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

Permalink
https://escholarship.org/uc/item/3tx9595f

Journal
npj Computational Materials, 6(1)

ISSN
2057-3960

Authors
Miglio, A
Brousseau-Couture, V
Godbout, E
et al.

Publication Date
2020-12-01

DOI
10.1038/s41524-020-00434-z

Peer reviewed
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. 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, that delivers a theoretically unjustified value of the band gap in the standard approach, even with exact KS potential.

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, then twenty years later self-consistent (GW) and further improved by accurate vertex corrections from electron-hole excitations (GWeh). The latter methodology, at the forefront for band-gap computations, delivers a 2–10% accuracy, usually overestimating the experimental band gap. GW 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 (ZPR$_g$) for historical reasons.

The ZPR$_g$ had been examined 40 years ago, by Allen, Heine and Cardona (AHC) who clarified the early theories by Fan and Antonick. 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. It was later extended without caution to GaAs and a few other III–V semiconductors.

In this work, we present first-principles AHC ZPR$_g$ calculations beyond the adiabatic approximation, for 30 materials. Comparing with experimental band gaps, we show that adding ZPR$_g$ improves the GWeh first-principles band gap, and moreover, that the 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 ZPR$_g$ is not taken into account. By including it, the theoretical agreement with direct measurements of experimental 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 ZPR$_g$ using the AHC theory are very challenging, and only started one decade ago, on a case-by-case basis (see the Supplementary Note 3), usually relying on the adiabatic approximation, and without comparison with experimental data. An approach to the ZPR$_g$, alternative to the AHC one, relies on computations of the band gap at fixed, distorted geometries, for large supercells (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 hereafter referred to as KEKF, presents ASC 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.

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

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.
Although the adiabatic approximation had already been criticized by Poncé et al. in 2015\textsuperscript{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\textsubscript{g} of the full first-principles non-adiabatic AHC ZPR\textsubscript{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\textsuperscript{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\textsubscript{g}.

RESULTS

Zero-point renormalization: experiment vs. first principles

Figure 2 compares first-principles ZPR\textsubscript{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\textsubscript{lat}, has been added to fixed volume results from both non-adiabatic AHC (present calculations) and ASC methodologies\textsuperscript{34}. While for a few materials experimental ZPR\textsubscript{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 CdS, 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.

By contrast, the non-adiabatic AHC ZPR\textsubscript{g} (blue full circles) and experimental ZPR\textsubscript{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\textsubscript{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\textsubscript{g} and ASC ZPR\textsubscript{g} also differ by more than a factor of two for TiO\textsubscript{2} and MgO (see Supplementary Note 3), although no experimental ZPR\textsubscript{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\textsuperscript{8}, are represented, as well as their non-adiabatic AHC ZPR\textsubscript{g} corrected values.

For GWeh without ZPR\textsubscript{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\textsubscript{g} is added to the GWeh data (blue dots), a 4% agreement is obtained for 9 out of the 12 materials (8 if ZPR\textsubscript{lat} is not included).

For ZnO and ZnS, with a 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...
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 is similar in size to the G0W0 to GW correction, and it is a significant fraction of it also for Si, GaN and MgO. As mentioned earlier, G0W0 calculations are more time-consuming than G0W0 calculations, possibly even more time-consuming than GW calculations (although we have not attempted to make a fair comparison). Thus, for materials containing light elements, first row and second row (e.g., Al, AIP) in the periodic table, G0W0 calculations are efficient target if not accompanied by ZPR calculations. Variance and accuracy of G0W0 calculations is discussed in the literature.

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 SrTiO3), and there are none in KEKF ASC set. The ZPR is 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% (TiO2). 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 originates, and explain our earlier observation that the ASC describes reasonably well the more relevant materials, but can fail badly for ionic materials. We argue that, for many polar materials, the ZPR 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 zone-center LO phonons and electrons close to the band edges, and the slow (non-adiabatic) response of the latter: the zone-center LO phonons and electrons close to the band edges.

Equations (1)–(3) 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 can be obtained with a perturbation treatment (Supplementary Note 5), giving

$$ZPR_{c}^{Fr} = \int dq (m_{n}^{\ast}(q))^{1/2} \frac{1}{2m_{n}^{\ast}(q)V_{BZK}} \frac{(\omega_{p}(q))^{2}}{e^{\omega_{p}(q)} - 1}.$$  

A similar expression exists for the valence band ZPR. 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 q = Γ (e.g., no phonon band structure.

For detailed explanations, see Supplementary Note 5. Such model covers all situations and still uses as input only long-wavelength parameters, that can be determined
calculation). Eq. (6) can be evaluated for all band extrema in our set of materials, irrespective of whether the extrema are located at $\Gamma$ 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-thirds 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\textsuperscript{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.

