Quantitative molecular orbital energies within a G0W0 approximation
Sharifzadeh, S.; Tamblyn, I.; Doak, P.; Darancet, P. T.; Neaton, J. B.

This publication could be one of several versions: author’s original, accepted manuscript or the publisher’s version. / La version de cette publication peut être l’une des suivantes : la version prépublication de l’auteur, la version acceptée du manuscrit ou la version de l’éditeur.
For the publisher’s version, please access the DOI link below. / Pour consulter la version de l’éditeur, utilisez le lien DOI ci-dessous.

Publisher’s version / Version de l’éditeur:
https://doi.org/10.1140/epjb/e2012-30206-0
The European Physical Journal B, 85, 9, 2012-09-24

NRC Publications Record / Notice d’Archives des publications de CNRC:
https://nrc-publications.canada.ca/eng/view/object/?id=7ea12fa5-2966-46ad-b0db-0ad003ec878b
https://publications-cnrc.canada.ca/fra/voir/objet/?id=7ea12fa5-2966-46ad-b0db-0ad003ec878b

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at https://nrc-publications.canada.ca/eng/copyright
READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L’accès à ce site Web et l’utilisation de son contenu sont assujettis aux conditions présentées dans le site https://publications-cnrc.canada.ca/fra/droits
LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D’UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n’arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.
Quantitative Molecular Orbital Energies within a $G_0W_0$ Approximation

Sahar Sharifzadeh,1 Isaac Tamblyn,1,2 Peter Doak,1 Pierre Darancet,1 and Jeffrey B. Neaton1,[1]

1Molecular Foundry, Lawrence Berkeley National Laboratory
2Physical and Life Sciences Directorate, Chemical Sciences Division, Lawrence Livermore National Laboratory

Using many-body perturbation theory within the $G_0W_0$ approximation, we explore routes for computing the ionization potential (IP), electron affinity (EA), and fundamental gap of three gas-phase molecules – benzene, thiophene, and (1,4) diamino-benzene – and compare with experiments. We examine the dependence of the IP on the number of unoccupied states used to build the dielectric function and the self energy, as well as the dielectric function plane-wave cutoff. We find that with an effective completion strategy for approximating the unoccupied subspace, and a converged dielectric function kinetic energy cutoff, the computed IPs and EAs are in excellent quantitative agreement with available experiment (within 0.2 eV), indicating that a one-shot $G_0W_0$ approach can be very accurate for calculating addition/removal energies of small organic molecules. Our results indicate that a sufficient dielectric function kinetic energy cutoff may be the limiting step for a wide application of $G_0W_0$ to larger organic systems.
INTRODUCTION

Organic molecules and assemblies are of considerable interest for next-generation photovoltaics \cite{1,3} and other energy conversion applications \cite{4,5}. Their performance and utility hinges on understanding and control of their spectroscopic properties, such as ionization potentials (IPs) in gas-phase and solid-state environments, and orbital energy level alignment at interfaces. Density functional theory (DFT) is a widely used computational framework for studying structural and electronic properties of materials. However, Kohn-Sham frontier orbital energies and energy differences within common approximations to DFT, such as the local density approximation (LDA) and generalized gradient approximations (GGAs), are known to dramatically underestimate these quantities \cite{6,8}. Recently, we have shown that accurate fundamental gaps for gas-phase and solid-state organic molecules \cite{9}, and frontier orbital energies for an organic/metal interface \cite{10} ((1,4) diamino-benzene on Au(111)) may be computed with many-body perturbation theory within the GW approximation \cite{11}. For the latter, we found that the calculation must be adequately converged with respect to addition/removal energies of the isolated components, i.e. molecule and substrate. In this article, building on prior work \cite{10,12–20}, we explore the extent to which we may obtain accurate IPs and electron affinities (EAs) of gas-phase molecules using a $G_0W_0$ approximation.

