First-principles study of excitonic effects in Raman intensities

Yannick Gillet,* Matteo Giantomassi, and Xavier Gonze
Université Catholique de Louvain, Institute of Condensed Matter and Nanosciences, Nanoscopic Physics, Chemin des Étoiles 8, bte L7.03.01, 1348 Louvain-la-Neuve, Belgium

(Received 23 April 2013; revised manuscript received 11 July 2013; published 17 September 2013)

The *ab initio* prediction of Raman intensities for bulk solids usually relies on the hypothesis that the frequency of the incident laser light is much smaller than the band gap. However, when the photon frequency is a sizable fraction of the energy gap, or higher, resonance effects appear. In the case of silicon, when excitonic effects are neglected, the response of the solid to light increases by nearly three orders of magnitude in the range of frequencies between the static limit and the gap. When excitonic effects are taken into account, an additional tenfold increase in the intensity is observed. We include these effects using a finite-difference scheme applied on the dielectric function obtained by solving the Bethe-Salpeter equation. Our results for the Raman susceptibility of silicon show stronger agreement with experimental data compared with previous theoretical studies. For the sampling of the Brillouin zone, a double-grid technique is proposed, resulting in a significant reduction in computational effort.

DOI: 10.1103/PhysRevB.88.094305 PACS number(s): 78.30.Am, 71.15.Qe

I. INTRODUCTION

Raman spectroscopy is widely used to characterize materials by means of their vibrational fingerprint. The dependence of the Raman intensity on the frequency of the incident light is well known. It is, for example, used to amplify the Raman response, resulting in the appearance of a resonance phenomenon when the frequency of the exciting light is close to the electronic transitions.[1] Unlike for molecules[2,3] and for graphene,[4,5] the first-principles prediction of the frequency dependence of the Raman intensity of crystalline systems has received little attention.[1]

The Raman intensity is related to the derivative of the macroscopic dielectric function with respect to collective atomic displacements. Different first-principles formalisms have been proposed for the computation of such a dielectric function. These formalisms often trade computational speed for predictive power, or vice versa. In the present study, a method that provides an accurate description of the dielectric properties of material was chosen in order to establish the importance of different physical effects, and in particular, excitonic effects.

Within the static limit (vanishing light frequency), the dielectric response can be computed with density-functional theory (DFT)[6–8] followed by density-functional perturbation theory (DFPT).[9–11] Although DFT is plagued by the well-known band-gap problem,[8] its prediction of the static dielectric tensor is reasonably accurate (to within 5–10%) except when the gap is very small.[12] Subsequent computation of the derivative of the dielectric tensor with respect to an atomic displacement can be performed by using either finite differences[13] or the $2n + 1$ theorem of perturbation theory.[14,15] Such a methodology has been applied in numerous studies.[16,17] As an example, more than 200 Raman spectra are provided in the WURM database.[18,19]

When the excitation frequency is comparable to the gap, DFT becomes unreliable for the prediction of the dielectric response. Not only does the proximity of the resonance increase the need to rely on an accurate band gap, but excitonic effects also drastically modify the optical properties of most semiconductors.[20] The band-gap correction is usually treated within the GW approximation of many-body perturbation theory (MBPT), while the Bethe-Salpeter equation (BSE) is the method of choice to introduce excitonic effects.[21,22] To the best of our knowledge, the BSE has not yet been used to compute Raman intensities of solids. The purpose of the present work is to compute the Raman intensities using a finite-difference approach that combines multiple BSE results performed for different atomic displacements.

Excitonic effects can also be addressed within the framework of time-dependent density functional theory.[22–25] This approach, which is computationally cheaper, also allows one to include excitonic effects, with an accuracy that depends on the choice of the exchange-correlation kernel. Recent studies shows interesting agreement with experiment for the macroscopic dielectric function (see, e.g., Ref. 26). This route is not pursued in the present study. Instead we rely on the best theoretical approach available today to compute the frequency-dependent Raman intensities, and we examine its predictive power in comparison with experimental data.

We chose to study silicon, for which the experimental frequency-dependent enhancement factor is particularly strong. The available data cover the frequency range between 1.8 and 3.8 eV,[27] and the experimental value of the direct gap is at 3.4 eV. Due to the high symmetry of silicon, there is only one Raman-active phonon mode, whose eigenvector is determined by symmetry.

In Sec. II of this article, the theoretical basis needed for the computation of the resonant Raman intensities is described, taking into consideration the main equations of the MBPT in the GW and BSE frameworks. Section III describes the numerical procedure. In Sec. IV, the problem associated with the slow convergence of results with respect to the sampling of the Brillouin zone is analyzed. Section V presents the theoretical results, including excitonic effects. Finally, in Sec. VI, theoretical and experimental results for the silicon Raman intensity are compared.

 ©2013 American Physical Society

II. THE COMPUTATION OF RESONANT RAMAN INTENSITIES FOR SOLIDS

The scattering efficiency (time average of the power radiated into a unit solid angle) of the phonon of frequency ω_m for a photon of frequency ω_i is defined as[15]

$$I = (\omega_i - \omega_m)^4 |\mathbf{e}_o \cdot \boldsymbol{\alpha}^m \cdot \mathbf{e}_i|^2 \frac{n_m + 1}{2\omega_m}, \quad (1)$$

with $\mathbf{e}_o$ and $\mathbf{e}_i$ the outgoing and ingoing polarization of the light, and n_m the phonon occupation factor:

$$n_m = \frac{1}{e^{\omega_m/kT} - 1}. \quad (2)$$

The complete field-theoretic expression for the Raman susceptibility $\alpha^m(\omega)$ is presented in Ref. 20. It includes six terms, in which the frequencies ω_m and ω_i are combined in different denominators, giving resonant as well as antiresonant contributions. In the following calculations, we will use the quasistatic approximation, which neglects the dynamical effects due to the phonons. Mathematically, this approximation is well-justified[28] when

$$\omega_m \ll |\omega_i - \omega_{\text{gap}} + i\eta|, \quad (3)$$

with ω_{gap} the frequency corresponding to the direct band gap and η the lifetime broadening of the gap.