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\textsuperscript{24} is avoided, either using the ASC methodology or using the non-adiabatic AHC methodology. As Fröhlich and Feynman have cautioned us\textsuperscript{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 $q$ is the small wavevector of the fluctuation, see Fig. 5a “LO phonon potential” part. Without such long-wavelength potential, the
electron at the minimum of the conduction band has a Bloch type wavefunction, with an envelope phase factor characterized by the wavevector $k$, 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”\(^\text{119}\). A wavepacket is formed, by combining Bloch wavefunctions with similar lattice periodic functions but slightly different wavevectors ($k_q, k_q - q, k_q - q$, etc), coming from a small interval of energy $\Delta \epsilon \propto q^2 \alpha^2 / (\Delta L)_S^2$, see Fig. 5a “Wavefunction” part. This nonquadratic effect is actually illustrated in Fig. 4 of ref. \textsuperscript{23} (see the frozen-phonon eigenvalues), as well as in Fig. 2 of ref. \textsuperscript{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, $\Delta \epsilon \Delta \tau \geq h$, hence $\Delta \tau \geq h / \Delta \epsilon \Delta L^2$. For long wavelengths, the characteristic time diverges. As soon as $\Delta \tau$ is larger than the phonon characteristic time $t_{\text{ph}} \sim 1 / \omega_{\text{LO}}$, 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\(^\text{26}\). 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\(^\text{25}\). 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\(^\text{50}\) 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\(^\text{51}\) has been used and the pseudopotentials have been taken from the PseudoDojo project\(^\text{24}\). For diamond, BN-zb, and AlN-wz, results reported here come from a previously published work\(^\text{19}\), 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 closely closes at such parameter, for GaP, as at such parameter the conduction band presents unphysical quasi-degenerate valleys, and for TiO\textsubscript{2}, as the GGA-PBE predicted structure is unstable\(^\text{23}\). For these, we have used the experimental lattice parameter. The case of SrTiO\textsubscript{3} is specific and will be explained later. Density-functional perturbation theory\(^\text{54,55}\) 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 $\Sigma$ appears due to the electron–phonon interaction, with Fan and Debye–Waller contributions at the lowest order of perturbation\(^\text{55}\):

$$
\Sigma_0(\omega) = \Sigma_{\text{fan}}(\omega) + \Sigma_{\text{DW}}.
$$

(7)

The Hartree atomic unit system is used throughout ($\hbar=m_e=e=1$). An electronic state is characterized by $k$, its wavevector, and $n$, its band index, $\omega$ 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\(^\text{55}\):

$$
\Sigma_{\text{fan}}(\omega) = \frac{1}{N_0} \sum_{\mathbf{k}\mathbf{q}} \sum_{n} \left| \langle \mathbf{k} + \mathbf{q} | | \mathbf{n} \rangle \right|^2 
\times \left[ \frac{1 - f_{\mathbf{k}+\mathbf{q},\mathbf{n}}}{\omega - \epsilon_{\mathbf{k}+\mathbf{q},\mathbf{n}} - \omega_{\mathbf{q}} - i\gamma} + \frac{f_{\mathbf{k}+\mathbf{q},\mathbf{n}}}{\omega - \epsilon_{\mathbf{k}+\mathbf{q},\mathbf{n}} + \omega_{\mathbf{q}} + i\gamma} \right].
$$

(8)

In this expression, contributions from phonon modes with harmonic phonon energy $\omega_{\mathbf{q}}$ are summed for all branches $j$, and wavevectors $\mathbf{q}$, in the entire Brillouin zone (BZ). The limit for infinite number $N_0$ of wavevectors (homogeneous sampling) is implied. Contributions from transitions to electronic states $| \mathbf{k} + \mathbf{q} \rangle$ with KS-DFT electron energy $\epsilon_{\mathbf{k}+\mathbf{q},\mathbf{n}}$ and occupation number $f_{\mathbf{k}+\mathbf{q},\mathbf{n}}$ (1 for valence, 0 for conduction, at $T=0 K$) are summed for all bands $\mathbf{n}$ (valence and conduction). $\omega_{\mathbf{q}} = \hbar^2 q^2 / 2m $ is the
self-consistent change of potential due to the q-phonon\(^{26}\). Limit of this expression for vanishing positive \(\eta\) is implied for the Debayle-Waller self-energy, \(\Sigma_{\omega}^{\text{DW}}\), we refer to the literature\(^{21,25}\).

In the non-adiabatic AHC approach, the ZPR is obtained directly from the real part of the self-energy, Eq. (7), evaluated at \(\omega = \epsilon_{\omega}^{\text{ZPR}}\):

\[
\text{ZPR}^{\text{ZPR}}_{\omega} = \frac{\pi \epsilon_{\omega}}{\omega (\omega + \epsilon_{\omega}^{\text{ZPR}})}.
\]

(9)

If the adiabatic approximation is made, the phonon frequencies \(\epsilon_{\omega}\) are considered small with respect to eigenenergy differences in the denominator of Eq. (8) and are simply dropped, while a finite \(\eta\), usually 0.1 eV, is kept. With a vanishing \(\eta\), 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 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 ZPR\(_{\text{expt}}\) and our non-adiabatic AHC ZPR\(_{\text{th}}\) values. The following phenomena may also play a role: (1) the rigid-ion approximation\(^{21,25}\), (2) the nonquadratic behavior of the eigenergies with collective displacements of the nuclei, in reference to the ASC\(^{32}\); (3) the reliance on GGA-PBE eigenergies 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 ZPR\(_{\text{expt}}\) beyond the AHC picture.

As an example, in KEK\(^{54}\), 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 KEK\(^{54}\). 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 ZPR\(_{\text{th}}\) 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 (Cambridge 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 (Cambridge 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., Kottani, 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).
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 SrTiO₃ 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