PHYSICAL REVIEW B **83**, 115103 (2011)

First-principles *GW* calculations for fullerenes, porphyrins, phtalocyanine, and other molecules of interest for organic photovoltaic applications

X. Blase, C. Attaccalite, and V. Olevano

Institut Néel, CNRS and Université Joseph Fourier, B.P. 166, F-38042 Grenoble Cedex 09, France and European Theoretical Spectroscopy Facility (ETSF), F-38042 Grenoble, France

(Received 17 November 2010; published 4 March 2011)

We evaluate the performances of *ab initio GW* calculations for the ionization energies and highest occupied molecular orbital-lowest unoccupied molecular orbital gaps of 13 gas phase molecules of interest for organic electronic and photovoltaic applications, including the C_{60} fullerene, pentacene, free-base porphyrins and phtalocyanine, PTCDA, and standard monomers such as thiophene, fluorene, benzothiazole, or thiadiazole. Standard G_0W_0 calculations, that is, starting from eigenstates obtained with local or semilocal functionals, significantly improve the ionization energy and band gap as compared to density functional theory Kohn-Sham results, but the calculated quasiparticle values remain too small as a result of overscreening. Starting from Hartree-Fock-like eigenvalues provides much better results and is equivalent to performing self-consistency on the eigenvalues, with a resulting accuracy of 2%–4% as compared to experiment. Our calculations are based on an efficient Gaussian-basis implementation of *GW* with explicit treatment of the dynamical screening through contour deformation techniques.

DOI: 10.1103/PhysRevB.83.115103 PACS number(s): 71.15.Ap, 71.15.Qe, 71.20.Rv

I. INTRODUCTION

The flexibility in the synthesis of novel molecules and polymers is an important advantage of organic photovoltaics as compared to the inorganic route.[1,2] Despite a rather limited quantum efficiency, the possibility of tailoring the solubility, crystallinity, and electronic properties of the building molecular units offers much to improve upon the actual best cells, such as those based on the combination of acceptor fullerene derivatives and derivatives of polythiophene as donors.[3,4] In particular, it has been shown that there are strong correlations between the "band offsets" at the donor/acceptor interface and the open circuit voltage or the driving force for separating the hole and electron of the photoinduced excitons.[5,6] The ability to tune the electronic affinity and ionization energy of the donor and acceptor molecules, under the constraint that sunlight absorption should be kept as large as possible, is a current and intense field of research.[7–10] There is, therefore, much interest in developing efficient quantum simulation methods, which allow us to provide the spectroscopic and optical properties of standard molecules with both a reasonable computer cost and accuracy.

For isolated molecules, an excellent trade-off between computer cost and accuracy of the calculations of the ionization energy and electronic affinity can be found with the so-called delta self-consistent field approach using hybrid functionals such as PBE0 and B3LYP obtained by the admixture of a fixed amount of Fock exchange.[11,12] However, these techniques can not be used for extended systems such as bulk semiconductors and molecules deposited on a surface or in a solution, and the percentage of Fock exchange needed for obtaining good results with these functionals is expected to change from isolated molecules to bulk systems. For the same reasons, the "Kohn-Sham" ionization energies, electronic affinities, and band gaps as obtained from the eigenvalues of the Hamiltonian may be certainly improved with hybrid functionals as compared to (semi)local ones, but again the amount of Fock exchange needed to get accurate results may change from one system to another.

A technique based on many-body perturbation theory (MBPT), namely, the *GW* approximation,[13–17] has shown excellent results for the evaluation of the band edges and band gaps of extended bulk systems.[18] Distinct from the perturbative techniques developed by the quantum chemistry community to build up correlations from the Hartree-Fock solution,[19] such an approach is generally derived from functional derivative techniques[13,20] yielding an exact (nonperturbative) set of self-consistent (closed) relations between the one-body Green's function G, the polarizability P, the dynamically screened Coulomb potential W, the "exchange and correlation" self-energy Σ, and the so-called vertex corrections Γ, which is related to the variation of the self-energy with respect to an external perturbation. In practice, neglect of vertex corrections leads to the so-called *GW* approximation for the self-energy, which can be loosely described as a generalization of the Hartree-Fock method by replacing the bare Coulomb potential with a dynamically screened Coulomb interaction. The ingredients needed to proceed through the *GW* calculations further pave the way to Bethe-Salpeter calculations[17] aimed at exploring optical absorption spectra as an alternative to time-dependent density functional theory (DFT). While decades of expertise exist for appraising the performances of the *GW* approximation in the case of extended bulk systems, the application of such MBPT approaches to organic molecules in the gas phase, and, in particular, molecules of interest for photovoltaic applications,[21–26] remains extremely scarce, a situation that can be mostly attributed to the associated computational cost for molecules such as fullerene derivatives or porphyrins containing several dozens of atoms. As a result, an understanding of the merits of such an approach in the case of organic molecular systems, as compared to well-established quantum chemistry techniques, is still in its infancy.

We present in this paper a *GW* study of the quasiparticle properties of 13 of the most standard molecules involved in organic electronic and photovoltaic devices, including

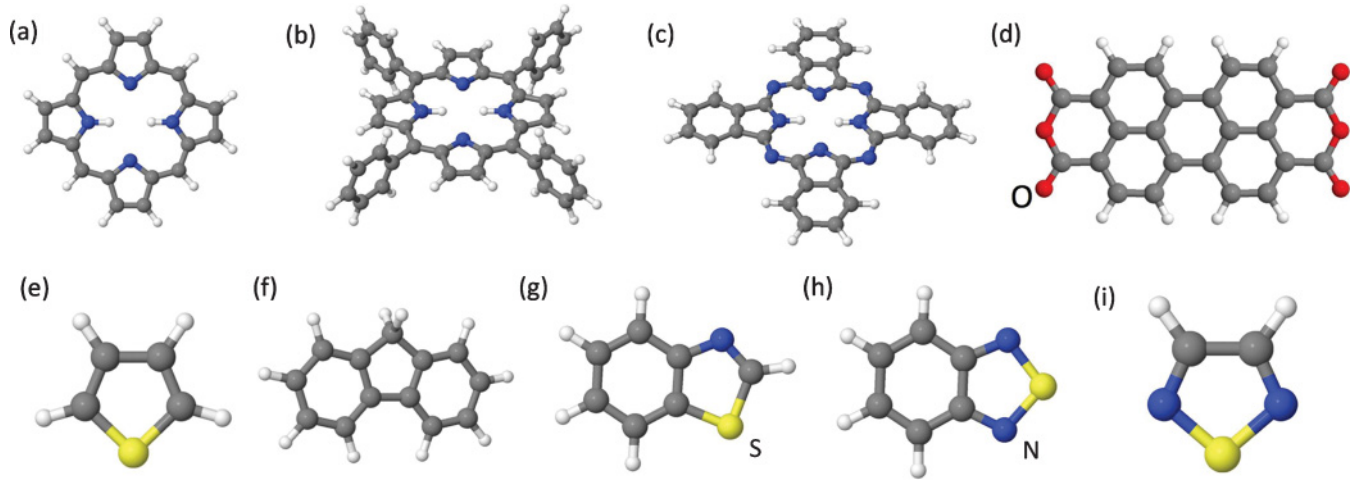

FIG. 1. (Color online) Symbolic representation of (a) 21H, 23H-porphine (H_2P), (b) tetraphenylporphyrin (H_2TPP), (c) phtalocyanine (H_2Pc), (d) 3,4,9,10-perylene tetracarboxylic acid dianydride (PTCDA), (e) thiophene, (f) fluorene, (g) benzothiazole, (h) 2,1,3-benzothiadiazole, and (i) 1,2,5-thiadiazole. Small white atoms are hydrogen atoms, gray atoms are carbon atoms, while red, blue, and yellow atoms are oxygen, nitrogen, and sulfur atoms, respectively.

