Strong renormalization of the electronic band gap due to lattice polarization in the 
\textit{GW} formalism

Silvana Botti\textsuperscript{1,2} and Miguel A. L. Marques\textsuperscript{1,2}

\textsuperscript{1} Institut Lumière Matière, UMR5306 Université Lyon 1–CNRS, Université de Lyon 69622 Villeurbanne cedex, France
\textsuperscript{2} European Theoretical Spectroscopy Facility

(Dated: May 11, 2014)

The self-consistent \textit{GW} band gaps are known to be significantly overestimated. We show that this overestimation is, to a large extent, due to the neglect of the contribution of the lattice polarization to the screening of the electron-electron interaction. To solve this problem, we derive within the \textit{GW} formalism a generalized plasmon-pole model that accounts for lattice polarization. The resulting \textit{GW} self-energy is used to calculate the band structures of a set of binary semiconductors and insulators. The lattice contribution always decreases the band gap. The shrinkage increases with the size of the longitudinal-transverse optical splitting and it can represent more than 15\% of the band gap in highly polar compounds, reducing the band-gap percentage error by a factor of three.

The causes for these errors were attributed to the lack of vertex corrections\textsuperscript{6–7} or to the underscreening by the random-phase approximation\textsuperscript{11} (RPA), and in both cases a simple empirical correction was adopted to improve the theoretical calculations (such as scaling the final self-energy $\Sigma_\text{xc}$ by about 80\%, obtaining the so-called 0.8$\Sigma$ approximation\textsuperscript{12}).

The study of electronic correlation and its effects on the band-structure of solids has a very long tradition. A much less studied subject is, instead, the influence of phonons on the quasi-particle spectrum. Only now, in fact, the high level of accuracy achieved by electronic structure calculations makes the phonon contributions larger than the theoretical precision.

There exist two contributions to the band gap coming from phonons: i) The electron-phonon coupling, that can be handled according to the Allen-Heine-Cardona theory\textsuperscript{15} using second-order perturbation theory within the harmonic and adiabatic approximations. One obtains two terms of the same order to be included in the \textit{GW} self-energy, known as the Fan and Debye-Waller terms. ii) The phonon contribution to the frequency-dependent dielectric function, that can be a sizable component of the total electron screening in polar compounds.

The electron-phonon coupling terms are responsible for the strong reduction of the band gap of diamond. In fact, in that case the zero-point contribution to the band gap was proved by Giustino \textit{et al}\textsuperscript{14} to be 0.6 eV. They observed that the example of diamond is an extreme case, while for most semiconductors and insulators the zero-point renormalization is as small as 10–50 meV\textsuperscript{14} and can be safely neglected. The Fan and Debye-Waller terms have also been used to determine the temperature dependence of the band gap of semiconductors\textsuperscript{14}.

In this work we want to focus on the second, usually overlooked, contribution of phononic nature: the screening due to the coupling of the electro-magnetic field and the oscillating dipoles\textsuperscript{17,18}. Of course, this contribution exists in polar materials only (and thus it is nonexistent in diamond). We should emphasize that by “polar” we mean any material that has non-negligible Born effective charges, leading to a measurable longitudinal-optical (LO) and transverse-optical (TO) phonon splitting. Most semiconducting and insulating materials known actually fit into this category. Indeed, polar materials are char-
acterized by LO and TO infrared-active phonons, whose excitation induces macroscopic electric fields in the crystal. These fields contribute, together with the electronic screening, to the total screening of the Coulomb potential, and their strength is proportional to the bond ionicity.

Already in 2005 Bechstedt et al. observed that the lattice polarization affects significantly the dressing of the quasi-particles in ionic materials, inducing large corrections to the band gap, that they calculated using a model electronic dielectric function and a static approximation for the lattice contribution to the screening. Unfortunately, it was already observed in Ref. [17] that that model always overestimates the band-gap correction. In this Letter, we take a step further and develop a fully ab initio framework to include the lattice contributions to the screening within the GW scheme. The basic ingredients are, besides the standard RPA screening of GW, the LO and TO frequencies of the infrared active phonon modes. We then calculate the electronic band structure of a set of polar materials applying this new scheme, and compare the results with experimental data and with many-body results including only electronic correlations.

