

Finite electric-field approach to evaluate the vertex correction for the screened Coulomb interaction in the quasiparticle self-consistent GW method

Hirofumi Sakakibara,^{1,*} Takao Kotani,¹ Masao Obata,² and Tatsuki Oda²

¹*Department of Applied Mathematics and Physics, Tottori University, Tottori 680-8552, Japan*

²*Institute of Science and Engineering, Kanazawa University, Kanazawa 920-1192, Japan*



(Received 27 September 2019; revised manuscript received 10 April 2020; accepted 15 April 2020; published 15 May 2020)

We apply the quasiparticle self-consistent GW method (QSGW) to slab models of ionic materials—LiF, KF, NaCl, MgO, and CaO—under electric field. Then we obtain the optical dielectric constants $\epsilon_\infty(\text{Slab})$ from the differences of the slopes of the electrostatic potential in the bulk and vacuum regions. Calculated $\epsilon_\infty(\text{Slab})$ show very good agreement with experiments. For example, we have $\epsilon_\infty(\text{Slab}) = 2.91$ for MgO, in agreement with the experimental value $\epsilon_\infty(\text{Experiment}) = 2.96$. This is in contrast to $\epsilon_\infty(\text{RPA}) = 2.37$, which is calculated in the random-phase approximation for the bulk MgO in QSGW. After we explain the difference between the quasiparticle-based perturbation theory and the Green's-function-based perturbation theory, we interpret the large difference $\epsilon_\infty(\text{Slab}) - \epsilon_\infty(\text{RPA}) = 2.91 - 2.37$ as the contribution from the vertex correction of the proper polarization, which determines the screened Coulomb interaction W . Our result encourages the theoretical development of the self-consistent G_0W approximation along the line of QSGW self-consistency, as was performed by Shishkin, Marsman, and Kresse [*Phys. Rev. Lett.* **99**, 246403 (2007)].

DOI: [10.1103/PhysRevB.101.205120](https://doi.org/10.1103/PhysRevB.101.205120)

I. INTRODUCTION

The quasiparticle self-consistent GW method (QSGW) is one of the most reliable methods to determine the one-particle effective Hamiltonian, which describes the independent-particle picture or the quasiparticle (QP) picture, for treating electric excitations of materials [1–3]. Other competitive methods, such as the Heyd-Scuseria-Ernzerhof (HSE) functional [4] and the Tran-Blaha-09 functional [5], may work well in many systems, although we may need to use material-dependent parameters [6]. In contrast, QSGW is virtually parameter-free and gives reliable descriptions for a wide range of materials, not only metals and semiconductors but also transition-metal oxides, type-II superlattices, and $4f$ systems [7–10]. Since heterogeneous mixtures of materials are used in current technologies, it is worthwhile to develop QSGW further as a tool to treat the electronic structures of such materials, where methods including such material-dependent parameters are hardly applicable.

We now summarize the idea of QSGW. First, let us recall the GW approximation (GWA): GWA is a perturbation method applicable to many-body Hamiltonians divided as $H = H_0 + (H - H_0)$. Then we construct the bare Green's function as $G_0 = 1/(\omega - H_0)$, which is used to give the self-energy $\Sigma(\mathbf{r}, \mathbf{r}', \omega)$ in the standard procedure of GWA. In principle, we can apply GWA to any choice of the one-body Hamiltonian H_0 , such as that in the local density approximation/generalized gradient approximation

(LDA/GGA), or even in LDA+ U . Thus the main problem of GWA is how to determine H_0 . QSGW is a method to determine H_0 from the standpoint of self-consistent perturbation.

However, QSGW has a shortcoming in that it leads to a systematic overestimation of the exchange effects. This results in slightly larger band gaps in QSGW for materials. Faleev, van Schilfgaarge, and Kotani [1–3] suggested that the overestimation can be eliminated by performing improved QSGW calculations, taking into account the enhancement of the screening effect due to the electron-hole correlation in the evaluation of the screened Coulomb interaction W . This is based on the theoretical consideration combined with the observation that the calculated optical dielectric constant ϵ_∞ in the random-phase approximation (RPA) in QSGW is $\sim 20\%$ smaller than that in experiments for certain materials [3,11].

Such an improved calculation was performed by Shishkin, Marsman, and Kresse, who included the enhancement of the screening effect [12]. The enhancement is via the vertex correction for the proper polarization P , which determines $W = v/(1 - vP)$, where v denotes the Coulomb interaction. They approximately include the lowest-order vertex correction due to the electron-hole correlation; see Eq. (15) in Ref. [13]. Their results are theoretically quite satisfactory in the sense that both band gaps and ϵ_∞ , which are calculated simultaneously and self-consistently without parameters as in HSE, are in agreement with experiments. For example, the calculated values $\epsilon_\infty = 2.96$ and the band gap $E_G = 8.12$ eV for MgO are in agreement with the experiments, 2.95 and 7.83 eV, respectively. See scGW (e-h) in Table I in Ref. [12]. Furthermore, based on these theoretical analyses, we can introduce QSGW80 to avoid the very expensive computational costs of the method by Shishkin *et al.*. QSGW80 is just a simple

*Corresponding author: sakakibara@tottori-u.ac.jp

hybridization, 80% QSGW + 20% GGA, which includes the enhancement of the screening effectively. The hybridization in QSGW80 is very different from the hybridization in HSE, where the mixing ratio α between GGA and the Hartree-Fock approximation strongly affects the final results. QSGW80 works well to describe experimental band gaps [14]. The performance of QSGW80 was systematically examined by Deguchi *et al.* in Ref. [8], where we can see that both the calculated band gaps and the effective masses are in good agreement with experiments. QSGW80 is successfully used for practical applications, such as to the type-II superlattice of InAs/GaSb [15,16].