the C_{60} fullerene, the free-base 21H,23H-porphine (H_2P), tetraphenylporphyrin (H_2TPP), phtalocyanine (H_2Pc), and the 3,4,9,10-perylene tetracarboxylic acid dianydride (PTCDA) (see Fig. 1). We also study anthracene, tetracene, and pentacene, π-conjugated molecules of interest for organic electronics (although not of interest for optical applications), and for which experimental band-gap data are available. Finally, the tiophene, fluorene, benzothiazole, 2,1,3-benzothiadiazole, and 1,2,5-thiadiazole monomers, building blocks of common donor polymers, are also investigated.[27,28] Our results suggest that, while the standard non-self-consistent G_0W_0 calculations based on Kohn-Sham eigenstates with (semi)local functionals certainly improve on the DFT results, the G_0W_0 ionization energy and highest occupied molecular orbital-lowest unoccupied molecular orbital (HOMO-LUMO) gap remain underestimated as compared to experiment. A simple partial self-consistency on the eigenvalues only, or the use of Hartree-Fock-like eigenvalues in a one-shot G_0W_0 calculation, allows us to obtain much improved results. We show that, in particular, these simple schemes lead to an average error of $\sim$0.3 eV for the ionization energies and 0.1–0.2 eV for the band gaps.

Our paper is organized as follows. In Sec. II, we briefly describe our implementation of the GW formalism within a Gaussian basis, including details about the evaluation of the Coulomb matrix elements. In Sec. III, our results for the ionization energy and HOMO-LUMO gap of selected molecules are presented and compared to existing experimental results. The importance of a simple self-consistency on the eigenvalues is discussed. Section IV describes a simplified non-self-consistent approach based on an approximate perturbative Hartree-Fock starting point for building the Green's function and screened Coulomb potential. We conclude in Sec. V.

II. METHODOLOGY

Our code is based on a Gaussian-basis implementation of the GW formalism and builds on a previous implementation of calculating the inverse dielectric matrix using numerical strictly localized orbitals.[29] To avoid dealing with a numerical basis, this implementation now expands the needed two-point operators (bare and screened Coulomb potentials, susceptibilities, etc.) on an "auxiliary" Gaussian basis composed of one-center atomiclike orbitals, with real spherical harmonics for the angular part and a radial dependence composed of Gaussian functions. The use of such an auxiliary basis, commonly implemented in several DFT quantum chemistry codes to express the charge density for ground- or excited-state[30] calculations, allows us to greatly speed up the evaluation of, e.g., the Coulomb matrix elements. We discuss these points in the following subsections.

A. General formalism

With the notations of Ref. 31, we introduce for any two-point function $f(\mathbf{r},\mathbf{r}')$ the $\langle f\rangle$ and $[f]$ matrices in the auxiliary basis related through

$$[f]_{\mu,\nu} = \int\int d\mathbf{r}\, d\mathbf{r}'\ \mu^*(\mathbf{r}) f(\mathbf{r},\mathbf{r}')\nu(\mathbf{r}'), \langle f\rangle = S\langle f\rangle S,$$
$$f(\mathbf{r},\mathbf{r}') = \sum_{\mu,\nu} \mu(\mathbf{r})\langle f\rangle_{\mu,\nu}\nu^*(\mathbf{r}'),$$

where μ and ν are elements of the basis and S is the overlap matrix. The standard Dyson equation relating the dynamically screened Coulomb potential $W(\omega)$ to the bare Coulomb potential (v) can then be written as

$$\langle W(\omega)\rangle = \langle v\rangle + \langle v\rangle[\chi^0(\omega)]\langle W(\omega)\rangle$$

with χ^0 the unscreened (free-electron) susceptibility

$$[\chi^0(\omega)]_{\mu,\nu} = \sum_{\text{spins}}\sum_{i}^{\text{occ}}\sum_{j}^{\text{unocc}} \langle\phi_i|\mu|\phi_j\rangle\langle\phi_j|\nu|\phi_i\rangle \times\left(\frac{1}{\omega+\varepsilon_i-\varepsilon_j+i\delta}-\frac{1}{\omega-\varepsilon_i+\varepsilon_j-i\delta}\right),$$

where $\delta = 0^+$. The input (ϕ_i, ε_i) are one-body eigenstates and related eigenvalues traditionally taken as the Kohn-Sham solutions of a ground-state DFT calculation. In this paper, we start with a standard DFT local density approximation (LDA) calculation but, as discussed in the following, this may not constitute the best starting point for molecular systems. The knowledge of the dynamical screened Coulomb potential $W(\omega)$ allows us to build the nonlocal and energy-dependent self-energy operator Σ, which accounts for exchange and correlation in the present quasiparticle formalism[13] and reads as

$$\Sigma^{GW}(\mathbf{r},\mathbf{r}'|E) = \frac{i}{2\pi}\int d\omega\, e^{i0^+\omega} G(\mathbf{r},\mathbf{r}'|E+\omega) W(\mathbf{r},\mathbf{r}'|\omega),$$
$$G(\mathbf{r},\mathbf{r}'|\omega) = \sum_n \phi_n(\mathbf{r})\phi_n^*(\mathbf{r}')/(\omega - \varepsilon_n \pm i\delta),$$

where the time-ordered Green's function G is again built from the (ϕ_i, ε_i) eigenstates. The sign of the δ infinitesimal ensures that the occupied (unoccupied) states correspond to poles in the fourth (second) quadrants. Again, the choice of the "best" input (ϕ_i, ε_i) for the building of G will be discussed in the following.

