, is also clearly present in MgO and LiF. However, additional structures, mirroring the electronic DOS, are also present. They originate from occupied as well as unoccupied bands, for both electron and hole self-energies. By contrast, the role of the phonon dispersion is minor.

The real part of the Fröhlich Hamiltonian self-energy has also an inverse square root behavior, on the other side of the LO phonon frequency threshold, compared to the imaginary part. This feature is also present in first-principles calculations. The Debye-Waller self-energy, which complements the Fan self-energy, gives a global shift, impacting the zero-point renormalization of eigenenergies.

To correctly obtain the inverse square root behavior in first-principles self-energies, the numerical integration over the phonon wave vectors has to be done carefully, and its convergence monitored. Numerical broadening of the denominator present in the Fan self-energy helps to reach convergence, but affects the accuracy of the prediction. Although schemes to overcome such convergence problems start to appear [65,70], we have simply performed the summation over a very fine grid, for a small broadening.

Our first-principles eigenenergy shifts range from 207 meV for the CBM of MgO to 751 meV for the VBM of LiF. Simple Fröhlich Hamiltonian estimates are too low although they account for a large fraction of such shift. There is cancellation between the Debye-Waller self-energy (not taken into account in the Fröhlich Hamiltonian) and the Fan self-energy. This cancellation is especially strong between the unoccupied part of the Fan self-energy and the Debye-Waller self-energy, in the case of the VBM for both MgO and LiF.

We analyze spectral functions from both cumulant and Dyson-Migdal approaches, using both the Fröhlich Hamiltonian and full DFT. Since the quasiparticle peak and the satellite location are both badly predicted from Dyson-Migdal theory even at low Fröhlich coupling constants, it is safe to argue that the shape of spectral function from the Dyson-Migdal approach can never be trusted for IR-active materials.

Our conclusions should apply to a large class of IR-active materials. In particular, the polar LO phonon contribution will give the biggest part of the zero-point renormalization and the spectral line shape. It would be worth to examine more materials, and find a rule for when the Fröhlich Hamiltonian might give a reasonably accurate estimate of the electron-phonon quasiparticle shift. For the conduction-band minimum, Fröhlich alone accounts for 4/5 of the CBM shift for LiF, but only 2/3 of the CBM shift for MgO. By contrast, the case of diamond is a remarkable example of a large QP shift without polar phonons. Its band edge zero-point renormalization (−330 meV for the indirect gap, and −416 meV for the direct gap) [57,58] has a magnitude similar to those of the IR-active materials studied here.

Finally, angle-resolved photoemission (ARPES), although difficult in insulators, could provide a nice test of some of the predicted spectral functions in this paper. In particular, consider the MgO VBM shown in Fig. 14. We predict a quasiparticle carrying 20% of the spectral weight, and clear phonon satellites. These may be accessible to experiment. By contrast, consider the VBM of LiF, shown in Fig. 15. The prediction is a completely blurred quasiparticle ($Z = 0.008$), with a broad peak rather than distinct satellites. However, this prediction is outside the trust range of our approach. The α values for LiF are in the range 4 to 15, as opposed to 1.6 to 5.5 for MgO. Experiment [90] shows that a hole at the top of the valence band of LiF self-traps in a local distortion similar to an F_2^- ion. This trapped polaron is known as a V_K center [91–93]. The present simple Hamiltonian (e.g., without anharmonicities) is insufficient to yield such a bound state. The perturbative spectral function correctly tells us that there is negligible weight in the quasiparticle peak. However, for holes at the top of the valence band in materials like LiF and NaCl, perturbation theory is not enough.

ACKNOWLEDGMENTS

We acknowledge helpful discussions with Y. Gillet, A. S. Mishchenko, S. Poncé, J. J. Rehr, and J. S. Zhou, as well as a careful reading of a preliminary version of the manuscript by S. Poncé. This work has been supported by the Fonds de la Recherche Scientifique (FRS-FNRS Belgium) through the PdR Grant No. T.0238.13 - AIXPHO, and by the European Research Council (Project SEED, grant agreement 320971). Work at Stony Brook was supported by US DOE Grant No. DE-FG02-08ER46550. Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Equipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the Fonds de la Recherche Scientifique (FRS-FNRS Belgium) under Grant No. 2.5020.11.

APPENDIX A: FIRST-PRINCIPLES CALCULATIONS : TECHNICAL DETAILS

All first-principles calculations are done using ABINITv8.4.1 (main executable ABINIT as well as postprocessor ElectronPhononCoupling). The computations use density functional theory (DFT, ground state and electronic structure) and density functional perturbation theory (DFPT, phonons and electron-phonon coupling), as described in Refs. [81,82], with the PBE generalized gradient approximation (PBE-GGA) [94,95] for MgO, and local density approximation (LDA) for LiF. See Ref. [96] for more information about the basic theories. The effective masses have been computed directly using the scheme from Ref. [97]. The Allen-Heine-Cardona formalism is used for the computation of the Fan and Debye-Waller self-energies [47,56,61,98]. The summation over unoccupied bands is handled efficiently thanks to a reformulation in terms of a Sternheimer equation [99]. However, in order to treat correctly the dynamical self-energy, the contribution from a number of "active" bands is computed explicitly by a sum over states, as described in Ref. [60], see Eq. (15). Atomic masses from natural isotopic ratios have been used throughout, namely (in atomic mass units), $M_{\mathrm{Mg}} = 24.305$, $M_{\mathrm{O}} = 15.9994$, $M_{\mathrm{Li}} = 6.941$, and $M_{\mathrm{F}} = 18.9984032$.