While there are numerous studies benchmarking the $GW$ approximation against transport gaps in bulk inorganic solids \cite{11,21,24}, similar works for isolated molecular systems are less common, and while all works exhibit marked improvement over standard DFT approaches, there is some quantitative disagreement (see e.g., \cite{10,12,19,25}). For example, for gas-phase molecules, using an atom-centered basis set, it has been found that self-consistency in either the $GW$ eigenvalues \cite{12,14} or in both the eigenvalues and eigenvectors \cite{15} is essential for obtaining good agreement of computed molecular IP and EA with experiment. On the other hand, with a planewave basis set, it has been demonstrated that a more systematic representation of the dielectric matrix and Coulomb-hole (CH) term, $\Sigma_{\text{CH}}$, brings the $G_0W_0$-predicted IP and EA in closer agreement with experiment \cite{10,16,17}. Beyond the differences in their basis sets, these studies have differed in their representation of the dielectric matrix, the presence of a truncation scheme for the Coulomb interaction, and their approach for handling the empty states necessary to converge $\Sigma_{\text{CH}}$. As a consequence, the accuracy of different $GW$ approaches for the IP of gas-phase molecules remains an open question.

Here, we compute the $G_0W_0$ IP, EA, and fundamental gap (IP - EA) of three gas-phase molecules benzene (BEN), (1,4) diamino-benzene (BDA), and thiophene (TP), as shown in Fig. 1 and compare the computed IP and EA with measurements \cite{20,29}. We examine the dependence of the IP and fundamental gap on the number of unoccupied states used to build the dielectric function and the self energy, as well as the dielectric function $G$-space cutoff. We find that as our calculations approach convergence, the computed IPs and EAs are in excellent quantitative agreement with experiment (within 0.2 eV), indicating that $G_0W_0$ can be very accurate for calculating addition/removal energies of small organic molecules.
**METHODS**

Our GW calculations are performed using the BerkeleyGW package, following an established $G_0W_0$ approach. The self-energy, $\Sigma = iGW$, is computed as a first order correction to the Kohn-Sham DFT Hamiltonian. The quasiparticle states are taken from DFT within the GGA of Perdew, Burke, and Ernzerhof (PBE) and are expanded in a planewave basis set. The cutoff for the planewave expansion is 80 Ry for BEN and TP and 60 Ry for BDA, and is determined such that the DFT total energy is converged to $< 1$ meV/atom. The molecular geometry is optimized such that forces are less than 0.04 eV/Å. Norm conserving pseudopotentials are used, with 1, 4, 5, and 6 electrons explicitly treated as valence for H, C, N, and S, respectively.

Since periodic boundary conditions are imposed in our planewave DFT and subsequent GW calculations, the molecules are placed in a large supercell, chosen to be twice the size necessary to contain $\geq 99\%$ of their charge density. The supercell dimensions are $14 \times 8 \times 15 \, \text{Å}^3$ for BEN, $14 \times 9 \times 14 \, \text{Å}^3$ for TP, and $15 \times 15 \times 15 \, \text{Å}^3$ for BDA. In constructing the dielectric matrix and the self-energy at the $G_0W_0$ step, the Coulomb potential is truncated at half of the unit cell length in order to avoid spurious interactions between periodic images. The electrostatic potential at the surface of the supercell is computed at the DFT level and its average subtracted from the GW eigenvalues to obtain absolute energies and therefore, IPs and EAs.

The static inverse dielectric function ($\epsilon_{G,G'}^{-1}(q)$) is expanded in planewaves (with wavevector $G$), and a cutoff ($\epsilon^\text{cut}_G = |q + G|^2/2$), where $q$ is a wavevector. $\epsilon_{G,G'}^{-1}(q)$ is constructed as a sum over unoccupied states, which is truncated at a finite number of states, $N_e$, with energy $E(N_e)$. The dielectric function is extended to finite frequency with the generalized plasmon-pole (GPP) model of Hybertsen and Louie.

For the purposes of analysis, we define the self-energy operator as a sum of Fock exchange, screened exchange, and the CH terms. The screened exchange term, $\Sigma_{sx}$, requires an explicit sum over just occupied states; however, it is implicitly dependent on $N_e$ through the dielectric function. The CH term, $\Sigma_{ch}$, involves a sum which in principle must span the full unoccupied subspace, but in practice is also truncated at finite number of unoccupied states $N_e$, with corresponding energy $E(N_e)$. For simplicity, we set $N_e$ equal to $N_c$, subtract the matrix elements of the Fock operator, $\Sigma_X$ (which is independent of $N_e$ and $\epsilon$), from $\Sigma_{sx}$, and study the convergence behavior of $\Sigma_{sx,x}$ and $\Sigma_{ch}$ terms with respect to $N_e$ and $\epsilon^\text{cut}_G$.