This implementation is formally equivalent to that of Ref. 31, except that we go beyond the plasmon-pole model and proceed with the explicit calculation of the frequency integral for the correlation part of the self-energy $\Sigma_c^{GW} = \Sigma^{GW} - \Sigma_x$, with Σ_x the Fock operator. We use contour deformation techniques with an integration along the imaginary axis complemented by the evaluation of the poles in the first and third quadrant for states away from the band edges[16,32]:

$$\Sigma_c^{GW}(\mathbf{r},\mathbf{r}'|E) = \sum_n \phi_n(\mathbf{r})\phi_n^*(\mathbf{r}')\mathcal{V}_n(\mathbf{r},\mathbf{r}'|E)$$

with, introducing $\tilde{W} = W - v$, E_F the Fermi level, and θ the Heaviside step function,

$$\mathcal{V}_n(\mathbf{r},\mathbf{r}'|E) = \tilde{W}(\mathbf{r},\mathbf{r}'|\varepsilon_n - E)\,[\theta(E-\varepsilon_n) - \theta(E_F - \varepsilon_n)] - \int_0^{+\infty} \frac{d\omega}{\pi} \frac{E-\varepsilon_n}{(E-\varepsilon_n)^2 + \omega^2} \tilde{W}(\mathbf{r},\mathbf{r}'|i\omega).$$

A change of variable allows us to fold the smooth function $\tilde{W}(i\omega)$ onto the finite [0,1] interval, where the Gaussian quadrature with as little as 12 Gaussian points is sufficient to reach convergency. An analytically integrable tail is added (subtracted) to avoid instabilities with the integrand for $\omega \to 0$ when $E \to \varepsilon_n$.

The first-order perturbation theory self-energy correction to the DFT Kohn-Sham eigenvalues is extrapolated to the quasiparticle energies by a Taylor expansion, namely,

$$\varepsilon_n^{QP} = \varepsilon_n + Z_n \langle \phi_n | \Sigma^{GW}(\varepsilon_n) - V_{xc}^{\mathrm{LDA}} | \phi_n \rangle,$$

where Z_n is the renormalization factor defined as

$$1/Z_n = 1 - \left(\partial \Sigma^{GW}/\partial E\right)_{\varepsilon=\varepsilon_n}$$

with (ε_n, ϕ_n) the LDA Kohn-Sham eigenvalues and eigenstates in this case.

B. Gaussian basis

The auxiliary basis used to expand the two-point functions reads as $\mu(\mathbf{r}) = \exp(-\alpha r^2) r^l R_l^m(\hat{r})$ with $R_l^m(\hat{r})$ the real-spherical harmonics and $(\hat{r})$ the angular components of the $\mathbf{r}$ vector. It is computationally more efficient to work with the R_l^m instead of the standard Y_l^m complex harmonics with the following relation:

$$R_l^m(\hat{r}) = \begin{cases} \left[Y_l^m(\hat{r}) + (-1)^m Y_l^{-m}(\hat{r})\right]/\sqrt{2} & (\mathrm{m} > 0) \\ Y_l^m(\hat{r}) & (\mathrm{m} = 0) \\ \left[Y_l^{-m}(\hat{r}) - (-1)^m Y_l^m(\hat{r})\right]/\sqrt{2} & (\mathrm{m} < 0). \end{cases}$$

The products $r^l R_l^m(\hat{r})$ yield the standard expressions (x, y, z, xy, yz, x^2-y^2, etc.) for the p, d, etc., orbitals (within constant factors). We briefly recall that the main advantage of a Gaussian radial part (as compared to numerical or Slater-type orbitals) is that the product of two Gaussians centered on atoms A and B with decay coefficients α_1 and α_2 yields a Gaussian centered on $C = (\alpha_1 A + \alpha_2 B)/(\alpha_1 + \alpha_2)$ with a decay constant $\gamma = \alpha_1\alpha_2/(\alpha_1 + \alpha_2)$. Further, the $r^l R_l^m(\hat{r})$ can easily be "shifted" from one center to another with, for the sake of illustration,

$$(x - x_A)(y - y_A) = (x - x_C)(y - y_C) + (y_C - y_A)(x - x_C) + (x_C - x_A)(y - y_C) + \text{constant},$$

showing that a d_{xy} orbital centered on A can be easily expressed as a function of (s,p) and d_{xy} orbitals centered on C. Such trivial expressions allow us to express multi-center overlaps in terms of one-center integrals.

In this paper, our calculations start with a DFT calculation of the structural and electronic properties of the molecules of interest using the SIESTA package.[33] We use a double-ζ+polarization (DZP) basis[34] and standard norm-conserving pseudopotentials. Since the SIESTA package uses "numerical" orbitals, we first fit the numerical radial part by up to five contracted Gaussians[35] in order to exploit the relations briefly sketched above. As such, both the "ground-state" DFT basis and the auxiliary basis are based on Gaussians. Beyond the analytic character of the Gaussian basis, our choice was also motivated by the possibility of using eigenstates generated by standard chemistry codes with all electron approaches and/or hybrid functionals, possibly providing for some systems a better starting point for MBPT calculations (see discussion in the following). We labeled our code "FIESTA" as an attempt to extend the "SIESTA" package to excited-state properties.

Contrary to the plane-wave case, the auxiliary basis for the two-point response functions is larger than the ground-state basis. Following Kaczmarski and co-workers,[36] we typically adopt for the first row elements such as carbon, nitrogen, and oxygen, four s,p,d sets of Gaussian orbitals, that is, 36 orbitals per atom, while three s,p,d Gaussian orbitals are sufficient for hydrogen. We show in the following that such a basis is large enough for the studied organic systems. In the case of sulfur, f-channel orbitals are added. With such a basis, a typical G_0W_0 calculation with full dynamics for our largest molecule (H_2TPP) can be performed within one day on a single standard processor. Better timings and scaling may be obtained upon implementing the recently introduced techniques allowing

us to avoid summation over the conduction states,[24,37–39] or techniques decoupling the sum over valence and conduction states,[40] even though the number of unoccupied states is rather limited with standard DZP or larger TZDP basis.

The choice of the "optimal" α coefficients, controlling the localization of the basis orbitals, is a difficult question. Auxiliary bases have been implemented in many quantum chemistry codes in order to fit the charge density and speed up the calculation of the Coulomb integrals. The coefficients of the charge density on the auxiliary basis are optimized using identity rules,[41] but not the decay coefficients in the exponentials. Years of expertise in the quantum chemistry community yielded reliable auxiliary bases for the Periodic Table and numerous tests have shown that high precision can be obtained with such bases, provided that they are sufficiently large.

Since the auxiliary basis must project onto products of Kohn-Sham orbitals, optimized bases for all-electron calculations can not be straightforwardly used for GW calculations starting from ground-state calculations with pseudopotentials. The same guiding lines, however, can be followed. We adopt, in particular, the idea of a "tempered" basis,[42–44] suggesting that it is better to generate a chain of α parameters such that $\alpha_{i+1}/\alpha_i = \text{constant}$, rather than spreading them uniformly between $\alpha_{\min}$ and $\alpha_{\max}$. Such a scheme hinges on the fact that the overlap of two Gaussian orbitals is a function of their alpha coefficient ratio and that, maintaining a constant overlap between "adjacent" Gaussians, allows us to better span the associated Hilbert space.[44] As such, with the $\alpha_{\min}$, $\alpha_{\max}$, and the number of Gaussian per l-channel being chosen, the other Gaussian coefficients are automatically generated.