The lattice polarization plays an important role in the determination of the total screening of polar compounds, as it is evident from the difference between \( \varepsilon_\infty \), the electronic dielectric function at zero frequency, and \( \varepsilon_s \), the total static dielectric function including contributions from both electrons and infrared active phonons. The ratio between these quantities is related to the phonon frequencies at the center of the Brillouin zone, \( \omega_{\text{LO}} \) and \( \omega_{\text{TO}} \), through the Lyddane-Sachs-Teller relation:

\[
\frac{\omega_{\text{LO}}^2}{\omega_{\text{TO}}^2} = \frac{\varepsilon_s}{\varepsilon_\infty}. \quad (1)
\]

The standard treatment of optical phonons in crystals using the dynamical matrix implicitly assumes that the interatomic interactions are instantaneous. In the case of polar crystals, the long-range nature of the Coulomb interaction requires a proper account of retardation effects. It is easy to understand that the coupling of phonon waves and electromagnetic waves is effective only for \( q \to \mathbf{0} \), since the speed of sound is negligible if compared with the speed of light. By combining a continuous approximation for the description of the mechanical waves of the optical modes and Maxwell’s equations for electromagnetic waves one obtains the polariton dispersion curves. For a polar crystal the total dielectric function, which includes electronic and lattice polarization terms, then reads (here for simplicity in the scalar form):

\[
\varepsilon(q \to \mathbf{0}, \omega) = \varepsilon_e(q \to \mathbf{0}, \omega) + \varepsilon_{\text{lat}}(q \to \mathbf{0}, \omega), \quad (2)
\]

where \( \varepsilon_e(q \to \mathbf{0}, \omega) \) is the long wavelength limit of the electronic contribution to the dielectric function (calculated, e.g., in the RPA approximation), and \( \varepsilon_{\text{lat}} \) is defined as the contribution to the dielectric function due to lattice polarization. This latter quantity can be related to the electronic screening by a frequency-dependent generalization of the Lyddane-Sachs-Teller relation, that reads

\[
\varepsilon_{\text{lat}}(q \to \mathbf{0}, \omega) = \varepsilon_e(q \to \mathbf{0}, \omega) \frac{\omega_{\text{LO}}^2 - \omega_{\text{TO}}^2}{\omega_{\text{TO}}^2 - (\omega + \eta)^2}. \quad (3)
\]

In this last expression, the infinitesimal \( \eta \to 0^+ \).

In the following discussion, and to keep the notation of the equations at a reasonably simple level, we will consider only cubic crystals with two atoms in the unit cell. However, all the equations can be easily generalized to arbitrary unit cells with more than two atoms, i.e. more than one infrared active mode.

Equations (2) and (3) can be used directly within the GW framework without any further approximation. However, the very large difference between typical phonon (of the order of the meV) and electronic (of the order of the eV) frequencies makes the use of methods like the contour deformation unpractical. We will therefore resort to the popular plasmon-pole model of the GW theory in order to derive a practical framework to include the effects of lattice polarization.

For convenience, we now move to the reciprocal lattice \( G \) space and write \( \varepsilon \) as a \( G, G' \) matrix. The electronic plasmon-pole model for \( \varepsilon^{-1}_e(GG', \omega, q) \) reads:

\[
\varepsilon^{-1}_e(GG', \omega, q) = \delta_{GG'} + \frac{\Omega^2_{GG'}(q)}{\omega^2 - \omega_{GG'}(q)^2 - i\eta}. \quad (4)
\]

