npj | Computational Materials

www.nature.com/npjcompumats

ARTICLE **OPEN**

Predominance of non-adiabatic effects in zero-point renormalization of the electronic band gap

Anna Miglio[1,7], Véronique Brousseau-Couture[2,7], Emile Godbout[2], Gabriel Antonius[3,4,5], Yang-Hao Chan[3,4], Steven G. Louie[3,4], Michel Côté[2], Matteo Giantomassi[1] and Xavier Gonze[1,6]✉

Electronic and optical properties of materials are affected by atomic motion through the electron–phonon interaction: not only band gaps change with temperature, but even at absolute zero temperature, zero-point motion causes band-gap renormalization. We present a large-scale first-principles evaluation of the zero-point renormalization of band edges beyond the adiabatic approximation. For materials with light elements, the band gap renormalization is often larger than 0.3 eV, and up to 0.7 eV. This effect cannot be ignored if accurate band gaps are sought. For infrared-active materials, global agreement with available experimental data is obtained only when non-adiabatic effects are taken into account. They even dominate zero-point renormalization for many materials, as shown by a generalized Fröhlich model that includes multiple phonon branches, anisotropic and degenerate electronic extrema, whose range of validity is established by comparison with first-principles results.

npj Computational Materials (2020)6:167 ; https://doi.org/10.1038/s41524-020-00434-z

INTRODUCTION

The electronic band gap is arguably the most important characteristic of semiconductors and insulators. It determines optical and luminescent thresholds, but is also a prerequisite for characterizing band offsets at interfaces and deep electronic levels created by defects[1]. However, accurate band gap computation is a challenging task. Indeed, the vast majority of first-principles calculations relies on Kohn–Sham Density-Functional Theory (KS-DFT), valid for ground state properties[2], that delivers a theoretically unjustified value of the band gap in the standard approach, even with exact KS potential[3–5].

The breakthrough came from many-body perturbation theory, with the so-called GW approximation, first non-self-consistent (G_0W_0) by Hybertsen and Louie in 1986[6], then twenty years later self-consistent (GW)[7] and further improved by accurate vertex corrections from electron-hole excitations (GWeh)[8]. The latter methodology, at the forefront for band-gap computations, delivers a 2–10% accuracy, usually overestimating the experimental band gap. GWeh calculations are computationally very demanding, typically about two orders of magnitude more than G_0W_0, itself two orders of magnitude more time-consuming than KS-DFT calculations, roughly speaking.

Despite being state-of-the-art, such studies ignored completely the electron–phonon interaction. The electron–phonon interaction drives most of the temperature dependence of the electronic structure of semiconductors and insulators, but also yields a zero-point motion gap modification at $T = 0$ K, often termed zero-point renormalization of the gap ($\mathrm{ZPR_g}$) for historical reasons.

The $\mathrm{ZPR_g}$ had been examined 40 years ago, by Allen, Heine and Cardona (AHC)[9,10] who clarified the early theories by Fan[11] and Antoncik[12]. Their approach is, like the GW approximation, rooted in many-body perturbation theory, where, at the lowest order, two diagrams contribute, see Fig. 1, the so-called "Fan" diagram, with two 1st-order electron–phonon vertices and the "Debye–Waller" diagram, with one 2nd-order electron–phonon vertex. In the context of semi-empirical calculations, the AHC method was applied to Si and Ge, introducing the adiabatic approximation, in which the phonon frequencies are neglected with respect to electronic eigenenergy differences and replaced by a small but non-vanishing imaginary broadening[10]. It was later extended without caution to GaAs and a few other III–V semiconductors[13,14].

In this work, we present first-principles AHC $\mathrm{ZPR_g}$ calculations beyond the adiabatic approximation, for 30 materials. Comparing with experimental band gaps, we show that adding $\mathrm{ZPR_g}$ improves the GWeh first-principles band gap, and moreover, that the $\mathrm{ZPR_g}$ has the same order of magnitude as the G_0W_0 to GWeh correction for half of the materials (typically materials with light atoms, e.g., O, N ...) on which GWeh has been tested. Hence, the GWeh level of sophistication misses its target for many materials with light atoms, if the $\mathrm{ZPR_g}$ is not taken into account. By including it, the theoretical agreement with direct measurements of experimental $\mathrm{ZPR_g}$ is improved. This also demonstrates the crucial importance of phonon dynamics to reach this level of accuracy.

Indeed, first-principles calculations of the $\mathrm{ZPR_g}$ using the AHC theory are very challenging, and only started one decade ago, on a case-by-case basis[15–25] (see the Supplementary Note 3), usually relying on the adiabatic approximation, and without comparison with experimental data. An approach to the $\mathrm{ZPR_g}$, alternative to the AHC one, relies on computations of the band gap at fixed, distorted geometries, for large supercells[18,26–33] (see and the Supplementary Note 3). As the adiabatic approximation is inherent in this approach, we denote it as ASC, for "adiabatic supercell". A recent publication by Karsai, Engel, Kresse and Flage-Larsen[34], hereafter referred to as KEKF, presents ASC $\mathrm{ZPR_g}$ based on DFT values, as well as based on G_0W_0 values for 18 semiconductors, with experimental comparison for 9 of them. Both AHC and ASC methodologies have been recently reviewed[35].

---

[1] Institute of Condensed Matter and Nanosciences, UCLouvain, B-1348 Louvain-la-Neuve, Belgium. [2] Département de Physique et Regroupement Québécois sur les Matériaux de Pointe, Université de Montreal, C.P. 6128, Succursale Centre-Ville, Montréal H3C 3J7, Canada. [3] Department of Physics, University of California at Berkeley, Berkeley, CA 94720, USA. [4] Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [5] Département de Chimie, Biochimie et Physique, Institut de recherche sur l'hydrogène, Université du Québec à Trois-Rivières, Trois-Rivières G8Z 4M3, Canada. [6] Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Nobel St. 3, Moscow 143026, Russia. [7] These authors contributed equally: Anna Miglio, Véronique Brousseau-Couture. ✉email: xavier.gonze@uclouvain.be

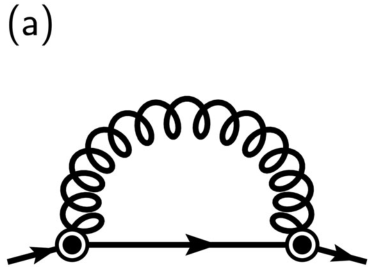

Fig. 1 Lowest-order diagrams for the contribution of the electron–phonon interaction to the zero-point renormalization of the gap ZPR$_g$. **a** Fan diagram, with two first-order screened electron–phonon vertices; **b** Debye-Waller diagram, with one second-order screened electron–phonon interaction vertex.

Although the adiabatic approximation had already been criticized by Poncé et al. in 2015[24] (Supplementary Note 5) as causing an unphysical divergence of the adiabatic AHC expression for infrared-active materials with vanishing imaginary broadening, thus invalidating all adiabatic AHC calculations except for non-infrared-active materials, the full consequences of the adiabatic approximation have not yet been recognized for ASC. We will show that the non-adiabatic AHC approach outperforms the ASC approach, so that the predictions arising from the mechanism by which the ASC approach bypasses the adiabatic AHC divergence problem for infrared-active materials are questionable. This is made clear by a generalized Fröhlich model with a few physical parameters, that can be determined either from first principles or from experimental data.