In this framework, the Raman susceptibility $\boldsymbol{\alpha}^m(\omega)$ for the phonon m is defined as

$$\alpha^m_{ij}(\omega) = \sqrt{\Omega_0} \sum_{\tau\beta} \frac{\partial \chi_{ij}(\omega)}{\partial R_{\tau\beta}} u^m_{\tau\beta} \quad (4)$$

with Ω_0 the unit-cell volume, χ_{ij} the macroscopic dielectric susceptibility, and $u^m_{\tau\beta}$ the eigendisplacement of phonon mode m of atom τ in direction β. In the present work, we neglect higher-order derivatives with respect to atomic displacements. The eigendisplacements are normalized as

$$\sum_{\tau,\beta} M_\tau u^m_{\tau\beta} u^n_{\tau\beta} = \delta_{mn}, \quad (5)$$

with M_τ the mass of atom τ.[29] For more details about the derivation of Eq. (1), we refer the reader to Refs. 15 and 28.

We define the Raman polarizability a as

$$a = \sqrt{\mu\Omega_0}\alpha \quad (6)$$

with μ the reduced mass (in the case of silicon, $\mu = M_{\text{Si}}/2$).

When the incoming frequency ω_i is close to the energy of the direct gap, there is a resonant process and the amplitude of $\boldsymbol{\chi}(\omega_i)$ and $\boldsymbol{\alpha}^m(\omega_i)$ can change by several orders of magnitude. The computation of the macroscopic dielectric susceptibility $\boldsymbol{\chi}(\omega_i)$ follows the standard procedure used in *ab initio* MBPT. Two steps are needed: the computation of the quasiparticle energies, followed by the computation of the dielectric response of the material. The quasiparticle amplitudes ψ_i^{QP} and the quasiparticle energies ϵ_i^{QP} are computed by solving the following equation:

$$\left(-\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r})\right)\psi_i^{\text{QP}}(\mathbf{r}) + \int d\mathbf{r}' \Sigma\left(\mathbf{r},\mathbf{r}';\omega = \epsilon_i^{\text{QP}}\right)\psi_i^{\text{QP}}(\mathbf{r}') = \epsilon_i^{\text{QP}}\psi_i^{\text{QP}}(\mathbf{r}), \quad (7)$$

with $V_{\text{ext}}(\mathbf{r})$ the electrostatic potential of the ions and $V_{\text{H}}(\mathbf{r})$ the Hartree potential originating from the electronic density $n(\mathbf{r})$. In Eq. (7), $\Sigma(\mathbf{r},\mathbf{r}';\omega)$ is the self-energy, which, in the so-called GW approximation, is given by

$$\Sigma(\mathbf{r},\mathbf{r}';\omega) = \frac{i}{2\pi}\int d\omega' G(\mathbf{r},\mathbf{r}',\omega + \omega')W(\mathbf{r},\mathbf{r}',\omega'), \quad (8)$$

where G is the Green's function and W is the screened Coulomb interaction.[22]

In the second part of the calculation, we include excitonic effects by working within the BSE framework.[22] In this framework, we introduce H, a two-particle Hamiltonian that describes the interaction between electrons and holes. In the transition space formed by products of two Kohn-Sham orbitals, the BSE Hamiltonian has the following block structure:

$$H = \begin{pmatrix} & |v'c'\mathbf{k}'\rangle & |c'v'\mathbf{k}'\rangle \\ |vc\mathbf{k}\rangle & R & C \\ |cv\mathbf{k}\rangle & -C^* & -R^* \end{pmatrix}, \quad (9)$$

where v, c, and $\mathbf{k}$ denotes the valence-band index, the conduction-band index, and the wave vector.

The resonant subblock R is Hermitian, and the coupling term C is symmetric. Due to the coupling subblocks that connect resonant and antiresonant transitions, the Bethe-Salpeter Hamiltonian is not Hermitian. This complicates the solution of the problem. In crystalline systems, however, the matrix elements of C are usually much smaller than the matrix elements of R. For this reason, the matrix elements of C are usually neglected when solving the Bethe-Salpeter problem in extended systems—the so called Tamm-Dancoff approximation (TDA).[30] This approximation is used in the rest of this work.

The matrix elements of the resonant block are given by

$$R_{(vc\mathbf{k}),(v'c'\mathbf{k}')} = H^{\text{diag}}_{(vc\mathbf{k}),(v'c'\mathbf{k}')} + H^{\text{exch},R}_{(vc\mathbf{k}),(v'c'\mathbf{k}')} + H^{\text{Coul},R}_{(vc\mathbf{k}),(v'c'\mathbf{k}')}, \quad (10)$$

where

$$H^{\text{diag}}_{(vc\mathbf{k}),(v'c'\mathbf{k}')} = (\epsilon_{c\mathbf{k}} - \epsilon_{v\mathbf{k}})\delta_{vv'}\delta_{cc'}\delta_{\mathbf{k}\mathbf{k}'}, \quad (11)$$

$$\begin{aligned} H^{\text{exch},R}_{(vc\mathbf{k}),(v'c'\mathbf{k}')} &= 2\langle vc\mathbf{k}|\bar{v}|v'c'\mathbf{k}\rangle \\ &= 2\iint \psi_{v\mathbf{k}}(\mathbf{r})\psi^*_{c\mathbf{k}}(\mathbf{r})\bar{v}(\mathbf{r}-\mathbf{r}')\psi^*_{v'\mathbf{k}'}(\mathbf{r}')\psi_{c'\mathbf{k}'}(\mathbf{r}')d\mathbf{r}'d\mathbf{r}, \end{aligned} \quad (12)$$

$$\begin{aligned} H^{\text{Coul},R}_{(vc\mathbf{k}),(v'c'\mathbf{k}')} &= -\langle vc\mathbf{k}|W|v'c'\mathbf{k}'\rangle \\ &= -\iint \psi_{v\mathbf{k}}(\mathbf{r})\psi^*_{v'\mathbf{k}'}(\mathbf{r})W(\mathbf{r},\mathbf{r}')\psi^*_{c\mathbf{k}}(\mathbf{r}')\psi_{c'\mathbf{k}'}(\mathbf{r}')d\mathbf{r}'d\mathbf{r}, \end{aligned} \quad (13)$$

with $\bar{v}$ the modified Coulomb potential, whose Fourier transform does not contain the $\mathbf{q} = \mathbf{0}$ component:

$$\bar{v}(\mathbf{q}) = \begin{cases} v(\mathbf{q}) & \text{if} \quad \mathbf{q} \neq \mathbf{0}, \\ 0 & \text{if} \quad \mathbf{q} = \mathbf{0}, \end{cases} \quad (14)$$

$v(\mathbf{r})$ is the standard Coulomb potential:

$$v(\mathbf{r}) = \frac{1}{|\mathbf{r}|}, \tag{15}$$

$W(\mathbf{r},\mathbf{r}')$ is the screened Coulomb potential:

$$W(\mathbf{r},\mathbf{r}') = \int d\mathbf{r}''\epsilon^{-1}(\mathbf{r},\mathbf{r}'')v(\mathbf{r}''-\mathbf{r}'), \tag{16}$$

and $\epsilon^{-1}(\mathbf{r},\mathbf{r}')$ is the inverse dielectric function. For the derivation of Eqs. (11)–(16), we refer the reader to Ref. 22.

