Fully self-consistent solution of the Dyson equation using a plane-wave basis set

Lin-Wang Wang

Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

(Received 5 July 2014; revised manuscript received 10 March 2015; published 23 March 2015)

Self-consistent solutions of the Dyson equation are obtained using a plane-wave basis set for seven small molecules. Such self-consistent solutions can help to unify the different GW self-consistent schemes, reduce the scatter of results in current GW calculations, and shed light on the true effects of GW self-consistency. Unlike other works of self-consistent GW calculations, in the present work the Green's function is expressed as a matrix under the plane-wave basis set. The algorithmic details which enable such calculations are presented. The ability to solve the full Green's function using a plane-wave basis set may open the door for future beyond-GW many-body perturbation theory calculations.

DOI: 10.1103/PhysRevB.91.125135 PACS number(s): 71.15.−m, 31.15.xm

I. INTRODUCTION

The GW method has been used as one of the most accurate methods to calculate the electronic structures of materials from bulk crystals to molecules [1]. However, there is a strong dependence of the traditional GW results on the initial input single-particle electron wave functions $\{\psi_i\}$ and eigenenergies $\{\epsilon_i\}$, especially for G_0W_0 calculations [2–5] (e.g., up to 1 eV for some oxide band gaps [6]), where the Green's function G and screened electron-electron interaction W are not updated. One way to solve this problem is to introduce self-consistency in the solution. However, there are different approaches to solve the GW problem self-consistently. One can divide these approaches into two major categories. In the first category, the Green's function is still described by an noninteracting Green's function using the eigenfunctions and eigenenergies of the single-particle orbital, e.g., $G_0(\omega) = \sum_i \psi_i(r_1)\psi_i^*(r_2)/(\omega + \mu - \epsilon_i \pm i\delta_i)$, although ψ_i and ϵ_i will be updated during the self-consistent iterations [7–11]. However, some of the self-consistence conditions could be a bit arbitrary since one can propose different self-consistent schemes [6]. Furthermore, in some cases, there can still be initial wave function dependence [12]. In the second category, on which we will focus, the Dyson equation is solved self-consistently, and the Green's function G can no longer be described by a noninteracting single-particle Green's function G_0. Schöne and Eguiluz [13] have solved the Dyson equation by expanding the Green's function G with the input single-particle orbital and used a truncation on the number of these orbitals. They used the Baym-Kadanoff formalism [14] to express the Dyson equation. Similarly, Kutepov, Savrasov, and Kotliar [15,16] also used the band states to expand the Green's function and iteratively solved the Dyson's equation in a bulk. The Matsubara frequency mesh [16] on the imaginary time and frequency axis is used for frequency integrations. Caruso, Scheffler *et al.* have used atomic orbitals to solve the Dyson equation [17,18]. So has the group of Thygesen [19,20]; they have solved the Dyson equation using localized atomic orbitals for 34 different molecules. The atomic orbital has also been used by van Leeuwen *et al.* [21,22] and Koval *et al.* [23] to solve the Dyson equation. Computationally, the atomic orbital has the advantage of being able to significantly reduce the dimensions of the problem while still being able to cover the important high-energy single-particle excitations. They are thus particularly suitable for isolated molecule systems, while the periodic crystals are traditionally solved with plane waves or full potential linearized augmented plane wave (FLAPW) methods. We notice that, for the methods starting with the plane-wave or FLAPW basis sets, the band states are often used to expand the Green's function G [13]. However, this could lead to issues related to the truncations of these band states [24,25]. In this work, we will use a plane-wave basis set to directly represent the Green's function matrix G without any truncation. Although doing so will significantly increase the computational cost, as we will show in this work, with the help of modern supercomputers, it is now possible to carry out such calculations. There could be other advantages for adopting this approach. For example, by representing the Green's function matrix in reciprocal and real space, the formalism becomes simpler. This might ease the step to adopt other formalisms beyond the GW approximation in the future.

One might ask why one should choose the Dyson equation as the self-consistent solution of the GW problem, given all the possible ways for the GW self-consistent calculations (i.e., an input equaling output criterion in an iterative procedure). Baym and Kadanoff [14,26] have shown that many conservation laws are preserved under the self-consistent solution of the Dyson equation. The same is true for the charge conservation law [27]. The Dyson equation is the variational minimum (or stationary point) solution of Klein's total energy expression [28] under the random phase approximation (RPA). This is like the Kohn-Sham equation is the variational minimum solution of the density functional theory (DFT) total energy. Furthermore, it has been shown that [29,30] under such a variational solution, the differences of the RPA total energies after adding or subtracting one electron equal the GW quasiparticle eigenenergies. Recently, there has been a surge of interest in using RPA for total energy calculations [29,31–33]. But many such calculations are based on the input (e.g., DFT) noninteracting single-particle Green's function G_0. The self-consistent solution of the Dyson equation is to find the electronic ground state of the total energy expression. As a result, one can, for example, use the Hellmann-Feynman theory to calculate the atomic force under the RPA total energy. Considering all these factors, it is not difficult to conclude that the Dyson equation as derived from the original GW formalism [34] is the most natural choice for the GW

self-consistent calculations. However, the cost is to represent the Green's function G as a full matrix. It can no longer be represented by a set of single-particle wave functions and eigenenergies.

The effects of self-consistency for GW calculations for homogeneous electron gas have been studied by Holm and Barth [35]. They found an overestimation of the free-electron bandwidth and the disappearance of the plasmon satellite structures in the spectral function due to self-consistent calculations. They thus concluded that the non-self-consistent G_0W_0 calculation is preferred unless the vortex correction is included, despite the fact that the self-consistent GW total energy is found to be rather accurate. Nevertheless, their conclusion is based on model metallic systems. It is thus interesting to test the self-consistent GW results on real and nonmetallic systems. Ku and Eguiluz [36] have calculated bulk Si and Ge using the self-consistent GW (sc-GW) method, and concluded that the self-consistency and core level should be used together as their effects can cancel out each other, although this conclusion has been contested by some later studies [24,25] due to the band state truncation issue. In terms of molecules, Caruso *et al.* [18] found that the self-consistency does not necessarily make the spectrum result worse than the non-self-consistent results. Thus, this could be the right time to revisit many of the related issues, especially if more accurate sc-GW calculations become available without the truncation issues.

We will use plane-wave pseudopotentials to solve the GW problem. There has been a long debate over the effects of ignoring the core levels and using the pseudo–valence wave functions in GW calculations. For example, Ku and Eguiluz [36] have claimed that the use of pseudopotentials will introduce a large error in the GW result. However, Delaney *et al.* [25] and Tiago *et al.* [24] have debated this result, and concluded that the pseudopotential should be good enough for most problems. More recently, Gomeze-Abal *et al.* [37] have revisited this problem. They concluded that the ignoring of core levels and the use of pseudo–wave functions might indeed have some effects on the final results. How to overcome these problems, e.g., by developing GW-appropriate pseudopotentials, will be beyond the scope of the current paper. For example, perhaps the effects of core level can be included in a core-level polarization model [38], and the effect of the pseudo–wave function might be corrected by introducing additional terms in the exchange integral [39,40]. Here, we would like to point out that for the light elements to be used in the current study, the pseudopotential introduced error should be small compared to the band gap, the highest occupied molecule orbital (HOMO), and the lowest unoccupied molecule orbital (LUMO) energies.