Although the full electronic and phonon band structures do not enter in this model, and the Debye-Waller diagram is ignored, for many materials it accounts for more than half the ZPR$_g$ of the full first-principles non-adiabatic AHC ZPR$_g$. As this model depends crucially on non-adiabatic effects, it demonstrates the failure of the adiabatic hypothesis, be it for the AHC or the ASC approach.

By the same token, we also show the domain of validity and accuracy of model Fröhlich large-polaron calculations based on the continuum hypothesis, that have been the subject of decades of research[36–42]. Such model Fröhlich Hamiltonian captures well the ZPR for about half of the materials in our list, characterized by their strong infrared activity, while it becomes less and less adequate for decreasing ionicity. In the present context, the Fröhlich large-polaron model provides an intuitive picture of the physics of the ZPR$_g$.

RESULTS

Zero-point renormalization: experiment vs. first principles

Figure 2 compares first-principles ZPR$_g$ with experimental values. As described in the "Methods" section, and in Supplementary Note 2, the correction due to zero-point motion effect on the lattice parameter, ZPR$_g^{lat}$, has been added to fixed volume results from both non-adiabatic AHC (present calculations) and ASC methodologies[34]. While for a few materials experimental ZPR$_g$ values are well established, within 5-10%, globally, experimental uncertainty is larger, and can hardly be claimed to be better than 25% for the majority of materials, see Supplementary Note 1. This will be our tolerance.

Let us focus first on the ASC-based results. For the 16 materials present in both KEKF and the experimental set described in the Supplementary Table I, the ASC vs. experimental discrepancy is more than 25% for more than half of the materials (KEKF KS-DFT ASC calculations are based on GGA-PBE, except for Si, Ge, GaAs, and CdSe, where the PBE0 hybrid functional has been used, thus the better score of Ge and GaAs for ASC calculations than for AHC calculations might be partly explained by this different KS-DFT functional). There is a global trend to underestimation by ASC, although CdTe is overestimated.

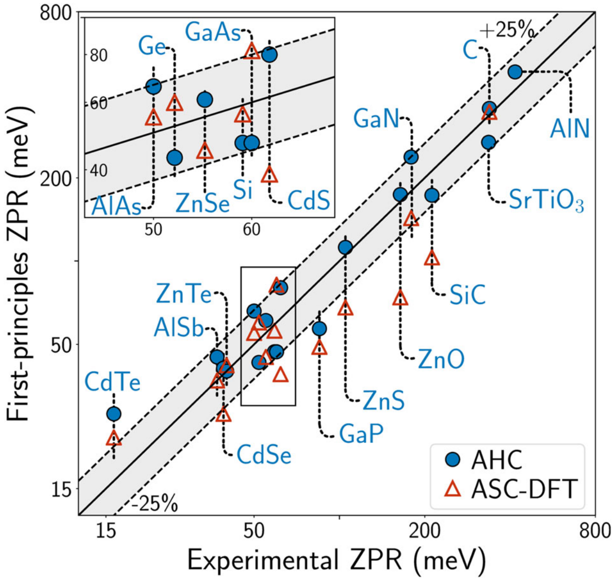

Fig. 2 Absolute values of first-principles band gap renormalization ZPR$_g$ compared with experimental ones. Blue full circles: present calculations, using non-adiabatic AHC, based on KS-DFT ingredients; red empty triangles: adiabatic supercell KS-DFT results [34]. Dashed lines: limits at which the smallest of both ZPR is 25% smaller than the largest one (note that the scales are logarithmic). For consistency, the computed lattice ZPR$_g^{lat}$ were added to both ASC and AHC. The majority of non-adiabatic AHC-based results fall within the 25% limits, while this is not the case for ASC-based ones. See numerical values in Supplementary Table II.

By contrast, the non-adiabatic AHC ZPR$_g$ (blue full circles) and experimental ZPR$_g$ agree with each other within 25% for 16 out of the 18 materials. The outliers are CdTe with a 43% overestimation by AHC, and GaP with a 33% underestimation. For none of these the discrepancy is a factor of two or larger. On the contrary, in the ASC approach, several materials show underestimation of the ZPR$_g$ by more than a factor of 2. The materials showing such large underestimation (CdS, ZnO, SiC) are all quite ionic, while more covalent materials (C, Si, Ge, AlSb, AlAs) are better described.

Therefore, Fig. 2 clearly shows that the non-adiabatic AHC approach performs significantly better than the ASC approach. AHC ZPR$_g$ and ASC ZPR$_g$ also differ by more than a factor of two for TiO$_2$ and MgO (see Supplementary Note 3), although no experimental ZPR$_g$ is available for these materials to our knowledge.

We now examine band gaps. Figure 3 presents the ratio between first-principles band gaps and corresponding experimental values, for 12 materials. The best first-principles values at fixed equilibrium atomic position, from GWeh[8], are represented, as well as their non-adiabatic AHC ZPR corrected values.

For GWeh without ZPR$_g$, a 4% agreement is obtained only for two materials (CdS and GaN). There is indeed a clear, albeit small, tendency of GWeh to overestimate the band gap value, except for the 3 materials containing shallow core d-electrons (ZnO, ZnS, and CdS) that are underestimated. By contrast, if the non-adiabatic AHC ZPR$_g$ is added to the GWeh data (blue dots), a 4% agreement is obtained for 9 out of the 12 materials (8 if ZPR$_g^{lat}$ is not included).

For ZnO and ZnS, with a final 10–12% underestimation, and CdS with a 5% underestimation, we question the GWeh ability to produce accurate fixed-geometry band gaps at the level obtained

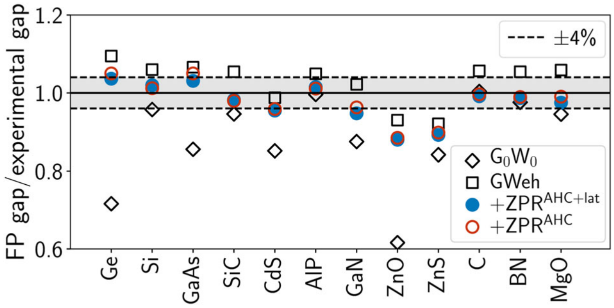

Fig. 3 Ratio between first-principles band gaps and experimental ones[64]. First-principles results without electron–phonon interaction[8] are based on G_0W_0 (diamonds) or on GWeh (squares). Blue full circles are the GWeh results to which AHC ZPR_g and lattice ZPR_g^{lat} have been added, while empty red circles are the GWeh results to which only AHC ZPR_g was added. GWeh usually overshoots, while adding the electron–phonon interaction gives better global agreement. See numerical values in Supplementary Table III.

for the other materials, due to the presence of rather localized 3d electrons in Zn and 4d electrons in CdS.

As a final lesson from Fig. 3, we note that for 4 out of the 12 materials (SiC, AlP, C, and BN), the ZPR_g is similar in size to the G_0W_0 to GWeh correction, and it is a significant fraction of it also for Si, GaN and MgO. As mentioned earlier, GWeh calculations are much more time-consuming than G_0W_0 calculations, possibly even more time-consuming than ZPR calculations (although we have not attempted to make a fair comparison). Thus, for materials containing light elements, first row and likely second row (e.g., AlP) in the periodic table, GWeh calculations miss their target if not accompanied by ZPR_g calculations. Variance and accuracy of G_0W_0 calculations is discussed in the litterature[43].