The dielectric susceptibility $\chi(\omega)$ and macroscopic dielectric function $\varepsilon(\omega)$ are then obtained from

$$\varepsilon(\omega) = 1 + 4\pi\chi(\omega) \tag{17}$$

$$= 1 - \lim_{\mathbf{q}\to 0} v(\mathbf{q})\langle P(\mathbf{q})|[(\omega + i\eta) - H]^{-1}F|P(\mathbf{q})\rangle, \tag{18}$$

where η is a broadening factor, F is taking into account the occupation numbers,

$$F = \begin{pmatrix} & |v'c'\mathbf{k}'\rangle & |c'v'\mathbf{k}'\rangle \\ |vc\mathbf{k}\rangle & 1 & 0 \\ |cv\mathbf{k}\rangle & 0 & -1 \end{pmatrix}, \tag{19}$$

and

$$P(\mathbf{q})_{(n_1n_2)} = \langle n_2|e^{i\mathbf{q}\cdot\mathbf{r}}|n_1\rangle \tag{20}$$

are the so-called oscillator matrix elements, where n_1 and n_2 are a shorthand notation for $vc\mathbf{k}$.

The random-phase approximation (RPA) is a simplification of the BSE approach, in which the exchange[31] and Coulomb terms, Eqs. (12) and (13), are neglected. In the RPA, the BSE Hamiltonian H is diagonal, and the spectrum is obtained directly as a simple sum over transitions between valence and conduction bands, weighted by the proper oscillator matrix elements. In such an independent-particle approximation, no excitonic effect is present. The importance of the excitonic effect on the optical spectrum is well known, with prominent peaks being created below the band gaps in most wide-gap insulators or semiconductors, and redistribution of the spectral weight.

III. NUMERICAL PROCEDURE

Calculations are performed using ABINIT.[32,33] The pseudopotential used to simulate the silicon atom is of the Troullier-Martins type used in the Teter parametrization.

The density functional theory–local density approximation (DFT-LDA) calculations are performed with a four-times-shifted $4 \times 4 \times 4$ Monkhorst-Pack grid to sample the Brillouin zone (BZ),[34] and a plane-wave basis set kinetic energy cutoff of 16 Ha. The theoretical lattice cell for silicon is 10.20 Bohr, which gives an error of 0.6% with respect to the experimental results (10.26 Bohr).[35] Using this theoretical lattice constant, the DFT-LDA indirect gap is 0.45 eV, while the direct gap is 2.52 eV.

Quasiparticle corrections are computed within the so-called *one-shot* GW or G_0W_0 approximation.[21] We use a cutoff energy of 8 Ha for the screening and 16 Ha for the self-energy matrix elements. An extrapolar energy of 3 Ha is used to reduce the number of bands needed to converge to 100 bands.[36] The computed GW corrections give a direct gap of 3.20 eV. These results are comparable to other GW results[37,38] and in good agreement with the experimental band gap of 3.4 eV.[39] During the computation of the BSE optical spectrum, the opening of the gap is simulated by a rigid scissor[40] with a value of 0.65 eV to reproduce the theoretical GW gap unless stated otherwise.

Convergence of the Bethe-Salpeter computations with respect to the BZ sampling is particularly difficult, and is discussed in the next section. The cutoff energies are 16 Ha for the wave functions and 3 Ha for the screening. The included bands range from the second to the ninth band. A broadening factor of 0.1 eV is used for the dielectric function.

The quasistatic approximation that is used extensively in this work is justified in the case of silicon since the lifetime broadening ($\approx$0.1 eV) is larger than the phonon frequency $\approx$0.065 eV.[28] In this quasistatic approximation, two-band as well as three-band contributions to the resonant Raman intensities are included, the latter coming from the matrix element changes due to changes in the wave functions produced by phonon-induced admixture of the two bands under consideration with a third band.[41]

To compute derivatives with the displacements, we add $h \times \sqrt{2}/2$ to the x position of the atom and $-h \times \sqrt{2}/2$ to the x position of the other atom for different values of h. The derivative is obtained by computing $\chi_{yz}(\omega)$ for $h = 0.01$ and 0 in the convergence studies and is obtained by computing $\chi_{yz}(\omega)$ for $h = 0.01$ and -0.01 for the final result.

We have analyzed the behavior of the G_0W_0 scissor shift, as a function of the frozen phonon amplitude. Because the Raman amplitude corresponds to a first-order derivative with respect to atomic displacement, see Eq. (4), we only have to consider the linear response. For nondegenerate eigenstates, due to the high symmetry of the crystals, such a derivative of the scissor shift with respect to atomic positions vanishes. For degenerate eigenenergies, linear variations of eigenenergies are present, but the mean variation vanishes over the set of degenerate states. Hence, the modification of the G_0W_0 corrections with respect to atomic displacement does not have any effect on the Raman intensity of silicon within the present formalism. Of course, this is a very specific situation. For the analysis of most other materials, the variation of the eigenenergies at linear order will have to be taken into account.

As implemented in ABINIT, the BSE gives the macroscopic dielectric function for a given q direction:

$$\varepsilon(\omega,\mathbf{q}) = \frac{\mathbf{q}^T\hat{\boldsymbol{\varepsilon}}\mathbf{q}}{\mathbf{q}^T\mathbf{q}}, \tag{21}$$

where $\hat{\boldsymbol{\varepsilon}}$ is the dielectric tensor.

Because of the symmetry of the tensor, the values of $\hat{\boldsymbol{\varepsilon}}$ for each pair of Cartesian coordinates may be obtained by computing the macroscopic dielectric function for six different directions q. The directions used in the calculations are (0,1,1), (1,0,1), (1,1,0), (1,0,0), (0,1,0), and (0,0,1).

The macroscopic dielectric function [Eq. (17)] is computed using the iterative Haydock technique.[42] The algorithm is terminated when a relative error of 1%, for the real and the imaginary part of ε, is achieved for each frequency in the frequency range under investigation.

IV. SAMPLING OF THE BRILLOUIN ZONE

As mentioned previously, achieving converged results with respect to the sampling of the Brillouin zone is a very difficult issue. In particular, we will show in this section that grids that are appropriate for obtaining a converged macroscopic dielectric function in the whole frequency range are not sufficiently dense for derivatives, such as the Raman intensities, at least in the resonance region.