One technical issue for solving the Dyson equation under the GW approximation is the frequency-space integration. This is a convolution in frequency, and there are different ways to carry it out. One approach, recently adopted by Koval *et al.* [23], is to calculate this convolution integral directly in real frequency space together with the use of spectral functions to avoid the singularity of $G(\omega)$. We will use the Fourier transformation to convert the functions to time space, then use the direct products in time space. To avoid the singularities in real frequency, imaginary frequency will be used. This is the approach used by Roja, Godby, and Needs [41] more than ten years ago, and it has also been used recently by several other groups [15,18]. Very often, the Matsubara time and frequency mesh [15] is used with an artificial temperature. One often extrapolates the final result from a series of artificial temperatures [15,22]. Recently, in a work by Caruso *et al.* [18], the scheme of McMahan *et al.* [42] was adopted to carry out this frequency integration without the use of artificial temperature. In this paper, we will introduce an alternative integration scheme, which also has high accuracy without the use of artificial temperature.

We will calculate 7 small molecules: Si_3, C_3, O_3, Al_2, SiH_2, HNO, and CHF. These molecules are chosen because their electron affinities are positive (LUMO level below the vacuum level). Thus, both experimental HOMO and LUMO levels exist. In contrast, for many of the small molecules studied previously (e.g., in Ref. [19]), only HOMO levels exist. We believe it is interesting to have both HOMO and LUMO levels since they might have very different characteristics. Besides comparing the G_0W_0 calculation with the full sc-GW results, we will also study how the self-consistent Green's function G is different from the single-particle noninteracting Green's function G_0. There is also a practical question to answer: Can the full GW equation be solved using current-day supercomputers by expressing the Green's function matrix directly using a plane-wave basis set without any additional truncations beyond the plane-wave kinetic energy cutoff?

II. THE BASIC FORMALISM

We will follow the "space-time" method on the imaginary $i\omega$ axis first used by Roja, Godby, and Needs [41]. Under this scheme, the Green's function will be solved along the imaginary axis $i\omega+\mu$ [to be denoted as $G(i\omega)$] in the ω complex plane, where μ is the electron Fermi energy (thus in our notation, both ω and μ are real numbers). The Dyson equation can be written as

$$G^{-1}(i\omega) = i\omega + \mu - H_0 - \Sigma(i\omega), \tag{1}$$

where G, H_0, and Σ are all matrices either represented in the real space r index, or reciprocal space q index. $H_0 = -\frac{1}{2}\nabla^2 + V(r) + \sum_l |\phi_l\rangle\langle\phi_l|$ is the noninteracting single-electron Hamiltonian, with $|\phi_l\rangle\langle\phi_l|$ being the nonlocal pseudopotential projector. The single-particle potential is calculated as $V(r) = \sum_R v_{at}(r-R) + \int \frac{\rho(r')}{|r-r'|} d^3r'$, where v_{at} is the local part of the atomic pseudopotential, and the ρ is the electron charge density. The $\Sigma(i\omega)$ is the self-energy term. For the ω-dependent matrices $X(i\omega)$ (e.g., G and Σ), they can also be represented in the time (τ) space as $X(i\tau)$. The transformation between these temporal duel representations is

$$\begin{aligned} X(i\tau) &= \frac{i}{2\pi}\int_{-\infty}^{\infty} X(i\omega)e^{i\omega\tau}d\omega, \\ X(i\omega) &= -i\int_{-\infty}^{\infty} X(i\tau)e^{-i\omega\tau}d\tau. \end{aligned} \tag{2}$$

The above equation is carried out separately for every element of the matrix. On the other hand, the r and q space

transformation can be carried out as

$$X(q_1,q_2,z) = \frac{1}{\Omega}\int X(r_1,r_2,z)e^{iq_1r_1}e^{-iq_2r_2}d^3r_1d^3r_2,$$
$$X(r_1,r_2,z) = \frac{1}{\Omega}\sum_{q_1,q_2} X(q_1,q_2,z)e^{-iq_1r_1}e^{iq_2r_2}, \quad (3)$$

where z can be either $i\omega$ or $i\tau$, and Ω is the volume of the periodic unit cell.

With the above definition, the electron charge density can be calculated as $\rho(r) = -iG(r,r,i\tau)|_{\tau\to 0^+}$. The correct total electron charge can be obtained by adjusting the Fermi energy μ. Within the GW approximation [34], the self-energy term Σ can be expressed as

$$\Sigma(r_1,r_2,i\tau) = iW(r_1,r_2,i\tau)G(r_1,r_2,i\tau), \quad (4)$$

where the screened Coulomb interaction W can be calculated in (q,$i\omega$) space as

$$W(q_1,q_2,i\omega) = \frac{4\pi}{q_1^2}\epsilon^{-1}(q_1,q_2,i\omega), \quad (5)$$

and the dielectric function ϵ is calculated as

$$\epsilon(q_1,q_2,i\omega) = \delta_{q_1,q_2} - \chi(q_1,q_2,i\omega)\frac{4\pi}{q_2^2}. \quad (6)$$

Finally the polarizability χ can be calculated as

$$\chi(r_1,r_2,i\tau) = -iG(r_1,r_2,i\tau)G(r_2,r_1,-i\tau). \quad (7)$$

The above equations form a close loop for the calculation of G. The key is the transformation of the matrix between the real space $(r_1,r_2,i\tau)$ representation and the reciprocal space $(q_1,q_2,i\omega)$ representation, and matrix inversion [Eqs. (1), (5)] in reciprocal space representation for each $i\omega$. A plane-wave (PW) energy cutoff $E_{\rm cut}$ is used to select the PW vector q_1, q_2 in $G(q_1,q_2,z)$ and $\Sigma(q_1,q_2,z)$. However, in the expressions of $W(q_1,q_2,z)$, $\chi(q_1,q_2,z)$, and $\epsilon(q_1,q_2,z)$, their q_1, q_2 are defined within an energy cutoff $E_{\rm cut2} = 4E_{\rm cut}$. This is because these matrices are the squares of the Green's function G [Eq. (7)]. In practice, a smaller $E_{\rm cut2}$ can often be used, much like in a conventional plane-wave DFT calculation. In the following, we will introduce numerical methods to calculate the above equations.

III. NUMERICAL TECHNIQUES

We first introduce an algorithm to numerically transform $G(q_1,q_2,i\omega)$ obtained from Eq. (1) to $G(q_1,q_2,i\tau)$ via Eq. (2). We will set up an exponential grid to discretize ω: $\omega_k = \mathrm{sgn}(k)\alpha_1(\beta_1^{|k|} - 1)$ for $k = -200$, 200, and the maximum ω equals to 3×10^6 hartrees, while the smallest interval $\omega_1 - \omega_0$ equals 2×10^{-4} hartrees. A similar exponential grid τ_k is used for τ, with $k = -20$, 20, and maximum τ is 200 hartree^{-1}, while the smallest interval $\tau_1 - \tau_0$ is 0.01 hartree^{-1}. The grid convergence has been tested to ensure that the resulting error in quasiparticle eigenenergy is less than 0.01 eV.