In this paper, we evaluate ϵ_∞ not in bulk calculations with approximations such as those used in Ref. [12], but by using slab models with a finite electric bias voltage. We treat five ionic materials: LiF, KF, NaCl, MgO, and CaO. We put a slab in the middle of the vacuum region in a supercell. The electric field is applied by the effective screening medium (ESM) method given by Otani and Sugino [17]. We obtain ϵ_∞ from the ratio of slopes of the electrostatic fields in the slab region and the vacuum region. Our approach is based on the self-consistent method, thus we do not need to utilize approximations such as those used in Ref. [12]. Since we explicitly treat the response to the bias, our method includes higher-order effects in a self-consistent manner.

Our findings are that the calculated ϵ_∞ in QSGW for the slab models are very close to experimental values. This is in contrast to the fact that ϵ_∞ in RPA of QSGW are generally $\sim 20\%$ smaller than experimental values. This indicates that the vertex correction at the level of derivative of the QSGW self-energy should ensure that W is in agreement with experiments. Our results are consistent with those shown in Table II of Ref. [12].

We can interpret the screening enhancement, represented by the enlargement of ϵ_∞ , as the size of the vertex correction for the proper polarization P . Note that the vertex correction we evaluate is not what is defined in Hedin's equation [18]. In that equation, we see $P = -iGG\Gamma$, that is, the vertex function Γ is for the correction to $P = -iGG$, where G denotes the Green's function. Instead, we evaluate Γ for $P = -iG_0G_0\Gamma$, where G_0 is the bare Green's function. To clarify the above theoretical point on Γ , we provide an extensive discussion in Sec. II. We explain the role of Γ in the two kinds of perturbation theories. In Sec. III, we discuss QSGW+ESM, which is an implementation of QSGW combined with ESM. The QSGW+ESM for slab models should be very useful not only for our purposes here, but also for cases in which the usual GGA+ESM have difficulties. In Sec. IV, we present our results for ϵ_∞ , which are then interpreted as the vertex correction. In Sec. IV B, we provide a rationale of QSGW80 followed by a summary.

II. QP-BASED VERSUS G-BASED PERTURBATION

To describe our motivation in this paper, we have to clarify the difference between quasiparticle-based perturbation (QbP) and Green's-function-based perturbation (GbP). QbP is based on Landau-Silin's QP theory, while GbP is based on Hedin's theory. To illustrate the difference between QbP and GbP, we present a narrowband model in the following. The model

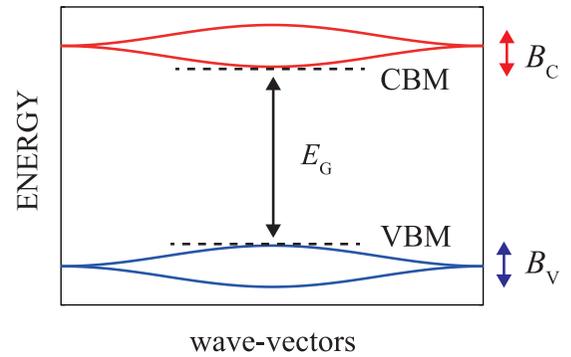


FIG. 1. Band structure of a narrowband model to illustrate the quasiparticle-based perturbation. CBM and VBM denote conduction-band minimum and valence-band maximum. The widths of bands B_C and B_V are smaller than the band gap E_G . As in the text, we expect well-defined QPs in this model.

represent situations in which we have a good QP picture (i.e., an independent-particle picture).

Recall that there are mainly two kinds of excitations in paramagnetic electronic systems, i.e., multiparticle excitations and collective excitations such as plasmons. The former is described by QPs interacting with each other. Note that plasmons are located at high energy because of the long-range Coulomb interaction [19]. These excitations can be hybridized, e.g., pseudoplasmons in silver, where one-particle excitations of $3d$ electrons are hybridized with plasmons.

A. Narrowband model to explain the QP-based perturbation

The QPs based on Landau-Silin's Fermi liquid theory were originally for metals [19]. However, the idea of QPs is rather easily applicable to insulators. We can consider QbP based on the QPs. To illustrate this, let us consider a narrowband model, a paramagnetic case given by a Hamiltonian H , which has a one-body term represented by finite numbers of Wannier functions in the primitive cell, and the Coulomb-like interaction $\frac{e^2}{\epsilon'|\mathbf{r}-\mathbf{r}'|}$, where ϵ' is a constant. We consider a case with the QPs shown in Fig. 1 given by the one-particle Hamiltonian H_0 .

Based on the perturbation, we expect that the lifetimes of all the electrons are infinite because the band gap E_G is large enough to prevent all the electrons from decaying into lower-energy electrons accompanying electron-hole pairs and holes as well. In other words, all bands are within the threshold of impact ionization [20]. Thus QPs described by $G_0 = 1/(\omega - H_0)$ should give well-defined one-particle excitations of the narrowband model.

This conclusion should be essentially kept even when we fully turn on the interaction as long as the following conditions are well-satisfied. First, excitons, which are binding states of electron-hole pairs, should be at only slightly lower energies than E_G . Second, plasmons should be located at high enough energies so that they are barely hybridized with the one-particle excitations. Under these assumptions, we have well-defined QPs. We can consider a path of adiabatic connection while maintaining the well-defined QPs given by H_0 since the QP spectrum is clearly separated from the other excitations. It can be written as $H_\lambda = H_0 + (H_\lambda - H_0)$, where λ is from

zero to unity, and where $H = H_{\lambda=1}$. Note that $(H_\lambda - H_0)$ must contain a λ -dependent one-body term.