To accelerate the convergence of BSE spectra, shifted homogeneous meshes are traditionally used. A symmetry-breaking shift allows one to sample more nonequivalent points and therefore leads to a more representative sampling of the band dispersion: with respect to nonshifted meshes, its presence lowers the computational load for an equivalent convergence criterion.

To ease the discussion, we introduce specific notations. The meshes are characterized by the number of divisions along each axis of the primitive cell in reciprocal space, namely n_1, n_2, and n_3. For a crystalline cubic structure, these three numbers are taken as equal ($n_1 = n_2 = n_3 = n_k$). The total number of points inside this mesh is therefore $N_k = n_1 n_2 n_3 = n_k^3$. All the points of the mesh are shifted by a certain vector characterized by three real numbers s_i between -0.5 and 0.5, $\mathbf{s} = (s_1,s_2,s_3)$. This shift is such that the point $(s_1/n_1,s_2/n_2,s_3/n_3)$ belongs to the shifted mesh. We use the notation $(n_1 \times n_2 \times n_3|\mathbf{s})$ to refer to such a mesh.

Figure 1 presents the macroscopic dielectric function (from the BSE) for different grids with increasing numbers of wave vectors, while Fig. 2 presents the corresponding Raman intensity, both with excitonic contributions (BSE) and without excitonic contributions (RPA). These grids, of size N_k, are shifted by the vector $\mathbf{s} = (0.11,0.21,0.31)$ in reciprocal space along a nonsymmetric direction.

As seen in Fig. 1, for the computation of the macroscopic dielectric function, the oscillations present in the $(10 \times 10 \times 10|\mathbf{s})$ case are progressively damped when the density of the mesh is increased and a $(16 \times 16 \times 16|\mathbf{s})$ grid gives converged results.

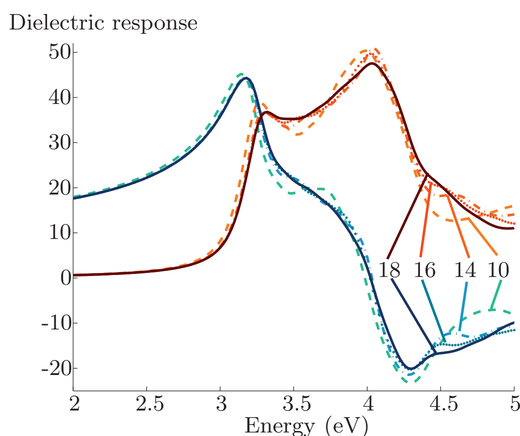

FIG. 1. (Color online) Convergence of $\varepsilon(\omega)$ (BSE) with respect to a traditional homogeneous sampling of the BZ. A shift $\mathbf{s}$ in a nonsymmetric direction is used (see text). The number indicated is n_k and the grid is therefore $(n_k \times n_k \times n_k|\mathbf{s})$. The imaginary part is given in blue while the real part is in orange-red. The full line corresponds to the finest grid that we have used, with $n_k = 18$. Oscillations appear for energies larger than 3.2 eV, but are damped with increasing n_k.

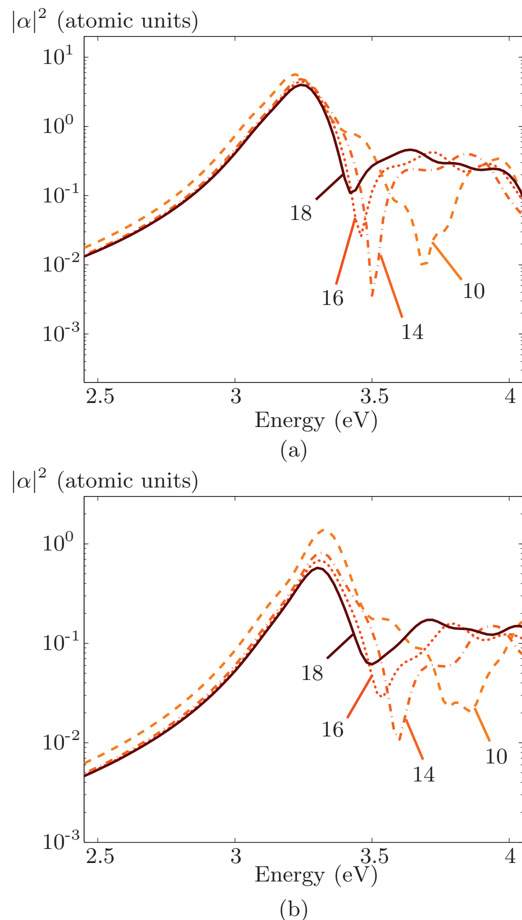

FIG. 2. (Color online) Convergence of $|\alpha|^2$ with respect to a traditional homogeneous sampling of the BZ for (a) BSE and (b) RPA. A shift in a nonsymmetric direction is used (see text). The number indicated is n_k and the grid is therefore $(n_k \times n_k \times n_k|\mathbf{s})$. The convergence is difficult to achieve for energies larger than 3.2 eV.

However, from Fig. 2, we observe that the convergence is much more difficult for the square of the Raman susceptibility, in the region beyond 3.2 eV (which corresponds to the optical gap). Important features still change going from $(16 \times 16 \times 16|\mathbf{s})$ to $(18 \times 18 \times 18|\mathbf{s})$. Interestingly, such a difficult convergence is present both with and without excitonic contributions, as exemplified by the upper and lower parts of Fig. 2.

With the method of shifted grids, convergence is not achievable, given our computational resources, beyond 3.2 eV. Indeed, the scaling of the method with respect to the sampling of the BZ is $O(N_k^2)$, with N_k the total number of k points in the full BZ. With $N_k = 18^3$, the convergence is not yet reached.

In contrast, convergence is much better below the gap value. In the next paragraphs, we first perform an analysis of the convergence for such frequencies, then turn to the higher-frequency part of the spectrum, for which we have developed a double-grid technique.