To carry out the Fourier transformation of Eq. (2) from $G(i\omega)$ to $G(i\tau)$, the $\int d\omega$ is carried out piecewise analytically for each interval $[\omega_k,\omega_{k+1}]$. To do this, for each matrix element (q_1,q_2) within the interval $[\omega_k,\omega_{k+1}]$, we first fit the $G(q_1,q_2,i\omega)$ using an expression $f_1(i\omega) = \frac{C_1}{i\omega - Z_1} + \frac{C_2}{i\omega - Z_2}$. Assuming $G_1 = G(q_1,q_2,i\omega_1)$, $G_2 = G(q_1,q_2,i\omega_2)$, and $G_3 = G(q_1,q_2,i\omega_3)$ for a given (q_1,q_2) element, where $\omega_1 = \omega_{k-1}, \omega_2 = \omega_k, \omega_3 = \omega_{k+1}$, then the complex C_1, C_2 and real Z_1, Z_2 can be obtained through the following equations:

$$I_1 = -(\omega_3-\omega_2)\omega_1^2G_1 - (\omega_1-\omega_3)\omega_2^2G_2 - (\omega_2-\omega_1)\omega_3^2G_3,$$
$$I_2 = i[-(\omega_3-\omega_2)\omega_1G_1 - (\omega_1-\omega_3)\omega_2G_2 - (\omega_2-\omega_1)\omega_3G_3],$$
$$I_3 = (\omega_3-\omega_2)G_1 + (\omega_1-\omega_3)G_2 + (\omega_2-\omega_1)G_3, \quad (8)$$

and

$$x = \frac{\mathrm{Re}(I_1)\mathrm{Im}(I_3) - \mathrm{Re}(I_3)\mathrm{Im}(I_1)}{\mathrm{Re}(I_2)\mathrm{Im}(I_3) - \mathrm{Re}(I_3)\mathrm{Im}(I_2)},$$
$$y = \frac{\mathrm{Re}(I_2)\mathrm{Im}(I_1) - \mathrm{Re}(I_1)\mathrm{Im}(I_2)}{\mathrm{Re}(I_3)\mathrm{Im}(I_2) - \mathrm{Re}(I_2)\mathrm{Im}(I_3)}. \quad (9)$$

Then the real Z_1 and Z_2 can be obtained from the following formulas:

$$Z_1 = \frac{x}{2} + \sqrt{\frac{x^2}{2} - y},$$
$$Z_2 = \frac{x}{2} - \sqrt{\frac{x^2}{2} - y}. \quad (10)$$

After Z_1, and Z_2 are obtained, the C_1 and C_2 can be obtained by simple linear equations using the corresponding analytical expressions of G_1 and G_2. After the Z_1, Z_2, C_1, C_2 are obtained, to carry out the integral within $[\omega_k,\omega_{k+1}]$, we need an analytical expression for

$$I = \int_{\omega_2}^{\omega_3}\frac{e^{i\omega\tau}}{i\omega - Z}d\omega. \quad (11)$$

We first define $t_1 = Z\tau - i\omega_2\tau$, $t_2 = Z\tau - i\omega_3\tau$. Then for large $|t_1|$ and $|t_2|$, say bigger than 8, one can use

$$I = -ie^{Z\tau}\left\{\frac{e^{-t_1}}{t_1}\sum_{n=0}^{N_1}\frac{n!}{(-t_1)^n} - \frac{e^{-t_2}}{t_2}\sum_{n=0}^{N_2}\frac{n!}{(-t_2)^2}\right\}. \quad (12)$$

Here N_1, N_2 are the order of terms which make the minima of $\frac{n!}{(-t_1)^n}$ and $\frac{n!}{(-t_2)^n}$, respectively. For $|t_1| < 8$ and $|t_2| < 8$, one can use

$$I = -ie^{Z\tau}[\ln(t_1/t_2) + \mathrm{Ein}(t_2) - \mathrm{Ein}(t_1)], \quad (13)$$

and the Ein(t) function is

$$\mathrm{Ein}(t) = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}t^k}{k\times k!}. \quad (14)$$

Using the above formulas, one can get the analytical ω integral within an interval $[\omega_1,\omega_2]$ for the $G(i\omega)$ to $G(i\tau)$ Fourier transform.

To test this integration scheme, we have used the noninteracting G_0 for the Si_3 molecule with LDA $\{\psi_i,\epsilon_i\}$, where its analytical expression in both $i\omega$ and $i\tau$ spaces are known. More specifically, the $G_0(q_1,q_2,i\omega)$ can be written down as

$$G_0(q_1,q_2,i\omega) = \sum_i\frac{\psi_i(q_1)\psi_i^*(q_2)}{i\omega - (\epsilon_i - \mu)}, \quad (15)$$

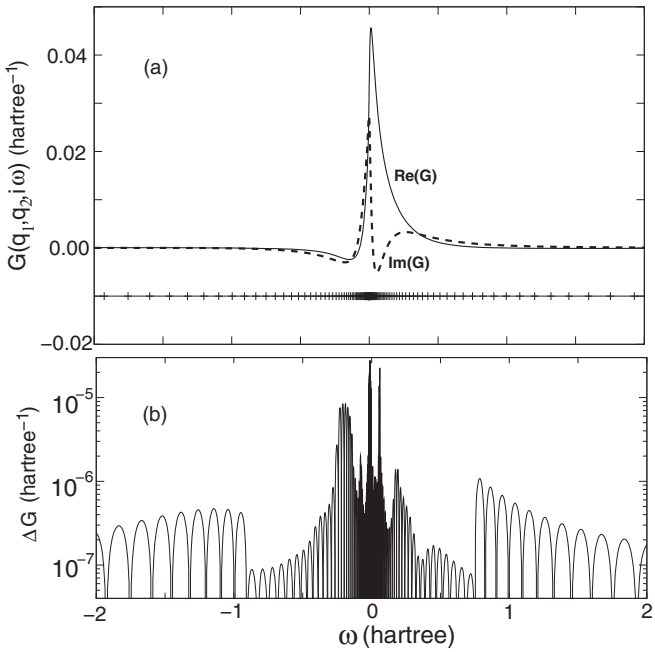

FIG. 1. The $G_0(q_1,q_2,i\omega)$ for an typical off diagonal (q_1,q_2) element (a), and (b) its fitting error within each interval $[\omega_k,\omega_{k+1}]$ using the analytical formula with coefficients obtained from Eqs. (8)–(10). The horizontal line with ticks in (a) indicates the ω_k positions.