Supplementary Table V gathers our full set of ZPR results, beyond those present in the ASC or experimental sets. It also includes 10 oxides, while the experimental set only includes three materials containing oxygen (ZnO, MgO, and $SrTiO_3$), and there are none in KEKF ASC set. The ZPR_g of the band gap for materials containing light elements (O or lighter) is between −157 meV (ZnO) and −699 meV (BeO), while, relatively to the experimental band gap, it ranges from −4.6% (ZnO) to −10.8% (TiO_2). This can hardly be ignored in accurate calculations of the gap.

Generalized Fröhlich model

We now come back to the physics from which the ZPR_g originates, and explain our earlier observation that the ASC describes reasonably well the more covalent materials, but can fail badly for ionic materials. We argue that, for many polar materials, the ZPR_g is dominated by the diverging electron–phonon interaction between zone-center LO phonons and electrons close to the band edges, and the slow (non-adiabatic) response of the latter: the effects due to comparatively fast phonons are crucial. This was already the message from Fröhlich[36] and Feynman[37], decades ago, initiating large-polaron studies. However, large-scale assessment of the adequacy of Fröhlich model for real materials is lacking. Indeed, the available flavors of Fröhlich model, based on continuum (macroscopic) electrostatic interactions, do not cover altogether degenerate and anisotropic electronic band extrema or are restricted to only one phonon branch, unlike most real materials[44–47].

In this respect, we introduce now a generalized Fröhlich model, gFr, whose form is deduced from first-principles equations in the long-wavelength limit (continuum approximation), see Supplementary Note 5. Such model covers all situations and still uses as input only long-wavelength parameters, that can be determined either from first principles (Supplementary Note 6) or from experiment (Supplementary Note 7). We then find the corresponding ZPR_g from perturbation theory at the lowest order, evaluate it for our set of materials, and compares its results to the ZPR_g obtained from full first-principles computations. The corresponding Hamiltonian writes (see Supplementary Note 5 for detailed explanations):

$$\hat{H}^{\mathrm{gFr}} = \hat{H}^{\mathrm{gFr}}_{\mathrm{el}} + \hat{H}^{\mathrm{gFr}}_{\mathrm{ph}} + \hat{H}^{\mathrm{gFr}}_{\mathrm{EPI}}, \tag{1}$$

with (i) an electronic part

$$\hat{H}^{\mathrm{gFr}}_{\mathrm{el}} = \sum_{\mathbf{k}n} \frac{k^2}{2m^*_n(\hat{\mathbf{k}})} \hat{c}^+_{\mathbf{k}n}\hat{c}_{\mathbf{k}n}, \tag{2}$$

that includes direction-dependent effective masses $m^*_n(\hat{\mathbf{k}})$, governed by so-called Luttinger parameters in case of degeneracy, electronic creation and annihilation operators, $\hat{c}^+_{\mathbf{k}n}$ and $\hat{c}_{\mathbf{k}n}$, with $\mathbf{k}$ the electron wavevector and n the band index, (ii) the multi-branch phonon part,

$$\hat{H}^{\mathrm{gFr}}_{\mathrm{ph}} = \sum_{\mathbf{q}j} \omega_{j0}(\hat{\mathbf{q}}) \hat{a}^+_{\mathbf{q}j}\hat{a}_{\mathbf{q}j}, \tag{3}$$

with direction-dependent phonon frequencies $\omega_{j0}(\hat{\mathbf{q}})$, phonon creation and annihilation operators, $\hat{a}^+_{\mathbf{q}j}$ and $\hat{a}_{\mathbf{q}j}$, with $\mathbf{q}$ the phonon wavevector and j the branch index, and finally (iii) the electron–phonon interaction part

$$\hat{H}^{\mathrm{gFr}}_{\mathrm{EPI}} = \sum_{\mathbf{q}j,\mathbf{k}n'n} g^{\mathrm{gFr}}(\mathbf{q}j,\mathbf{k}n'n)\hat{c}^+_{\mathbf{k}+\mathbf{q}n'}\hat{c}_{\mathbf{k}n}(\hat{a}_{\mathbf{q}j} + \hat{a}^+_{-\mathbf{q}j}). \tag{4}$$

The $\mathbf{k}$ and $\mathbf{q}$ sums run over the Brillouin zone. The sum over n and n' runs only over the bands that connect to the degenerate extremum, renumbered from 1 to n_{deg}. The generalized Fröhlich electron–phonon interaction[48,49] is

$$\begin{aligned} g^{\mathrm{gFr}}(\mathbf{q}j,\mathbf{k}n'n) = & \frac{i}{q}\frac{4\pi}{\Omega_0}\Big(\frac{1}{2\omega_{j0}(\hat{\mathbf{q}})V_{\mathrm{BvK}}}\Big)^{1/2}\frac{\hat{\mathbf{q}}\cdot\mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^{\infty}(\hat{\mathbf{q}})} \\ & \times \sum_m s_{n'm}(\hat{\mathbf{k}}')(s_{nm}(\hat{\mathbf{k}}))^*. \end{aligned} \tag{5}$$

This expression depends on the directions $\hat{\mathbf{k}}$, $\hat{\mathbf{q}}$, and $\hat{\mathbf{k}}'$ ($\mathbf{k}' = \mathbf{k} + \mathbf{q}$), but not explicitly on their norm (hence only long-wavelength parameters are used), except for the $\frac{1}{q}$ factor. The electron–phonon part also depends only on few quantities: the Born effective charges (entering the mode-polarity vectors $\mathbf{p}_j$), the macroscopic dielectric tensor ϵ^{∞}, and the phonon frequencies ω_{j0}, the primitive cell volume Ω_0, the Born-von Karman normalization volume V_{BvK} corresponding to the $\mathbf{k}$ and $\mathbf{q}$ samplings. The s tensors are symmetry-dependent unitary matrices, similar to spherical harmonics. Equations (1)–(5) define our generalized Fröhlich Hamiltonian. Although we will focus on its $T = 0$ properties within perturbation theory, such Hamiltonian could be studied for many different purposes (non-zero T, mobility, optical responses ...), like the original Fröhlich model, for representative materials using first-principles or experimental parameters.

The conduction band ZPR_c^{gFr} can be obtained with a perturbation treatment (Supplementary Note 5), giving

$$\begin{aligned} ZPR_c^{\mathrm{gFr}} = & -\sum_{jn}\frac{1}{\sqrt{2}\Omega_0 n_{\mathrm{deg}}}\int_{4\pi} d\hat{\mathbf{q}}(m^*_n(\hat{\mathbf{q}}))^{1/2} \\ & \times (\omega_{j0}(\hat{\mathbf{q}}))^{-3/2}\left(\frac{\hat{\mathbf{q}}\cdot\mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^{\infty}(\hat{\mathbf{q}})}\right)^2. \end{aligned} \tag{6}$$

A similar expression exists for the valence band ZPR_v^{gFr}. The few material parameters needed in Eq. (6) can be obtained from experimental measurements, but are most easily computed from first principles, using density-functional perturbation theory with calculations only at $\mathbf{q} = \Gamma$ (e.g., no phonon band structure calculation). Eq. (6) can be evaluated for all band extrema in our set of materials, irrespective of whether the extrema are located at Γ or other points in the Brillouin Zone (e.g., X for the valence band of many oxides, with anisotropic effective mass), whether they are degenerate (e.g., the three-fold degeneracy of the top of the valence band of many III–V or II–VI compounds), and irrespective of the number of phonon branches (e.g., 3 different LO frequencies for TiO_2, moreover varying with the direction along which $\mathbf{q} \to 0$).