A. The convergence below the band gap

To have a quantitative understanding of the convergence in the low-energy part, we use a Taylor expansion and give

TABLE I. Cauchy coefficients for ε and α within the BSE framework [see Eqs. (22) and (23)]. The grids used are $(n_k \times n_k \times n_k|\mathbf{s})$.

n_k	12	14	16	18
ε_0	$1.3276 \times 10^{+1}$	$1.3259 \times 10^{+1}$	$1.3256 \times 10^{+1}$	$1.3256 \times 10^{+1}$
C_2^ε	7.7792×10^{-1}	7.7571×10^{-1}	7.7527×10^{-1}	7.7543×10^{-1}
C_4^ε	5.2340×10^{-2}	5.2084×10^{-2}	5.2020×10^{-2}	5.2030×10^{-2}
C_6^ε	5.8317×10^{-3}	5.7806×10^{-3}	5.7638×10^{-3}	5.7605×10^{-3}
α_0	2.5167×10^{-2}	2.4464×10^{-2}	2.4159×10^{-2}	2.4027×10^{-2}
C_2^α	4.8710×10^{-3}	4.7337×10^{-3}	4.6719×10^{-3}	4.6460×10^{-3}
C_4^α	5.7303×10^{-4}	5.5503×10^{-4}	5.4689×10^{-4}	5.4347×10^{-4}
C_6^α	1.2462×10^{-4}	1.1894×10^{-4}	1.1618×10^{-4}	1.1491×10^{-4}

coefficients similar to the so-called Cauchy coefficients for the macroscopic dielectric function.[43] Since the function is even with respect to the frequency, we can expand the absolute value of the Raman tensor and the real part of the dielectric function with even powers of the frequency:

$$\alpha(\omega) = \alpha_0 + C_2^\alpha \omega^2 + C_4^\alpha \omega^4 + C_6^\alpha \omega^6, \tag{22}$$

$$\mathrm{Re}\varepsilon(\omega) = \varepsilon_0 + C_2^\varepsilon \omega^2 + C_4^\varepsilon \omega^4 + C_6^\varepsilon \omega^6. \tag{23}$$

The coefficients can be obtained by a least-square fitting of the finite-difference results until 1.5 eV (see Table I).

The results of the fit obtained with this technique are presented in Fig. 3. The range of validity of this fit goes beyond 2 eV. A fitting above 2 eV leads to oscillatory behavior in the very low-energy range: the four-term expansion in Eqs. (22) and (23) is not accurate enough to describe the higher-energy part.

Cauchy coefficients are already well-converged for the 14^3 grid (within a few percent for the first and second ones). However, such a fit does not correctly describe the resonance close to the gap energy.

B. The convergence above the band gap

As mentioned earlier, we analyze the convergence of the Raman intensities in both the BSE case [Fig. 2(a)] and in the RPA approximation [Fig. 2(b)]. Both the RPA and the BSE present similar difficulties to converge the final results. Hence, we can conclude that the convergence issue is *not* primarily due to the building up of the excitons that arises from the off-diagonal couplings, but is already present at the independent-particle level. Why the convergence rate in the case of Raman susceptibility is smaller than in the macroscopic dielectric function can be understood as follows. The imaginary part of the dielectric function, for a given wave-vector grid, is made of numerous broadened Dirac delta contributions, each of which corresponds to one transition from a valence band to a conduction band. For such a spectrum to look smooth, the broadening should be comparable to the typical spacing between delta functions. In contrast, the frequency-dependent Raman intensity is obtained by differentiating the dielectric function. Hence, the Raman intensity evolution corresponds to the superposition of a large number of *derivatives* of broadened delta functions, whose oscillatory character is much stronger than the broadened Dirac functions. This is reflected at the level of the Raman intensity.

Having identified the problem, we design another strategy for sampling the BZ, which largely reduces the computational burden and memory requirements. In the same spirit as Ref. 44, but with a rather different implementation, we introduce a double-grid technique.

We perform a set of BSE calculations, indexed by the label i, each with the same number of points in the BZ forming a "coarse" grid, differing by their shift $\mathbf{s}_i$:

$$\{\mathbf{k}\}_i = (n_k \times n_k \times n_k|\mathbf{s}_i). \tag{24}$$

The shifts $\mathbf{s}_i$ are chosen in order to obtain a homogeneous sampling of the subspace between the coarse points:

$$\{\mathbf{s}_i\} = (n_{\mathrm{div}} \times n_{\mathrm{div}} \times n_{\mathrm{div}}|\mathbf{h}) \tag{25}$$

with n_{div} the number of subdivisions in each direction and $\mathbf{h} = (1/2,1/2,1/2)$. A 2D schematic representation is illustrated in Fig. 4.

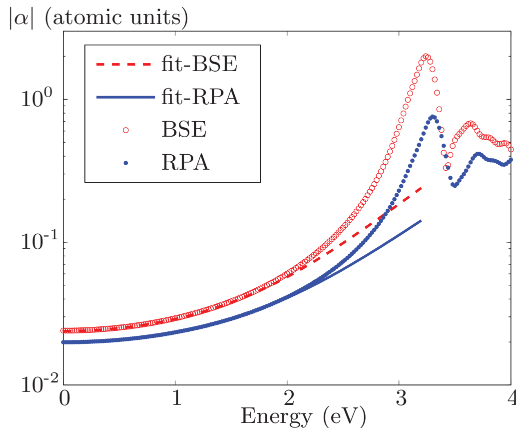

FIG. 3. (Color online) Comparison between the absolute value of the Raman tensor obtained from the BSE and RPA, and their Cauchy polynomial fits in the low-frequency part of the spectrum. Coefficients have been obtained by fitting the data up to 1.5 eV for a uniformly shifted $18 \times 18 \times 18$ k-grid.

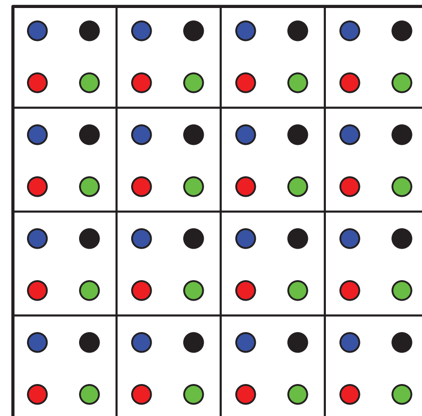

FIG. 4. (Color online) Schematic representation of the sampling of the BZ by a double-grid technique in two dimensions, with $n_k = 4$ and $n_{\mathrm{div}} = 2$ (see text for notations). One BSE computation is done on each grid (represented by four different colors in this picture). The final result is obtained by averaging the $n_{\mathrm{div}}^2 = 4$ different computations.