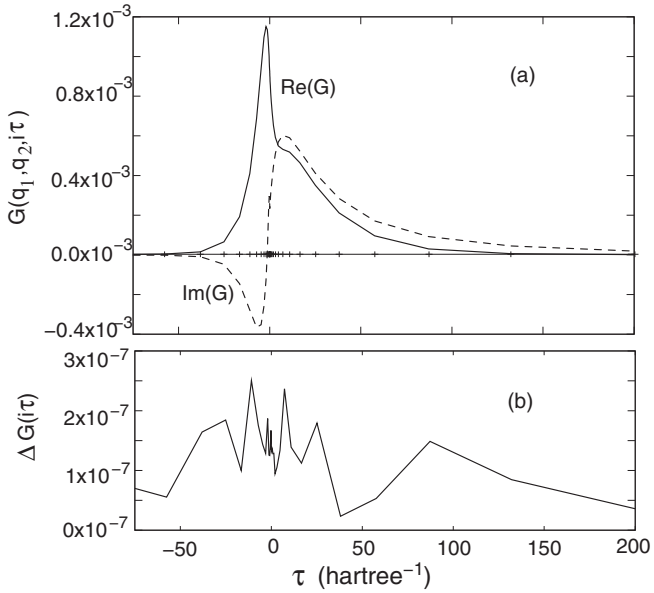

FIG. 2. The numerically integrated $G_0(q_1,q_2,i\tau)$ from the $G_0(q_1,q_2,i\omega)$ in Fig. 1(a); the error of the numerically integrated $G_0(q_1,q_2,i\tau)$ when compared with the analytical expression of Eq. (16) (b). The ticks on the horizontal axis in (a) indicate the τ_k positions.

while its analytical expression in $i\tau$ space is

$$G_0(q_1,q_2,i\tau) = i\sum_{i,\epsilon_i<\mu}\psi_i(q_1)\psi_i^*(q_2)e^{\tau(\epsilon_i-\mu)}, \quad \text{for} \quad \tau>0,$$
$$= -i\sum_{i,\epsilon_i>\mu}\psi_i(q_1)\psi_i^*(q_2)e^{\tau(\epsilon_i-\mu)}, \quad \text{for} \quad \tau<0. \tag{16}$$

The $G_0(q_1,q_2,i\omega)$ for a typical off-diagonal element ($|q_1|=1.43$ a.u., $|q_2|=1.66$ a.u., and q_1, q_2 are not in the same direction) is illustrated in Fig. 1(a) together with its grid points ω_k, while the error of $f_1(i\omega)$ fitting is shown in Fig. 1(b). We have also calculated the average relative fitting error using the formula $\{\sum_{\omega,q_1,q_2}|G_0(q_1,q_2,i\omega)-G_{\rm fit}(q_1,q_2,i\omega)|^2/\sum_{\omega,q_1,q_2}|G_0(q_1,q_2,i\omega)|^2\}^{1/2}$, where $G_{\rm fit}$ is the fitted G_0 using $f_1(i\omega)=\frac{C_1}{i\omega-Z_1}+\frac{C_2}{i\omega-Z_2}$, and ω is taken at the center of the interval $[\omega_k,\omega_{k+1}]$ where the error is maximum as shown in Fig. 1. The average error so calculated for the Si_3 system is 5.2×10^{-5}, similar to the one shown in Fig. 1(b). The numerically transformed $G_0(q_1,q_2,i\tau)$ together with its grid points τ_k are shown in Fig. 2(a), and its error when compared with the analytical expression Eq. (16) is shown in Fig. 2(b). As we can see, the errors are rather small, typically 10^{-4} times smaller than its absolute values.

To carry out the ω to τ Fourier transformation, massively parallel processing can be used to distribute the q_2 into different computer processors. Typically, we have the number of processors in the same order as the number of q_2. After $G(q_1,q_2,i\tau)$ is obtained, it is Fourier-transformed into $G(r_1,r_2,i\tau)$ using fast Fourier transformation (FFT) one τ_k at a time, so there is no need to store the full $G(r_1,r_2,i\tau)$ for all τ.

After $G(r_1,r_2,i\tau)$ is obtained for each τ_k, the $\chi(r_1,r_2,i\tau_k)$ for this τ_k is calculated using Eq. (7), and it is followed by a $(r_1,r_2,i\tau_k)$ to $(q_1,q_2,i\tau_k)$ transformation to store $\chi(q_1,q_2,i\tau_k)$ for all τ_k. After the values for all τ_k are obtained, a time Fourier transformation for all elements (q_1,q_2) is carried out to obtain $\chi(q_1,q_2,i\omega)$. Note, in the $i\omega$ representations of $\chi(i\omega)$, $\epsilon(i\omega)$, and $W(i\omega)$, we have used a different ω grid ω'_k with $k=-50,50$ points, and $\omega'_{50}=200$ hartrees, and $\omega'_1-\omega'_0=2\times10^{-3}$ hartrees. This is allowed since these functions decay much faster than that of $G(i\omega)$. For each matrix element (q_1,q_2), in transforming $\chi(i\tau)$ to $\chi(i\omega)$ we have represented $\chi(i\tau)$ within the interval $[\tau_k,\tau_{k+1}]$ with an analytical expression of $f_2(\tau)=[C_3+C_4(\tau-\tau_k)]\exp[-\beta_2(\tau-\tau_k)]$. An analytical expression is then used to represent the interval integral $-i\int_{\tau_k}^{\tau_{k+1}}f_2(\tau)\exp(-i\omega\tau)d\tau$ of Eq. (2). The resulting $\chi(q_1,q_2,i\omega)$ is used in Eq. (6) and Eq. (5) to get $W(q_1,q_2,i\omega)$. The inversion of the matrix $\epsilon(q_1,q_2,i\omega)$ for each $i\omega'_k$ value is done with SCALAPACK with additional parallelization on the k index.

The $W(q_1,q_2,i\omega)$ is then transformed into $W(q_1,q_2,i\tau)$. For this Fourier transformation, a linear expression $C_5+C_6(\omega-\omega'_k)$ is used to represent $W(i\omega)$ within interval $[\omega'_{k+1},\omega'_k]$. With $W(q_1,q_2,i\tau)$ obtained, we can now transform it to $W(r_1,r_2,i\tau_k)$ at each τ_k point, together with $G(r_1,r_2,i\tau_k)$, to get $\Sigma(r_1,r_2,i\tau_k)$ according to Eq. (4). After $\Sigma(r_1,r_2,i\tau_k)$ for each τ_k is calculated, it is immediately transformed and stored as $\Sigma(q_1,q_2,i\tau_k)$. After $\Sigma(q_1,q_2,i\tau_k)$ for all τ_k are calculated, it is Fourier-transformed to $\Sigma(q_1,q_2,i\omega_k)$ to be used in Eq. (1). In this transformation, the $\Sigma(q_1,q_2,i\tau)$ within $[\tau_k,\tau_{k+1}]$ is represented as $f_3(\tau)=C_7+C_8(\tau-\tau_k)$. Through the iterations over the loops of Eqs. (1) to (7), the potential $V(r)$ is updated with a Kerker potential mixing [43].

IV. RESULTS AND DISCUSSION

The molecules studied are listed in Table I, along with their atomic positions. These atomic positions are obtained starting from the literature-reported experimental atomic positions, followed by density functional theory (DFT) atomic relaxations using the generalized gradient approximation (GGA) PBE functional.

To calculate these systems with our sc-GW method, the molecules are placed in periodic boxes about 10 to 12 Å in size. To avoid dipole-dipole interactions for the Coulomb interaction and exchange integrals from the periodic imaging molecules, a special technique is used to truncate the range of the unscreened electron-electron interaction $1/|r-r'|$. This truncated interaction is then Fourier-transformed back to q space as $v(q)$. This $v(q)$ is then used to replace the $4\pi/q^2$ in Eqs. (5) and (6), which can avoid the possible correlation effects between periodic imaging molecules (e.g., van der Waals interactions). The Poisson equation $\int \frac{\rho(r')}{|r-r'|}d^3r'$ is solved using a double-box technique where a box twice the size of the original box is used to place the $\rho(r')$ at the center, then a longer range truncated $1/|r-r'|$ can be used to calculate the Hartree potential via an FFT technique. The convergence of this procedure is tested using LDA calculations with different box sizes to ensure the finite box size results are similar to the large box limit results. Note that we could also use postprocess extrapolation to get the converged results without using the above truncation techniques. In such a scheme, several box sizes will be used to carry out the calculations [44].