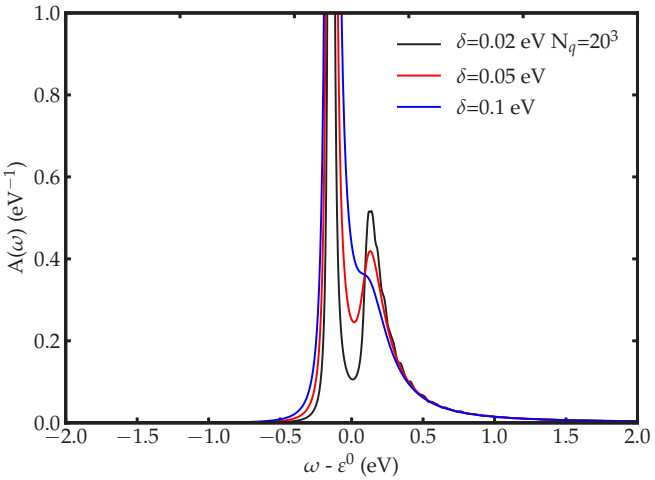

FIG. 16. Spectral function (Dyson-Migdal) of the MgO CBM self-energy with a 20^3 q grid and decreasing $\delta = 0.1$, 0.05, and 0.02 eV. The separation between the quasiparticle peak and the satellite is still not complete with the smallest δ value.

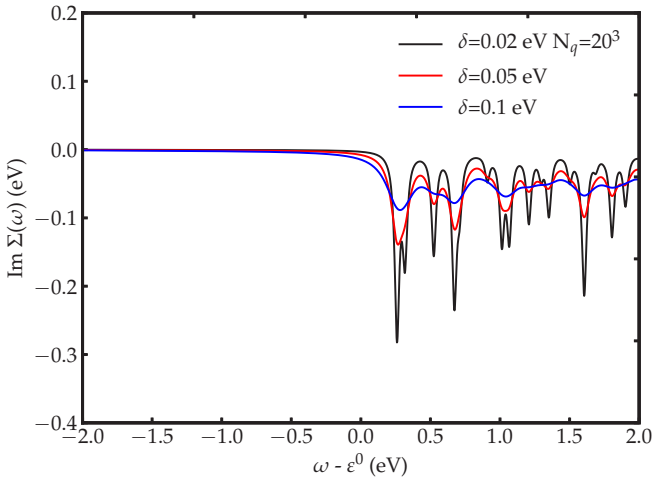

FIG. 17. The imaginary part of the MgO CBM self-energy with a 20^3 q grid and decreasing $\delta = 0.1$, 0.05, and 0.02 eV.

Norm-conserving pseudopotentials for MgO were taken from the Pseudo-dojo table (version 0.3) [100] generated using ONCVPSP [101], while for LiF Troullier-Martins pseudopotentials have been used. For both MgO and LiF, we used a 8^3 Monkhorst-Pack grid sampling of the Brillouin zone for the electronic wave functions, and a 50-Ha kinetic energy plane-wave cutoff. The summation over unoccupied bands for the dynamical self-energy is made over six conduction bands for MgO and three conduction bands for LiF, corresponding to a range of more than 10 eV above the CBM. The difference between the dynamical denominator (including the phonon frequency) and the static denominator (with only the electronic eigenenergy differences) is smaller than 1% above these bands, hence the summation over states can be safely replaced by the Sternheimer equation beyond them. Note that the replacement of an explicit sum over states, which includes a physical infinitesimal imaginary η, by a static contribution from the Sternheimer equation [60], destroys the Kramers-Kronig relations (23) and (24). Of course, this has no influence on the real and imaginary parts directly computed in the low-energy region of the conduction bands or in the valence bands.

The phonon wave-vector sampling needed to get converged self-energies and spectral functions is a delicate issue, already mentioned in Ref. [61]. The most difficult case is the CBM of MgO, since this band disperses quite strongly, as can be deduced from its low effective mass. In order to get the imaginary part of the self-energy, the wave-vector sampling ought to sample electronic eigenenergies such that their differences are not larger than the LO phonon frequency. This is hard to achieve when the effective mass is small. As described in Ref. [60], see in particular Eq. (16), a broadening factor δ is used to smooth the computed functions. Too large a δ value prevents distinguishing characteristic features of the self-energy or spectral function, which may typically appear at scales of the LO phonon frequency.

Figure 16 presents the DM spectral function of the CBM of MgO with a 20^3 sampling, and δ values 0.1, 0.05, and 0.02 eV. Larger δ smoothes the function, but also modifies its shape: the clear and physical separation between the quasiparticle peak and the satellite is washed out.

The values of the self-energy at the CBM are also strongly affected, as can be seen in Table III. From the numbers in this table, one can appreciate the convergence of the real part of the self-energy as the inverse of N_q, mentioned in Sec. VII, mathematically derived in Ref. [60]. This convergence study also highlights why the ZPR of the MgO CBM obtained in the present work, namely −207 meV, is quite different from the same quantity presented in Table I of Ref. [61], $\Sigma^{\rm dyn}(\epsilon^0) = -143$ meV, although the same software and pseudopotentials have been used.

A 0.01 eV broadening appears maximum for a meaningful investigation of the different spectral functions and self-energies. The self-energy is even noisier than the spectral function (see Fig. 17). Going from a 20^3 sampling, that delivers the meaningless Fig. 18, to our best 96^3 grid, Fig. 9, requires increasing by two orders of magnitude the computational effort, introducing possible computer memory problems, and still leaves some noise.