Figure 4 compares the band gap ZPR from the first-principles non-adiabatic AHC methodology and from the generalized Fröhlich model. The 30 materials can be grouped into five sets, based on their ionicity: 11 materials containing oxygen, rather ionic, for which the Born effective charges and the ZPR are quite large, 6 materials containing chalcogenides, also rather ionic, 4 materials containing nitrogen and 5 III-V materials, less ionic, and 4 materials from group-IV elements, non-ionic, except SiC.

For oxygen-based materials, the ZPR ranges from 150 meV to 700 meV, and the gFr model captures this very well, with less than 25% error, with only one exception, BeO. The chalcogenide materials are also reasonably well described by the gFr model, capturing at least two-third of the ZPR. Globally their ZPR is smaller (note the logarithmic scale).

For the nitride materials and for SiC, the gFr captures about 50% of the quite large ZPR (between 176 meV and 406 meV). The adequacy of the gFr model decreases still with the III–V materials and the three non-ionic IV materials. In the latter case, the vanishing Born effective charges result in a null ZPR within the gFr model (these three materials are omitted from Fig. 4).

For the oxides and chalcogenides, the ZPR is thus dominated by the zone-center parameters (including the phonon frequencies), and the physics corresponds to the one of the large-polaron picture[37], namely, the slow electron motion is correlated to a phonon cloud that dynamically adjusts to it. This physics is completely absent from the ASC approach. Even for nitrides, the gFr describes a significant fraction of the ZPR.

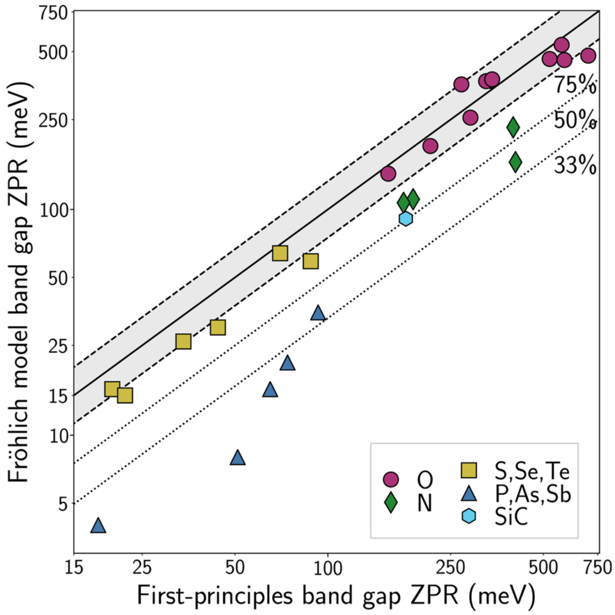

Fig. 4 Zero-point renormalization of the band gap: comparison between first-principles AHC values and generalized Fröhlich model. See Eq. (6). Data for 27 materials are represented. Non-ionic group IV materials are omitted. The markers identify materials of similar ionicity: oxides (purple circles), other II–VI materials containing S, Se, or Te (yellow squares), nitrides (green diamonds), other III-V materials, containing P, As, or Sb (dark blue triangles), and group IV material SiC (light blue hexagon). Dashed lines: limits at which the smallest of both ZPR is 25% smaller than the largest one. Dotted lines: limits at which the Fröhlich model ZPR misses, respectively, 50% and 67% of the AHC values (note that the scales are logarithmic). See numerical values in Supplementary Table V.

A perfect agreement between the non-adiabatic AHC first-principles ZPR and the generalized Fröhlich model ZPR is not expected. Indeed, differences can arise from different effects: lack of dominance of the Fröhlich electron–phonon interaction in some regions of the Brillouin Zone, departure from parabolicity of the electronic structure (obviously, the electronic structure must be periodic so that the parabolic behavior does not extend to infinity), interband contributions, phonon band dispersion, incomplete cancellation between the Debye–Waller and the acoustic phonon mode contribution.

It is actually surprising to see that for so many materials, the generalized Fröhlich model matches largely the first-principles AHC results. Anyhow, as a conclusion for this section, for a large number of materials, we have validated, a posteriori and from first principles, the relevance of large-polaron research based on Fröhlich model despite the numerous approximations on which it relies.

DISCUSSION

We focus on the mechanism by which the AHC divergence of the ZPR in the adiabatic case for infrared-active materials[24] is avoided, either using the ASC methodology or using the non-adiabatic AHC methodology. As Fröhlich and Feynman have cautioned us[36,37], and already mentioned briefly in previous sections, the dynamics of the “slow” electron is crucial in this electron–phonon problem.

In the ASC approach, the bypass of this divergence can be understood as follows, see Fig. 5a. Consider a long-wavelength fluctuation of the atomic positions, frozen in time. At large but finite wavelength, the potential is periodically lowered in some regions of space and increased in some other regions of space, in an oscillatory manner with periodicity $\Delta L \propto 1/q$, where $\mathbf{q}$ is the small wavevector of the fluctuation, see Fig. 5a “LO phonon potential” part. Without such long-wavelength potential, the

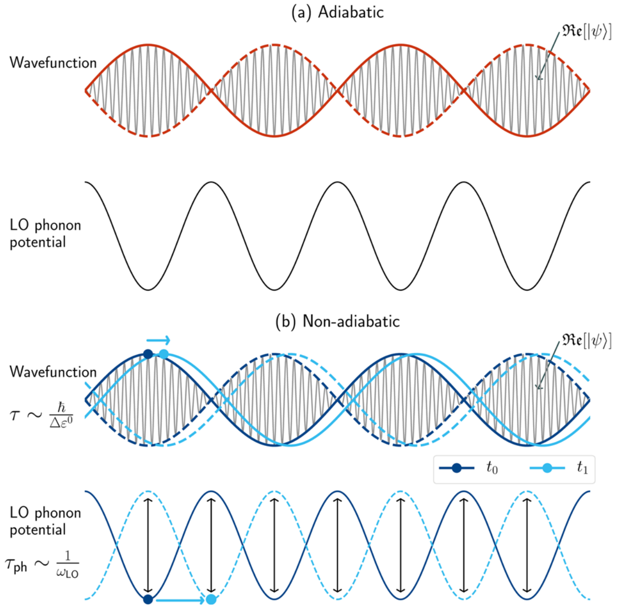
(a) Adiabatic
Wavefunction
$\mathfrak{Re}[|\psi\rangle]$
LO phonon potential
(b) Non-adiabatic
Wavefunction
$\tau \sim \frac{\hbar}{\Delta\varepsilon^0}$
$\mathfrak{Re}[|\psi\rangle]$
t_0
t_1
LO phonon potential
$\tau_{\mathrm{ph}} \sim \frac{1}{\omega_{\mathrm{LO}}}$

Fig. 5 Schema of the long-wave phonon-induced potential and corresponding electronic wavefunction. Real part (oscillating with lattice periodicity) and envelope are represented. **a** Adiabatic case, **b** non-adiabatic, time-dependent case. In the non-adiabatic case, the electron does not have the time to adjust to the change of potential, see text.