In this narrowband model, QbP should be suitable; QPs are interacting with each other by $(H_{\lambda=1} - H_0)$. As for the proper polarization function, we can take all the noninteracting two-body QP excitations correctly by $P_0 = -iG_0G_0$. Based on this QbP, the G_0W_0 approximation $\Sigma = iG_0W_0$, where $W_0 = v/(1 - vP_0)$, is physically justified. That is, it describes how the motion of QPs in G_0 is perturbed by the dynamical self-interaction given by W_0 in RPA. Most of the first-principles calculations in the literature are implicitly based on this QbP, while some of the latest literature [21,22] is based on GbP as explained in Sec. II C. QSGW is a method to determine H_0 self-consistently in QbP.

B. Vertex function Γ_{QbP} in QP-based perturbation

To improve G_0W_0 , we may include electron-hole correlation. The corrections replace $P_0 = -iG_0G_0$ with P , where we include the correlation via the Bethe-Salpeter equation (ladder diagrams). That is, we include the two-body spectrum in the proper polarization accurately. This enables us to use $W = 1/(1 - vP)$ instead of W_0 , resulting in the G_0W approximation as $\Sigma = iG_0W$. In this paper, we concentrate on the G_0W approximation as in the case of scGW (e-h) in Ref. [12]. In QbP, we thus define the vertex function Γ_{QbP} for W as $P = -iG_0G_0\Gamma_{\text{QbP}}$. Roughly speaking, $P = P_0\Gamma_{\text{QbP}}$. We give the numerical evaluation for this Γ_{QbP} via the evaluation of ϵ_∞ as shown in Sec. IV A.

To go beyond the $\Sigma = iG_0W$ approximation here, we need to take into account three-particle intermediate states. However, this is not theoretically straightforward because of a double-counting problem in that $\Sigma = iG_0W$ already partially takes into account such states. We will need to construct theories of the three-particle problem without double-counting along the lines of first-principles calculations. We do not address this problem in this paper.

C. G-based perturbation and vertex function Γ_{GbP}

Let us consider how we can apply GbP to the narrowband model. In contrast to G_0 in QbP, the one-body Green function G has complex meanings. We have an imaginary part of G at high energies, representing QPs hybridized with plasmons (plasmarens). Because of the sum rule for $\text{Im}[G]$, the QP parts are suppressed by the Z -factor. Thus $P = -iGG$ do not contain the two-body noninteracting excitations with the correct weight, in contrast to the case $P_0 = -iG_0G_0$.

In principle, this problem is corrected by including the vertex function Γ_{GbP} in Hedin's equation to determine the one-particle Green's function $G(1, 2)$ [18]. Because Hedin's equation is theoretically rigorous, we expect $-iGG\Gamma_{\text{GbP}} \approx P_0 = -iG_0G_0$ in the model; see the discussion of Sec. II A where P_0 gives a good approximation for the model. That is, contributions related to the collective excitations and renormalization factors Z in $-iGG$ should be virtually eliminated by the factor Γ_{GbP} . However, such numerical calculations can be computationally very demanding [23]. A similar discussion of Z -factor cancellation is also seen when we multiply G_0 by W . That is, we should have $G_0W \approx GW\Gamma_{\text{GbP}}$ since QbP

correctly treats the model. Our analysis here is consistent with Takada's analysis based on the Ward identity [24].

QbP should be superior to GbP in materials such as semiconductors. In contrast to GbP, QbP is quite simple and physically convincing. We should not be confused by the similarity of QbP and GbP. In this paper, we evaluate Γ_{QbP} . In the following, we calculate the enhancement of the screening effect. Then we evaluate the size of the ratio P/P_0 from a comparison between calculated ϵ_∞ in RPA and ϵ_∞ in the slab models. This ratio gives the size of Γ_{QbP} .

III. QSGW COMBINED WITH THE EFFECTIVE SCREENING MEDIUM METHOD

To calculate ϵ_∞ , we put a slab in the middle of the vacuum region in a supercell. ϵ_∞ is calculated from the difference between the slopes in the vacuum region and in the slab region under small bias voltage. The supercell we use is detailed at the beginning of Sec. IV. In such calculations, we can obtain ϵ_∞ beyond the bulk calculation in RPA, as we explain in the next paragraph. That is, we can obtain ϵ_∞ including the effect of the vertex correction.

To explain how the effect is included, let us first consider slab calculations in the case of GGA. We first perform self-consistent calculation under zero bias. Then we perform self-consistent calculation under finite bias (theoretically, it should be infinitesimally small). Then we have the difference of the electron density $\delta n(\mathbf{r})$ between the two calculations. Simultaneously, we have a corresponding response of the one-particle potential given as $\delta V(\mathbf{r}) = \int d^3r' v(\mathbf{r} - \mathbf{r}')\delta n(\mathbf{r}') + \frac{\partial V_{\text{xc}}^{\text{GGA}}}{\partial n(\mathbf{r})}\delta n(\mathbf{r})$. The last term is the difference in the exchange-correlation (xc) potential caused by $\delta n(\mathbf{r})$ self-consistently. That is, the derivative $\frac{\delta V(\mathbf{r})}{\delta n(\mathbf{r})}$ contains the contribution of the xc kernel $f_{\text{xc}} = \frac{\partial V_{\text{xc}}^{\text{GGA}}}{\partial n(\mathbf{r})}$. Under the bias, we can obtain ϵ_∞ from the ratio of slopes of the electrostatic potential in the vacuum region and in the slab region. It should contain the contribution from f_{xc} , which is identified as the vertex correction in GGA.

