JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

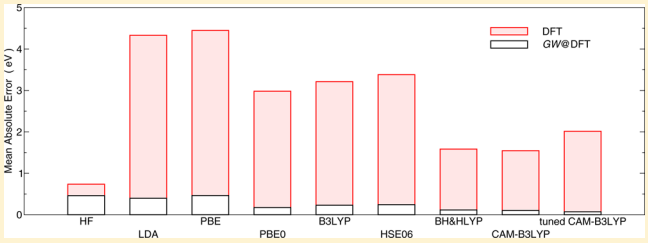
Benchmarking the Starting Points of the *GW* Approximation for Molecules

Fabien Bruneval*,† and Miguel A. L. Marques‡

†CEA, DEN, Service de Recherches de Métallurgie Physique, F-91191 Gif-sur-Yvette, France

‡Université de Lyon, F-69000 Lyon, France and LPMCN, CNRS, UMR 5586, Université Lyon 1, F-69622 Villeurbanne, France

ABSTRACT: The *GW* approximation is nowadays being used to obtain accurate quasiparticle energies of atoms and molecules. In practice, the *GW* approximation is generally evaluated perturbatively, based on a prior self-consistent calculation within a simpler approximation. The final result thus depends on the choice of the self-consistent mean-field chosen as a starting point. Using a recently developed *GW* code based on Gaussian basis functions, we benchmark a wide range of starting points for perturbative *GW*, including Hartree−Fock, LDA, PBE, PBE0, B3LYP, HSE06, BH&HLYP, CAM-B3LYP, and tuned CAM-B3LYP. In the evaluation of the ionization energy, the hybrid functionals are clearly superior results starting points when compared to Hartree−Fock, to LDA, or to the semilocal approximations. Furthermore, among the hybrid functionals, the ones with the highest proportion of exact-exchange usually perform best. Finally, the reliability of the frozen-core approximation, that allows for a considerable speed-up of the calculations, is demonstrated.



1. INTRODUCTION

In theoretical condensed matter physics, the *GW* approximation[1] has been extremely successful in predicting the band gap of materials.[2−4] However, besides the pioneering works of Shirley and Martin,[5] the *GW* approximation was not used for atoms and molecules until very recently. In the past few years, there has been a blooming literature about *GW* calculations for gas-phase molecules.[6−17] Unfortunately the published results show large deviations between themselves. The present situation is further complicated because of the numerous ways of performing *GW* calculations and the diverse approximations used in practice. First, the basis set convergence of *GW* calculations is known to be an important issue[14,15,18] and could explain the scattering of the published results. Furthermore, the accuracy of the pseudopotential approximation, often used in this context, has been constantly questioned in the past decade.[19−21] Finally, *GW* calculations are generally performed perturbatively, based on a prior self-consistent mean-field calculations,[3] a procedure customary called G_0W_0. This leads to a vast variety of possible choices for the starting mean-field calculations: Hartree−Fock (HF), density-functional theory (DFT)[22] within the local, semilocal, or hybrid flavors.

There is nowadays a stringent need for rationalizing the available data. It would therefore be desirable to conduct unambiguous calculations that could be accurately reproduced by any other research group. The Gaussian basis-set code, named MOLGW, was recently developed[15] to address this goal. The philosophy behind this development is to completely disregard efficiency in order to have the highest possible reliability. Therefore, there is basically no technical approximation besides the choice of the basis set for the wave functions.

In this Article, our purpose is (i) to produce highly converged ionization energies that can be used in the future as a benchmark and (ii) to address the issue of the starting mean-field approximation used to initiate perturbative *GW* calculations. Besides the usual approximations [HF, local density approximation (LDA), generalized gradient approximation (PBE)[23]], we also used a wide variety of hybrid functionals that incorporate different fractions of (screened) exact-exchange: PBE0,[24] B3LYP,[25] HSE06,[26] BH&HLYP,[27] CAM-B3LYP,[28] and tuned CAM-B3LYP.[29]

The remainder of this article is organized as follows: In section 2 we briefly introduce the methods used in the our calculations, with an emphasis on the basis-set convergence issue. In section 3 we evaluate the performance of the different starting points for predicting the ionization energy of molecules; and finally in section 4 we discuss the reasons for the scattering of the values in the literature.