TABLE I. The atomic coordinations of the small molecules calculated in this work. The coordinates are relaxed using DFT PBE functional. The $E_{\rm cut}$ is the plane-wave cutoff energy used for the GW calculation. N_q is the resulting number of plane waves, while $N_r = n_1 n_2 n_3$ is the total number of grid points in real space, where n_1, n_2, n_3 are the real space grids of the supercell in 1, 2, 3 directions, respectively.

Systems	Atoms	x (Å)	y (Å)	z (Å)	$E_{\rm cut}$ (Ry)	N_q	N_r
Si_3	Si	0	0	0	30	6272	55296
	Si	0	2.817	0			
	Si	1.664	1.409	0			
C_3	C	0	0	0	50	7210	46080
	C	1.302	0	0			
	C	−1.302	0	0			
O_3	O	0	0	0	50	8484	57600
	O	−1.284	0.016	0			
	O	0.593	1.139	0			
Al_2	Al	0	0	0	30	6072	46080
	Al	2.838	0	0			
SiH_2	Si	0	0	0	30	6290	43200
	H	−1.412	−0.606	0			
	H	−0.586	1.420	0			
HNO	N	0	0	0	50	8484	57600
	O	1.215	−0.080	0			
	H	−0.409	−0.998	0			
CHF	C	0	0	0	50	8484	57600
	H	−1.009	−0.528	0			
	F	−0.373	1.278	0			

The plane-wave cutoff energies $E_{\rm cut}$ are listed in Table I, along with the resulting number of plane waves N_q and number of real space grid points N_r. We can see that N_q is around 5000 to 8000, while N_r can be about 57 000 (because plane-wave vectors are within a sphere defined by $E_{\rm cut}$ while N_r is defined by the full FFT grid points). Thus $G(q_1,q_2,i\omega)$ could be an 8000×8000 matrix for each $i\omega$ point, requiring 1 GB of memory for each $i\omega$ point. The number of real space grid points N_r is much larger. Hence we cannot store the real space matrix like $G(r_1,r_2,i\tau)$ for all $i\tau$ points. Instead, we only store the $G(r_1,r_2)$ matrix for one $i\tau$ as discussed above. This $G(r_1,r_2,i\tau)$ is obtained by spatial FFT from $G(q_1,q_2,i\tau)$ for the same $i\tau$. Note, we can store all the $G(q_1,q_2,i\tau)$ and $G(q_1,q_2,i\omega)$ in memory (for all $i\tau$ points and $i\omega$ points). Also note that, unlike the conventional GW calculations where only a limited number of conduction bands are used in the Green's function expression, here there is no such cutoff. The full matrix, hence in a sense, all the conduction band states, are used in the G expression. In our calculation, norm-conserving pseudopotentials are used. As discussed in the introduction, the pseudopotential-introduced GW error is relatively small [24,25,37]. This should be particularly true for our molecules consisting of mostly light elements where the semicore effect is small [39]. The calculations are carried out on the Titan supercomputer at the Oak Ridge Leadership Computing Facility using about 50 000 CPU processors. It takes about a few hours to finish one molecule calculation.

We first test the self-consistent convergence of the iterations. During the iterations, the single-electron potential is mixed with results from previous steps using the conventional Kerker mixing scheme [43], while the self-energy terms $\Sigma(q_1,q_2,i\omega)$ are used directly in the next iteration without any mixing. We found that, with the local density approximation (LDA) ψ_i, ϵ_i as the inputs, the iteration converges typically in 10 to 20 steps. Figure 3 shows the convergence of the Si_3, HNO, and Al_2 molecules. Note, to judge the convergence, we have used the eigenvalues of $H_0+\Sigma(i\omega=0)$. The eigenvalues ϵ_i'(iter) of all the occupied states plus one unoccupied state are used, where "iter" is the iteration number. The convergence shown in Fig. 3(a) is measured as $\sum_i |\epsilon_i'(\text{iter})-\epsilon_i'(\text{converged})|/N$, where N is the number of states i in the summation. We can see that the system converged in about 15 iterations. We have also measured this self-consistent convergence directly by the change of the Green's function as $\{\sum_{\omega,q_1,q_2}|G_{\rm iter}(q_1,q_2,i\omega)-G_{\rm iter-1}(q_1,q_2,i\omega)|^2/\sum_{\omega,q_1,q_2}|G_{\rm iter}(q_1,q_2,i\omega)|^2\}^{1/2}$. This is shown in Fig. 3(b). As we can see, the convergence measured this way has similar rate as it is measured by the eigenenergies. We also found that the final results are independent of the initial input wave functions and eigenenergies. Figure 3(c) shows the convergence between an initial LDA ψ_i, ϵ_i result and a initial Hartree-Fock (HF) ψ_i, ϵ_i result measured by $\sum_i|\epsilon_i'(\text{iter,LDA})-\epsilon_i'(\text{iter,HF})|/N$. Despite the large initial difference, they converge into the same final result after 15 iterations. Note, here, the initial HF result is first calculated using the same GW program with the screening in W turned off.

In the introduction, we asked how the full Green's function is different from the noninteracting Green's function of Eq. (15), which is used in many current self-consistent (category 1) or non-self-consistent GW calculations. The first

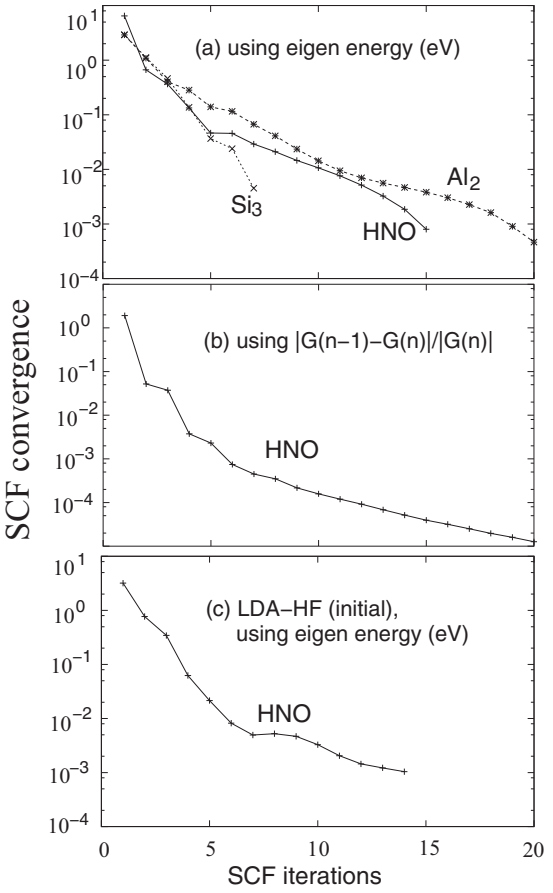

FIG. 3. (a) The self-consistent field (SCF) convergence of the *GW* calculations. The convergence is measured by the convergence of the eigenvalues of $H_0 + \Sigma(\omega = 0)$. See main text for more details. (b) The convergence using the difference of the Green's function between adjacent steps. (c) The convergence between the LDA initial wave functions and eigenenergies and the HF initial wave functions and eigenenergies.