This is essentially the same in QSGW. Recall that the self-energy in QSGW, denoted as $V_{\text{xc}}^{\text{QSGW}}(\mathbf{r}, \mathbf{r}')$, is a static nonlocal potential replacing $V_{\text{xc}}^{\text{GGA}}$. The derivative of the one-particle potential is given as $\delta V(\mathbf{r}, \mathbf{r}') = \int d^3r'' v(\mathbf{r} - \mathbf{r}'')\delta n(\mathbf{r}'') + \delta V_{\text{xc}}^{\text{QSGW}}(\mathbf{r}, \mathbf{r}')$, where the last term plays the role of $\frac{\partial V_{\text{xc}}^{\text{GGA}}}{\partial n(\mathbf{r})}\delta n(\mathbf{r})$. Note that $\delta V_{\text{xc}}^{\text{QSGW}}(\mathbf{r}, \mathbf{r}')$ is determined self-consistently, although it is not as simply given as $\frac{\partial V_{\text{xc}}^{\text{QSGW}}}{\partial n(\mathbf{r})}\delta n(\mathbf{r})$. Our calculations include the contribution of $\delta V_{\text{xc}}^{\text{QSGW}}(\mathbf{r}, \mathbf{r}')$ self-consistently as in the case of GGA. Our method is similar to the method used to solve the Bethe-Salpeter equation in Ref. [25].

Our QSGW+ESM is implemented in a first-principles package ECALJ [8,26], which is based on a mixed-basis method, namely the augmented plane wave (APW) and muffin-tin (MT) orbital method (the PMT method) [27–30]. The PMT method is an all-electron full-potential method that uses not only the APW basis in the LAPW method, but also the MT orbitals in the LMTO method simultaneously in the expansion of eigenfunctions. It also uses the local orbital basis [31]. In addition to the PMT method, we had implemented the QSGW method [8,29]. In PMT, we use very localized

untuned MTOs, which contain a damping factor $\exp(-\kappa r)$, where κ are fixed to be 1/bohr and/or 2/bohr, together with low-cutoff APWs (≤ 3 Ry). We do not need empty spheres since the APWs can handle vacuum regions of slab models. The charge density is represented in the following three-component representation: “smooth part,” “true part within MT,” and “counterpart within MT,” as in the case of the PAW method [32]. In contrast to the other *GW* methods, which require the Wannier-interpolation technique to make band plots in the whole Brillouin zone, we can make band plots easily without resorting to that technique [8]. In the following, after an explanation of the general theory of ESM, we show how to implement ESM in the PMT method.

A. The electrostatic potential in the ESM method

We apply the ESM method [17] to slab models under an external electric field. We treat a supercell with a periodic boundary condition where we have a slab with periodicity in the xy plane. The slab is at the middle of the supercell. The position in the cell is specified by $\mathbf{r} = (\mathbf{r}_{\parallel}, z)$. The planes at $z = -z_0$ and $z = z_0$ are the left and right ends of the supercells. The electrostatic potential is calculated from the charge density in the supercell assuming two electrodes are at $z = \pm z_0$ (we set $z_1 = z_0$ in Fig. 1 of Ref. [17]) to apply voltage to the supercell. In summary, the ESM in DFT is formulated from the total energy minimization, however this is not true in QSGW since QSGW itself is not formulated from the total energy minimization. After we obtain the following, we use Eq. (2) to determine the electrostatic potential. We use it even in QSGW.

Let us start from the energy functional of DFT in the ESM. It is written as

$$E[n] = E^{\text{kin}}[n] + E^{\text{xc}}[n] + E^{\text{es}}[n] + E^{\text{app}}[n]. \quad (1)$$

Here, we have kinetic energy $E^{\text{kin}}[n]$, xc energy $E^{\text{xc}}[n]$, and electrostatic energy $E^{\text{es}}[n]$. In addition, the last term is the applied electrostatic term $E^{\text{app}}[n] = \int d^3r V^{\text{app}}(\mathbf{r})[n(\mathbf{r}) + n_{\text{N}}(\mathbf{r})]$, where $n(\mathbf{r})$ and $n_{\text{N}}(\mathbf{r})$ are the electron density and the charge density of nuclei, respectively; $V^{\text{app}}(\mathbf{r})$ is a linear function of z , representing the external field.

In ESM, we enforce the periodicity in the supercell for the electrostatic potential. Thus we use $V^{\text{app}}(\mathbf{r})s(\mathbf{r})$ instead of $V^{\text{app}}(\mathbf{r})$, where we introduce a support function $s(\mathbf{r})$, which is unity for most of the regions, but it is going to be zero at $z = -z_0$ and $z = z_0$. It is different from unity only near the boundaries, $z \approx -z_0$ or $z \approx z_0$. Thus the potential $V^{\text{app}}(\mathbf{r})s(\mathbf{r})$ recovers the periodicity of the supercell. A constant can be added to $V^{\text{app}}(\mathbf{r})$ so that it maintains smooth periodicity over $z = \pm z_0$. As long as we use a large enough vacuum region, we have few electrons near the boundaries. Thus the choice of $s(\mathbf{r})$ is irrelevant.

A key in ESM is that we use the Green function $\bar{v}(\mathbf{r}, \mathbf{r}')$ for the electrostatic energy $E^{\text{es}}[n]$ instead of the Coulomb interaction $v(\mathbf{r} - \mathbf{r}')$ in the usual GGA calculations. As in Ref. [17], $\bar{v}(\mathbf{r}, \mathbf{r}')$ contains not only the Coulomb interaction $v(\mathbf{r} - \mathbf{r}')$ but also the effects due to the polarization of virtual electrodes, which are at $z = \pm z_1$ (we use $z_0 = z_1$ in our calculations here). Polarization of the slab occurs when the electrostatic potential remains constant at the electrodes.

Corresponding to $V^{\text{app}}(\mathbf{r})s(\mathbf{r})$, we use $s(\mathbf{r})\bar{v}(\mathbf{r}, \mathbf{r}')s(\mathbf{r}')$ instead of $\bar{v}(\mathbf{r}, \mathbf{r}')$ in practice. Then we have the well-defined Kohn-Sham total energy while maintaining the periodic boundary condition for a given $V^{\text{app}}(\mathbf{r})$.