2. METHODS

2.1. G_0W_0 Quasiparticle Energy. The *GW* approximation to the self-energy arises from a first-order perturbation with respect to the screened Coulomb interaction *W*. This screened Coulomb interaction is obtained from the Hartree-only response. It is sometimes also called direct random-phase approximation (RPA) screening, as opposed to the full RPA that would also include exchange response. The *GW* self-energy

Received: September 26, 2012
Published: December 3, 2012

   dx.doi.org/10.1021/ct300835h | J. Chem. Theory Comput. 2013, 9, 324−329

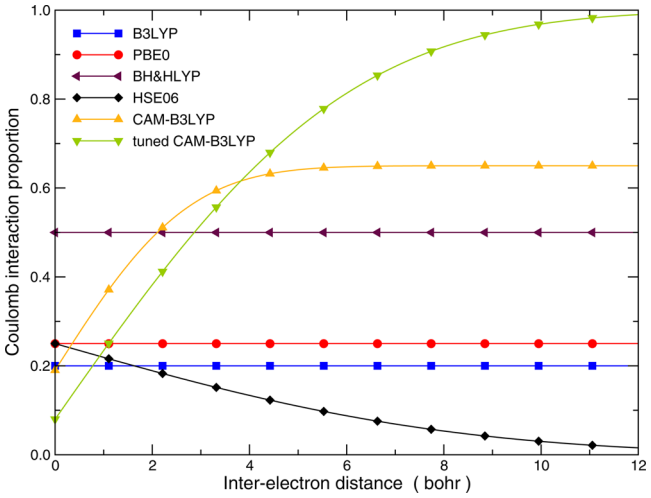


Figure 1. Coulomb interaction proportion $\alpha + \beta\ \mathrm{erf}(\omega r)$ effectively used by different hybrid functionals of DFT as a function of the interelectron distance r. On this plot, HF would be an horizontal line at 1.0 and LDA, PBE would be an horizontal line at 0.0. Hybrid functionals offer a wide variety of choice in between these two extremes: full range, short-range, or long-range.

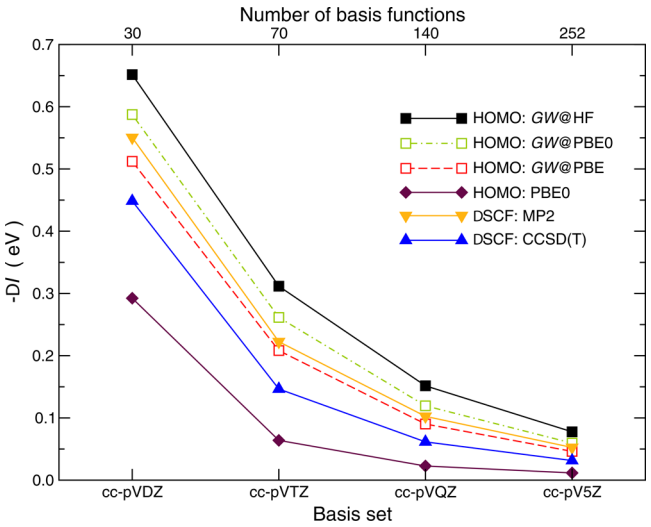


Figure 2. Basis set convergence of the error in the HOMO energy of the carbon monoxide molecule (or minus the ionization energy) for different techniques and approximations. The CCSD(T) and MP2 values were obtained through total energy differences.

is then obtained by frequency convolution of the Green's function G with the screened interaction W.

In the Green's function theory, the poles of the Green's function have a physical meaning: they correspond to the vertical electron addition or removal energies. In particular, the pole arising from highest occupied molecular orbital (HOMO) can be directly interpreted as minus the ionization energy. Conversely, the pole corresponding to the lowest unoccupied molecular orbital (LUMO) accounts for minus the electron affinity.