We adopt the basis proposed by Kaczmarski and coworkers,[36] namely Gaussians with localization parameters of (0.2,0.5,1.25,3.2) a.u. for the (s,p,d) channels of C, O, and N atoms, and Gaussians with $\alpha = (0.1,0.4,1.5)$ a.u. for hydrogen. As shown in Table I, in the case of anthracene, H_2P porphyrin, and C_{60}, changing the $\alpha_{\min}$ and $\alpha_{\max}$ values, or increasing the number of Gaussians in the basis, does not significantly change the results. The case of C_{60} shows, however, that reducing the number of Gaussians to three per l-channel yields a significant error on the band gap. We will show in the following that the results obtained with this implementation compare rather well with previous available calculations based on another Gaussian basis, plane waves (PWs), or a combination of Gaussians, PWs, and Wannier functions.

TABLE I. Evolution of the ionization (IE) and band-gap energies of selected molecules as a function of the carbon auxiliary basis, changing the number (ng) of Gaussians per l-channel, the $\alpha_{\min}$ and $\alpha_{\max}$ coefficients. Results are in eVs.

Auxiliary basis	Anthracene		H_2P		C_{60}	
ng (in $\alpha_{\min} \rightarrow \alpha_{\max}$)	IE	Gap	IE	Gap	IE	Gap
3 in 0.2 → 3.2	6.83	6.02	6.49	4.67	7.21	4.08
4 in 0.2 → 3.2	6.89	6.15	6.56	4.79	7.29	4.44
5 in 0.2 → 3.2	6.86	6.14	6.52	4.76	7.30	4.37
4 in 0.15 → 3.2	6.89	6.15	6.52	4.74	7.40	4.47
5 in 0.15 → 3.2	6.82	6.06	6.56	4.77	7.29	4.36
5 in 0.15 → 3.5	6.83	6.08	6.51	4.75	7.28	4.33

We conclude this section related to the auxiliary basis by mentioning an important numerical aspect related to the overcompleteness of the generated nonorthogonal Gaussian basis. While the basis on a given atom can be easily orthogonalized using, e.g., a Gram-Schmidt procedure, the overlap between the most diffuse orbitals on adjacent atoms tends also to be rather large, yielding an overlap S matrix that is "nearly singular." Following the strategy developed in the case of product basis,[40,45] we rotate our basis to that of the eigenvectors of the overlap S matrix from which we remove the eigenvectors with eigenvalue smaller than typically 10^{-5}. In the case of the auxiliary basis, such a truncation does not reduce significantly the size of the basis, but avoids the potential numerical instability associated with inverting the nearly singular S matrix and the amplification of errors associated with the $\langle v \rangle = S^{-1}[v]S^{-1}$ transformation (see above). The cost of rotating the Coulomb and $\langle \phi_i|\beta|\phi_j \rangle$ matrix elements from the original one-center auxiliary basis (β) to the (filtered)

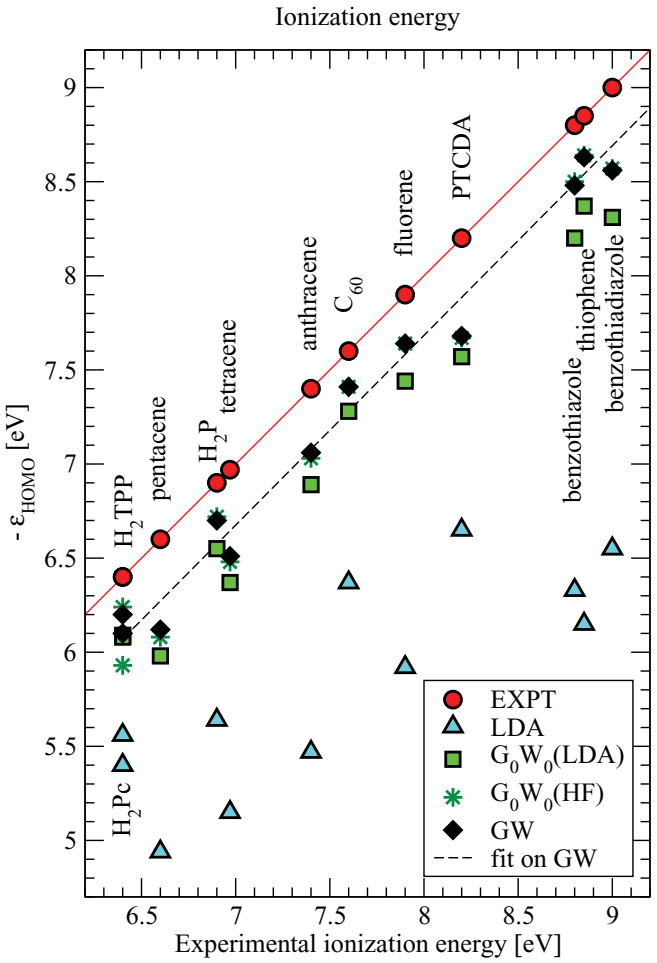

FIG. 2. (Color online) Experimental and theoretical ionization energies in electronvolts. Red circles: experimental values; light blue triangles up: LDA Kohn-Sham HOMO energy; green squares: non-self-consistent G_0W_0(LDA) value; black diamonds: GW value with self-consistency on the eigenvalues; green stars: non self-consistent $G_0W_0(\text{HF}_{\text{diag}})$ (see text). The black dashed line is a least-squares fit of the GW results. The figure has been formatted so as to preserve the same physical scale on both axes.

TABLE II. Ionization energies in eV as obtained from the Kohn-Sham eigenvalues (LDA-KS), from non-self-consistent G_0W_0(LDA) calculations, from a GW calculation with self-consistency on the eigenvalues (GW), and from a non-self-consistent G_0W_0 calculation starting from Hartree-Fock-like eigenvalues [$G_0W_0(\mathrm{HF}_{\mathrm{diag}})$, see text]. MAE is the absolute mean error in eV. The average error in percent as compared to the experiment is indicated in parentheses.