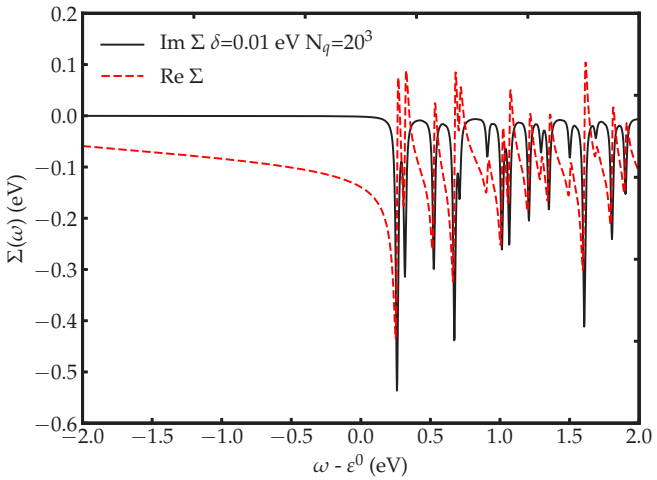

FIG. 18. The MgO CBM self-energy with $\delta = 0.01$ eV and a 20^3 q grid. Full black line: imaginary part; dotted red line: real part.

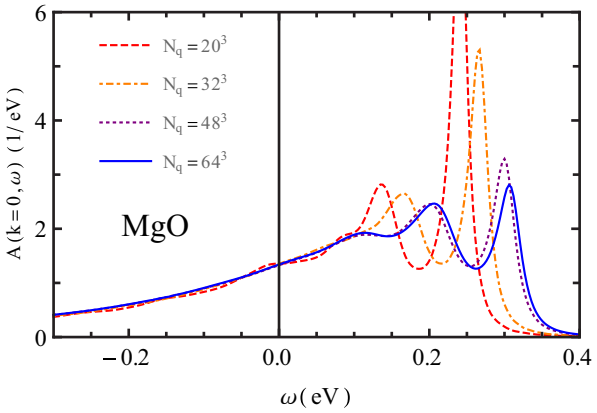

FIG. 19. Cumulant spectral function of the MgO VBM with $\delta = 0.01$ eV and 20^3, 32^3, 48^3, and 64^3 q grids.

The convergence parameters also impact the cumulant spectral functions, although they are more easily converged than the self-energy or the DM spectral function. The cumulant spectral functions corresponding to the VBM of MgO with $\delta = 0.01$ eV and 20^3, 32^3, 48^3, and 64^3 grids are presented in Fig. 19.

APPENDIX B: DIFFERENT FLAVORS OF CUMULANTS

The calculations presented in this paper have been performed using the retarded cumulant. Another approach that has been used in the literature is the time-ordered (t-O) cumulant. At temperature $T = 0$, the t-O Green's function yields also a simple expression [cf. Eq. (1)],

$$A(k,\omega) = |\Im m G_{\text{t-O}}(k,\omega)|/\pi, \tag{B1}$$

since the $T = 0$ t-O function has the same imaginary part as $G_R(k,\omega)$, except for a sign change at $\omega = \mu$. In diagrammatic perturbation theory the t-O Green's function is normally computed at $T = 0$. However, there are two main advantages to the retarded representation. (i) The retarded Green's function enters more naturally at $T > 0$ through analytic continuation of Matsubara Green's functions to the real frequency axis [2]. (ii) Previous t-O versions [102] typically make an additional approximation, by considering only one of the terms of the t-O Green's function:

$$\begin{aligned} G_{\text{t-O}}(k,t) &= -i\langle c_k(t)c_k^+(0)\rangle\theta(t) + i\langle c_k^+(0)c_k(t)\rangle\theta(-t) \\ &= G^>(k,t)\theta(t) + G^<(k,t)\theta(-t), \end{aligned} \tag{B2}$$

where $G^<$ is the lesser Green's function and $G^>$ the greater Green's function. For the hole, only the $G^<$ term is kept and a cumulant expansion $G(t) = G_0(t)e^{C(t)}$ is written for $t < 0$. An analogous expansion can be written for $t > 0$. These expansions are more accurately described as lesser (hole) or greater (electron) cumulant expansions. This explains discrepancies in the literature between retarded and the so-called time-ordered representations, despite Eqs. (1) and (B1) indicating the spectral functions should be the same (and second order approximations to the self-energy should yield very similar results). The retarded version, having only one θ prefactor, treats greater and lesser contributions on the same footing and naturally leads to a cumulant Green's function with satellites on both sides of the chemical potential.

As a result of (ii), the spectral function of the t-O cumulant is also not necessarily normalized to 1 (for more details, see Ref. [102]). However, the discrepancies disappear in core level spectra, where the cumulant approach, in its time-ordered version, is justified. Here, we will also see that for large gaps, the retarded cumulant and other versions yield only minor differences.

We first notice that there are separate contributions to the Fan self-energy from the intermediate states $|\mathbf{k}+\mathbf{q}n'\rangle$ in the conduction bands (labelled "un" for unoccupied) and in the valence band (labelled "oc" for occupied) [103]. No matter whether the initial state $|\mathbf{k}n\rangle$ is from the valence or the conduction band, both contributions occur in Eq. (6),

$$\Sigma^{\text{Fan}}(\mathbf{k}n,\omega) = \Sigma^{\text{Fan}}_{\text{un}}(\mathbf{k}n,\omega) + \Sigma^{\text{Fan}}_{\text{oc}}(\mathbf{k}n,\omega). \tag{B3}$$

Explicit equations are

$$\Sigma^{\text{Fan}}_{\text{un}}(\mathbf{k}n,\omega) = \frac{1}{N_{\mathbf{q}}}\sum_{\mathbf{q}j}^{\text{BZ}}\sum_{n'}^{\text{unocc}} \frac{\left|\langle \mathbf{k}n|H_j^{(1)}|\mathbf{k}+\mathbf{q}n'\rangle\right|^2}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'} - \omega_{\mathbf{q}j} + i\eta}, \tag{B4}$$

for the intermediate unoccupied state contribution to the self-energy, and

$$\Sigma^{\text{Fan}}_{\text{oc}}(\mathbf{k}n,\omega) = \frac{1}{N_{\mathbf{q}}}\sum_{\mathbf{q}j}^{\text{BZ}}\sum_{n'}^{\text{occ}} \frac{\left|\langle \mathbf{k}n|H_j^{(1)}|\mathbf{k}+\mathbf{q}n'\rangle\right|^2}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'} + \omega_{\mathbf{q}j} + i\eta}, \tag{B5}$$

for the intermediate occupied state contribution to the retarded self-energy.