In practice, the Green's function used for the GW approximation is generally obtained perturbatively through an initial self-consistent (generalized) Kohn–Sham (GKS) calculation.[30] The GKS framework encompasses not only the local and semilocal approximations to the exchange-correlation potential but also fully nonlocal approximations, such as HF and hybrid functionals. Once the GKS Hamiltonian has been diagonalized, the GW self-energy is build using the GKS wave functions and eigenvalues. If one assumes that the difference between the GW Hamiltonian and the GKS Hamiltonian is small enough, the GW quasiparticle energies ε_i^{GW} can be then calculated as a first-order perturbation with respect to the GKS eigenvalues $\varepsilon_i^{\mathrm{GKS}}$:

$$\varepsilon_i^{GW} = \varepsilon_i^{\mathrm{GKS}} + Z_i\langle i|\Sigma_x + \Sigma_c^{GW}(\varepsilon_i^{\mathrm{GKS}}) - v_{xc}^{\mathrm{GKS}}|i\rangle \quad (1)$$

where Σ_x stands for the exact-exchange operator, $\sum_c^{GW}$ stands for the GW correlation self-energy operator, and v_{xc}^{GKS} is the GKS exchange-correlation potential. The renormalization factor Z_i,

$$Z_i = \left(1 - \left.\frac{\partial\langle i|\Sigma_c^{GW}(\varepsilon)|i\rangle}{\partial\varepsilon}\right|_{\varepsilon=\varepsilon_i^{\mathrm{GKS}}}\right)^{-1} \quad (2)$$

is due to the dependence of Σ_c^{GW} on the quasiparticle energy ε_i^{GW}. The spin variable has been dropped for convenience, since all the presented calculations are spin-restricted.

All the GW energies reported here use the quasiparticle eq 1. As the construction is perturbative, the final result depends of course on the GKS scheme selected as a starting point. The outcome of this procedure will be named GW@HF, for G_0W_0 based on HF inputs; GW@PBE, for G_0W_0 based PBE inputs, etc.

2.2. Generalized Kohn–Sham Starting Points. In the GKS formalism, the exchange-correlation potential is allowed to be fully nonlocal. The exchange-correlation approximations under scrutiny here can be written in the form:

$$v_{xc}^{\mathrm{GKS}}(\mathbf{r},\mathbf{r}') = \alpha\Sigma_x(\mathbf{r},\mathbf{r}') + \beta\Sigma_x^{\omega\mathrm{LR}}(\mathbf{r},\mathbf{r}') + v_{xc}^{\mathrm{semilocal}}(\mathbf{r}) \quad (3)$$

The exchange-correlation potential has a semilocal contribution $v_{xc}^{\mathrm{semilocal}}$ that only depends on the density and its gradient. The exact-exchange operator Σ_x is included with a fixed proportion α and the long-range exchange operator $\sum_x^{\omega\mathrm{LR}}$,

$$\Sigma_x^{\omega\mathrm{LR}}(\mathbf{r},\mathbf{r}') = -\sum_{i\mathrm{occ}} \varphi_i(\mathbf{r})\varphi_i^*(\mathbf{r}')\frac{\mathrm{erf}(\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} \quad (4)$$

with a given multiplicative factor β. The range of this interaction is controlled by the inverse radius ω.

Figure 1 illustrates the proportion of the Coulomb interaction that is effectively used to calculated the exact-exchange for hybrid functionals. The local and semilocal form of the exchange-correlation potential, such as LDA[31] or PBE,[23] have $\alpha = \beta = 0$. At the opposite end, the HF approximation has $\alpha = 1$, $\beta = 0$, and no $v_{xc}^{\mathrm{semilocal}}$. In between these two extremes, the available hybrid functionals propose different choices for α and β. The full range hybrids have $\beta = 0$ and α varying from 0.20 for B3LYP[25] to 0.50 for BH&HLYP,[27] passing through 0.25 for PBE0.[24]