	Ionization energy				
	LDA-KS	G_0W_0(LDA)	GW	$G_0W_0(\mathrm{HF}_{\mathrm{diag}})$	Experiment
Anthracene	5.47	6.89	7.06	7.03	7.4[a]
Tetracene	5.15	6.37	6.51	6.48	6.97[a]
Pentacene	4.94	5.98	6.12	6.08	6.6[a]
C_{60}	6.37	7.28	7.41	7.41	7.6[a]
PTCDA	6.65	7.57	7.68	7.67	8.2[b]
H_2P	5.64	6.55	6.70	6.72	6.9[a]
H_2TPP	5.40	6.09	6.20	6.24	6.4[a]
H_2Pc	5.56	6.08	6.10	5.93	6.4[c]
Thiophene	6.15	8.37	8.63	8.64	8.85[a]
Fluorene	5.92	7.44	7.64	7.64	7.9[a]
Benzothiazole	6.33	8.20	8.48	8.50	8.8[a]
Thiadiazole	7.22	9.65	9.89	9.90	10.1[d]
Benzothiadiazole	6.55	8.31	8.56	8.57	9.0[a]
MAE	1.83(23%)	0.47(6%)	0.30(3.8%)	0.31(4.0%)	

[a]Reference 48.
[b]Reference 21.
[c]Reference 49.
[d]Reference 50.

S eigenvectors basis scales as N^3 and represents a marginal part of the CPU time.

C. Coulomb matrix elements

An important ingredient is the evaluation of the Coulomb matrix elements between two auxiliary basis orbitals localized on two different atoms. Exploiting the properties of the Fourier transform (FT) of Gaussian-based orbitals, namely,

$$\mathrm{FT}\left[e^{-\alpha r^2} r^l R_l^m(\hat{r})\right] = C e^{-\gamma q^2} q^l R_l^m(\hat{q}) \qquad (1)$$

with $\gamma = 1/4\alpha$ and $(\hat{r},\hat{q})$ the angular components of the (**r**,**q**) vectors in direct and reciprocal space, respectively (C is a constant), the Coulomb matrix elements reduce to a sum of terms built from the product of one-center overlaps of three real-spherical harmonics $\langle R_l^m R_{l'}^{m'} | R_L^M \rangle$ [related to Gaunt coefficients with $|l-l'| \leqslant L \leqslant (l+l')$] times radial integrals $I(l,l';L)$ of the form

$$I(l,l';L) = \int_0^\infty dq\ e^{-\zeta q^2} q^\mu J_\nu(-\beta q^2).$$

The $\langle R_l^m R_{l'}^{m'} | R_L^M \rangle$ factors are pretabulated. The oscillatory behavior of the Bessel function of the first kind J_ν makes the direct numerical evaluation rather unstable. We prefer to notice that $I(l,l';L)$ is straightforwardly related to the ${}_1F_1$ confluent hypergeometric functions,[46] which, for the needed (l,l') values, can be expressed in terms of simple functions such as the error function (erf) or the Dawson integral $F(z) = \sqrt{\pi}\ \exp(-z^2)\ \mathrm{erf}i(z)/2$, with $\mathrm{erf}i(z) = \mathrm{erf}(iz)/i$, for which rapidly convergent serial expressions exist.[47] This is an important advantage of the auxiliary-basis approach that the evaluation of the off-site Coulomb matrix elements is not a costly part of the present GW implementation.

III. RESULTS

A. Ionization energies

We start by exploring the ionization energy of our selected molecules. While experimental data for the electronic affinity of molecules in the gas phase are scarce, accurate measured ionization energies are much more common.[48] Experimental ionization energies are represented by red circles in Fig. 2 and are given in the last column of Table II. The DFT-LDA ionization energies, as obtained from the opposite sign of the Kohn-Sham highest occupied (HOMO) energy level, are clearly much too small, with an average error of 1.83 eV or 23% (see blue triangles in Fig. 2). Very similar results are obtained using the HOMO energy value as obtained with a semilocal functional such as PBE.[51]

We now turn to G_0W_0(LDA) calculations, that is, non-self-consistent calculations with the Green's function and screened Coulomb potential directly built from the LDA Kohn-Sham eigenstates and eigenvalues. The analysis of the results (column 3 in Table I and green squares in Fig. 2) shows that the ionization energies are greatly improved, with an average error of 0.47 eV, that is, a much reduced 6% error.

Even though they are in much better agreement with experiment than LDA or PBE, the discrepancies are still sizable. As shown in the following, part of the problem originates in that the "starting" LDA HOMO-LUMO gap is dramatically too small for isolated molecules, inducing a large overscreening.[52] To avoid using some arbitrary scissor operator to open the HOMO-LUMO gap in calculating the susceptibility, we instead perform a restricted self-consistency by reinjecting the corrected eigenvalues in G and W up to convergency. As a matter of fact, no more than three or four iterations are needed to reach convergency within 0.01 eV.

Such an approximation is labeled as GW in the following. This is not a full self-consistent approach, as the eigenstates are not updated, with the advantage that it keeps the computational costs reasonable.

The analysis of the results (fourth column in Table II and black diamonds in Fig. 2) clearly shows that the self-consistency on the eigenvalues improves the results for the ionization energy, reducing the average error from 0.47 eV (6%) to 0.30 eV (or 3.8%). Such a discrepancy is still sizable, but much better than that obtained from the LDA Kohn-Sham HOMO energy. An interesting observation is that the final GW ionization energies gather much closer to a straight line (dotted black line on Fig. 2) parallel to the first diagonal (red "experimental" line) than the LDA data, which are much more spread. From a pragmatical point of view, this means that the band offset between two molecules will strongly benefit from cancellation of errors in GW as compared to LDA. In particular, the remaining error (∼0.2 eV) on the ionization energy for C_{60}, the most standard acceptor, is nearly identical to the error on the ionization energies of porphyrins and phtalocyanines, which are commonly used donors. We now show that self-consistency, even though limited to updating the eigenvalues only, has an even larger effect on the magnitude of the HOMO-LUMO gaps.

B. HOMO-LUMO gaps

Due to the lack of experimental values for the electronic affinity, experimental quasiparticle HOMO-LUMO gaps (red circles in Fig. 3) are scarce, so we plot our results as a function of our "best" calculated HOMO-LUMO gaps, namely, the GW ones. In the case of C_{60}, anthracene, tetracene, and pentacene, for which experimental data are available, we observe as expected that the LDA HOMO-LUMO gap (blue triangles) is too small. This is well known in the case of bulk semiconductors, but here the discrepancy is much larger, with an average error of ∼4.1 eV or 71%.