**RESULTS**

Convergence of the dielectric matrix

Fig. 1 summarizes our calculated IPs for BEN, TP, and BDA as a function of two parameters, $\epsilon^\text{cut}_G$ and $N_e$. The IP is defined here such that a positive value indicates a bound electron. The IP increases significantly, by about 0.5 eV, as either parameter is increased (taken towards convergence) for all three molecules; in contrast, the fundamental gap, IP-EA, converges rapidly to within 0.1 eV for $E(N_e) > 2$ Ry and $\epsilon^\text{cut}_G > 4$ Ry.
As noted previously [33], the interdependence of \( N_c \) and \( \epsilon_G^{\text{cut}} \) can lead to a “false convergence” of the IP with respect to the dielectric cutoff at small fixed \( N_c \). For all three molecules at \( E(N_c) = 2 \text{ Ry} \) and \( \epsilon_G^{\text{cut}} = 4 \text{ Ry} \), the IP is apparently converged to within 0.1 eV; however, if \( E(N_c) \) is increased to 6 Ry, the IP varies by 0.3 – 0.4 eV as \( \epsilon_G^{\text{cut}} \) is raised from 4 to 24 Ry. For all three molecules studied, this “false convergence” subsides for \( E(N_c) \sim 6 \text{ Ry} \) above the vacuum level (corresponding to \( N_c \sim 3000 \) for BEN and TP, and \( \sim 5000 \) for BDA within our supercells); the computed IP is unaffected by an increase of \( \epsilon_G^{\text{cut}} \) for values greater than 12 Ry for \( E(N_c) \geq 6 \text{ Ry} \). However, the IP is still quite sensitive to \( N_c \), as we will discuss further below.

For fixed \( N_c \), both \( \Sigma_{\text{SX-X}} \) and \( \Sigma_{\text{CH}} \) also appear converged (to within 0.1 eV) for \( \epsilon_G^{\text{cut}} \geq 12 \text{ Ry} \), as shown in Fig. 2 for the highest occupied molecular orbitals (HOMOs) of BEN, TP, and BDA. Interestingly, for low \( N_c \), the variation of \( \Sigma_{\text{SX-X}} \) and \( \Sigma_{\text{CH}} \) with \( \epsilon_G^{\text{cut}} \) ranges from 2 – 100 times larger than the corresponding variation of the IP. Thus, \( \Sigma_{\text{SX-X}} \) and \( \Sigma_{\text{CH}} \) are evidently less prone to “false convergence” at low \( N_c \) than the IP. Since both \( \Sigma_{\text{SX-X}} \) and \( \Sigma_{\text{CH}} \) depend on \( \epsilon^{-1} \), but with opposite sign [32], their sum (which determines the IP) is less sensitive to an underconverged dielectric function.

While Fig. 1 and Fig. 2 suggest that an \( \epsilon_G^{\text{cut}} \geq 12 \text{ Ry} \) is sufficient for a precision of 0.1 eV or better in the IP for fixed \( N_c \), they also highlight the fact that the self-energy corrections are more sensitive to \( N_c \) when the high energy Fourier components of \( \epsilon^{-1} \) are well-described. Fig. 1 shows a variation in IP of > 1 eV as \( E(N_c) \) grows from 2 Ry to greater than 6 Ry. Fig. 2 indicates that the \( \Sigma_{\text{CH}} \) term is responsible for this variation, as \( \Sigma_{\text{SX}} \) appears converged within 0.2 eV for a dielectric matrix described with \( E(N_c) \geq 6 \text{ Ry} \). This implies that for the molecules and supercells under study, for \( \epsilon_G^{\text{cut}} \geq 12 \text{ Ry} \) and \( E(N_c) \geq 6 \text{ Ry} \), the only remaining convergence issue in the calculation is the sum over the unoccupied subspace. We now discuss the different strategies for converging this sum.
FIG. 2. For BEN, TP, and BDA HOMOs: $\Sigma_{sx-x}$ (solid lines) and $\Sigma_{CH}$ (dashed lines) as a function $N_c$ and $\epsilon_c^{cut}$ for $E(N_c) = 25, 50, 80$, and $105$ eV. The legend follows Fig. 1.

![Graphs showing $\Sigma_{sx-x}$ and $\Sigma_{CH}$ for BEN, TP, and BDA HOMOs as a function of $N_c$.](image)

FIG. 3. a) $\Sigma_{CH}$ as a function of number of bands for the BDA HOMO for both static COHSEX and $G_0W_0$. The static COHSEX result for $N_c \to \infty$ is indicated with a horizontal dotted line. b) The $G_0W_0$ IP with the CH term extrapolated to infinite $N_c$ using fitting techniques and the static remainder approach.