The two parameters \( \Omega^2_{GG'}(q) \) and \( \omega_{GG'}(q) \) can be fixed for example by calculating \( \varepsilon^{-1}_e(GG', \omega, q) \) at \( \omega = 0 \) and \( \omega = \omega_p \), where \( \omega_p \) is the plasmon frequency. In order to determine the total screening, in the case of \( q \to \mathbf{0} \), one has to evaluate \( \varepsilon^{-1}_e(GG', \omega) \) by inverting the sum of the two matrices \( \varepsilon_e(q \to \mathbf{0}, \omega) + \varepsilon_{\text{lat}}(q \to \mathbf{0}, \omega) \). The way to proceed is to apply the matrix equality:

\[
(G + H)^{-1} = G^{-1} - \frac{1}{1 + g} G^{-1} H G^{-1}, \quad (5)
\]

where

\[
g = \text{Tr} \{ H G^{-1} \}, \quad (6)
\]

and \( G \) is an invertible matrix. In our case the matrix \( G \) is \( \varepsilon_e(GG', \omega, q) \) and \( H \) is \( \varepsilon_{\text{lat}}(G \to 0, \omega) \). A useful simplification comes from the fact that \( H \) has only one non-zero element, the head of the matrix \( \varepsilon_{\text{lat}}(00, q \to \mathbf{0}, \omega) \). In the following we will remove the explicit dependence on \( q \) of the dielectric matrix elements as it is clear that the lattice polarization contribution exists only for \( q \to \mathbf{0} \). Using (5) in (4) it is easy to show that

\[
\varepsilon^{-1}_{GG'}(\omega) = \varepsilon^{-1}_e(GG', \omega) - \frac{\varepsilon^{-1}_e(00, \omega) + \varepsilon^{-1}_{GG'}(\omega)}{\varepsilon^{-1}_e(00, \omega)} \frac{\omega_{\text{LO}}^2 - \omega_{\text{TO}}^2}{\omega_{\text{TO}}^2 - (\omega + i\eta)^2}. \quad (7)
\]

If we now replace \( \varepsilon^{-1}_{GG'}(\omega) \) by the standard plasmon-pole model we obtain our final expression for
The matrix elements of the wings have very similar expressions, thanks to the simplification of $\epsilon^{-1}_{e00}(\omega)$ appearing at the denominator in (7):

$$
\epsilon^{-1}_{G0} = \frac{\Omega_G^2}{\omega^2 - (\omega_{G0} - i\eta)^2} - \epsilon^{-1}_{e00}(\omega_{LO} - \omega_{TO}^2) - \omega_{LO} - (\omega + i\eta)^2 \right),
$$  \hspace{1cm} (10)

with an analogous expression for the $0G$ terms.

The other matrix elements of $\epsilon^{-1}_{G0}(\omega)$ are more complicated: beside the extra poles at $\omega = \pm \omega_{G0}$, there appear terms involving poles at $\omega = \pm \omega_{G0}$, $\omega = \pm \omega_{G0}$, and $\omega = \pm \omega_{G0}/\epsilon_{\infty}$.

We can then insert (11) and (7) in the expression for the screened Coulomb interaction $W = \epsilon^{-1}_{e1} v$, where $v$ is the bare Coulomb interaction. This allows to evaluate analytically the convolution integral of the self-energy

$$
\Sigma(q, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dw' e^{i\eta w'} G(q, \omega + \omega') W(q, \omega'),
$$  \hspace{1cm} (11)

remembering that

$$
G(r, r', \omega) = \sum_{k,j} \frac{\phi_{k,j}(r) \phi_{k,j}^*(r')}{\omega - \varepsilon_{k,j} + i\eta \text{sgn}(\varepsilon_{k,j} - \mu)},
$$  \hspace{1cm} (12)

where $\phi_{k,j}$ are Kohn-Sham or quasi-particle orbitals, $\varepsilon_{k,j}$ are the corresponding energies, and $\mu$ is the Fermi level.