The G_0W_0(LDA) HOMO-LUMO gaps (green squares) significantly improve with respect to LDA. Comparing to available G_0W_0(LDA) data for this class of aromatic molecules, our calculated 6.15 eV HOMO-LUMO gap for anthracene compares well with the 5.97 eV values of Niehaus and co-workers, despite the differences in basis and the treatment of dynamical effects.[53] Our G_0W_0(LDA) 4.79 and 4.23 eV HOMO-LUMO gaps for the H_2P and H_2TPP free-base porphyrins, respectively, compare well with the 5 and 4.39 eV plane-wave results of Palummo and co-workers.[23] Similarly, our G_0W_0(LDA) 4.44 eV band gap for C_{60} is in good agreement with the real-space grid formulation of Tiago and co-workers,[22] yielding a band gap of 4.36 eV. Such comparisons certainly underline the reliability of our Gaussian-basis implementation. Our 4.53 eV band gap for PTCDA is, however, smaller than the 4.9 eV band gap found with a previous plane-wave GW calculation.[21,58,61]

Overall, we remark on a systematic underestimation of the G_0W_0(LDA) HOMO-LUMO gap with respect to the experiment, with an average error for our test molecules of ∼0.75 eV or 13%. This contrasts with the case of bulk systems for which the results of G_0W_0(LDA) are generally in much better agreement with experimental values. Such a behavior can be analyzed by noticing that building the polarizabilities and screened Coulomb potential with LDA eigenvalues leads to a significant overscreening.[52] This induces too large a correlation correction $G(W - V^C)$ to the Hartree-Fock HOMO-LUMO gap.

Even though it is much better than the Kohn-Sham HOMO-LUMO gap obtained with, e.g., the B3LYP functional[54] (see empty down triangles in Fig. 3), it is desirable to improve the results. Following the simple scheme introduced above, performing self-consistency on the eigenvalues in G and W, the GW HOMO-LUMO gap is further increased to reach much better agreement with experiment. The MAE is now reduced to 0.22 eV or 3.8% for our four test molecules. In the case of C_{60}, which is the most standard acceptor in organic photovoltaic cells, the excellent agreement with experiment for the band-gap value is rather satisfactory. It is interesting to note further that the MAE of 0.22 eV for HOMO-LUMO gaps is close to the 0.3 eV MAE obtained for the ionization energies, suggesting that the electronic affinity is quite well reproduced on average.

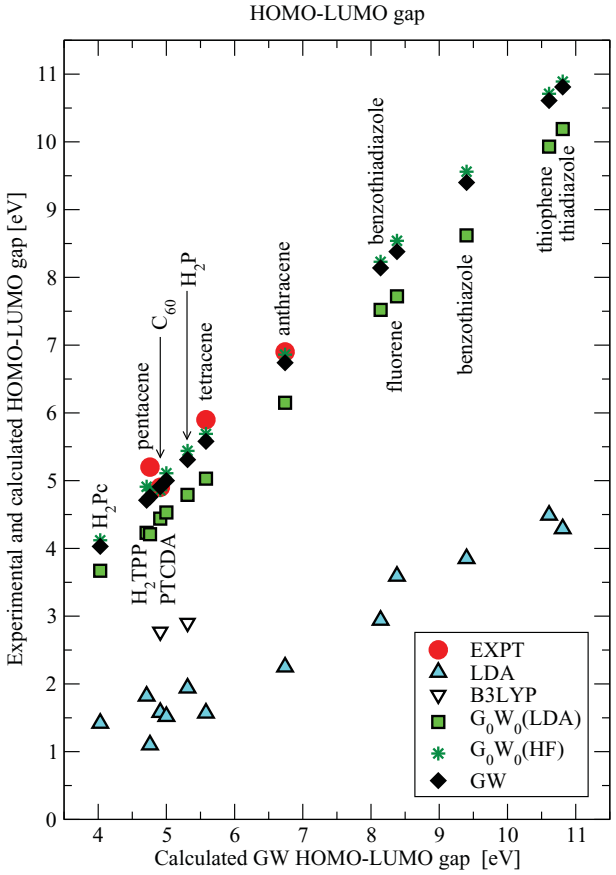

FIG. 3. (Color online) Experimental and theoretical HOMO-LUMO gaps in electronvolts. Red circles: experimental values; light blue triangles up: LDA Kohn-Sham HOMO-LUMO gap; green squares: non-self-consistent G_0W_0(LDA) value; black diamonds: GW value with self-consistency on the eigenvalues; green stars: non-self-consistent $G_0W_0(HF_{diag})$ values (see text). The two down-pointing empty triangles are B3LYP/6-31G(d) HOMO-LUMO gap values from Refs. 54–57 for C_{60} and the H_2P porphin.

TABLE III. HOMO-LUMO gap in eV as obtained from the Kohn-Sham eigenvalues (LDA-KS), non-self-consistent G_0W_0(LDA) calculations, a GW calculation with self-consistency on the eigenvalues (GW), and a non-self-consistent G_0W_0 calculation starting from Hartree-Fock-like eigenvalues [$G_0W_0(\mathrm{HF_{diag}})$, see text]. MAE is the absolute mean error in eV for the anthracene, tetracene, pentacene, and C_{60} cases for which experimental band-gap data are available. The average error in percent as compared to the experiment is indicated in parentheses.

	HOMO-LUMO gap				
	LDA-KS	G_0W_0(LDA)	GW	$G_0W_0(\mathrm{HF_{diag}})$	Experiment
Anthracene	2.25	6.15	6.74	6.86	6.9[a]
Tetracene	1.57	5.03	5.58	5.69	5.9[a]
Pentacene	1.10	4.21	4.76	4.86	5.2[a]
C_{60}	1.58	4.44	4.91	5.08	4.9[a]
MAE	4.10 (71%)	0.76 (13%)	0.22 (3.8%)	0.10 (2%)	
PTCDA	1.52	4.53	5.0	5.11	
H_2P	1.94	4.79	5.31	5.44	
H_2TPP	1.82	4.23	4.71	4.91	
H_2Pc	1.42	3.67	4.03	4.12	
Thiophene	4.49	9.93	10.61	10.71	
Fluorene	3.59	7.72	8.38	8.54	
Benzothiazole	3.85	8.62	9.40	9.56	
Thiadiazole	4.29	10.19	10.81	10.89	
Benzothiadiazole	2.94	7.52	8.14	8.23	

[a]Reference 48.

IV. A SIMPLE NON-SELF-CONSISTENT G_0W_0 APPROACH BASED ON HARTREE-FOCK-LIKE EIGENVALUES

We conclude this study by exploring a simple non-self-consistent G_0W_0 scheme starting from an "ansatz" Hartree-Fock (HF) calculation obtained by removing the exchange-correlation contribution to the LDA eigenvalues and adding the diagonal part of the exchange operator in the LDA basis, namely,

$$\epsilon_n^{\mathrm{HF}} = \epsilon_n^{\mathrm{LDA}} + \langle \phi_n^{\mathrm{LDA}} | \Sigma_x - V_{xc}^{\mathrm{LDA}} | \phi_n^{\mathrm{LDA}} \rangle,$$

where Σ_x and V_{xc}^{LDA} are the Fock and (semi)local exchange-correlation operators. We label this very simple scheme $G_0W_0(\mathrm{HF_{diag}})$. This approximation was tested by Hahn, Schmidt, and Bechstedt[62] in the case of three small molecules (silane, disilane, and water), arguing as we do that the Kohn-Sham eigenvalues are not a good starting point to evaluate the time-ordered Green's function and the screened potential. Such an approach is also a variation on the G_0W_0(HF) scheme recently introduced in Ref. 52, which was shown to yield the best ionization energies for small molecules. With increasing size and number of electrons, the role of correlations in the self-energy is expected to become more important; using Hartree-Fock eigenstates (eigenvalues) as a starting point for the much larger systems we study may, in principle, not be better than using (semi)local functionals for generating the starting eigenstates. This is what we now explore.