The minimization of $E[n]$ with respect to $n(\mathbf{r})$ gives the Kohn-Sham potential $V(\mathbf{r})$ as

$$V(\mathbf{r}) = \int d^3r' \bar{v}(\mathbf{r}, \mathbf{r}') [n(\mathbf{r}') + n_{\text{N}}(\mathbf{r}')] + V^{\text{app}}(\mathbf{r}) + V^{\text{xc}}(\mathbf{r}). \quad (2)$$

Hereafter, we omit $s(\mathbf{r})$ for simplicity.

In QSGW [29], we cannot derive its fundamental equation from the energy minimization. Thus the formulation of QSGW+ESM is not exactly along the same lines as mentioned above. However, we can use the one-particle potential of Eq. (2) in the self-consistent cycle, where $V^{\text{xc}}(\mathbf{r})$ is replaced by a static version of the self-energy [1]. Thus, in principle, it is straightforward to perform QSGW+ESM.

B. ESM in the PMT method

In the PMT method, the electron density (and also the charge density) is represented by the three-component formalism described in Ref. [30], originally introduced by Soler and Williams [33–35]. At first, space is divided into MT regions and interstitial regions. Then electron density is represented by three components as $n = \{n_0(\mathbf{r}), \{n_{1,a}(\mathbf{r})\}, \{n_{2,a}(\mathbf{r})\}\}$, where a is the index of atomic sites in the primitive cell. Following Ref. [30], this is simply expressed as $n = n_0 \oplus n_1 \ominus n_2$. The zeroth component $n_0(\mathbf{r})$ is the spatially smooth functions expanded in analytic functions, i.e., plane waves, Gaussians, and smooth Hankel functions [29]. The first components $n_{1,a}(\mathbf{r})$ are the true electron density within MT at \mathbf{R}_a . The second components $n_{2,a}(\mathbf{r})$ are the counterpart, i.e., the projection of $n_0(\mathbf{r})$ into the MT at \mathbf{R}_a . $n_0(\mathbf{r})$ and $n_{2,a}(\mathbf{r})$ are identical within MT at \mathbf{R}_a up to a given angular momentum cutoff in their spherical harmonics expansion.

We can obtain the charge density $n^{\text{Zcv}} = n_0^{\text{Zcv}} \oplus n_1^{\text{Zcv}} \ominus n_2^{\text{Zcv}}$ by adding the ion-core contribution to n . Then we apply the multipole transformation clearly defined in Ref. [30], yielding $\bar{n}_0^{\text{Zcv}} \oplus \bar{n}_1^{\text{Zcv}} \ominus \bar{n}_2^{\text{Zcv}}$ as shown in Eqs. (28)–(30) in Ref. [30]. The transformation results in $\bar{n}_0^{\text{Zcv}}(\mathbf{r})$, $\bar{n}_{1,a}^{\text{Zcv}}(\mathbf{r})$, and $\bar{n}_{2,a}^{\text{Zcv}}(\mathbf{r})$ all having the same multipole in each MT site at \mathbf{R}_a , although the physically observable density remains unchanged. The first components $\bar{n}_{1,a}^{\text{Zcv}}(\mathbf{r})$, which are unchanged by the transformation, are the sum of the ion-core charge density and $n_{1,a}(\mathbf{r})$.

From the smooth density $\bar{n}_0^{\text{Zcv}}(\mathbf{r})$, we can calculate the electrostatic potential as $V_0^{\text{es}}(\mathbf{r}) = \int d^3r' \bar{v}(\mathbf{r}, \mathbf{r}') \bar{n}_0^{\text{Zcv}}(\mathbf{r}') + V^{\text{app}}(\mathbf{r})$. This gives a correct interstitial part of the potential $V_0^{\text{es}}(\mathbf{r})$ calculated from the charge density. The values of $V_0^{\text{es}}(\mathbf{r})$ at the MT boundaries are used to determine the electrostatic potential within MTs.

We can apply the usual procedure to determine the electrostatic potential within MTs. In each MT, we have the first and second components $\bar{n}_{1,a}^{\text{Zcv}}(\mathbf{r})$ and $\bar{n}_{2,a}^{\text{Zcv}}(\mathbf{r})$, which have the same multipole. With the condition that the electrostatic potential is zero at the MT boundary, we can calculate the potential

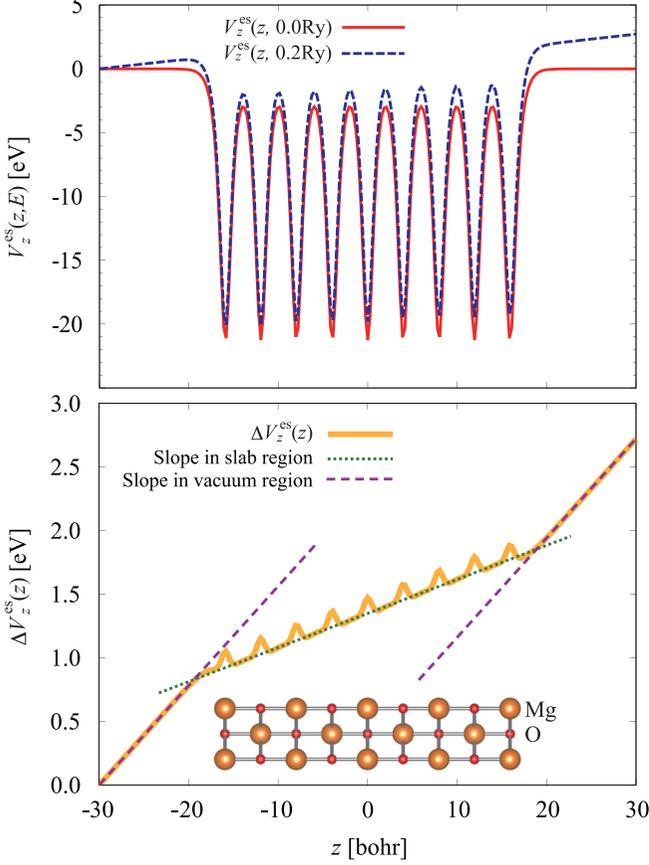


FIG. 2. A slab (18 atoms per cell) is placed in the middle of a supercell (60 a.u. width along the z axis, which is perpendicular to the slab), with electrodes at the left and right ends. In the top panel, we show $V_z^{es}(z, E)$ for $E = 0.2$ and 0.0 Ry. In the bottom panel, we show their difference $\Delta V_z^{es}(z)$. From the ratio of two slopes of ΔV_z^{es} in the slab region (green) and in the vacuum region (violet), we obtain $\epsilon_\infty(\text{Slab})$. We achieve better numerical accuracy by using $\Delta V_z^{es}(z)$ instead of $V_z^{es}(z, 0.2 \text{ Ry})$ directly.