Thanks to Eqs. (23) and (24), the retarded cumulant of Eq.(21) can be rewritten as

$$\begin{aligned} C(\mathbf{k}n,t) = &\int_{-\infty}^{\infty} \beta(\mathbf{k}n,\omega)e^{-i\omega t}\Re e\frac{1}{(\omega+i\delta)^2}d\omega \\ &- it\Re e\Sigma^{\text{Fan}}(\mathbf{k}n,\varepsilon_{\mathbf{k}n}) + \frac{\partial\Re e\Sigma^{\text{Fan}}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}}. \end{aligned} \tag{B6}$$

In the t-O cumulant approach of Ref. [9], Eqs. (B6) and (22) change as follows for the electrons:

$$\begin{aligned} C_e^{\text{t-O}}(\mathbf{k}n,t) = &\int_{\mu-\varepsilon_{\mathbf{k}n}}^{\infty} \beta_e^{\text{t-O}}(\mathbf{k}n,\omega)e^{-i\omega t}\Re e\frac{1}{(\omega+i\delta)^2}d\omega \\ &- it\Sigma^{\text{Fan}}(\mathbf{k}n,\varepsilon_{\mathbf{k}n}) + \frac{\partial\Sigma^{\text{Fan}}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}}, \end{aligned} \tag{B7}$$

$$\beta_e^{\text{t-O}}(\mathbf{k}n,\omega) = \frac{1}{\pi}\left|\Im m\,\Sigma^{\text{Fan}}_{\text{un}}(\mathbf{k}n,\omega+\varepsilon_{\mathbf{k}n})\right|. \tag{B8}$$

Note the reduced range of the integral, as well as the selection of part of the self-energy in the β factor. By contrast, the contributions that are either constant in time or linear in time are computed from the whole self-energy. Unlike the retarded cumulant, this version of the t-O cumulant does not vanish at $t = 0$, nor does its time-derivative, which means that the spectral function is not normalized to 1, and its first moment

is changed by the dynamical contribution. Corresponding expressions for the holes are also presented in Ref. [102]:

$$C_h^{\text{t-O}}(\mathbf{k}n,t) = \int_{-\infty}^{\mu-\varepsilon_{\mathbf{k}n}} \beta_h^{\text{t-O}}(\mathbf{k}n,\omega) e^{-i\omega t} \Re e \frac{1}{(\omega+i\delta)^2} d\omega - it\,\Sigma^{\text{Fan}}(\mathbf{k}n,\varepsilon_{\mathbf{k}n}) + \frac{\partial \Sigma^{\text{Fan}}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}}, \tag{B9}$$

$$\beta_h^{\text{t-O}}(\mathbf{k}n,\omega) = \frac{1}{\pi}\left|\Im m\ \Sigma_{\text{oc}}^{\text{Fan}}(\mathbf{k}n,\omega+\varepsilon_{\mathbf{k}n})\right|. \tag{B10}$$

Our computations for MgO and LiF have such large band gaps that the differing limits of integration in Eqs. (B6), (B7), and (B9) have negligible consequences. Similarly, the function $\beta(\mathbf{k},n=c)$ is almost identical to $\beta_e^{\text{t-O}}$ and $\beta(\mathbf{k},n=v)$ is almost identical to $\beta_h^{\text{t-O}}$ in the relevant range of ω integration where the denominator ω^2 is small. Also, the imaginary part of the self-energy and its derivative, evaluated at the unrenormalized energy, are usually small (numerically, this might not be the case due to the use of a broadening factor, see, for example, Fig. 4). This completes the comparison between Eqs. (B6), (B7), and (B9).

Different approximations give rise to other representations, described in Ref. [102], where the full Fan self-energy and its derivative are replaced in the last two terms of Eqs. (B7) and (B9), by their unoccupied or occupied counterparts, as in Ref. [22] [see Eq. (36)]. The normalization of the spectral function is again one. While the *retarded* cumulant weights (or the t-O versions introduced here) can be decomposed in their hole and electron factors, following Eqs. (B3) and (26),

$$Z_{\mathbf{k}n}^R = Z_{\mathbf{k}n}^{\text{un}} Z_{\mathbf{k}n}^{\text{oc}}, \tag{B11}$$

only one of these factors is included in the t-O cumulant weights of Refs. [9,10,22] and [102] [Eq. (3.64)]. However, as can be judged by the closeness to unity of $Z_{\mathbf{k}n}^e$ in the VBM case and $Z_{\mathbf{k}n}^h$ in the CBM case (see Table II), in MgO and LiF, the normalization defect is very small. The smallest of these weights is at least 0.988. This lack of impact of the unoccupied states on the VBM self-energy, and of the occupied states on the CBM self-energy, can be traced back to the large ratio between the electronic gap and the largest phonon frequency. This might not be true for small-gap semiconductors. Therefore the t-O approximations present in the literature seem safe for wide-gap semiconductors, but the retarded approach is in general preferable.

[1] L. D. Landau, The theory of a fermi liquid, Sov. Phys. JETP **3**, 920 (1957).

[2] R. M. Martin, L. Reining, and D. M. Ceperley, *Interacting Electrons. Theory and Computational Approaches* (Cambridge University Press, Cambridge, 2016).

[3] M. L. Cohen and T. K. Bergstresser, Band structures and pseudopotential form factors for fourteen semiconductors of the diamond and zinc-blende structures, Phys. Rev. **141**, 789 (1966).