For the sake of comparison, we have studied the two small carbon-based conjugated molecules C_2H_2 and C_2H_4, which were investigated by Rostgaard and co-workers within their full G_0W_0(HF) scheme. The present $G_0W_0(\mathrm{HF_{diag}})$ treatment increases the ionization energy by 3.48 and 3.80 eV for C_2H_4 and C_2H_2, respectively, as compared to the LDA values. Such corrections compare well with the 3.61 and 3.90 eV values obtained within the full G_0W_0(HF) scheme of Rostgaard and co-workers (as compared to DFT/PBE), emphasizing the reliability of this simplified approximation.

As compiled in Tables II and III (column 5) and in Figs. 2 and 3 (green stars), we do find as well that a single shot $G_0W_0(\mathrm{HF_{diag}})$ calculation provides results that are in good agreement with the full GW calculations with self-consistency on the eigenvalues. In particular, the $G_0W_0(\mathrm{HF_{diag}})$ calculations yield much better results than the G_0W_0(LDA) scheme. Such a conclusion agrees with that of Rostgaard and co-workers, concluding that, for small isolated molecules, the full G_0W_0(HF) scheme actually outperforms a full self-consistent GW calculation where both eigenstates and eigenvalues are updated.[52]

Within the present $G_0W_0(\mathrm{HF_{diag}})$ approach, the MAE on the ionization energies as compared to experiment is 0.31 eV, in good agreement with the 0.4 eV result of Ref. 52 for small molecules. Such an agreement indicates that the present $G_0W_0(\mathrm{HF_{diag}})$ implementation captures most of the features of a full G_0W_0(HF) approach, suggesting that LDA and HF eigenfunctions may not be too different for this set of molecules, a conclusion often discussed in the literature. Further, the error on the band gap, averaged on the calculated values for anthracene, tetracene, pentacene, and C_{60}, for which precise experimental data are available, is found to be as small as 0.1 eV (2% error). Such values compare very well with accurate quantum chemistry calculations with a scheme, the GW formalism, which can be applied both to finite-size and extended systems, and allows us to obtain not only the band edges, or frontier orbitals, but also the full quasiparticle spectrum (see Ref. 61).

V. CONCLUSIONS

We have explored the performances of several GW approximations for the calculation of the ionization energy and HOMO-LUMO gap of 13 "large" molecules of interest for photovoltaic applications, including C_{60}, free-base porphyrins and phtalocyanine, PTCDA, and standard donor monomers such as thiophene. Our calculations are based on a Gaussian-basis implementation with full dynamical effects through contour deformation techniques. Due to the dramatic error on the HOMO-LUMO gaps obtained with (semi)local functionals, we find that the standard non-self-consistent G_0W_0 calculations based on input LDA eigenstates perform rather poorly, in particular in evaluating the HOMO-LUMO gaps. A simple self-consistency on the eigenvalues used to build G and W provides much better results. As an even simpler scheme, a non-self-consistent $G_0W_0(\mathrm{HF_{diag}})$ starting from Hartre-Fock-like eigenvalues provides equivalent results. Both the GW and $G_0W_0(\mathrm{HF_{diag}})$ approaches provide ionization energies with a mean average error within ∼0.3 eV (∼4%) of the experiment. Concerning the HOMO-LUMO gaps, with a limited number of experimental data, the same GW and $G_0W_0(\mathrm{HF_{diag}})$ approaches yield a mean average error of 0.1–0.2 eV (2%–4%), in much better agreement than the 4.1 eV (71%) error within DFT/LDA, but also in significantly better agreement than the 0.76 eV (13%) error within the "standard" G_0W_0(LDA) approach. The possibility of performing GW calculations for molecules comprising several dozens of atoms with reasonable computer time and accuracy, with a scheme allowing us to obtain the full quasiparticle spectrum of both finite-size and extended systems, opens the way to the investigation of organic photovoltaic systems with techniques that may possibly compete with well-established quantum chemical approaches.

Note added in Proof: In a recent work, $G_0W_0(LDA)$ calculations for anthracene, starting from similar SIESTA LDA eigenstates, but with the full product-basis for describing the two point operators within the GW scheme, yield an ionization energy and band gap in excellent agreement with our own results. See: D. Foerster, P. Koval, and D. Sánchez-Portal, cond-mat, arXiv:1101.2065.

ACKNOWLEDGMENTS

X.B. is indebted to Marc Cassida for suggesting numerous references, Pascal Quemerais for pointing out the relations between the radial integrals involved in the Coulomb matrix elements evaluation and the ${}_1F_1$ confluent hypergeometric functions, Julian Gale for discussions on the Gaussian fit of the radial part of numerical orbitals, and Brice Arnaud for suggesting techniques to stabilize the imaginary axis integration in the contour deformation approach. Calculations were performed on the CIMENT platform in Grenoble thanks to the Nanostar RTRA project.