**Convergence of the Coulomb-hole term of the self-energy**

The slow convergence of the $\Sigma_{CH}$ term, for a converged value of $\epsilon_c^{cut}$, with respect to $N_c$ can be seen in Fig. 3a for the BDA HOMO. $\Sigma_{CH}$ varies by more than 2 eV for $N_c \in [500; 5000]$ and shows a finite slope of $10^{-4}$ eV/$N_c$ at $N_c = 5000$. Moreover, this same slow convergence behavior can be seen with a static CH and screened exchange method (static COHSEX) for which a full evaluation (shown as dashed line) does not require a sum of empty states. Comparison of our dynamic and static calculations suggests that the $N_c$ dependence of $\Sigma_{CH}$ comes from both static and dynamical correlation terms. The static COHSEX CH term is still 0.2 eV away from the exact solution at $N_c = 5000$, and has a different slope than the full dynamical $\Sigma_{CH}$.

The slow convergence of $\Sigma_{CH}$ with $N_c$ has been addressed with different strategies in prior work. Here, we examine three different approaches for extrapolating the CH term to infinite $N_c$ and examine their consequence for the IP: i) fitting $\Sigma_{CH}(N_c)$ for a given orbital with an analytical form, and calculating its limit when $N_c \to \infty$ (see, e.g., [11]); ii) fitting the dynamical $\Sigma_{CH}(N_c)$ to a functional form determined from the corresponding static COHSEX
term \([10]\); and \(iii\) approximating the correction to the dynamical CH term based on completing the unoccupied subspace within the static COHSEX approximation, \textit{i.e.} the static remainder (SR) approach \([36]\).

Kang and Hybertsen applied a fitting scheme to \(\Sigma_{\text{CH}}\) to obtain the valence band maximum of TiO\(_2\) and found a 0.2 eV range in predicted values for two different functional forms for the fit \([41]\). We take a similar approach and consider the following four functional forms for the dynamical \(\Sigma_{\text{CH}}(N_c)\):

\[
\begin{align*}
\Sigma_{\text{CH}}(N_c) &\simeq \alpha + \beta N_c^{-\frac{1}{2}}, \\
\Sigma_{\text{CH}}(N_c) &\simeq \alpha + \beta N_c^{-1}, \\
\Sigma_{\text{CH}}(N_c) &\simeq \alpha + \beta N_c^{-\frac{3}{2}}, \\
\Sigma_{\text{CH}}(N_c) &\simeq \alpha + \beta e^{-\frac{N_c}{\gamma}},
\end{align*}
\]

where \(\alpha\), \(\beta\), and \(\gamma\) are fitting parameters. In practice, we find that good fits (P value < 0.005) can be consistently obtained using any of these forms.

We also fit the partial sum \(\Sigma_{\text{CH}}(N_c)\) computed within static COHSEX such that \(\alpha\) is the numerically exact closed form value of the static CH (\(\Sigma_{\text{CH}}^{\text{static}}(\infty)\)). More precisely, the static CH term, \(\Sigma_{\text{CH}}^{\text{static}}(N_c)\), is fit to Eq. (1) with \(\beta\) and \(\gamma\) as fitting parameters. The dynamical \(\Sigma_{\text{CH}}(N_c)\) is then fit to Eq. (1) with \(\gamma\) fixed and \(\alpha\) and \(\beta\) as fitting parameters. Here, we are assuming that the same functional form describing the static \(\Sigma_{\text{CH}}\) also describes the dynamical case.

Lastly, we apply the SR correction defined in Ref. \([36]\) where

\[
\Sigma_{\text{CH}}(N_c \to \infty) \simeq \Sigma_{\text{CH}}(N_c) + \frac{1}{2} \left[ \Sigma_{\text{CH}}^{\text{static}}(\infty) - \Sigma_{\text{CH}}^{\text{static}}(N_c) \right].
\]

In Fig. 3b), we report the computed IPs of BDA using all five extrapolation techniques described above. Because we are far from convergence in \(N_c\), the fitting procedure \((i)\) is much less favorable than found by Ref. \([41]\) both by its error with respect to experiments and its range of uncertainty: the assigned functional form can produce predicted IPs ranging from 5.8 to 7.2 eV. More importantly, the computed IP is very sensitive to the number of bands initially used. The best fit to the static COHSEX result for \(\Sigma_{\text{CH}}\), \((ii)\), results in IPs that monotonically increase with the number of bands used in the fit, and appears to be converging towards the SR result.