generated by the difference between the first and second components.

Thus we finally have the electrostatic potential $V^{es}(\mathbf{r})$ represented in the three-component formalism. With this potential, we can perform self-consistent calculations for the slab models.

IV. RESULTS

A. Optical dielectric constants via the slab model

In Fig. 2, we illustrate our treatments in the slab models for five NaCl-structure ionic materials, where we use $\pm z_0 = \pm 30$ a.u. We use slabs made of nine layers, with 18 atoms in the supercell. We use experimental lattice constants of bulk materials without relaxation of atomic positions. The electrostatic potential $V_z^{es}(z, E)$ is the average of $V_0^{es}(\mathbf{r})$ in the xy plane under the bias voltage E . We plot the cases of $E = 0.2$ and 0.0 Ry. We show $\Delta V_z^{es}(z) = V_z^{es}(z, 0.2 \text{ Ry}) - V_z^{es}(z, 0.0 \text{ Ry})$ in the bottom panel in Fig. 2. $\Delta V_z^{es}(z, E)$ changes linearly as a function of z in the vacuum region and

TABLE I. Calculated optical dielectric constant ϵ_∞ . Values in the column labeled “RPA” are in bulk calculations with local field correction (LFC). Values in the column labeled “RPA(noLFC)” are without LFC. Values in the column labeled “Slab” are calculated from the slab models in the setting of Fig. 2. Ratios $\eta = \frac{\epsilon_\infty(\text{RPA})}{\epsilon_\infty(\text{Slab})}$ and $\gamma = \frac{\epsilon_\infty(\text{Slab})-1}{\epsilon_\infty(\text{RPA})-1}$ are calculated just simply from the values of $\epsilon(\text{QSGW, RRA})$ and $\epsilon(\text{QSGW, Slab})$.

		RPA				Experiments	
		(noLFC)	RPA	Slab	η	γ	[36–38]
LiF	GGA	2.04	1.95	2.01	0.86	1.40	1.96
	QSGW	1.73	1.67	1.94			
KF	GGA	2.16	1.96	1.94	0.90	1.26	1.85
	QSGW	1.79	1.68	1.86			
NaCl	GGA	2.70	2.33	2.42	0.83	1.42	2.34
	QSGW	2.13	1.92	2.31			
MgO	GGA	3.17	2.96	3.09	0.81	1.39	2.96
	QSGW	2.50	2.37	2.91			
CaO	GGA	3.94	3.59	3.68	0.81	1.38	3.33
	QSGW	2.88	2.68	3.31			

in the slab region. From the ratio of two slopes of ΔV_z^{es} in the slab region and in the vacuum region, we obtain $\epsilon_\infty(\text{Slab})$.

Our main results are ϵ_∞ calculated from slab models in QSGW, $\epsilon_\infty(\text{QSGW, Slab})$, shown in Table I. Note that $\epsilon_\infty(\text{Slab})$ contains the effect of vertex corrections based on QbP (see Sec. II), because changes in the self-energy caused by the bias E are self-consistently taken into account. The numerical reliability of our calculations is estimated to be $\lesssim 1\%$. See the Supplemental Material for computational details [39]. In Table I, we also show bulk values $\epsilon_\infty(\text{RPA})$. To obtain them, we first perform self-consistent calculations in QSGW for bulk materials. Then we calculate ϵ_∞ in the random-phase approximation (RPA) with/without local field correction (LFC). We also show ϵ_∞ in GGA together.

The QSGW values are in good agreement with experiments. For example, $\epsilon_\infty(\text{QSGW, Slab}) = 1.94$ for LiF is in surprisingly good agreement with $\epsilon_\infty(\text{Experiment}) = 1.96$. In contrast, $\epsilon_\infty(\text{QSGW, RPA}) = 1.67$ is much smaller than $\epsilon_\infty(\text{QSGW, Slab}) = 1.94$. These are generally true in all other materials. We see that the ratios $\eta = \epsilon_\infty(\text{QSGW, RPA})/\epsilon_\infty(\text{QSGW, Slab})$ in Table I are ~ 0.8 . This is consistent with Ref. [3], where ϵ_∞ for ZnO, Cu₂O, MnO, and NiO are presented. From the standpoint of estimating the enhancement factors (\approx vertex Γ) of the proper polarization, we may consider the ratios $\gamma = \frac{\epsilon_\infty(\text{Slab})-1}{\epsilon_\infty(\text{RPA})-1}$. As shown in Table I, $\gamma \sim 1.4$. Since $\epsilon_\infty(\text{QSGW, Slab})$ is in very good agreement with $\epsilon_\infty(\text{experiment})$, we can say that the vertex correction for bulk should illustrate the difference between $\epsilon_\infty(\text{RPA})$ and $\epsilon_\infty(\text{experiment})$ very well, where the vertex correction is calculated at the level of the functional derivative of the self-energy in QSGW; see Sec. III.