[1] A. C. Mayer *et al.*, Mater. Today **10**, 28 (2007).
[2] *Organic Photovoltaics: Concepts and Realization*, Springer Series in Material Science, edited by C. Brabec, V. Dyakonov, J. Parisi, and N. S. Sariciftci (Springer, New York, 2003).
[3] M. Reyes-Reyes, K. Kim, and D. Carroll, Appl. Phys. Lett. **87**, 083506 (2005).
[4] C. J. Brabec *et al.*, Adv. Mater. (Weinheim, Ger.) **22**, 3839 (2010).
[5] M. C. Scharber *et al.*, Adv. Mater. (Weinheim, Ger.) **18**, 789 (2006).
[6] F. B. Kooistra *et al.*, Org. Lett. **9**, 551-554 (2007).
[7] S. Park *et al.*, Nat. Photonics **3**, 297 (2009).
[8] H-Y. Chen *et al.*, Nat. Photonics **3**, 649 (2009).
[9] R. Schueppel *et al.*, Phys. Rev. B **77**, 085311 (2008).
[10] F. Lincker *et al.*, Adv. Funct. Mater. **18**, 3444 (2008).
[11] A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).
[12] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
[13] L. Hedin, Phys. Rev. **139**, A796 (1965).
[14] G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. Lett. **45**, 290 (1980); Phys. Rev. B **25**, 2867 (1982).
[15] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[16] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).
[17] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[18] W. G. Aulbur, L. Jonsson, and J. W. Wilkins, in *Solid State Physics*, edited by H. Ehrenreich (Academic, Orlando, 1999), Vol. 54, p. 1.
[19] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover, New York, 1996).
[20] P. C. Martin and J. Schwinger, Phys. Rev. **115**, 1342 (1959).
[21] N. Dori, M. Menon, L. Kilian, M. Sokolowski, L. Kronik, and E. Umbach, Phys. Rev. B **73**, 195208 (2006).
[22] M. L. Tiago, P. R. C. Kent, R. Q. Hood, and F. A. Reboredo, J. Chem. Phys. **129**, 084311 (2008).
[23] M. Palummo *et al.*, J. Chem. Phys. **131**, 084102 (2009).
[24] P. Umari, G. Stenuit, and S. Baroni, Phys. Rev. B **79**, 201104(R) (2009).
[25] Y. Ma, M. Rohlfing, and C. Molteni, Phys. Rev. B **80**, 241405(R) (2009).
[26] G. Stenuit, C. Castellarin-Cudia, O. Plekan, and V. Feyer *et al.*, Phys. Chem. Chem. Phys. **12**, 10812 (2010).
[27] O. Inganäs *et al.*, Adv. Mater. (Weinheim, Ger.) **22**, E100 (2010).
[28] The availability of experimental results for organic molecules in the gas phase is also an important factor directing our choice.
[29] X. Blase and P. Ordejón, Phys. Rev. B **69**, 085111 (2004).
[30] R. Bauernschmitt, Marco Häser, Oliver Treutler, and Reinhart Ahlrichs, Chem. Phys. Lett. **264**, 573 (1997).
[31] Michael Rohlfing, Peter Krüger, and Johannes Pollmann, Phys. Rev. B **52**, 1905 (1995).
[32] B. Farid, in *Electron Correlation in the Solid State*, edited by N. H. March (World Scientific, Singapore, 1999), p. 217, and references therein.
[33] José M. Soler *et al.*, J. Phys. Condens. Matter **14**, 2745 (2002).
[34] In good agreement with the results of Ref. 52 in the case of small molecules, we have verified in the case of the polyacenes that a triple-ζ+ double polarization (TZDP) basis for expending the Kohn-Sham orbitals increases the quasiparticle correction to the

ionization energy by less than 0.07 eV as compared to the DZP basis adopted in this paper.

[35]The cutoff radius used by the SIESTA package is chosen large enough so that the "normalization spilling" induced by truncating the Gaussian tails is less than 10^{-10}.

[36]M. S. Kaczmarski, Y. C. Ma, and M. Rohlfing, Phys. Rev. B **81**, 115433 (2010).

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

[38]F. Giustino, M. L. Cohen, and S. G. Louie, Phys. Rev. B **81**, 115105 (2010).

[39]J. A. Berger, L. Reining, and F. Sottile, Phys. Rev. B **82**, 041103 (2010).

[40]D. Foerster and P. Koval, J. Chem. Phys. **131**, 044103 (2009).

[41]For example, see A. Ipatov *et al.*, THEOCHEM **762**, 179 (2006).

[42]C. M. Reeves and M. C. Harrison, J. Chem. Phys. **39**, 11 (1963).

[43]K. Ruedenberg, R. C. Raffenetti, and R. D. Bardo, in *Energy, Structure and Reactivity*, proceedings of the 1972 Boulder Conference on Theoretical Chemistry (Wiley, New York, 1973).

[44]For a recent analysis, see I. Cherkes, S. Klaiman, and N. Miseyev, Int. J. Quantum Chem. **109**, 2996 (2009).

[45]F. Aryasetiawan and O. Gunnarsson, Phys. Rev. B **49**, 16214 (1994).

[46]I. S. Gradshteyn and I. M. Ryzhik, in *Tables of Integrals, Series and Products*, edited by A. Jeffrey and D. Zwillinger, 7th ed. (Academic, New York, 2007).

[47]W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge University Press, New York, 1996), p. 252.

[48]Most of the experimental data are taken from the NIST chemistry webbook at [http://webbook.nist.gov/chemistry/]. The band gaps are calculated as the difference of the experimental ionization energy and electronic affinity. We adopt the most recent experimental data. The provided ionization energies are the vertical ones.

[49]J. Berkowitz, J. Chem. Phys **70**, 2819 (1979). Previously calculated values using ΔSCF and transition state (TS) techniques yield 6.52 and 6.55 eV, respectively. See M.-S. Liao and S. Scheiner, J. Chem. Phys. **114**, 9780 (2001).

[50]T. Pasinzki, M. Krebsz, and G. Vass, J. Mol. Struct. **966**, 85 (2009).

[51]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[52]C. Rostgaard, K. W. Jacobsen, and K. S. Thygesen, Phys. Rev. B **81**, 085103 (2010); K. Kaasbjerg and K. S. Thygesen, *ibid.* **81**, 085102 (2010).

[53]T. A. Niehaus, M. Rohlfing, F. Della Sala, A. Di Carlo, and Th. Frauenheim, Phys. Rev. A **71**, 022508 (2005).

[54]The B3LYP/6-31G(d) HOMO-LUMO gap of C_{60} was found to be 2.77 eV in Refs. 55 and 56 and that of the H_2P porphyrin was shown to be 2.9 eV in Ref. 57.

[55]M. K. Shukla and J. Leszczynski, Chem. Phys. Lett. **428**, 317 (2006).

[56]Z. Zhang *et al.*, J. Phys. Chem. C **112**, 19158 (2008).

[57]K. A. Nguyen, P. N. Day, and R. Pachter, J. Chem. Phys. **110**, 9135 (1999).

[58]We observe that our largest discrepancies with plane-wave-based G_0W_0(LDA) calculations occurs for rather large systems. Further work is needed to know if such discrepancies stem from the Gaussian basis or possibly from the difficulty in converging plane-wave calculations in terms of unit cell and dielectric matrix size or number of conduction bands. Tests on benzene with the PWSCF/Yambo packages (Refs. 59 and 60) show that too small a cutoff on the dielectric matrix size, or the summation over a too limited number of conduction bands, open the band gap as a result both of underscreening (unconverged W) and by neglecting correlation with higher states (unconverged G).

[59]P. Giannozzi *et al.*, J. Phys. Condens. Matter **21**, 395502 (2009).

[60]A. Marini, C. Hogan, M. Grüning, and D. Varsano, Comput. Phys. Commun. **180**, 1392 (2009).

[61]Apart from the band gap, the opening up to 1.5 eV for the HOMO/HOMO-1 energy difference in PTCDA (instead of the 0.7 eV LDA value) and up to 1.2 eV for the HOMO/HOMO-2 energy difference in H_2P (instead of the 0.83 eV LDA value) are in excellent agreement with the plane-wave G_0W_0(LDA) calculations of Refs. 21 and 23.

[62]P. H. Hahn, W. G. Schmidt, and F. Bechstedt, Phys. Rev. B **72**, 245425 (2005).