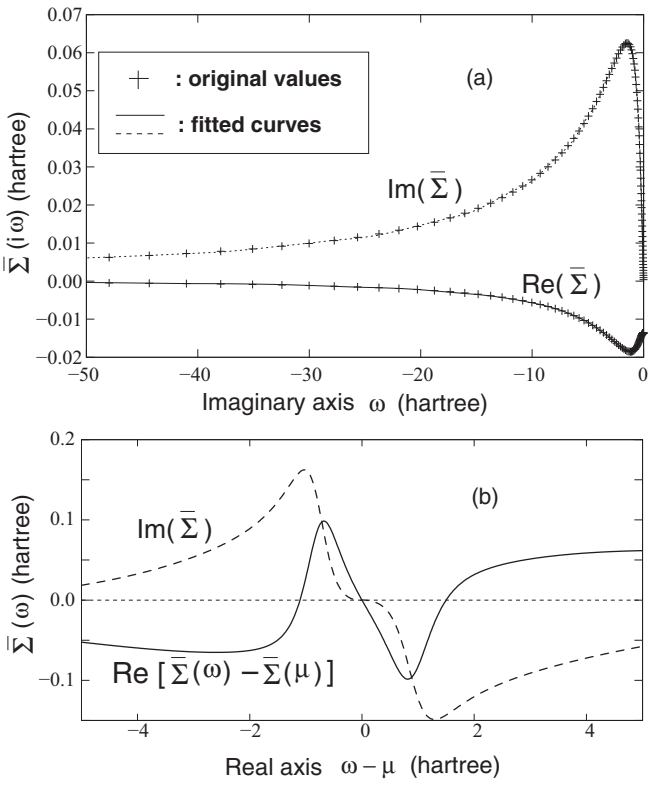

FIG. 4. The expectation value $\bar{\Sigma}(i\omega) = \langle\psi_i'(0)|\Sigma(i\omega)|\psi_i'(0)\rangle$ for $i = 6$ on the imaginary axis (a). The crosses are the directly calculated values, and the continued lines are the analytical fitting curve; the analytically extended $\bar{\Sigma}(\omega)$ on real ω axis (b).

difference is the non-Hermitian part of the self-energy term Σ, which is ignored in the construction of G_0 in Eq. (15). The non-Hermitian part of Σ can be as large as the Hermitian part, as indicated by the imaginary part of its expectation value shown in Fig. 4(a). Its amplitude can be around 1 eV. Besides, both its real and imaginary parts vary significantly with ω, while for the G_0 it is assumed to be a constant ϵ_i. Besides the eigenenergy ω dependence, the $H(i\omega) = H_0 + \Sigma(i\omega) = i\omega + \mu - G^{-1}(i\omega)$ also has an $i\omega$ dependence for its eigenvector wave functions, while in G_0 these eigen–wave functions are assumed to be $i\omega$ independent. To show how large this wave function ω dependence is, we have first hermitized $H(i\omega)$ as $H'(i\omega) = \frac{1}{2}[H(i\omega) + H^T(i\omega)]$, then diagonalized $H'(i\omega)$ to get $\psi_i'(i\omega)$ for different ω. To check whether $\psi_i'(i\omega)$ is changing with ω, we have calculated $d_i(i\omega) = 1 - \sum_j |\langle\psi_i'(i\omega)|\psi_j'(i\omega = 0)\rangle|^4$. The larger the change, the larger is the value of d_i (0 means no change). The $d_i(i\omega)$ for the first 12 states from the Si_3 molecule are shown in Fig. 5. As can be seen, for all the occupied states ($i \leqslant 6$), d_i is rather small. Only for some high conduction band states, d_i become large. This means the eigen–wave functions have minimum ω dependence. We thus conclude that while the eigenenergy ω dependence and the imaginary parts of the eigenenergies are both large, which can render Eq. (15) invalid, the wave function ω dependence is small. This suggests a direction for future approximations of the Green's functions if the full Green's function is not to be used.

We next study the quasiparticle eigenenergies. As shown in Fig. 5, for the HOMO and LUMO states, their $\psi_i'(i\omega)$ do not change significantly with ω in the imaginary axis. Thus we can also assume they do not change much along the real ω axis based on their analytic extension properties. As a result,

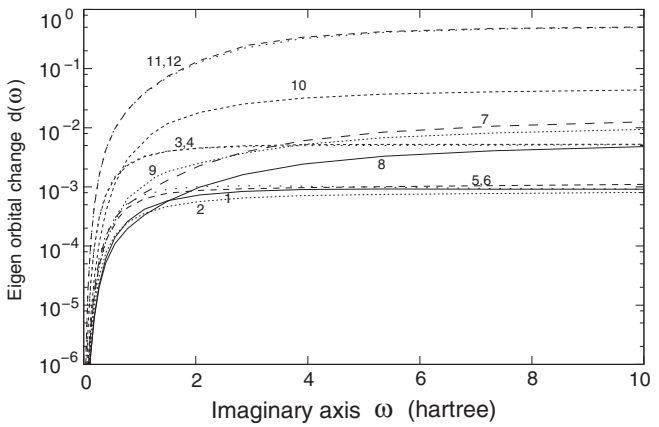

FIG. 5. The spread of the eigenvectors along the imaginary $i\omega$ axis. The $d_i(i\omega)$ is defined in the text. The numbers are the state index i's.

TABLE II. The quasiparticle lowest unoccupied molecule orbital (LUMO) and highest occupied molecule orbital (HOMO) energies, and comparison with the experimental EA and IE (in units of eV). The mean averaged error (MAE) (taken as the root-mean-square average of the errors) for the calculated values are measured against the experimental results. The experimental EA and IE are taken from Refs. [45,46]. For the G_0W_0 calculation, the input single-particle wave functions and eigenenergies are from LDA calculations.

System		LDA	G_0W_0	GW	Exp.
Si_3	LUMO	−4.75	−2.88	−3.01	−2.4
	HOMO	−5.39	−8.17	−7.72	−8.0
C_3	LUMO	−5.69	−2.76	−2.67	−2.0
	HOMO	−8.01	−11.46	−11.91	−13.0
O_3	LUMO	−6.44	−3.05	−3.46	−2.10
	HOMO	−8.14	−12.32	−12.63	−12.53
Al_2	LUMO	−3.72	−1.81	−1.20	−1.46
	HOMO	−4.07	−6.40	−5.93	−5.4
SiH_2	LUMO	−4.09	−1.54	−1.04	−1.12
	HOMO	−5.90	−9.50	−9.28	−8.92
HNO	LUMO	−4.70	−0.25	−0.29	−0.34
	HOMO	−5.62	−9.96	−10.40	−10.1
CHF	LUMO	−4.40	−0.19	−0.19	−0.54
	HOMO	−5.89	−10.22	−10.52	−10.06
MAE	err	3.64	0.65	0.59	