There are two categories of range-separated hybrids: the short-ranged hybrids and the long-ranged hybrids. Among the short-ranged hybrids, the HSE06 functional[26] is very popular nowadays in the study of solids. In this case, the long-range exchange precisely compensates the full-range exchange $\alpha = -\beta = 0.25$. Furthermore, the value $\omega = 0.11$ bohr^{-1} was obtained by optimization on a set of molecules.[26] As representatives for the long-ranged hybrid category, we picked up CAM-B3LYP[28] ($\alpha = 0.19$, $\beta = 0.46$, and $\omega = 0.33$ bohr^{-1}) and the tuned CAM-

dx.doi.org/10.1021/ct300835h | J. Chem. Theory Comput. 2013, 9, 324–329

B3LYP[29] (α = 0.0799, β = 0.9201, and ω = 0.15 bohr^{-1}). At long-range, the exchange-correlation potential should behave as the full exact-exchange. Tuned CAM-B3LYP was specifically designed to capture this feature, as $\alpha + \beta = 1$.

These approximations were implemented in the exchange-correlation library LIBXC[32] that is used by MOLGW.

2.3. Basis Set Convergence. Our implementation of the perturbative *GW* method, or G_0W_0, in MOLGW is based on Cartesian Gaussian basis functions.[15] The basis functions $\phi(\mathbf{r})$ are defined as a contraction of M atom-centered Gaussian functions with the same generalized angular momentum (n_x,n_y,n_z):

$$\phi(\mathbf{r}) = (x - A_x)^{n_x}(y - A_y)^{n_y}(z - A_z)^{n_z} \sum_{i=1}^{M} C_i e^{-\zeta_i(\mathbf{r}-\mathbf{A})^2} \quad (5)$$

for an atom located at coordinate $\mathbf{A} = (A_x,A_y,A_z)$. With Cartesian Gaussian basis functions, the number of *d*-like orbitals is 6 instead of 5 for pure Gaussian functions, the number of *f*-like orbitals is 10 instead of 7, etc.

In the present study, we use the coefficients C_i and ζ_i from the Dunning basis set family,[33,34] as can be obtained from a web-available database.[35,36] These basis sets form a series of increasingly accurate basis sets labeled cc-pVXZ, where X can be D, T, Q, and 5 for respectively double-, triple-, quadruple-, and quintuple-zeta basis sets. The number of polarization functions and the maximum angular momentum also increase with increasing *X*. The Dunning basis sets are particularly suited for extrapolation to the complete basis set limit. Even though the Dunning basis sets were primarily designed for pure Gaussian basis sets, we use them with Cartesian basis functions. This simply adds some computational effort and without changing considerably the final result, but computational efficiency is clearly not our goal here. The Cartesian Gaussian Coulomb integrals are obtained from the library LIBINT.[37]

Figure 2 shows the typical convergence of the HOMO energy with size of the basis set. All small molecules studied here exhibit the same behavior, so we only display the example of carbon monoxide CO. As previously reported,[14,15] the convergence of the G_0W_0 HOMO energy in Gaussian basis is very slow for small molecules. Regardless of the starting point (*GW*@HF, *GW*@PBE0, *GW*@PBE), the accuracy of 0.2 eV is reached only at the level of a cc-pVQZ basis set, which means as many as 140 basis functions for the CO molecule. Furthermore, the convergence of the G_0W_0 HOMO energy is much slower than the PBE0 HOMO energy. However, it is in line with the convergence rate of correlated methods that depend on virtual orbitals, such as MP2[38] and coupled-cluster including double and triple excitations (perturbatively) [CCSD(T)].[39]

It is striking to note that for all the molecules of the present study and as exemplified for CO in Figure 2, the HOMO energy within the G_0W_0 approximation convergences from above. Conversely, the ionization energies ($I = -\varepsilon_{\mathrm{HOMO}}$) converge from below. This statement will help us rationalizing the comparison of our data with the published data.[11,14,16] Note also that all calculations presented in the next section are performed within the cc-pVQZ basis set. This implies that the calculated G_0W_0 HOMO energies then lie at most 0.2 eV above the complete basis set limit, which sets the error bar in our benchmark.