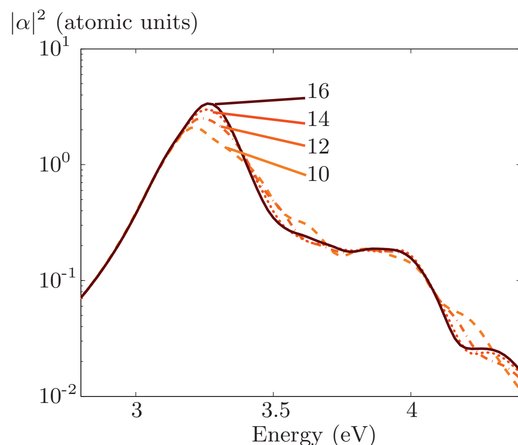

FIG. 5. (Color online) Convergence of $|\alpha|^2$ with respect to the sampling of the Brillouin zone obtained with the double-grid technique. The number indicates n_k for the "coarse" sampling of the Brillouin zone in each direction. A total of 64 points are sampled in the subspace between the coarse points.

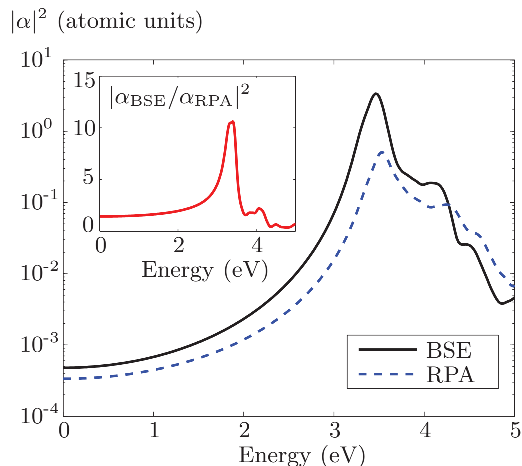

FIG. 6. (Color online) Comparison of the frequency evolution of $|\alpha|^2$ (in atomic units) with Bethe-Salpeter and RPA formalisms. The inset shows the ratio between the two curves. The ratio goes up to 10 at the level of the direct band gap (3.2 eV).

With this technique, the macroscopic dielectric function is obtained by averaging the different results computed on the "coarse" grids:

$$\varepsilon^{\mathrm{av}}(\omega) = \frac{1}{n_{\mathrm{div}}^3} \sum_i \varepsilon(\omega|\{\mathbf{k}\}_i), \tag{26}$$

where $\varepsilon(\omega|\{\mathbf{k}\}_i)$ is the macroscopic dielectric function obtained for the "coarse" computation with the grid $\{\mathbf{k}\}_i$, Eq. (24).

Figure 5 presents the results obtained with different coarse mesh samplings (different n_k), averaging over 64 calculations ($n_{\mathrm{div}} = 4$ is kept constant). Of course, when n_k becomes very large, the Raman spectrum must tend to the same spectrum as without this double-grid technique. But the computational effort is largely reduced. Indeed, the residual fluctuation when going from $n_k = 14$ to 16 can be seen to be rather small already. In the RPA case, $n_k = 16$ with $n_{\mathrm{div}} = 4$ corresponds exactly to a $[64 \times 64 \times 64|(1/2,1/2,1/2)]$ uniform grid, which would be untractable in the BSE case. It is worth stressing that in the current method, we can take advantage of symmetries to reduce the number of "coarse" grid calculations, since some meshes are equivalent. For example, the number of required computations with $n_{\mathrm{div}} = 4$ falls to 20 for the case in which an atom is displaced and to 10 for the equilibrium position.

V. ANALYSIS OF THE THEORETICAL RESULTS

In this section, we analyze in more detail the importance of excitonic effects on the Raman spectrum. A comparison between BSE and RPA results is reported in Fig. 6. Note how the excitonic effects amplify the Raman intensity by more than an order of magnitude in the band-gap region. Much smaller amplifications are observed for low frequencies.

Since the integral of the imaginary part of the dielectric susceptibility is related to the plasmon frequency ω_p (the so-called f-sum rule),[45,46]

$$\int_0^\infty \omega\, \mathrm{Im}\{\varepsilon_{ij}(\omega)\} d\omega = \frac{1}{2}\pi\omega_p^2\delta_{ij}, \tag{27}$$

and since ω_p does not depend on atomic positions, the integral of the imaginary part of the Raman susceptibility vanishes:

$$\int_0^\infty \omega\, \mathrm{Im}\{\alpha_{ij}(\omega)\} d\omega = 0. \tag{28}$$

Accordingly, negative and positive regions are present in Fig. 7. On the basis of Eq. (28), we can see that the difference between the BSE and RPA results for the Raman intensity is due to the lowering of the energy and the amplification of the main peak of the imaginary part of the Raman susceptibility.

In the approximation for which the atomic displacement induces a global rigid shift of all the conduction bands with respect to all valence bands, with energy $\Delta\varepsilon = \varepsilon_{c\mathbf{k}} - \varepsilon_{v\mathbf{k}}$ (or, alternatively, if one transition dominates), the derivative with respect to an atomic displacement is related to the derivative

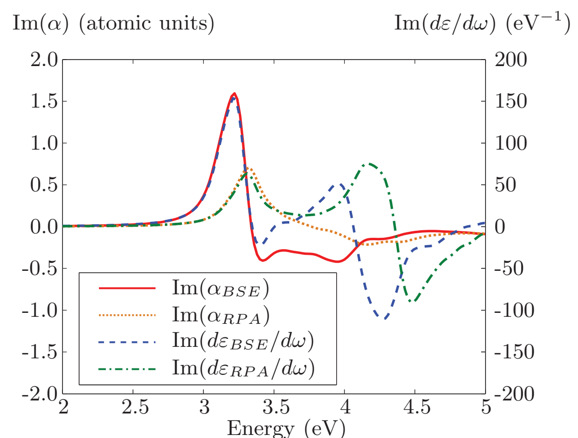

FIG. 7. (Color online) Comparison of the imaginary part of the Raman susceptibility and of the derivative of the susceptibility with respect to frequency for the Bethe-Salpeter equation (BSE) and for the random-phase approximation (RPA). These computations are achieved with the double-grid technique for $n_k = 16$ and $n_{\mathrm{div}} = 4$.

with respect to frequency by

$$\left.\frac{\partial \chi}{\partial \tau}\right|_{\omega} \approx \left.\frac{\partial \chi}{\partial \Delta\varepsilon}\right|_{\omega} \frac{\partial \Delta\varepsilon}{\partial \tau} = -\left.\frac{\partial \chi}{\partial \omega}\right|_{\omega} \frac{\partial \Delta\varepsilon}{\partial \tau}. \tag{29}$$

This relation shows that the amplitude of the Raman effect will follow the variation of the band structure with the atomic displacement.[1]

As represented in Fig. 7, this approximation is only valid at the onset of absorption. In this range of energies, the curves are qualitatively on top of each other. This shows that the transition corresponding to the energy gap dominates in this range of energies. For higher energies, however, this approximation is not valid since each band can contribute differently from other bands in the Raman susceptibility.