[4] C. N. Berglund and W. E. Spicer, Photoemission studies of copper and silver: Experiment, Phys. Rev. **136**, A1044 (1964).

[5] R. Y. Koyama and N. V. Smith, Photoemission properties of simple metals, Phys. Rev. B **2**, 3049 (1970).

[6] A. Altland and B. Simons, *Condensed Matter Field Theory, Second Ed.* (Cambridge University Press, Cambridge, 2010).

[7] S. Moser, L. Moreschini, J. Jaćimović, O. S. Barišić, H. Berger, A. Magrez, Y. J. Chang, K. S. Kim, A. Bostwick, E. Rotenberg, L. Forró, and M. Grioni, Tunable Polaronic Conduction in Anatase TiO_2, Phys. Rev. Lett. **110**, 196403 (2013).

[8] P. Steiner, H. Höchst, and S. Hüfner, Simple metals, in *Photoemission in Solids II*, edited by L. Ley and M. Cardona (Springer, Berlin, 1979), Chap. 7, pp. 349–372.

[9] F. Aryasetiawan, L. Hedin, and K. Karlsson, Multiple Plasmon Satellites in Na and Al Spectral Functions from Ab Initio Cumulant Expansion, Phys. Rev. Lett. **77**, 2268 (1996).

[10] M. Guzzo, G. Lani, F. Sottile, P. Romaniello, M. Gatti, J. J. Kas, J. J. Rehr, M. G. Silly, F. Sirotti, and L. Reining, Valence Electron Photoemission Spectrum of Semiconductors: Ab Initio Description of Multiple Satellites, Phys. Rev. Lett. **107**, 166401 (2011).

[11] C. Verdi, F. Caruso, and F. Giustino, Origin of the crossover from polarons to Fermi liquids in transition metal oxides, Nat. Commun. **8**, 15769 (2017).

[12] A. B. Migdal, Interaction between electrons and lattice vibrations in a normal metal, Sov. Phys. JETP **34**, 996 (1958).

[13] H. Y. Fan, Temperature dependence of the energy gap in monatomic semiconductors, Phys. Rev. **78**, 808 (1950).

[14] H. Y. Fan, Temperature dependence of the energy gap in semiconductors, Phys. Rev. **82**, 900 (1951).

[15] D. Dunn, Electron-phonon interactions in an insulator, Can. J. Phys. **53**, 321 (1975).

[16] O. Gunnarsson, V. Meden, and K. Schönhammer, Corrections to Migdal's theorem for spectral functions: A cumulant treatment of the time-dependent Green's function, Phys. Rev. B **50**, 10462 (1994).

[17] S. M. Story, J. J. Kas, F. D. Vila, M. J. Verstraete, and J. J. Rehr, Cumulant expansion for phonon contributions to the electron spectral function, Phys. Rev. B **90**, 195135 (2014).

[18] We use here the original non-self-consistent formulation of Migdal, as also reported in Ref. [19], see, e.g., Eq. (7.20) and Table 7.1, column E_{TD}/ω_0. This formulation might be improved by introducing self-consistency on the electronic eigenenergies (i.e., replacing G_0 by the self-consistent G), or even by introducing a simple self-consistent shift of the whole band following the one of the bottom of the band, as described, e.g., by Hedin, see Eqs. (21) and (22) in Ref. [21]. In the cumulant method as applied here, no self-consistency is needed to obtain polaron quasiparticle energies in reasonable agreement with reference data from Ref. [26].

[19] G. D. Mahan, *Many-Particle Physics* (Kluwer, New York, 2000).

[20] C.-O. Almbladh and L. Hedin, *Handbook of Synchrotron Radiation* (North-Holland, Amsterdam, 1983).

[21] L. Hedin, On correlation effects in electron spectroscopies and the GW approximation, J. Phys.: Condens. Matter **11**, R489 (1999).

[22] B. Gumhalter, V. Kovač, F. Caruso, H. Lambert, and F. Giustino, On the combined use of GW approximation and cumulant expansion in the calculations of quasiparticle spectra: The paradigm of Si valence bands, Phys. Rev. B **94**, 035103 (2016).

[23] L. Hedin, Effects of recoil on shake-up spectra in metals, Phys. Scr. **21**, 477 (1980).

[24] G. D. Mahan, Phonon-broadened optical spectra: Urbach's rule, Phys. Rev. **145**, 602 (1966).

[25] H. Fröhlich, Electrons in lattice fields, Adv. Phys. **3**, 325 (1954).

[26] A. S. Mishchenko, N. V. Prokof'ev, A. Sakamoto, and B. V. Svistunov, Diagrammatic quantum Monte Carlo study of the Fröhlich polaron, Phys. Rev. B **62**, 6317 (2000).

[27] A. S. Kheifets, V. A. Sashin, M. Vos, E. Weigold, and F. Aryasetiawan, Spectral properties of quasiparticles in silicon: A test of many-body theory, Phys. Rev. B **68**, 233205 (2003).

[28] J. Lischner, D. Vigil-Fowler, and S. G. Louie, Physical Origin of Satellites in Photoemission of Doped Graphene: An Ab Initio GW Plus Cumulant Study, Phys. Rev. Lett. **110**, 146801 (2013).

[29] M. Guzzo, J. J. Kas, L. Sponza, C. Giorgetti, F. Sottile, D. Pierucci, M. G. Silly, F. Sirotti, J. J. Rehr, and L. Reining, Multiple satellites in materials with complex plasmon spectra: From graphite to graphene, Phys. Rev. B **89**, 085425 (2014).

[30] G. Mahan and W. Plummer, *Many-body effects in Photoemission. Handbook of Surface Science* (Elsevier, Amsterdam, 2000), Vol. 2, Chap. 14, pp. 953–987.