we can approximate $\psi_i'(i\omega)$ with $\psi_i'(0)$ [which is $\psi_i'(i\omega = 0)$]. To get the corresponding quasiparticle energy, we can find the ω solution of $\langle\psi_i'(0)|G^{-1}(\omega)|\psi_i'(0)\rangle = 0$ (the pole of G). This requires us to get $\bar{\Sigma}_i(\omega) = \langle\psi_i'(0)|\Sigma(\omega)|\psi_i'(0)\rangle$ on the real axis. From the above calculations, we can get $\bar{\Sigma}_i(i\omega)$ on the imaginary axis. To analytically extend it to the real axis, we have followed the procedure from Ref. [41]. An analytical expression $\sum_l C_l/(i\omega - Z_l)$ (with typically three l terms and complex C_l and Z_l) is used to fit the $\bar{\Sigma}_i(i\omega)$ on the imaginary axis, with the fitting accuracy shown in Fig. 4(a). Then the same analytical expression is used to obtain the values of $\bar{\Sigma}_i(\omega)$ on the real axis, as shown in Fig. 4(b). By testing different numbers of l, we found this procedure very reliable in obtaining $\bar{\Sigma}_i(\omega)$ on the real axis for ω within 1 or 2 hartrees from μ. The quasiparticle energy ϵ_i on the real axis can be obtained by solving the equation $\omega + \mu = \epsilon_i(0) + \mathrm{Re}[\bar{\Sigma}_i(\omega) - \bar{\Sigma}_i(0)]$; here $\epsilon_i(0)$ is the eigenenergy of $H'(0)$. The resulting ϵ_i can be compared with the experimental electron affinity (EA) and ionization energy (IE) [45,46] as shown in Table II. We see that the GW results agree well with the experimental results, but the GW root-mean-square (rms) error is similar to the G_0W_0 results. However, we do not see a systematic increase of the GW band gap compared to the G_0W_0 band gap. This is different from the conclusions based on self-consistent GW calculations by only updating the eigenenergies in G and W using the G_0 formalism of Eq. (15) [9,47–48]. Our result is similar to Ref. [19], where they found the mean average errors (MAE) of sc-GW, G_0W_0 (HF-initial), and G_0W_0 (PBE-initial) as 0.5, 0.4, and 0.5 eV, respectively (taken from the HOMO levels of 34 molecules). In our case, the MAE of sc-GW and G_0W_0 (LDA-initial) are 0.6 and 0.7 eV, respectively (taken from HOMO and LUMO levels of the 7 molecules studied).

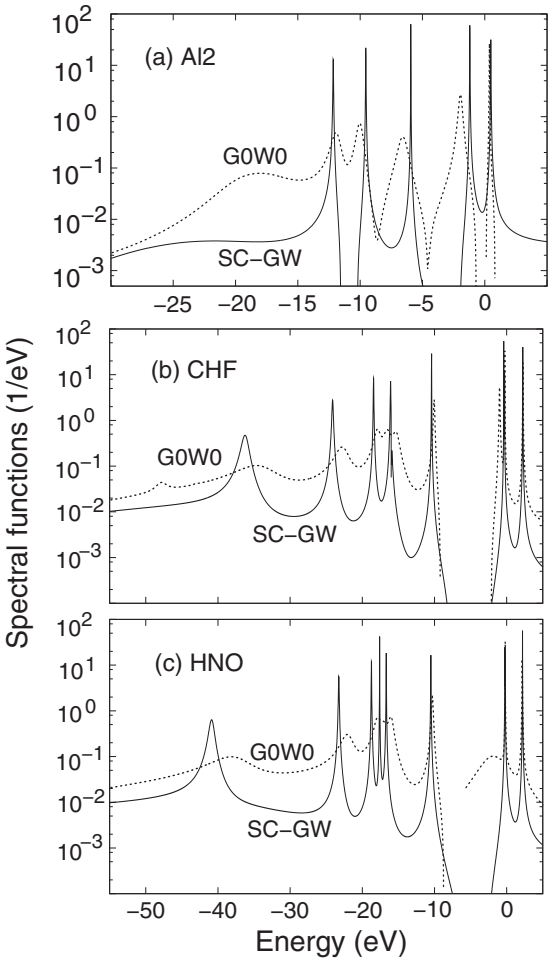

FIG. 6. The spectral functions as calculated from Eq. (19). The dotted lines are for G_0W_0 results starting with LDA input wave functions and eigenenergies, and the solid lines are for sc-GW results. The Fermi energy μ for Al_2, CHF, and HNO are around −3, −5, and −5 eV, respectively.

In the work of Holm and von Barth [35], it is the valence band bandwidth and spectral functions which are evaluated and compared to experiment, not the band gap. Here, we also calculate the spectral functions of the molecules. The spectral functions can be defined as

$$A(\omega) = 1/\pi\,\mathrm{Tr}[\mathrm{Im}G(\omega)]. \tag{17}$$

Here "Tr" is the trace of the matrix G. Thus, it can also be expressed as

$$A(\omega) = \frac{1}{\pi}\sum_i \mathrm{Im}\langle\psi_i'(0)|\frac{1}{\omega - \mu - H_0 - \Sigma(\omega)}|\psi_i'(0)\rangle. \tag{18}$$

If we further assume $\psi_i'(0)$ are a set of good eigenvectors in the energy range in which we are interested, then the expectation values under $\psi_i'(0)$ can be applied to the denominator, and the above equation can be approximated as

$$A(\omega) = \frac{1}{\pi}\sum_i \mathrm{Im}\left[\frac{1}{\omega - \mu - \epsilon_i(0) - \bar{\Sigma}_i(\omega) + \bar{\Sigma}_i(0)}\right]. \tag{19}$$

Here $\bar{\Sigma}_i(\omega) = \langle\psi_i'(0)|\Sigma(\omega)|\psi_i'(0)\rangle$ can be obtained from the above analytical extension from imaginary axis to real axis, as shown in Fig. 4. The resulting $A(\omega)$ for Al_2, CHF, and HNO are shown in Fig. 6. The fully self-consistent GW results for $A(\omega)$ are compared with the non-self-consistent G_0W_0 results. A few features are worth noticing. First, the widths of the valence bands, counted from the lowest occupied valence state to the highest occupied valence state, are slightly wider in sc-GW than G_0W_0 results, similarly to what is found in the homogeneous electron gas [35]. Unfortunately, for these systems, there are no experimental data for comparison. Second, in the G_0W_0 result, one can indeed see some satellite peaks (e.g., the peak around -20 eV in Al_2, and the small peak around -48 in CHF). Such satellite peaks disappear in sc-GW calculations. It will be interesting to see whether such satellite peaks are real in future experiments. Lastly, we notice that the $A(\omega)$ peaks in sc-GW are much sharper than the ones in G_0W_0. This conclusion is in contrast with the results found in the case of homogeneous electron gas [35], but is consistent with the results found in Ref. [18]. The sharpness of $A(\omega)$ peaks for molecules is expected, especially for HOMO and LUMO states. Note that the sc-GW spectral peak widths for deep levels (e.g, the ones near -40 eV in CHF and HNO) are significantly larger than the ones for HOMO. This is expected, as for such large excitation energies, there could be other excitation modes involving two shallow levels being excited to continuous spectrum of the vacuums. Hence different excitation modes can be mixed together, reducing the quasiparticle lifetime and increasing the peak width. To describe such excitation properly, one needs to describe all the continuous states accurately, which highlights the advantages of plane-wave basis sets.