VI. COMPARISON WITH EXPERIMENTAL DATA

Figure 8 shows the *ab initio* results for the Raman susceptibility of silicon, and compares these results with the experimental data obtained by Compaan and Trodahl,[27] who measured Raman intensity as a function of the frequency in silicon. In this figure, the theoretical results are obtained with a scissor value of 0.85 eV that reproduces the experimental gap at 0 K (3.4 eV) instead of 0.65 eV, which reproduces the theoretical GW gap (3.2 eV).

In terms of absolute value, the polarizability $a_{\mathrm{BSE}} = 19.75$ Å^2 obtained at 1.1 eV compares reasonably well with the experimental data of 23 ± 4 Å^2. The RPA value, $a_{\mathrm{RPA}} = 15.87$ Å^2, does not match the experimental value. This confirms the need to correctly describe excitonic effects even for energies well below the gap.

We can distinguish three different regions with different behavior: the low-energy (preresonance) region, from 2 to 3.2 eV, the band-gap region from 3.2 to 3.5 eV, and the higher-energy region beyond 3.5 eV.

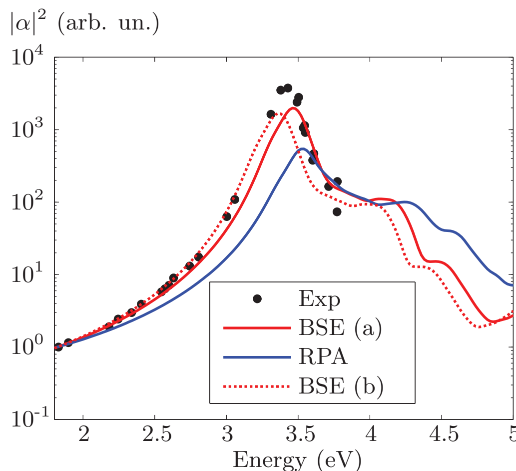

FIG. 8. (Color online) Comparison of the theoretical and experimental Raman susceptibility of silicon. The experimental gap is reproduced by choosing a scissor value of 0.85 eV. A rigid energy shift of −0.1 eV is applied to results called "BSE (a)" in order to obtain "BSE (b)." The vertical values are given in arbitrary units in Ref. 27. Therefore, these data are fitted to the BSE and to the RPA by means of a multiplicative factor fixed to pass through the first experimental point.

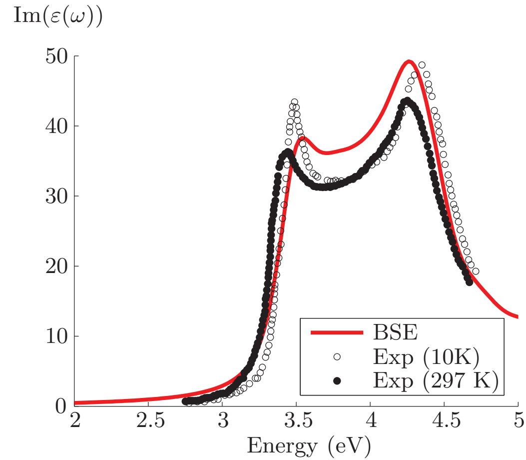

FIG. 9. (Color online) Theoretical (BSE) and experimental absorption spectrum of silicon for 10 K and for 297 K.[47]

The Bethe-Salpeter method allows for the reproduction of the experimental amplification of the Raman susceptibility with the frequency. Agreement using this method is significantly better than the agreement obtained by the RPA. In the band-gap region, however, there is a discrepancy between the theoretical and the experimental maximal Raman susceptibility. The BSE maximum is nevertheless still closer to the experimental maximum than the RPA maximum. Moreover, it is important to note that the BSE theoretical results obtained in this work are valid only at low temperature, whereas the experimental data are measured at 300 K. The effect of the temperature on the absorption spectrum is illustrated in Fig. 9, where the first absorption peak at 10 K is brought to lower energies at 297 K, in agreement with the gap narrowing. On the basis of this observation, we can expect an improvement in the agreement if temperature effects are included in the *ab initio* calculations. As a first approximation, the effect of temperature on the gap energy can be mimicked by a rigid shift of the Raman intensity curve toward lower energies, as shown in Fig. 8 for data called "BSE (b)." With this correction, the agreement is highly improved in the low-energy part of the Raman susceptibility. The amplification factor and postresonance are, however, not significantly improved. We attribute the disagreement to the different approximations we have performed so far, in particular within the BSE framework and the quasistatic approximation.

Our *ab initio* approach is not able to describe the higher-energy region as well as the lower-energy region. Without the temperature correction, the BSE theoretical results underestimate the evolution of intensities in the preresonance region. The inclusion of temperature leads to a partial improvement in the overall agreement, except for the postresonance region. The slope of decrease in the postresonance region is in agreement with the experimental slope while the intensity value is underestimated.

The present approach relies on several approximations, whose roles still need to be analyzed. We have neglected, among others, the phonon frequency in the "quasistatic" approach, the quadratic response with respect to atomic

displacements, the self-consistency with the *GW* approximation, the non-Hermitian coupling within the BSE, the frequency dependence of the BSE Hamiltonian, indirect transitions, and the additional effects due to phonons (the experimental data have been obtained at room temperature). The latter effects could be included using a method similar to the method presented by Marini *et al.*[48] Calculations including all these effects would require computational resources unavailable to us at present.

Despite these effects, in all regions, the BSE results are in better agreement with experimental data than the RPA results. This clearly indicates the importance of excitonic effects for an accurate *ab initio* description of Raman spectra.

Note also that silicon might possess specific characteristics that induce the good agreement observed here with the present method and associated approximations. Such a good agreement might not be observed for materials with different characteristics, such as a lower crystalline symmetry, the presence of stronger spin-orbit coupling, the presence of some ionicity, the presence of multiple types of atoms, etc.

VII. CONCLUSIONS AND PERSPECTIVES

A technique for the *ab initio* study of resonant Raman intensity has been proposed and applied to silicon. The technique relies on finite differences of the macroscopic dielectric function evaluated for distorted geometries, and it includes excitonic effects by solving the Bethe-Salpeter equation. We found that convergence of the Raman spectrum with respect to the number of wave vectors used to sample the Brillouin zone is problematic. To tackle this problem, we proposed a double-grid averaging process that significantly improves convergence while keeping computational effort at a reasonable level.