The frequency integration (11) is performed using the residue theorem, in strict analogy with the procedure followed for the electronic screening alone, with the only difference that more than one pole is present in the generalized plasmon-pole model that we derived. Once again, the full result is rather cumbersome, but a relatively simple expression is obtained for the lattice contribution to the matrix elements of $\Sigma$ when only the head of the matrix $\epsilon^{-1}_{G0}(q \to 0, \omega)$ is retained:

$$
\langle \phi_{k',ij} | \Sigma_{lat} | \phi_{k,j} \rangle = \frac{-4\pi}{V} \frac{\omega_{LO}^2 - \omega_{TO}^2}{\epsilon_{\infty}} \sum_j \theta(\mu - \varepsilon_{k',ij}) (\omega - \varepsilon_{k',ij})^2 - (\omega_{LO} - \omega_{ TO}^2) - \omega_{LO} - (\omega + i\eta)^2 \right),
$$  \hspace{1cm} (13)

with $V$ the volume of the unit-cell and $\tilde{p}_{k',ij}(q \to 0) = \lim_{q \to 0} \int_{-\infty}^{\infty} \rho_{k',ij}(r) \exp(-iqr) \phi_{k'q,j}(r)$. Note that this term only contributes to the occupied states due to the presence of the Heaviside function $\theta(\mu - \varepsilon_{k',ij})$.

We implemented in ABINIT\textsuperscript{23} the complete contribution of the lattice polarization to $\Sigma$, i.e., the contribution of the head, wings and body of the dielectric matrix. However, only the poles at $\omega = \pm \omega_{LO}$ can contribute significantly to the integral, as the coefficient in front of the other poles is always extremely small. For all systems that we studied, the contribution of the head of the matrix alone is enough to describe the correction to the band structures within less than 1 meV, but this may not be the case for non-cubic systems.

Our calculations start by the determination of the standard Kohn-Sham ground-state, using the LDA and norm-conserving pseudopotentials. The LDA Kohn-Sham states are then used to calculate $GW@LDA$ band structures and as a starting point for QPscGW runs. There are many different versions of $GW$, and consequently there are many different ways of using $\Sigma_{lat}$ in the context of $GW$. We chose as starting point a converged QPscGW calculation, as this is in our opinion the most accurate level of theory available in ABINIT. However, one can not use $\Sigma_{lat}$ in a self-consistent scheme, as this would induce contributions of the lattice polarization in

| Material | $\omega_{LO}$ | $\omega_{TO}$ | $GW@LDA$ | $GW^{lat}$ | QPscGW | $scGW^{lat}$ | $\exp$ |
|----------|---------------|---------------|-----------|------------|---------|--------------|---|
| LiF**24** | 2.99 | 1.39 | 13.24 | 12.05 | 15.81 | 13.69 | 14.20 |
| LiCl**24** | 1.74 | 0.87 | 8.60 | 7.97 | 10.28 | 9.05 | 9.4 |
| NaCl**24** | 1.21 | 0.75 | 7.73 | 7.14 | 9.52 | 8.37 | 8.5 |
| MgO**25** | 3.29 | 1.82 | 6.97 | 6.38 | 8.94 | 7.71 | 7.7 |
| AlP**25** | 2.28 | 2.01 | 2.32 | 2.26 | 2.79 | 2.70 | 2.49 |
| AlAs**30** | 1.84 | 1.65 | 1.83 | 2.34 | 2.26 | 2.23 |
| GaAs**30** | 1.30 | 1.22 | 1.16 | 1.15 | 1.52 | 1.46 | 1.52 |
| MAE | 0.60 | 1.04 | 0.74 | 0.18 |
| MAPE | 11.5% | 16.3% | 9.4% | 3.0% |
MgO, AlP, AlAs, GaAs. These are all polar structures of a set of binary compounds: LiF, LiCl, NaCl, and MgO, and to 30 Ha for the remaining compounds, and the \textit{k}-point grids were a $6 \times 6 \times 6$ Monkhorst-Pack. This allowed for a convergence in the total energy to better than 0.1 eV. Finally, and in order to limit the effect of other possible sources of imprecision, we decided to use for these calculations experimental lattice constants and experimental LO and TO phonon frequencies at $\Gamma$ (see Table I). Of course the method can be easily made fully \textit{ab initio} by calculating such quantities within DFT.