electron at the minimum of the conduction band has a Bloch type wavefunction, with an envelope phase factor characterized by the wavevector $\mathbf{k}_c$ multiplying a lattice-periodic function. Its density is lattice periodic. With such long-wavelength potential, as a function of the amplitude of the atomic displacements, the corresponding electronic eigenenergy changes first quadratically (as the average of the lowering and increase of potential for this Bloch wavefunction forbids a linear behavior except in case of degeneracy), but for larger amplitudes, it behaves linearly, as the electron localizes in the lowered potential region and the minimum of the potential is linear in the amplitude of the atomic displacements. This is referred to as "nonquadratic coupling"[32]. A wavepacket is formed, by combining Bloch wavefunctions with similar lattice periodic functions but slightly different wavevectors ($\mathbf{k}_c$, $\mathbf{k}_c + \mathbf{q}$, $\mathbf{k}_c - \mathbf{q}$, etc), coming from a small interval of energy $\Delta\epsilon^0 \propto q^2 \propto 1/(\Delta L)^2$, see Fig. 5a "Wavefunction" part. This nonquadratic effect is actually illustrated in Fig. 4 of ref. [23] (see the frozen-phonon eigenvalues), as well as in Fig. 2 of ref. [32].

By contrast, in the time-dependent case, as illustrated in Fig. 5b, the wavepacket will require a characteristic time $\Delta\tau$, to form or to displace. This will be given by the Heisenberg uncertainty relation, $\Delta\epsilon^0\Delta\tau \gtrsim \hbar$, hence $\Delta\tau \gtrsim \hbar/\Delta\epsilon^0 \propto \Delta L^2$. For long wavelengths, the characteristic time diverges. As soon as $\Delta\tau$ is larger than the phonon characteristic time $\tau_{\mathrm{ph}} \sim 1/\omega_{\mathrm{LO}}$, the "slow" electron will lag behind the phonon, and the static or adiabatic picture described above is no longer valid.

In all adiabatic approaches, either AHC or ASC, the electron is always supposed to have the time to adjust to the change of potential, in contradiction with the time-energy uncertainty principle. Furthermore, the adiabatic AHC approach only considers the quadratic region for the above-mentioned dependence of eigenvalues with respect to amplitudes of displacements. This results in a diverging term[24]. At variance with the AHC case, the ASC approach samples a whole set of amplitudes, including the onset of the asymptotic linear regime, in which case the divergence does not build up. However, this ASC picture does not capture the real physical mechanism that prevents the divergence to occur, the impossibility for the electron to follow the phonon dynamics, that we have highlighted above. By contrast, such physical mechanism is present both in the non-adiabatic AHC approach and in the (generalized) Fröhlich model: the "slow" electron does not follow adiabatically (instantaneously) the atomic motion. The divergence of the adiabatic AHC is indeed avoided in the non-adiabatic picture by taking into account the non-zero phonon frequencies.

Thus, the ASC avoids the adiabatic AHC divergence for the wrong reason, which explains its poor predictive capability for the more ionic materials emphasized by Fig. 2. To be clear, we do not pretend the nonquadratic effects are all absent, but the non-adiabatic effects have precedence, at least for materials with significant infrared activity, and the nonquadratic localization effects will be observed only if the electrons have the time to physically react. The shortcomings of the ASC approach are further developed in the Supplementary Discussion. As a consequence of such understanding, all the results obtained for strongly infrared-active materials using the adiabatic frozen-phonon supercell methodology should be questioned.

For non-infrared-active materials, the physical picture that we have outlined, namely the inability of slow electrons to follow the dynamics of fast phonons, is still present, but does not play such a crucial role: the electron–phonon interaction by itself does not diverge in the long-wavelength limit as compared to the infrared-active electron–phonon interaction, see Eq. (5), only the denominator of the Fan self-energy diverges, which nevertheless results in an integrable ZPR[24]. In such case, neglecting non-adiabatic effects, as in the ASC approach, is only one among many approximations done to obtain the ZPR.

Beyond the discovery of the predominance of non-adiabatic effects in the zero-point renormalization of the band gap for many materials, in the present large-scale first-principles study of this effect, we have established that electron–phonon interaction diminishes the band gap by 5% to 10% for materials containing light atoms like N or O (up to 0.7 eV for BeO), a decrease that cannot be ignored in accurate calculations of the gap. Our methodology, the non-adiabatic Allen-Heine-Cardona approach, has been validated by showing that, for nearly all materials for which experimental data exists, it achieves quantitative agreement (within 25%) for this property.

We have also shown that most of the discrepancies with respect to experimental data of the (arguably) best available methodology for the first-principles band-gap computation, denoted GWeh, originate from the first-principles zero-point renormalization: after including it, the average overestimation from GWeh nearly vanishes. There are some exceptions, materials in which transition metals are present, for which the addition of zero-point renormalization worsens the agreement of the band gap. For the latter materials, we believe that the GWeh approach is not accurate enough.

METHODS

First-principles electronic and phonon band structures

Calculations have been performed using ABINIT[50] with norm-conserving pseudopotentials and a plane-wave basis set. Supplementary Table I provides calculation parameters: plane-wave kinetic cut-off energy, electronic wavevector sampling in the BZ, and largest phonon wavevector sampling in the BZ. For most of the materials, the GGA-PBE exchange-correlation functional[51] has been used and the pseudopotentials have been taken from the PseudoDojo project[52]. For diamond, BN-zb, and AlN-wz, results reported here come from a previously published work[24], where the LDA has been used, with other types of pseudopotentials.

The calculations have been performed at the theoretical optimized lattice parameter, except for Ge for which the gap closes at such parameter, for GaP, as at such parameter the conduction band presents unphysical quasi-degenerate valleys, and for TiO_2, as the GGA-PBE predicted structure is unstable[53]. For these, we have used the experimental lattice parameter. The case of $SrTiO_3$ is specific and will be explained later.

Density-functional perturbation theory[54–56] has been used for the phonon frequencies, dielectric tensors, Born effective charges, effective masses, and electron–phonon matrix elements.

First-principles calculations of zero-point renormalization

We first detail the method used for the AHC calculations. In the many-body perturbation theory approach, an electronic self-energy Σ appears due to the electron–phonon interaction, with Fan and Debye–Waller contributions at the lowest order of perturbation[35]:

$$\Sigma_{\mathbf{k}n}(\omega) = \Sigma^{\mathrm{Fan}}_{\mathbf{k}n}(\omega) + \Sigma^{\mathrm{DW}}_{\mathbf{k}n}. \tag{7}$$

The Hartree atomic unit system is used throughout ($\hbar = m_e = e = 1$). An electronic state is characterized by $\mathbf{k}$, its wavevector, and n, its band index, ω being the frequency. These two contributions correspond to the two diagrams presented in Fig. 1.