3. RESULTS: IONIZATION ENERGIES

Recently, Rostgaard, Jacobsen, and Thygesen[11] proposed a benchmark of 34 small closed-shell molecules in order to test the different flavors of the *GW* approximation. Among other results, they included data calculated at the *GW*@HF and *GW*@PBE levels. These data should be readily comparable with the more recently published results of Caruso and co-workers[16] and of Ren et al.[40] Unfortunately, it turns out that the discrepancy between the aforementioned studies is surprisingly large: the difference in the ionization energy of the F_2 molecule is larger than 1.7 eV! Furthermore, the differences cannot be attributed to the sole change in the relaxed geometries of the molecules. Considering the magnitude of the uncertainty, it is certainly difficult to extract a conclusive benchmark from these data.

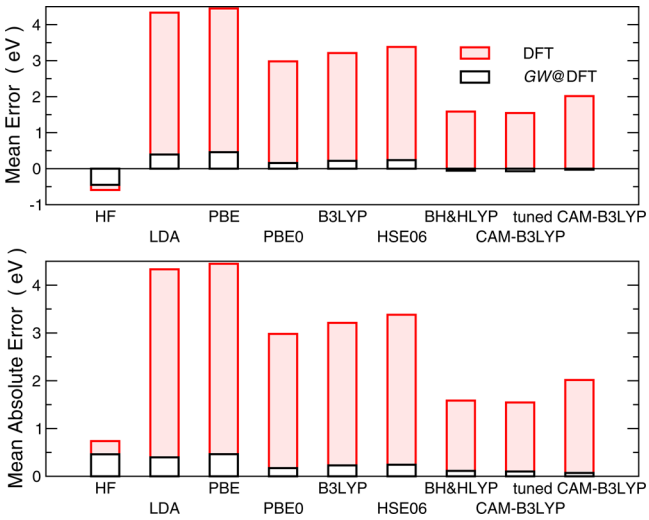


Figure 3. Deviation of the HOMO energies with respect to the CCSD(T) total energy difference. The upper panel shows the mean-error (ME), and the lower panel, the mean-absolute error (MAE) for the 34 molecule set. The shaded bar corresponds to the HOMO energy of the GKS starting point, and the white bar shows the corresponding *GW*@GKS HOMO energy.

Therefore, we propose to reconsider the same 34 molecule benchmark with our own implementation of the *GW* approximation. As underlined in ref 15, the code MOLGW has been specifically written with the aim of removing all the technical approximations that usually affect *GW* results. In particular, the frequency convolution in the *GW* self-energy is performed analytically thanks to the solution of the RPA equations.[8] We also avoid using auxiliary basis set for the expansion of the Coulomb operator.[6] In summary, there is no further approximation besides the choice of the basis set for the wave functions.

The 34 molecule set is extremely convenient since it forms a subset of the famous G2/97 test set.[41] We employ here the relaxed geometry at the MP2 level within the 6-31G(d) basis set, as published in ref 41. As we also use a widespread basis set, namely the Dunning cc-pVQZ, our calculations should be straightforward to reproduce.

In order to evaluate the quality of each starting point, we prefer not to compare with experimental value as it is customary.[11,16] The experimental values indeed include several physical effects that the *GW* HOMO energies do not have. First, *GW* energies are vertical transitions, whereas the

Table 1. G_0W_0 HOMO Energy of the 34 Molecules Employing Different Starting Points with the cc-pVQZ Basis Set[a]