In Table I we report the calculated band gaps for the selected binary compounds using the different \textit{GW} schemes. As it is well known, \textit{GW}@LDA underestimates the band gaps of \textit{sp} compounds by slightly more than 10\% while \textit{QPscGW} overestimates them by about 10\%. Turning on the correction due to lattice polarization cuts down the error of these techniques by a factor of three, bringing the mean absolute percentage error to 3\%. These results clearly prove that the overestimation of the band gap by \textit{QPscGW} is not mainly due to a deficient treatment of electronic correlation, but to a deficient description of the screening of the medium. We remark that the inclusion of the lattice contribution to the screening in a \textit{GW}@LDA calculation deteriorates the agreement with experiments, leading to a mean absolute percentage error of 16\%. This shows that the good performance of perturbative \textit{GW} calculations for \textit{sp} polar materials is due to a partial cancellation of errors: the underestimation of the band gap opening is compensated by the neglect of the band gap shrinkage due to the lattice polarization.

In Fig. 1 we display the band structures of LiF obtained with the different approximations under study. We can observe that the inclusion of the lattice polarization in the screening produces a rigid shift downwards of the conduction bands, with negligible effects on the band widths and band dispersions. While the same qualitative behavior is observed for all the materials under study here, we cannot exclude that other effects may be present in more complex materials.

Finally, we would like to make two remarks: (i) The coupling between electrons and phonons in our model is indirect and comes through Maxwell equations. It is therefore completely unrelated to electron-phonon coupling. (ii) To take into account all phonon contributions to the electronic band structure, both the lattice contribution to the screening and the \textit{Fan} and \textit{Debye-Waller} (and possibly higher order) terms of the Allen-Heine-Cardona theory\textsuperscript{15} should be included. It is true that for the materials studied here one does not expect a large contribution from the electron-phonon coupling, but it is not inconceivable to find a strong ionic compound with strong electron-phonon coupling. Note that both contributions will tend to decrease the purely electronic gap.

In conclusion, we developed a fully \textit{ab initio} \textit{GW} framework that includes the effects of the screening owing to the polarization of the lattice. Within this framework we show that the overestimation of the band gaps by restricted self-consistent \textit{GW} techniques is due to the neglect of this contribution, and we manage to bring the error in the calculated band gaps to a mere 3\%. The lattice contribution decreases the band gap by a factor that increases with the size of the longitudinal-transverse optical splitting and it can represent more than 15\% of the band gap in highly polar compounds. These results call for a reexamination of many theoretical calculations of the quasi-particle spectrum for polar materials, including for many oxides, nitrides, etc. that are important in several fields of technology.