The Bethe-Salpeter results are in better agreement with experimental results than those results obtained without excitonic effects (RPA). For laser energies in the band-gap region or lower, the agreement is excellent if one takes into account the small rigid shift that is needed to align the first peak of the theoretical dielectric absorption, as well as the experimental shift for the same temperature as the Raman spectrum. The agreement is still not perfect, however, in the high-energy region. This may be attributed to many different effects, which will be examined in future works.

ACKNOWLEDGMENTS

Y.G. and M.G. wish to acknowledge the financial support of the Fonds National de la Recherche Scientifique (FNRS, Belgium). Y.G. is supported by an FRS-FNRS Research Fellowship. The authors would like to thank Yann Pouillon and Jean-Michel Beuken for their valuable technical support and help with the test and build system of ABINIT, and Nicola Thrupp for her careful reading of the final version of the article. Computational resources have been provided by the supercomputing facilities of the Université Catholique de Louvain (CISM/UCL) and the Consortium des Equipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the Fonds de la Recherche Scientifique de Belgique (FRS-FNRS).

*yannick.gillet@uclouvain.be

[1] W. Weber and R. Merlin, in *Raman Scattering in Materials Science*, edited by W. Weber and R. Merlin (Springer, Berlin, 2000).
[2] G. Stock, C. Woywod, W. Domcke, T. Swinney, and B. S. Hudson, J. Chem. Phys. **103**, 6851 (1995).
[3] M. Dresselhaus, G. Dresselhaus, R. Saito, and A. Jorio, Phys. Rep. **409**, 47 (2005).
[4] L. Malard, M. Pimenta, G. Dresselhaus, and M. Dresselhaus, Phys. Rep. **473**, 51 (2009).
[5] P. Venezuela, M. Lazzeri, and F. Mauri, Phys. Rev. B **84**, 035433 (2011).
[6] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[7] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[8] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).
[9] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett. **58**, 1861 (1987).
[10] X. Gonze and C. Lee, Phys. Rev. B **55**, 10355 (1997).
[11] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).
[12] A true semiconductor might be described as a metal by DFT, e.g., with usual local or semilocal functionals.[49,50]
[13] S. Baroni and R. Resta, Phys. Rev. B **33**, 5969 (1986).
[14] M. Veithen, X. Gonze, and P. Ghosez, Phys. Rev. Lett. **93**, 187401 (2004).
[15] M. Veithen, X. Gonze, and P. Ghosez, Phys. Rev. B **71**, 125107 (2005).
[16] P. Hermet, G. Fraysse, A. Lignie, P. Armand, and P. Papet, J. Phys. Chem. C **116**, 8692 (2012).
[17] J. Zhang, X. F. Li, B. Xu, and K. L. Yao, J. Alloy Compd. **509**, 4929 (2011).
[18] R. Caracas and E. Bobocioiu, WURM database, http://www.wurm.info.
[19] R. Caracas and E. Bobocioiu, Am. Mineral. **96**, 437 (2011).
[20] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors: Physics and Material Properties* (Springer, Berlin, 2010).
[21] W. Aulbur, L. Jönsson, and J. Wilkins, Solid State Phys. **54**, 1 (1999).
[22] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[23] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
[24] S. Botti, F. Sottile, N. Vast, V. Olevano, L. Reining, H.-C. Weissker, A. Rubio, G. Onida, R. Del Sole, and R. W. Godby, Phys. Rev. B **69**, 155112 (2004).
[25] E. Gross and K. Burke, in *Time-Dependent Density Functional Theory*, edited by M. Marques, C. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. Gross, Lecture Notes in Physics Vol. 706 (Springer, Berlin, 2006).
[26] S. Sharma, J. K. Dewhurst, A. Sanna, and E. K. U. Gross, Phys. Rev. Lett. **107**, 186401 (2011).
[27] A. Compaan and H. J. Trodahl, Phys. Rev. B **29**, 793 (1984).
[28] J. B. Renucci, R. N. Tyte, and M. Cardona, Phys. Rev. B **11**, 3885 (1975).

[29] Atomic units are used throughout this paper. One atomic unit for α corresponds to $7.62 \times 10^9 \mathrm{m}^{1/2}/\mathrm{kg}^{1/2}$.

[30] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).

[31] The exchange contributions, Eq. (12), are often referred to as being local-field effects.

[32] X. Gonze, G.-M. Rignanese, M. Verstraete, J.-M. Beuken, Y. Pouillon, R. Caracas, F. Jollet, M. Torrent, G. Zerah, M. Mikami, Ph. Ghosez, J.-Y. Veithen, M. Raty, V. Olevano, F. Bruneval, L. Reining, R. Godby, G. Onida, D. Hamann, and D. Allan, Z. Kristallogr. **220**, 558 (2005).

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

[34] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[35] R. W. G. Wyckoff, *Crystal Structures* (Wiley, New York, 1963).

[36] F. Bruneval and X. Gonze, Phys. Rev. B **78**, 085125 (2008).

[37] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).

[38] B. Arnaud and M. Alouani, Phys. Rev. B **62**, 4464 (2000).

[39] Landolt-Bornstein, *Zahlenwerte und Funktionen aus Naturwissenschaften und Technik* (Springer-Verlag, Berlin, 1982), Vol. III, Pt. 23a.

[40] Z. H. Levine and D. C. Allan, Phys. Rev. Lett. **63**, 1719 (1989).

[41] M. A. Renucci, J. B. Renucci, R. Zeyher, and M. Cardona, Phys. Rev. B **10**, 4309 (1974).

[42] R. Haydock, Comput. Phys. Commun. **20**, 11 (1980).

[43] M. A. L. Cauchy, *Mémoire sur la dispersion de la lumière* (Calve, Prague, 1836).

[44] D. Kammerlander, S. Botti, M. A. L. Marques, A. Marini, and C. Attaccalite, Phys. Rev. B **86**, 125203 (2012).

[45] D. L. Johnson, Phys. Rev. B **9**, 4475 (1974).

[46] D. Pines, *Elementary Excitations in Solids* (W.A. Benjamin, New York, 1963).

[47] G. E. Jellison and F. A. Modine, Phys. Rev. B **27**, 7466 (1983).

[48] A. Marini, Phys. Rev. Lett. **101**, 106405 (2008).

[49] R. W. Godby and R. J. Needs, Phys. Rev. Lett. **62**, 1169 (1989).

[50] P. Ghosez, X. Gonze, and R. W. Godby, Phys. Rev. B **56**, 12811 (1997).