This is in contrast to the case of GGA. For example, consider the case of LiF. The difference $\epsilon_\infty(\text{GGA, Slab}) - \epsilon_\infty(\text{GGA, RPA}) = 2.01 - 1.95 = 0.06$ is very small. The difference originates from the xc kernel f_{xc} in the density functional perturbation theory. This is consistent with the

results in Ref. [40], where the authors explicitly evaluate f_{xc} in GGA for bulk materials. Note that $\epsilon_{\infty}(\text{GGA, Slab}) = 2.01$ is slightly larger than $\epsilon_{\infty}(\text{experiment}) = 1.96$; this is true for all other materials. We can see that the contributions of vertex corrections f_{xc} do not necessarily improve agreement; $\epsilon_{\infty}(\text{GGA, Slab})$ yields poorer agreement with $\epsilon_{\infty}(\text{experiment})$ than $\epsilon_{\infty}(\text{GGA, RPA})$.

B. Rationale for QSGW80

Our result in Sec. IV A shows that the vertex correction should be included in the proper polarization P to obtain ϵ_{∞} in agreement with experiments. We have to use such P in the QSGW self-consistent cycle. Such improved QSGW self-consistency can be identified as a self-consistent method in the G_0W approximation on the basis of QbP. Shishkin *et al.* [12] provide a method based on this idea. However, the computational requirements of their method are too costly for application to a wide range of materials. In fact, although their method was applied to calculate ionization potentials in Ref. [45], it was not satisfactory because the calculations were performed using a combination of simple materials (bulk calculations) and supercell calculations in GGA. Furthermore, there is no information on how to treat transition-metal oxides such as LaMnO_3 in the method of Shishkin *et al.* We have to develop such an improved QSGW method that is applicable to a wide range of materials, and this method must have computational efficiency and theoretical validity.

With regard to efficiency, we can consider a hybridization method between QSGW and the density functional xc [14]. In QSGW80, which is a simple hybridization (80% QSGW + 20% GGA), we can see that it works well for a wide range of materials. Our present results support the QSGW80 method, which uses only 80% of the QSGW self-energy. We can identify QSGW80 as a simplification of the method in Ref. [12]. Reference [46] presents another approximation at the level of QSGW80 for the vertex correction in QSGW, resulting in similarly good agreement with experiments.

Let us examine how QSGW80 is justified for materials calculated here. This is shown by the fact that $\eta = \epsilon_{\infty}(\text{QSGW, RPA})/\epsilon_{\infty}(\text{QSGW, Slab})$ in Table I are approximately 80%, and they show *little material dependency*. Thus we expect that QSGW80 can mimic QSGW with the vertex corrections. The too large screened-exchange effect is reduced by a factor 0.8, including adding a 0.2 GGA term so as to keep the total size of the xc term. In Table II, we show band gaps in QSGW and QSGW80 for materials treated here. The band gaps are systematically too large in QSGW in comparison with experimental values [8], while QSGW80 is in rather better agreement with experimental values. In Ref. [8], we checked the performance of the QSGW80 for a range of materials. As in the case of Ref. [12], QSGW80 is theoretically reasonable in the sense that the band gaps are improved by using the corrected W . To go beyond QSGW80, we have to develop methods to take the vertex correction into W , as was done in Ref. [12] in a simple manner. Considering the fact that QSGW80 works well, as shown in Ref. [8], we may expect simple methods to represent the vertex correction

TABLE II. Calculated band gaps (eV) of bulk materials. In QSGW80, we show self-consistent results with the hybrid xc potential, 80% QSGW + 20% GGA. QSGW80nosc specifies one-shot calculations with the hybrid potentials after QSGW 100% self-consistent calculations. QSGW80nosc is slightly larger because it is not fully self-consistent under such an xc potential. See Ref. [8].

	Experiments [41–44]	QSGW	QSGW80	QSGW80nosc	GGA
LiF	13.6	16.04	14.53	14.85	9.52
KF	10.9	11.78	10.53	10.82	6.43
NaCl	8.6	9.51	8.55	8.76	5.37
MgO	7.77	8.86	7.91	8.10	4.86
CaO	7.1	7.45	6.57	6.74	3.69

by a scalar factor or by a limited number of parameters. The vertex correction can be relatively insensitive to materials, thus we expect that some simple method might be available.

V. SUMMARY

To clarify the importance of quasiparticle self-consistency in QSGW, we have explained the quasiparticle-based perturbation in Sec. II. We emphasized the importance of self-consistency in the G_0W approximation, and then we obtained the quasiparticles (independent particles) given by H_0 and the interaction between the quasiparticles given by W . In addition, the vertex correction in QbP was introduced.

We have performed QSGW calculations for slab models under electric field by means of the ESM method. The calculated ϵ_{∞} are in good agreement with experimental values. Compared with ϵ_{∞} in the bulk calculation in the RPA, we evaluated the size of the vertex corrections as a functional derivative of the static self-energy in QSGW. Our results on ϵ_{∞} validate the method by Shishkin, Marsman, and Kresse [12]. As a simplified substitution of their method, we examined the performance of QSGW80 [8] for materials treated here. The QSGW+ESM method developed for the calculations should be useful even for other purposes, such as bias-dependent spin susceptibility in material theory, as well as practical device applications and material designs.

To go beyond the usual QSGW, we should develop an improved QSGW method in the G_0W approximation, whereas we should use accurate W by including the vertex correction. We would then have virtually the best division of $H = H_0 + (H - H_0)$, where H_0 gives the optimum independent-particle picture.

ACKNOWLEDGMENTS

T.K. gratefully acknowledges the support from JSPS KAKENHI under Grant No. 17K05499. We are also grateful for the computing time provided by the Research Institute for Information Technology (Kyushu University). H.S. gratefully acknowledges the computing resources provided by the supercomputer system in RIKEN (HOKUSAI) and the supercomputer system in ISSP (sekirei).