[31] B. Holm and F. Aryasetiawan, Self-consistent cumulant expansion for the electron gas, Phys. Rev. B **56**, 12825 (1997).

[32] F. Caruso and F. Giustino, The GW plus cumulant method and plasmonic polarons: Application to the homogeneous electron gas, Eur. Phys. J. B **89**, 238 (2016).

[33] D. Vigil-Fowler, S. G. Louie, and J. Lischner, Dispersion and line shape of plasmon satellites in one, two, and three dimensions, Phys. Rev. B **93**, 235446 (2016).

[34] F. Giustino, Electron-phonon interactions from first principles, Rev. Mod. Phys. **89**, 015003 (2017).

[35] B. A. McDougall, T. Balasubramanian, and E. Jensen, Phonon contribution to quasiparticle lifetimes in Cu measured by angle-resolved photoemission, Phys. Rev. B **51**, 13891 (1995).

[36] A. Eiguren, B. Hellsing, F. Reinert, G. Nicolay, E. V. Chulkov, V. M. Silkin, S. Hüfner, and P. M. Echenique, Role of Bulk and Surface Phonons in the Decay of Metal Surface States, Phys. Rev. Lett. **88**, 066805 (2002).

[37] E. W. Plummer, J. Shi, S.-J. Tang, E. Rotenberg, and S. D. Kevan, Enhanced electronphonon coupling at metal surfaces, Prog. Surf. Sci. **74**, 251 (2003).

[38] M. Hengsberger, D. Purdie, P. Segovia, M. Garnier, and Y. Baer, Photoemission Study of a Strongly Coupled Electron-Phonon System, Phys. Rev. Lett. **83**, 592 (1999).

[39] S. LaShell, E. Jensen, and T. Balasubramanian, Nonquasiparticle structure in the photoemission spectra from the Be(0001) surface and determination of the electron self-energy, Phys. Rev. B **61**, 2371 (2000).

[40] T. Valla, A. V. Fedorov, P. D. Johnson, and S. L. Hulbert, Many-Body Effects in Angle-Resolved Photoemission: Quasiparticle Energy and Lifetime of a Mo(110) Surface State, Phys. Rev. Lett. **83**, 2085 (1999).

[41] F. Reinert, B. Eltner, G. Nicolay, D. Ehm, S. Schmidt, and S. Hüfner, Electron-Phonon Coupling and Its Evidence in the Photoemission Spectra of Lead, Phys. Rev. Lett. **91**, 186406 (2003).

[42] J. J. Lee, F. T. Schmitt, R. G. Moore, S. Johnston, Y.-T. Cui, W. Li, M. Yi, Z. K. Liu, M. Hashimoto, Y. Zhang, D. H. Lu, T. P. Devereaux, D.-H. Lee, and Z.-X. Shen, Interfacial mode coupling as the origin of the enhancement of T_c in FeSe films on $SrTiO_3$, Nature (London) **515**, 245 (2014).

[43] S. N. Rebec, T. Jia, C. Zhang, M. Hashimoto, D.-H. Lu, R. G. Moore, and Z.-X. Shen, Coexistence of Replica Bands and Superconductivity in FeSe Monolayer Films, Phys. Rev. Lett. **118**, 067002 (2017).

[44] C. Chen, J. Avila, E. Frantzeskakis, A. Levy, and M. C. Asensio, Observation of a two-dimensional liquid of Fröhlich polarons at the bare $SrTiO_3$ surface, Nat. Commun. **6**, 8585 (2015).

[45] R. Yukawa, K. Ozawa, S. Yamamoto, H. Iwasawa, K. Shimada, E. F. Schwier, K. Yoshimatsu, H. Kumigashira, H. Namatame, M. Taniguchi, and I. Matsuda, Phonon-dressed two-dimensional carriers on the ZnO surface, Phys. Rev. B **94**, 165313 (2016).

[46] C. Cancellieri, A. S. Mishchenko, U. Aschauer, A. Filippetti, C. Faber, O. S. Barisic, V. A. Rogalev, T. Schmitt, N. Nagaosa, and V. N. Strocov, Polaronic metal state at the $LaAlO_3/SrTiO_3$ interface, Nat. Commun. **7**, 10386 (2016).

[47] P. B. Allen and V. Heine, Theory of the temperature dependence of electronic band structures, J. Phys. C **9**, 2305 (1976).

[48] A. Marini, Ab-Initio Finite Temperature Excitons, Phys. Rev. Lett. **101**, 106405 (2008).

[49] F. Giustino, S. G. Louie, and M. L. Cohen, Electron-Phonon Renormalization of the Direct Band Gap of Diamond, Phys. Rev. Lett. **105**, 265501 (2010).

[50] E. Cannuccia and A. Marini, Effect of the Quantum Zero-Point Atomic Motion on the Optical and Electronic Properties of Diamond and Trans-Polyacetylene, Phys. Rev. Lett. **107**, 255501 (2011).

[51] E. Cannuccia and A. Marini, Zero point motion effect on the electronic properties of diamond, trans-polyacetylene and polyethylene, Eur. Phys. J. B **85**, 320 (2012).

[52] B. Monserrat, N. D. Drummond, and R. J. Needs, Anharmonic vibrational properties in periodic systems: Energy, electron-phonon coupling, and stress, Phys. Rev. B **87**, 144302 (2013).

[53] H. Kawai, K. Yamashita, E. Cannuccia, and A. Marini, Electron-electron and electron-phonon correlation effects on the finite-temperature electronic and optical properties of zinc-blende GaN, Phys. Rev. B **89**, 085202 (2014).

[54] B. Monserrat, N. D. Drummond, C. J. Pickard, and R. J. Needs, Electron-Phonon Coupling and the Metallization of Solid Helium at Terapascal Pressures, Phys. Rev. Lett. **112**, 055504 (2014).