V. CONCLUSIONS

We have demonstrated that, with the help of supercomputers, it is now possible to solve the Dyson equation self-consistently with a plane-wave basis set. The space-time scheme on the imaginary frequency axis is used. Some numerical techniques to solve this problem have been presented. In particular, a method is introduced to carry out the numerical Fourier transformation for the Green's function from frequency space to time space. The self-consistent iteration converges well within about 15 iterations, and different initial single-particle wave function and eigenenergy inputs result in the same final solution. Seven small molecules are calculated. We found that for their HOMO and LUMO quasiparticle eigenenergies, the sc-GW results have errors similar to those of the G_0W_0 results when compared to experiments. For the spectral functions, we found that sc-GW results have sharper peaks than the G_0W_0 results, while the satellite peaks in some of the G_0W_0 results disappear in the sc-GW results. The sc-GW gives slightly wider valence band width than the G_0W_0 results. For the self-energy term, we found that while its expectation values have large imaginary parts and frequency dependence, the eigenvectors of the Hamiltonian (after being hermitized) have much less frequency dependence. This perhaps points out a way for a future Green's function approximation that maintains the noninteracting Green's function form of Eq. (15), but with frequency-dependent and complex eigenenergies $\epsilon_i(\omega)$.

ACKNOWLEDGMENTS

We like to thank Prof. Rex Godby for helpful discussions. This work was supported by the Director, Office of Science (SC), Basic Energy Science (BES), Materials Science and Engineering Division (MSED), of the US Department of Energy (DOE) under Contract No. DE-AC02-05CH11231 through the Materials Theory program. It used resources of the National Energy Research Scientific Computing Center (NERSC) and Oak Ridge Leadership Computing Facility (OLCF) with the computational time allocated by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) project.

[1] W. G. Aulbur, L. Jonsson, and J. W. Wilkins, Solid State Phys. **54**, 1 (2000).
[2] P. Rinke, A. Qteish, J. Neugebauer, C. Freysoldt, and M. Scheffler, New J. Phys. **7**, 126 (2005).
[3] F. Fuchs, J. Furthmüller, F. Bechstedt, M. Shishkin, and G. Kresse, Phys. Rev. B **76**, 115109 (2007).
[4] T. Körzdörfer and N. Marom, Phys. Rev. B. **86**, 041110(R) (2012).
[5] N. Marom, F. Caruso, X. Ren, O. T. Hofmann, T. Körzdörfer, J. R. Chelikowsky, A. Rubio, M. Scheffler, and P. Rinke, Phys. Rev. B **86**, 245127 (2012).
[6] L. I. Bendavid and E. A. Carter, Top. Curr. Chem. **347**, 47 (2014).
[7] S. Lany, Phys. Rev. B **87**, 085112 (2013).
[8] H. Jiang, R. I. Gomez-Abal, P. Rinke, and M. Scheffler, Phys. Rev. Lett. **102**, 126403 (2009).
[9] M. Shishkin and G. Kresse, Phys. Rev. B **75**, 235102 (2007).
[10] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).
[11] R. Sakuma, T. Miyake, and F. Aryasetiawan, Phys. Rev. B **78**, 075106 (2008).
[12] P. Liao and E. A. Carter, Phys. Chem. Chem. Phys. **13**, 15189 (2011).
[13] W.-D. Schone and A. G. Eguiluz, Phys. Rev. Lett. **81**, 1662 (1998).
[14] G. Baym and L. P. Kadanoff, Phys. Rev. **124**, 287 (1961).
[15] A. Kutepov, S. Y. Savrasov, and G. Kotliar, Phys. Rev. B **80**, 041103 (2009).
[16] A. Kutepov, K. Haule, S. Y. Savrasov, and G. Kotliar, Phys. Rev. B **85**, 155129 (2012).
[17] F. Caruso, P. Rinke, X. Ren, M. Scheffler, and A. Rubio, Phys. Rev. B **86**, 081102 (2012).
[18] F. Caruso, P. Rinke, X. Ren, A. Rubio, and M. Scheffler, Phys. Rev. B **88**, 075105 (2013).
[19] C. Rostgaard, K. W. Jacobsen, and K. S. Thygesen, Phys. Rev. B **81**, 085103 (2010).
[20] M. Strange, C. Rostgaard, H. Hakkinen, and K. S. Thygesen, Phys. Rev. B **83**, 115108 (2011).

[21] N. E. Dahlen and R. van Leeuwen, J. Chem. Phys. **122**, 164102 (2005).
[22] A. Stan, N. E. Dahlen, and R. van Leeuwen, Europhys. Lett. **76**, 298 (2006).
[23] P. Koval, D. Foerster, and D. Sánchez-Portal, Phys. Rev. B **89**, 155417 (2014).
[24] M. L. Tiago, S. Ismail-Beigi, and S. G. Louie, Phys. Rev. B **69**, 125212 (2004).
[25] K. Delaney, P. Garcia-Gonzalez, A. Rubio, P. Rinke, and R. W. Godby, Phys. Rev. Lett. **93**, 249701 (2004).
[26] G. Baym, Phys. Rev. **127**, 1391 (1962).
[27] A. Schindlmayr, Phys. Rev. B **56**, 3528 (1997).
[28] A. Klein, Phys. Rev. **121**, 950 (1961).
[29] Y. M. Niquet, M. Fuchs, and X. Gonze, Phys. Rev. A **68**, 032507 (2003).
[30] L. W. Wang, Phys. Rev. B **82**, 115111 (2010).
[31] J. Harl and G. Kresse, Phys. Rev. Lett. **103**, 056401 (2009).
[32] J. Jung, P. Garcia-Gonzalez, J. F. Dobson, and R. W. Godby, Phys. Rev. B **70**, 205107 (2004).
[33] D. Pines, *Elementary Excitations in Solids* (Benjamin/Cummings, London, 1963).
[34] L. Hedin and S. Lundqvist, Solid State Phys. **23**, 1 (1969).
[35] B. Holm and U. von Barth, Phys. Rev. B **57**, 2108 (1998).
[36] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).
[37] R. Gomez-Abal, X. Li, M. Scheffler, and C. Ambrosch-Draxl, Phys. Rev. Lett. **101**, 106404 (2008).
[38] E. L. Shirley, X. Zhu, and S. G. Louie, Phys. Rev. B **56**, 6648 (1997).
[39] B. Lee, L.-W. Wang, C. D. Spataru, and S. G. Louie, Phys. Rev. B **76**, 245114 (2007).
[40] B. Lee, A. Canning, and L. W. Wang, J. Appl. Phys. **103**, 113713 (2008).
[41] H. N. Rojas, R. W. Godby, and R. J. Needs, Phys. Rev. Lett. **74**, 1827 (1995).
[42] A. K. McMahan, K. Held, and R. T. Scalettar, Phys. Rev. B **67**, 075108 (2003).
[43] G. Kresse and J. Furthmuller, Phys. Rev. B **54**, 11169 (1996).
[44] C. Freysoldt, P. Eggert, P. Rinke, A. Schindlmayr, and M. Scheffler, Phys. Rev. B **77**, 235428 (2008).
[45] J. C. Rienstra-Kiracofe, G. S. Tschumper, and H. F. Schaefer III, Chem. Rev. **102**, 231 (2002).
[46] NIST Chemistry Webbook, http://webbook.nist.gov/chemistry.
[47] M. P. Surh, H. Chacham, and S. G. Louie, Phys. Rev. B **51**, 7464 (1995).
[48] M. Rohlfing, P. Kruger, and J. Pollmann, Phys. Rev. B **56**, R7065(R) (1997).