The SR method gives the best results, with predicted IP values within 0.1 eV for \(N_c \in [500; 5000]\). The results of the SR method are particularly remarkable in the sense that when using this procedure, less unoccupied states are needed to converge the CH term than the dielectric matrix (respectively 500 and 5000 for BDA).

**Comparison with experiment**

Table 1 shows the \(G_0W_0\) IP and EA for BEN, TP, and BDA, along with experimental values. For all molecules, we use \(E(N_c) = 6\) Ry and \(\epsilon_{\text{cut}}^G = 24\) Ry, and \(\Sigma_{\text{CH}}\) is extrapolated to infinite number of bands via SR \([36]\). Our
TABLE I. $G_0W_0$ IP for BEN, TP, and BDA in eV. The calculations are performed with $\epsilon_{\text{G}}^{\text{cut}}$ or 24 Ry, with $E(N_c)$ fixed at 6 Ry, and the static remainder correction applied.

| Molecule | BEN | TP | BDA |
|----------|-----|----|-----|
| IP Theory | 9.4 | 9.0 | 7.3 |
| IP experiment | 9.24 [26] | 8.86 [26] | 7.34 [27, 28] |
| EA Theory | -0.92 | -0.94 | -0.90 |
| EA experiment | -1.1 [29] | —– | —– |

$G_0W_0$ results are in excellent agreement with experiment, within 0.2 eV for IP of all three molecules and the EA of BEN. Our predictions agree well with previous planewave-based $G_0W_0$ studies [17–19] for BEN, but differ somewhat quantitatively with with other $G_0W_0$ results obtained using localized basis sets for TP [13] and BDA [20].

CONCLUSIONS

With use of unoccupied states that span $\sim 6$ Ry in energy, an $\epsilon_{\text{G}}^{\text{cut}}$ greater than or equal to 12 Ry, and the static remainder approach to correct for the finite number of empty states in $\Sigma_{\text{CH}}$, we obtain converged values for the $G_0W_0$-calculated IP and EA of three organic molecules in the gas-phase. The predicted IPs and EAs agree to within 0.2 eV with available experiment. Our results indicate that $G_0W_0$ provide quantitatively accurate addition/removal energies for small organic molecules. We find that a limiting step to these calculations is the large $\epsilon_{\text{G}}^{\text{cut}}$ required for convergence. Thus, extrapolation techniques for $\epsilon_{\text{G}}^{\text{cut}}$ will be increasingly valuable for describing larger systems, such as metal/organic molecule interfaces.

ACKNOWLEDGMENT

The authors acknowledge fruitful discussions with J. Deslippe, G. Samsonidze, D. Strubbe (UC Berkeley), X. Blase (Institute Neel), and Leecor Kronik (Weizmann Institute). Work at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DEAC02-05CH11231. P.T.D. was funded by the Helios Solar Energy Research Center, which is supported by the Director, Office of Science, Office of Basic Energy Sciences of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Computer simulations were performed using NERSC. S.S. and P.D acknowledge funding from the National Science Foundation through the Network for Computational Nanotechnology (NCN) and I.T. acknowledges financial support from NSERC.

* jhnathan@lbl.gov

[1] M. Scharber, D. Mhlbcher, M. Koppe, P. Denk, C. Waldhief, A. Heeger, C. Brabec, Advanced Materials 18(6), 789 (2006)
[38] F. Giustino, M.L. Cohen, S.G. Louie, Physical Review B 81(11), 115105 (2010)
[39] J.A. Berger, L. Reining, F. Sottile, Physical Review B 82(4), 041103 (2010)
[40] L. Reining, G. Onida, R.W. Godby, Physical Review B 56(8), R4301 (1997)
[41] W. Kang, M.S. Hybertsen, Phys. Rev. B 82, 085203 (2010)
[42] E. Van Veen, Chemical Physics Letters 41(3), 535 (1976), for TP, the 1-1' shape resonance, which corresponds to the EA of a molecule in its first excited vibrational state, has been reported to be -1.17 eV. The ground state EA must be lower in energy than this value.