Approximating the electronic Green's function by its non-interacting KS-DFT counterpart without electron–phonon interaction, gives the standard result for the $T = 0$ K retarded Fan self-energy[35]:

$$\begin{aligned}\overset{\mathrm{Fan}}{\underset{\mathbf{k}n}{\Sigma}}(\omega) = {} & \frac{1}{N_{\mathbf{q}}}\sum_{\mathbf{q}j}^{\mathrm{BZ}}\sum_{n'}|\langle\mathbf{k}+\mathbf{q}n'|H^{(1)}_{\mathbf{q}j}|\mathbf{k}n\rangle|^2 \\ & \times\left[\frac{1-f_{\mathbf{k}+\mathbf{q}n'}}{\omega-\varepsilon_{\mathbf{k}+\mathbf{q}n'}-\omega_{\mathbf{q}j}+i\eta}+\frac{f_{\mathbf{k}+\mathbf{q}n'}}{\omega-\varepsilon_{\mathbf{k}+\mathbf{q}n'}+\omega_{\mathbf{q}j}+i\eta}\right].\end{aligned} \tag{8}$$

In this expression, contributions from phonon modes with harmonic phonon energy $\omega_{\mathbf{q}j}$ are summed for all branches j, and wavevectors $\mathbf{q}$, in the entire Brillouin Zone (BZ). The limit for infinite number $N_{\mathbf{q}}$ of wavevectors (homogeneous sampling) is implied. Contributions from transitions to electronic states $\langle\mathbf{k}+\mathbf{q}n'|$ with KS-DFT electron energy $\varepsilon_{\mathbf{k}+\mathbf{q}n'}$ and occupation number $f_{\mathbf{k}+\mathbf{q}n'}$ (1 for valence, 0 for conduction, at $T = 0$ K) are summed for all bands n' (valence and conduction). The $H^{(1)}_{\mathbf{q}j}$ is the self-consistent change of potential due to the $\mathbf{q}j$-phonon[35]. Limit of this expression for vanishing positive η is implied. For the Debye-Waller self-energy, $\Sigma^{\mathrm{DW}}_{\mathbf{k}n}$, we refer to the litterature[21,35].

In the non-adiabatic AHC approach, the ZPR is obtained directly from the real part of the self-energy, Eq. (7), evaluated at $\omega = \varepsilon_{\mathbf{k}n}$[35]:

$$\mathrm{ZPR}^{\mathrm{AHC}}_{\mathbf{k}n} = \Re e \Sigma_{\mathbf{k}n}(\omega = \varepsilon_{\mathbf{k}n}). \tag{9}$$

If the adiabatic approximation is made, the phonon frequencies $\omega_{\mathbf{q}j}$ are considered small with respect to eigenenergy differences in the denominator of Eq. (8) and are simply dropped, while a finite η, usually 0.1 eV, is kept. With a vanishing η, the adiabatic AHC ZPR at band edges diverges for infrared-active materials, see Supplementary Note 5.

Summing the Fan and Debye-Waller self-energies, and working also with the rigid-ion approximation for the Debye-Waller contribution delivers the non-adiabatic AHC ZPR, given explicitly in Eqs. (16) and (17) of ref. [24]. We do not work with the approach called dynamical AHC, also mentioned in this work[24]. It corresponds to Eq. (166) of ref. [35]. Both non-adiabatic and dynamical AHC ZPR flavors were studied e.g., in ref. [23], but the comparison with the diagrammatic quantum Monte Carlo results for the Fröhlich model, see e.g., Fig. 1 of ref. [25], is clearly in favor of the non-adiabatic AHC approach. Actually, this also constitutes a counter-argument to the claim by Cannuccia and Marini[17], that band theory might not apply to carbon-based nanostructures. The cumulant expansion results for the spectral function[25] demonstrates that the dynamical AHC spectral function is unphysical, with only one wrongly placed satellite. The physical content of the present non-adiabatic AHC theory, focusing on the crucial role of the LO phonons, is very different from the physical analysis based on the dynamical AHC theory[17].

The imaginary smearing of the denominator in the ZPR computation is 0.01 eV, except for SiC, where 0.001 eV is used. Other technical details are similar to previous studies by some of ours[24,25].

The dependence of the electronic structure on zero-point lattice parameter corrections is computed from

$$\mathrm{ZPR}^{\mathrm{lat}}_{\mathbf{k}n} = \varepsilon_{\mathbf{k}n}(\{\mathbf{R}_i^{T=0}\}) - \varepsilon_{\mathbf{k}n}(\{\mathbf{R}_i^{\mathrm{fix}}\}), \tag{10}$$

where the lattice parameters $\{\mathbf{R}_i^{\mathrm{fix}}\}$ minimize the Born-Oppenheimer energy without phonon contribution, while $\{\mathbf{R}_i^{T=0}\}$ minimizes the free energy that includes zero-point phonon contributions, see Supplementary Note 2.

At variance with the AHC approach, in the ASC the temperature-dependent average band edges (here written for the bottom of the conduction band) are obtained from

$$\langle \varepsilon_c(T) \rangle = Z_I^{-1} \sum_m \exp(-\beta E_m) \langle \varepsilon_c \rangle_m, \tag{11}$$

where $\beta = k_B T$ (with k_B the Boltzmann constant), T the temperature, Z_I the canonical partition function among the quantum nuclear states m with energies E_m ($Z_I = \sum_m \exp(-\beta E_m)$), and $\langle \varepsilon_c \rangle_m$ the band edge average taken over the corresponding many-body nuclear wavefunction. At zero Kelvin, this gives an instantaneous average of the band edge value over zero-point atomic displacements, computed while the electron is NOT present in the conduction band (or hole in the valence band) thus suppressing all correlations between the phonons and the added (or removed) electron.

Convergence of the calculations

As previously noted[24,25], the sampling of phonon wavevectors in the Brillouin zone is a delicate issue, and has been thoroughly analyzed in Sec. IV.B.2 of ref. [24]. In particular, for infrared-active materials treated with the non-adiabatic effects, at the band structure extrema, a $N_{\mathbf{q}}^{-1/3}$ convergence of the value is obtained. We have taken advantage of the knowledge acquired[24] to accelerate the convergence by three different methodologies. In the first one, we fit the $N_{\mathbf{q}}^{-1/3}$ behavior for grids of different sizes, and extrapolate to infinite $N_{\mathbf{q}}$. In the second one, we estimate the missing contribution to the integral around $\mathbf{q} = 0$, at the lowest order, see Supplementary Methods B, using ingredients similar to those needed for the generalized Fröhlich model, except the effective masses. For ZnO and $\mathrm{SrTiO_3}$, a third correcting scheme, further refining the region around the band edge with an extremely fine grid, is used.

The special case of $\mathrm{SrTiO_3}$

While phonons in most materials in this study are well addressed within the harmonic approximation, this is not the case for $\mathrm{SrTiO_3}$. This material is found in the cubic perovskite structure at room temperature, and undergoes a transition to a tetragonal phase below 150 K, characterized by tilting of the $\mathrm{TiO_6}$ octahedra[57]. Within our first-principles scheme in the adiabatic approximation, the cubic phase remains unstable with respect to tilting of the octahedra. Quantum fluctuations of the atomic positions actually play a critical role in stabilizing the cubic phase at high temperature[58], as well as suppressing the ferroelectric phase at low temperature[59,60].

Since $\mathrm{SrTiO_3}$ is however a material for which large polaron effects are clearly identified[46], we decided to study it as well. We addressed the anharmonic stabilization problem using the state-of-the-art TDep methodology[61]. We used VASP molecular dynamics to generate 40 configurations in a 2 × 2 × 2 cubic cell of STO at 300 K, producing 20,000 steps with 2 fs per step and sampling the 40 configurations out of the last 5000 steps. Then we computed the forces with ABINIT and performed TDep calculations with the ALAMODE code[62]. Our calculation stabilizes the acoustic phonon branches and yields a phonon band structure in good agreement with experimental data.