[55] B. Monserrat and R. J. Needs, Comparing electron-phonon coupling strength in diamond, silicon, and silicon carbide: First-principles study, Phys. Rev. B **89**, 214304 (2014).

[56] S. Poncé, G. Antonius, P. Boulanger, E. Cannuccia, A. Marini, M. Côté, and X. Gonze, Verification of first-principles codes: Comparison of total energies, phonon frequencies, electron–phonon coupling and zero-point motion correction to the gap between ABINIT and QE/Yambo, Comput. Mater. Sci. **83**, 341 (2014).

[57] G. Antonius, S. Poncé, P. Boulanger, M. Côté, and X. Gonze, Many-Body Effects on the Zero-Point Renormalization of the Band Structure, Phys. Rev. Lett. **112**, 215501 (2014).

[58] S. Poncé, G. Antonius, Y. Gillet, P. Boulanger, J. Laflamme Janssen, A. Marini, M. Côté, and X. Gonze, Temperature

dependence of electronic eigenenergies in the adiabatic harmonic approximation, Phys. Rev. B **90**, 214304 (2014).

[59] B. Monserrat, E. A. Engel, and R. J. Needs, Giant electron-phonon interactions in molecular crystals and the importance of nonquadratic coupling, Phys. Rev. B **92**, 140302 (2015).

[60] S. Poncé, Y. Gillet, J. Laflamme Janssen, A. Marini, M. Verstraete, and X. Gonze, Temperature dependence of the electronic structure of semiconductors and insulators, J. Chem. Phys. **143**, 102813 (2015).

[61] G. Antonius, S. Poncé, E. Lantagne-Hurtubise, G. Auclair, X. Gonze, and M. Côté, Dynamical and anharmonic effects on the electron-phonon coupling and the zero-point renormalization of the electronic structure, Phys. Rev. B **92**, 085137 (2015).

[62] M. Friedrich, A. Riefer, S. Sanna, W.G. Schmidt, and A. Schindlmayr, Phonon dispersion and zero-point renormalization of $LiNbO_3$ from density-functional perturbation theory, J. Phys.: Condens. Matter **27**, 385402 (2015).

[63] B. Monserrat and D. Vanderbilt, Temperature Effects in the Band Structure of Topological Insulators, Phys. Rev. Lett. **117**, 226801 (2016).

[64] A. Molina-Sanchez, M. Palummo, A. Marini, and L. Wirtz, Temperature-dependent excitonic effects in the optical properties of single-layer MoS_2, Phys. Rev. B **93**, 155435 (2016).

[65] J. P. Nery and P. B. Allen, Influence of Fröhlich polaron coupling on renormalized electron bands in polar semiconductors: Results for zinc-blende GaN, Phys. Rev. B **94**, 115135 (2016).

[66] C. E. P. Villegas, A. R. Rocha, and A. Marini, Anomalous thermal dependence of the band gap in black phosphorus, Nano Lett. **16**, 5095 (2016).

[67] W. A. Saidi, S. Poncé, and B. Monserrat, Temperature dependence of the energy levels of methylammonium lead iodide perovskite from first principles, J. Phys. Chem. Lett. **7**, 5247 (2016).

[68] J. T. Devreese and A. S. Alexandrov, Fröhlich polaron and bipolaron: Recent developments, Rep. Prog. Phys. **72**, 066501 (2009).

[69] D. Emin, *Polarons* (Cambridge University Press, Cambridge, 2012).

[70] W. R. L. Lambrecht, C. Bhandari, and M. van Schilfgaarde, Lattice polarization effects on the screened Coulomb interaction W of the GW approximation, Phys. Rev. Materials **1**, 043802 (2017).

[71] J. Sjakste, N. Vast, M. Calandra, and F. Mauri, Wannier interpolation of the electron-phonon matrix elements in polar semiconductors: Polar-optical coupling in GaAs, Phys. Rev. B **92**, 054307 (2015).

[72] C. Verdi and F. Giustino, Fröhlich Electron-Phonon Vertex from First Principles, Phys. Rev. Lett. **115**, 176401 (2015).

[73] R. Kubo, Generalized cumulant expansion method, J. Phys. Soc. Jpn. **17**, 1100 (1962).

[74] D. C. Langreth, Singularities in the x-ray spectra of metals, Phys. Rev. B **1**, 471 (1970).

[75] L. Hedin, Properties of electron self-energies and their role in electron spectroscopies, Nucl. Instrum. Methods Phys. Res., Sect. A **308**, 169 (1991).

[76] J. J. Kas, J. J. Rehr, and L. Reining, Cumulant expansion of the retarded one-electron Green function, Phys. Rev. B **90**, 085112 (2014).

[77] J. S. Zhou, J. J. Kas, L. Sponza, I. Reshetnyak, M. Guzzo, C. Giorgetti, M. Gatti, F. Sottile, J. J. Rehr, and L. Reining, Dynamical effects in electron spectroscopy, J. Chem. Phys. **143**, 184109 (2015).

[78] F. Caruso, H. Lambert, and F. Giustino, Band Structures of Plasmonic Polarons, Phys. Rev. Lett. **114**, 146404 (2015).

[79] M. Z. Mayers, M. S. Hybertsen, and D. R. Reichman, Description of quasiparticle and satellite properties via cumulant expansions of the retarded one-particle Green's function, Phys. Rev. B **94**, 081109 (2016).

[80] J. McClain, J. Lischner, T. Watson, D. A. Matthews, E. Ronca, S. G. Louie, T. C. Berkelbach, and G. Kin-Lic Chan, Spectral functions of the uniform electron gas via coupled-cluster theory and comparison to the GW and related approximations, Phys. Rev. B **93**, 235139 (2016).