- [1] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, *Phys. Rev. Lett.* **93**, 126406 (2004).
- [2] M. van Schilfgaarde, T. Kotani, and S. Faleev, *Phys. Rev. Lett.* **96**, 226402 (2006).
- [3] T. Kotani, M. van Schilfgaarde, and S. V. Faleev, *Phys. Rev. B* **76**, 165106 (2007).
- [4] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **124**, 219906 (2006).
- [5] F. Tran and P. Blaha, *Phys. Rev. Lett.* **102**, 226401 (2009).
- [6] J. He and C. Franchini, *Phys. Rev. B* **86**, 235117 (2012).
- [7] F. Bruneval and M. Gatti, in *First Principles Approaches to Spectroscopic Properties of Complex Materials*, edited by C. Di Valentin, S. Botti, and M. Cococcioni (Springer, Berlin, 2014), Vol. 347, pp. 99–135.
- [8] D. Deguchi, K. Sato, H. Kino, and T. Kotani, *Jpn. J. Appl. Phys.* **55**, 051201 (2016).
- [9] H. Okumura, K. Sato, and T. Kotani, *Phys. Rev. B* **100**, 054419 (2019).
- [10] Y. Lee, T. Kotani, and L. Ke, [arXiv:2002.12417](https://arxiv.org/abs/2002.12417) [*Phys. Rev. B* (to be published)].
- [11] C. Bhandari, M. van Schilfgaarde, T. Kotani, and W. R. L. Lambrecht, *Phys. Rev. Mater.* **2**, 013807 (2018).
- [12] M. Shishkin, M. Marsman, and G. Kresse, *Phys. Rev. Lett.* **99**, 246403 (2007).
- [13] F. Bruneval, F. Sottile, V. Olevano, R. Del Sole, and L. Reining, *Phys. Rev. Lett.* **94**, 186402 (2005).
- [14] A. N. Chantis, M. van Schilfgaarde, and T. Kotani, *Phys. Rev. Lett.* **96**, 086405 (2006).
- [15] J. Otsuka, T. Kato, H. Sakakibara, and T. Kotani, *Jpn. J. Appl. Phys.* **56**, 021201 (2017).
- [16] A. Sawamura, J. Otsuka, T. Kato, and T. Kotani, *J. Appl. Phys.* **121**, 235704 (2017).
- [17] M. Otani and O. Sugino, *Phys. Rev. B* **73**, 115407 (2006).
- [18] L. Hedin and S. Lundqvist, *Effects of Electron-Electron and Electron-Phonon Interactions on the One-Electron States of Solids* (Oxford University Press, New York, 1969), Vol. 12.
- [19] D. Pines and P. Nozieres, *The Theory of Quantum Liquid. Vol. I* (Benjamin, New York, 1966).
- [20] T. Kotani and M. van Schilfgaarde, *Phys. Rev. B* **81**, 125201 (2010).
- [21] A. L. Kutepov, *Phys. Rev. B* **94**, 155101 (2016).
- [22] A. L. Kutepov, *Phys. Rev. B* **95**, 195120 (2017).
- [23] F. Bechstedt, K. Tenelsen, B. Adolph, and R. Del Sole, *Phys. Rev. Lett.* **78**, 1528 (1997).
- [24] Y. Takada, *Mol. Phys.* **114**, 1041 (2016).
- [25] N. L. Nguyen, H. Ma, M. Govoni, F. Gygi, and G. Galli, *Phys. Rev. Lett.* **122**, 237402 (2019).
- [26] The ECALJ package is available from <https://github.com/tkotani/ecalj/>.
- [27] T. Kotani and M. van Schilfgaarde, *Phys. Rev. B* **81**, 125117 (2010).
- [28] T. Kotani and H. Kino, *J. Phys. Soc. Jpn.* **82**, 124714 (2013).
- [29] T. Kotani, *J. Phys. Soc. Jpn.* **83**, 094711 (2014).
- [30] T. Kotani, H. Kino, and H. Akai, *J. Phys. Soc. Jpn.* **84**, 034702 (2015).
- [31] D. Singh, *Phys. Rev. B* **43**, 6388 (1991).
- [32] P. E. Blochl, *Phys. Rev. B* **50**, 17953 (1994).
- [33] J. M. Soler and A. R. Williams, *Phys. Rev. B* **40**, 1560 (1989).
- [34] J. M. Soler and A. R. Williams, *Phys. Rev. B* **42**, 9728 (1990).
- [35] J. M. Soler and A. R. Williams, *Phys. Rev. B* **47**, 6784 (1993).
- [36] N. Ashcroft, N. Mermin, and D. Wei, *Solid State Physics: Revised Edition* (Cengage Learning, Asia, 2016).
- [37] M. J. L. Sangster and A. M. Stoneham, *Philos. Mag. B* **43**, 597 (1981).
- [38] J. Jacobson and E. Nixon, *J. Phys. Chem. Solids* **29**, 967 (1968).
- [39] See Supplemental Material at <https://github.com/hskkbr/supplement4esmpaper> for computational detail and minimum inputs of LiF, KF, NaCl, MgO, and CaO.
- [40] 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).
- [41] D. Roessler and W. Walker, *J. Phys. Chem. Solids* **28**, 1507 (1967).
- [42] J. E. Eby, K. J. Teegarden, and D. B. Dutton, *Phys. Rev.* **116**, 1099 (1959).
- [43] D. M. Roessler and W. C. Walker, *Phys. Rev.* **159**, 733 (1967).
- [44] R. Whited, C. J. Flaten, and W. Walker, *Solid State Commun.* **13**, 1903 (1973).
- [45] A. Grüneis, G. Kresse, Y. Hinuma, and F. Oba, *Phys. Rev. Lett.* **112**, 096401 (2014).
- [46] W. Chen and A. Pasquarello, *Phys. Rev. B* **92**, 041115(R) (2015).