Sources of discrepancies between experiment and theory

The anharmonic corrections to phonon frequencies are not the only reasons for potential differences between the experimental $\mathrm{ZPR_g}$ and our non-adiabatic AHC $\mathrm{ZPR_g}$ values. The following phenomena may also play a role: (1) the rigid-ion approximation[18,21]; (2) the nonquadratic behavior of the eigenenergies with collective displacements of the nuclei, in reference to the ASC[32]; (3) the reliance on GGA-PBE eigenenergies and eigenfunctions, instead of more accurate (e.g., GW) ones[29,34]; (4) self-trapping effects, overcoming the quantum fluctuations, yielding small polarons[63]. There is still little knowledge about each of these effects when correctly combined to predict the $\mathrm{ZPR_g}$ beyond the AHC picture.

As an example, in KEKF[34], the difference between the ASC-PBE and the ASC-GW was argued to be only a few meV, but a more careful look at their values show that it is often bigger than 10% of the ASC-PBE. Unfortunately, the convergence of the ASC-GW results with respect to supercell size could not be convincingly achieved by KEKF[34]. It remains to be seen whether a non-adiabatic AHC treatment based on GW matrix elements would differ by such relative ratio, see the Supplementary Discussion. Altogether, it would be hard to claim more than 25% accuracy with respect to experimental data, from our non-adiabatic AHC $\mathrm{ZPR_g}$ calculations. Together with the experimental uncertainties, this explains our choice for the 25% accuracy comparative limit used in Fig. 2.

DATA AVAILABILITY

The numerical data used to create all the figures in the main text have been collected in the Supplementary Tables.

CODE AVAILABILITY

ABINIT is available under GNU General Public Licence from the ABINIT web site (http://www.abinit.org).

Received: 6 June 2020; Accepted: 30 September 2020;
Published online: 06 November 2020

REFERENCES

1. Yu, P. Y. & Cardona, M. *Fundamentals of Semiconductors*. 4th edn. (Springer-Verlag, Berlin, 2010).
2. Martin, R. M. *Electronic Structure: Basic Theory and Methods* (Cambrige University Press, Cambridge, 2004).
3. Perdew, J. P. & Levy, M. Physical content of the exact Kohn-Sham orbital energies: band gaps and derivative discontinuities. *Phys. Rev. Lett.* **51**, 1884–1887 (1983).
4. Sham, L. J. & Schlüter, M. Density-functional theory of the energy gap. *Phys. Rev. Lett.* **51**, 1888–1891 (1983).
5. Martin, R., Reining, L. & Ceperley, D. *Interacting Electrons. Theory and Computational Approaches* (Cambrige University Press, Cambridge, 2016).
6. Hybertsen, M. S. & Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B* **34**, 5390–5413 (1986).
7. van Schilfgaarde, M., Kotani, T. & Faleev, S. Quasiparticle self-consistent GW theory. *Phys. Rev. Lett.* **96**, 226402 (2006).
8. Shishkin, M., Marsman, M. & Kresse, G. Accurate quasiparticle spectra from self-consistent GW calculations with vertex corrections. *Phys. Rev. Lett.* **99**, 246403 (2007).