[81] X. Gonze, B. Amadon, P. M. Anglade, J. M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Côté, T. Deutsch, L. Genovese, Ph. Ghosez, M. Giantomassi, S. Goedecker, D. R. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M. J. T. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G. M. Rignanese, D. Sangalli, R. Shaltaf, M. Torrent, M. J. Verstraete, G. Zerah, and J. W. Zwanziger, ABINIT: First-principles approach to material and nanosystem properties, Comput. Phys. Commun. **180**, 2582 (2009).

[82] X. Gonze, F. Jollet, F. Abreu Araujo, D. Adams, B. Amadon, T. Applencourt, C. Audouze, J.-M. Beuken, J. Bieder, A. Bokhanchuk, E. Bousquet, F. Bruneval, D. Caliste, M. Côté, F. Dahm, F. Da Pieve, M. Delaveau, M. Di Gennaro, B. Dorado, C. Espejo, G. Geneste, L. Genovese, L. Gerossier, M. Giantomassi, Y. Gillet, D. R. Hamann, L. He, G. Jomard, J. Laflamme Janssen, S. Le Roux, A. Levitt, A. Lherbier, F. Liu, I. Lukacevic, A. Martin, C. Martins, M. J. T. Oliveira, S. Poncé, Y. Pouillon, T. Rangel, G.-M. Rignanese, A. H. Romero, B. Rousseau, O. Rubel, A. A. Shukri, M. Stankovski, M. Torrent, M. J. Van Setten, B. Van Troeye, M. J. Verstraete, Waroquiers D., J. Wiktor, B. Xue, A. Zhou, and J. W. Zwanziger, Recent developments in the ABINIT software package, Comput. Phys. Commun. **205**, 106 (2016).

[83] A. Marini, S. Poncé, and X. Gonze, Many-body perturbation theory approach to the electron-phonon interaction with density-functional theory as a starting point, Phys. Rev. B **91**, 224310 (2015).

[84] X. Gonze and C. Lee, Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory, Phys. Rev. B **55**, 10355 (1997).

[85] One can distinguish a physical η from a numerical δ, as in Eq. (16) of Ref. [60]. In the present paper, we consistently use δ to indicate the numerical broadening (Lorentzian). It replaces the η used in the different equations when the latter are numerically evaluated.

[86] S. N. Klimin, J. Tempere, and J. T. Devreese, All-coupling polaron optical response: Analytic approaches beyond the adiabatic approximation, Phys. Rev. B **94**, 125206 (2016).

[87] A. S. Mishchenko, N. Nagaosa, N. V. Prokof'ev, A. Sakamoto, and B. V. Svistunov, Optical Conductivity of the Fröhlich Polaron, Phys. Rev. Lett. **91**, 236401 (2003).

[88] N. A. Mecholsky, L. Resca, I. L. Pegg, and M. Fornari, Theory of band warping and its effects on thermoelectronic transport properties, Phys. Rev. B **89**, 155131 (2014).

[89] Note however that the lack of structure at the broad maximum of the spectral function for large α is also a side effect of

our numerical broadening of the self-energy, albeit quite small - of order $0.12\omega_{\mathrm{LO}}$. Indeed, at zero broadening, the spectral function is made of a sum of functions, each of which has one inverse-square divergence appearing at a multi-phonon threshold, as described at the end of Sec. IV. These divergencies are spaced by ω_{LO}, corresponding to functions differing by one phonon. For such a function corresponding to nLO phonons, the numerical broadening is n times $0.12\omega_{\mathrm{LO}}$, and thus exceeds the spacing of the multiphonon peaks beyond 8 phonons, which happens at the maximum of the spectral functions beyond $\alpha = 8$.

[90] T. O. Woodruff and W. Känzig, Paramagnetic resonance absorption of a V center in LiF, J. Phys. Chem. Solids **5**, 268 (1958).

[91] W. Känzig, Electron spin resonance of V_1-centers, Phys. Rev. **99**, 1890 (1955).

[92] T. G. Castner and W. Känzig, The electronic structure of V-centers, J. Phys. Chem. Solids **3**, 178 (1957).

[93] A. M. Stoneham, *Theory of Defects in Solids: Electronic Structure of Defects in Insulators and Semiconductors* (Oxford University Press, London, 2001).

[94] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys. Rev. Lett. **77**, 3865 (1996).

[95] J. P. Perdew, K. Burke, and M. Ernzerhof, Errata to Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)], Phys. Rev. Lett. **78**, 1396(E) (1997).

[96] R. M. Martin, *Electronic Structure: Basic Theory and Methods* (Cambridge University Press, Cambridge, 2004).

[97] J. Laflamme Janssen, Y. Gillet, S. Poncé, A. Martin, M. Torrent, and X. Gonze, Precise effective masses from density functional perturbation theory, Phys. Rev. B **93**, 205147 (2016).

[98] P. B. Allen and M. Cardona, Theory of the temperature dependence of the direct gap of germanium, Phys. Rev. B **23**, 1495 (1981).

[99] X. Gonze, P. Boulanger, and M. Côté, Theoretical approaches to the temperature and zero-point motion effects on the electronic band structure, Ann. Phys. (Berlin) **523**, 168 (2011).

[100] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G.-M. Rignanese, The PseudoDojo: Training and grading a 85 element optimized norm-conserving pseudopotential table, Comput. Phys. Commun., doi: 10.1016/j.cpc.2018.01.012.

[101] D. R. Hamann, Optimized norm-conserving Vanderbilt pseudopotentials, Phys. Rev. B **88**, 085117 (2013).

[102] J.S. Zhou, Theory of electron spectroscopy beyond the state-of-the-art. An improved description of fermion-plasmon coupling in Green's function calculations. Ph.D. thesis, U. Paris-Saclay, 2015.

[103] This notation differs from the one in Ref. [102]. We use the "un" and "oc" labels in replacement of "e" and "h," in order to avoid using the latter with two different senses.