We would like to thank Julien Vidal and Friedhelm Bechstedt for many fruitful discussions. We acknowledge financial support from the French ANR project ANR-12-BS04-0001-02. Computational resources were provided by GENCI (project x2011096017).
1. L. Hedin, Phys. Rev. 139, A796 (1965).
2. M. S. Hybertsen and S. G. Louie, Phys. Rev. B 34, 5390 (1986).
3. G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. 74, 601 (2002).
4. F. Bruneval, N. Vast, L. Reining, M. Izquierdo, F. Sirotti, and N. Barrett, Phys. Rev. Lett. 97, 267601 (2006).
5. M. Gatti, F. Bruneval, V. Olevano, and L. Reining, Phys. Rev. Lett. 99, 266402 (2007).
6. J. Vidal, S. Botti, F. Olsson, J-F. Guillemoles, and L. Reining, Phys. Rev. Lett. 104, 056401 (2010).
7. J. Vidal, F. Trani, F. Bruneval, M.A.L. Marques, and S. Botti Phys. Rev. Lett. 104, 136401 (2010).
8. S.V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. 93, 126406 (2004).
9. M. van Schilfgaarde, T. Kotani, and S.V. Faleev, Phys. Rev. Lett. 96, 226402 (2006).
10. F. Bruneval, N. Vast, and L. Reining, Phys. Rev. B 74, 045102 (2006).
11. A. Svane, N.E. Christensen, I. Gorczyca, M. van Schilfgaarde, A.N. Chantis, and T. Kotani, Phys. Rev. B 82, 115102 (2010).
12. A. Punya and W.R.L. Lambrecht, and M. van Schilfgaarde, Phys. Rev. B 84, 165204 (2011).
13. C.E. Patrick and F. Giustino, J. Phys. Cond. Mat. 24, 202201 (2012); W. Kang and M.S. Hybertsen, Phys. Rev. B 82, 085203 (2010).
14. F. Giustino, S.G. Louie, and M.L. Cohen, Phys. Rev. Lett. 105, 265501 (2010).
15. P. B. Allen and V. Heine, J. Phys. C 9, 2305 (1976); P. Lautenschlager, P. B. Allen, and M. Cardona, Phys. Rev. B 31, 2163 (1985); Lautenschlager, P. B. Allen, and M. Cardona, Phys. Rev. B 33, 5501 (1986).
16. J. Liu and L. H. Liu, J. Appl. Phys. 111, 083508 (2012).
17. F. Bechstedt, K. Seino, P. H. Hahn, and W. G. Schmidt, Phys. Rev. B 72, 245114 (2005).
18. F. Trani, J. Vidal, S. Botti, and M.A.L. Marques, Phys. Rev. B 82, 085115 (2010).
19. G. Grosso and G. Pastorì Parravicini, “Solid State Physics” (Academic Press, London, 2000).
20. X. Gonze and C. Lee, Phys. Rev. B 55, 10355 (1997).
21. S. Lebêgue, B. Arnaud, M. Alouani, and P. E. Bloechl, Phys. Rev. B 67, 155208 (2003).
22. R.W. Godby and R.J. Needs, Phys. Rev. Lett. 62, 1169 (1989).
23. K.S. Miller, Mathematics Magazine 54, 67 (1981).
24. X. Gonze, G.-M. Rignanese, M. Verstraete, J.-M. Beneken, Y. Pouillon, R. Caracas, F. Jollet, M. Torrent, G. Zerah, M. Mikami, P. Ghosez, M. Veithen, J.-Y. Raty, V. Olevano, F. Bruneval, L. Reining, R. Godby, G. Onida, D.R. Hamann, and D.C. Allan, Z. Kristallogr. 220, 558 (2005).
25. G. Dollings, H. G. Smith, R.M. Nicklow, R. Vigayaraghavan, M. K. Wilkinson, Phys Rev. 168, 970 (1968); M. Placentini, G. Lynch, and C. Olson, Phys. Rev. B 13, 5530 (1976).
26. M. Hass, J. of Phys. and Chem. of Solids 24, 1159 (1963); C.M. Randall, R.M. Fuller, and D.J. Montgomery, Solid State Commun. 2, 273 (1964); J.T. Lewis, A. Lehoczky, and C.V. Briscoe, Phys. Rev. B 161, 877 (1967); G. Baldini and B. Bosacchi, Phys. Status Solidi 38, 325 (1970).
27. G. Raunio and S. Rolandson, Phys.Rev. B 2, 2098 (1970); R.T. Poole, J. Liesegang, R.C.G. Leckey, and J.G. Jenkins, Phys. Rev. B 11, 5194 (1975); R.A. Bartells and A. Smith, Phys. Rev. B 7, 3885 (1973).
28. R. K. Singh and K. S. Upadhyaya, Phys. Rev. B 6, 1589 (1972); A.R. Oganov, M.J. Gillan, and G.D. Price, J. Chem. Phys. 118, 10174 (2003); Ü. Özgür, Ya. I. Alivov, C. Liu, A. Teke, M. A. Reschchikov, S. Dogan, V. Avrutin, S.-J. Cho, and H. Morkoç, J. Appl. Phys. 98, 041301 (2005).
29. “Group IV Elements, IV-IV and III-V Compounds: - Lattice Properties”, Landolt-Börnstein - Group III Condensed Matter, ed. O. Madelung, U. Rssler, M. Schulz, Vol. 41Ala, p. 1 (2001); Y.F. Tsay, A.J. Corey, and S.S. Mitra, Phys. Rev. B 12, 1354 (1975).
30. D. W. Palmer, www.semiconductors.co.uk, 2006.02.