9. Allen, P. B. & Heine, V. Theory of the temperature dependence of electronic band structures. *J. Phys. C* **9**, 2305–2312 (1976).
10. Allen, P. B. & Cardona, M. Theory of the temperature dependence of the direct gap of germanium. *Phys. Rev. B* **23**, 1495–1505 (1981).
11. Fan, H. Y. Temperature dependence of the energy gap in semiconductors. *Phys. Rev.* **82**, 900–905 (1951).
12. Antoncik, E. On the theory of temperature shift of the absorption curve in nonpolar crystals. *Cechoslovackij Fiziceskij Zurnal* **5**, 449–461 (1955).
13. Kim, C., Lautenschlager, P. & Cardona, M. Temperature dependence of the fundamental energy gap in GaAs. *Solid State Commun.* **59**, 797–802 (1986).
14. Zollner, S., Gopalan, S. & Cardona, M. The temperature dependence of the band gaps in InP, InAs, InSb, and GaSb. *Sol. State Commun.* **77**, 485–488 (1991).
15. Marini, A. Ab-initio finite temperature excitons. *Phys. Rev. Lett.* **101**, 106405 (2008).
16. Giustino, F., Louie, S. & Cohen, M. Electron-phonon renormalization of the direct band gap of diamond. *Phys. Rev. Lett.* **105**, 265501 (2010).
17. Cannuccia, E. & Marini, A. 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).
18. Gonze, X., Boulanger, P. & Côté, M. Theoretical approaches to the temperature and zero-point motion effects on the electronic band structure. *Annalen der Physik* **523**, 168–178 (2011).
19. Cannuccia, E. & Marini, A. Zero point motion effect on the electronic properties of diamond, trans-polyacetylene and polyethylene. *Eur. Phys. J. B* **85**, 320 (2012).
20. Kawai, H., Yamashita, K., Cannuccia, E. & Marini, A. 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).
21. Poncé, S. et al. 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. *Comp. Mat. Sci.* **83**, 341–348 (2014).
22. Poncé, S. et al. Temperature dependence of electronic eigenenergies in the adiabatic harmonic approximation. *Phys. Rev. B* **90**, 214304 (2014).
23. Antonius, G. et al. Dynamical and anharmonic effects on the electron-phonon coupling and the zero-point renormalization of the electronic structure. *Phys. Rev. B* **92**, 085137 (2015).
24. Poncé, S. et al. Temperature dependence of the electronic structure of semiconductors and insulators. *J. Chem. Phys.* **143**, 102813 (2015).
25. Nery, J. et al. Quasiparticles and phonon satellites in spectral functions of semiconductors and insulators: Cumulants applied to the full first-principles theory and the Fröhlich polaron. *Phys. Rev. B* **97**, 115145 (2018).
26. Capaz, R. B., Spataru, C. D., Tangney, P., Cohen, M. L. & Louie, S. G. Temperature dependence of the band gap of semiconducting carbon nanotubes. *Phys. Rev. Lett.* **94**, 036801 (2005).
27. Patrick, C. E. & Giustino, F. Quantum nuclear dynamics in the photophysics of diamondoids. *Nat. Commun.* **4**, 2006 (2013).
28. Monserrat, B., Drummond, N. & Needs, R. Anharmonic vibrational properties in periodic systems: energy, electron-phonon coupling, and stress. *Phys. Rev. B* **87**, 144302 (2013).
29. Antonius, G., Poncé, S., Boulanger, P., Côté, M. & Gonze, X. Many-body effects on the zeropoint renormalization of the band structure. *Phys. Rev. Lett.* **112**, 215501 (2014).
30. Monserrat, B., Drummond, N., Pickard, C. & Needs, R. Electron-phonon coupling and the metallization of solid helium at terapascal pressures. *Phys. Rev. Lett.* **112**, 055504 (2014).
31. Monserrat, B. & Needs, R. Comparing electron-phonon coupling strength in diamond, silicon, and silicon carbide: first-principles study. *Phys. Rev. B* **89**, 214304 (2014).
32. Monserrat, B., Engel, E. & Needs, R. Giant electron-phonon interactions in molecular crystals and the importance of nonquadratic coupling. *Phys. Rev. B* **92**, 140302(R) (2015).
33. Engel, E. A., Monserrat, B. & Needs, R. J. Vibrational renormalisation of the electronic band gap in hexagonal and cubic ice. *J. Chem. Phys.* **143**, 244708 (2015).
34. Karsai, F., Engel, M., Kresse, G. & Flage-Larsen, E. Electron–phonon coupling in semiconductors within the GW approximation. *New J. Phys.* **20**, 123008 (2018).
35. Giustino, F. Electron-phonon interactions from first principles. *Rev. Mod. Phys.* **89**, 015003 (2017).
36. Fröhlich, H. Electrons in lattice fields. *Adv. Phys.* **3**, 325–361 (1954).
37. Feynman, R. P. Slow electrons in a polar crystal. *Phys. Rev.* **97**, 660–665 (1955).
38. Mishchenko, A., Prokof'ev, N., Sakamoto, A. & Svistunov, B. Diagrammatic quantum Monte Carlo study of the Fröhlich polaron. *Phys. Rev. B* **62**, 6317–6336 (2000).
39. Mishchenko, A. S., Nagaosa, N., Prokof'ev, N., Sakamoto, A. & Svistunov, B. Optical conductivity of the Fröhlich polaron. *Phys. Rev. Lett.* **91**, 236401 (2003).
40. Devreese, J. & Alexandrov, A. Fröhlich polaron and bipolaron: recent developments. *Rep. Prog. Phys.* **72**, 066501 (2009).
41. Emin, D. *Polarons*. (Cambrige University Press, Cambridge, 2012).
42. Hahn, T., Klimin, S., Tempere, J., Devreese, J. T. & Franchini, C. Diagrammatic Monte Carlo study of Fröhlich polaron dispersion in two and three dimensions. *Phys. Rev. B* **97**, 134305 (2018).
43. Rangel, T. et al. Reproducibility in G0W0 calculations for solids. *Comput. Phys. Commun.* **255**, 107242 (2020).
44. Mahan, G. D. Temperature dependence of the band gap in CdTe. *J. Phys. Chem. Solid.* **26**, 751–756 (1965).
45. Trebin, H.-R. & Rössler, U. Polarons in the degenerate-band case. *Phys. Status Solid. B* **70**, 717–726 (1975).
46. Devreese, J. T., Klimin, S. N., van Mechelen, J. L. M. & van der Marel, D. Many-body large polaron optical conductivity in $SrTi_{(1-x)}Nb_xO_3$. *Phys. Rev. B* **81**, 125119 (2010).
47. Schlipf, M., Poncé, S. & Giustino, F. Carrier lifetimes and polaronic mass enhancement in the hybrid halide perovskite $CH_3NH_3PbI_3$ from multiphonon Fröhlich coupling. *Phys. Rev. Lett.* **121**, 086402 (2018).
48. Vogl, P. Microscopic theory of electron-phonon interaction in insulators or semiconductors. *Phys. Rev. B* **13**, 694–704 (1976).
49. Verdi, C. & Giustino, F. Fröhlich electron-phonon vertex from first principles. *Phys. Rev. Lett.* **115**, 176401 (2015).
50. Gonze, X. et al. Recent developments in the ABINIT software package. *Comput. Phys. Commun.* **205**, 106–131 (2016).
51. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
52. van Setten, M. et al. The pseudodojo: training and grading a 85 element optimized norm-conserving pseudopotential table. *Comput. Phys. Commun.* **226**, 39–54 (2018).
53. Montanari, B. & Harrison, N. M. Lattice dynamics of TiO_2 rutile: influence of gradient corrections in density functional calculations. *Chem. Phys. Lett.* **364**, 528–534 (2002).
54. Baroni, S., de Gironcoli, S., Corso, A. D. & Giannozzi, P. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.* **73**, 515–562 (2001).
55. Gonze, X. & Lee, C. Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys. Rev. B* **55**, 10355–10368 (1997).
56. Laflamme Janssen, J. et al. Precise effective masses from density functional perturbation theory. *Phys. Rev. B* **93**, 205147 (2016).
57. Guennou, M., Bouvier, P., Kreisel, J. & Machon, D. Pressure-temperature phase diagram of $SrTiO_3$ up to 53 GPa. *Phys. Rev. B* **81**, 054115 (2010).
58. Tadano, T. & Tsuneyuki, S. Self-consistent phonon calculations of lattice dynamical properties in cubic $SrTiO_3$ with first-principles anharmonic force constants. *Phys. Rev. B* **92**, 054301 (2015).
59. Müller, K. A. & Burkard, H. $SrTiO_3$: an intrinsic quantum paraelectric below 4 K. *Phys. Rev. B* **19**, 3593–3602 (1979).
60. Zhong, W. & Vanderbilt, D. Effect of quantum fluctuations on structural phase transitions in $SrTiO_3$ and $BaTiO_3$. *Phys. Rev. B* **53**, 5047–5050 (1996).
61. Hellman, O., Steneteg, P., Abrikosov, I. A. & Simak, S. I. Temperature dependent effective potential method for accurate free energy calculations of solids. *Phys. Rev. B* **87**, 104111 (2013).
62. Tadano, T., Gohda, Y. & Tsuneyuki, S. Anharmonic force constants extracted from first-principles molecular dynamics: applications to heat transfer simulations. *J. Phys. Condens. Matter* **26**, 225402 (2014).
63. Sio, W. H., Verdi, C., Poncé, S. & Giustino, F. Polarons from first principles, without supercells. *Phys. Rev. Lett.* **122**, 246403 (2019).
64. Cardona, M. & Thewalt, M. Isotope effects on the optical spectra of semiconductors. *Rev. Mod. Phys.* **77**, 1173–1224 (2005).

ACKNOWLEDGEMENTS

We acknowledge fruitful discussions with Y. Gillet and 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, the PdR Grant No. T.0103.19 - ALPS, the Fonds de Recherche du Québec Nature et Technologie (FRQ-NT), the Natural Sciences and Engineering Research Council of Canada (NSERC) under grants RGPIN-2016-06666. Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL), the Consortium des Equipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the FRS-FNRS under Grant No. 2.5020.11, the Tier-1 supercomputer of the Fédération Wallonie-Bruxelles, infrastructure funded by the Walloon Region under the grant agreement No. 1117545, as well as the Canadian Foundation for Innovation, the Ministère de l'Éducation des Loisirs et du Sport (Québec), Calcul Québec, and Compute Canada. This work was supported by the Center for Computational Study of Excited-State Phenomena in Energy Materials (C2SEPEM) at the Lawrence Berkeley National Laboratory, which is funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division under Contract No. DE-AC02-05CH11231, as part of the Computational Materials Sciences Program (advanced algorithms/codes) and by the National Science Foundation under grant DMR-1926004 (basic theory and formalism). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of

Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

AUTHOR CONTRIBUTIONS

A.M. has conducted calculations for most oxyde materials, with help from M.G. V.B.-C. has conducted calculations for most other materials. G.A. and Y.-H.C. have conducted calculations for $SrTiO_3$ and ZnO with supervision of S.G.L.. E.G. has conducted calculations of the lattice ZPR of 6 materials. X.G. has worked out the generalized Fröhlich model and perturbative treatment. X.G. and M.C. have supervised the work. All authors have contributed to the writing of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41524-020-00434-z.

Correspondence and requests for materials should be addressed to X.G.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020