	GW@										
starting point	HF	LDA	PBE	PBE0	B3LYP	HSE06	BH&HLYP	CAM-B3LYP	tuned CAM-B3LYP	CCSD(T)	exp
LiH	−8.20	−7.24	−7.07	−7.66	−7.53	−7.47	−7.91	−8.03	−8.07	−7.94	
Li_2	−5.36	−5.13	−5.12	−5.29	−5.23	−5.19	−5.30	−5.32	−5.38	−5.17	
LiF	−11.62	−10.61	−10.37	−10.93	−10.82	−10.89	−11.29	−11.49	−11.45	−11.51	
Na_2	−4.98	−4.91	−4.89	−4.97	−4.96	−4.91	−4.97	−4.98	−5.01	−4.82	
NaCl	−9.36	−8.56	−8.43	−8.82	−8.77	−8.70	−9.06	−9.15	−9.22	−9.13	−9.80
CO	−14.97	−13.63	−13.55	−14.00	−13.92	−13.92	−14.36	−14.26	−14.11	−14.05	
CO_2	−14.38	−13.45	−13.32	−13.68	−13.57	−13.59	−13.91	−13.91	−13.82	−13.78	
CS	−13.08	−10.97	−10.93	−11.43	−11.31	−11.33	−11.79	−11.69	−11.55	−11.45	
C_2H_2	−11.65	−11.10	−11.08	−11.27	−11.23	−11.21	−11.40	−11.41	−11.41	−11.42	−11.49
C_2H_4	−10.85	−10.39	−10.37	−10.53	−10.52	−10.48	−10.65	−10.67	−10.66	−10.69	−10.68
CH_4	−14.86	−14.07	−14.03	−14.30	−14.27	−14.23	−14.52	−14.53	−14.48	−14.40	−14.40[44]
CH_3Cl	−11.74	−11.02	−10.98	−11.21	−11.18	−11.15	−11.41	−11.43	−11.41	−11.41	−11.29
CH_3OH	−11.69	−10.70	−10.64	−10.97	−10.89	−10.88	−11.20	−11.22	−11.17	−11.08	−10.96
CH_3SH	−9.81	−9.18	−9.17	−9.36	−9.35	−9.30	−9.53	−9.55	−9.53	−9.49	
Cl_2	−12.01	−11.22	−11.16	−11.42	−11.38	−11.36	−11.63	−11.63	−11.59	−11.62	
ClF	−13.32	−12.43	−12.33	−12.61	−12.57	−12.55	−12.87	−12.85	−12.79	−12.82	−12.77
F_2	−16.59	−15.38	−15.19	−15.66	−15.56	−15.56	−16.00	−16.00	−15.84	−15.85	
HOCl	−11.83	−10.92	−10.85	−11.14	−11.08	−11.07	−11.37	−11.37	−11.32	−11.30	
HCl	−12.97	−12.37	−12.35	−12.54	−12.51	−12.48	−12.69	−12.72	−12.71	−12.74	
H_2O_2	−12.13	−11.12	−11.02	−11.38	−11.28	−11.29	−11.63	−11.65	−11.57	−11.49	−11.70
H_2CO	−11.56	−10.61	−10.51	−10.87	−10.77	−10.77	−11.10	−11.12	−11.07	−10.95	
HCN	−13.86	−13.27	−13.20	−13.44	−13.39	−13.38	−13.59	−13.61	−13.59	−13.64	
HF	−16.39	−15.62	−15.51	−15.81	−15.73	−15.72	−16.01	−16.10	−16.02	−16.09	−16.12
H_2O	−13.04	−12.22	−12.15	−12.44	−12.36	−12.36	−12.63	−12.69	−12.64	−12.64	
NH_3	−11.38	−10.53	−10.50	−10.78	−10.72	−10.70	−10.98	−11.00	−10.97	−10.92	−10.82
N_2	−16.48	−15.08	−14.98	−15.45	−15.33	−15.35	−15.80	−15.72	−15.57	−15.49	
N_2H_4	−10.78	−9.91	−9.87	−10.15	−10.09	−10.08	−10.36	−10.39	−10.34	−10.24	
SH_2	−10.67	−10.11	−10.10	−10.27	−10.26	−10.22	−10.42	−10.43	−10.42	−10.43	−10.50
SO_2	−13.12	−11.96	−11.83	−12.23	−12.15	−12.14	−12.55	−12.51	−12.41	−12.41	−12.50
PH_3	−10.79	−10.21	−10.21	−10.39	−10.39	−10.34	−10.55	−10.55	−10.53	−10.49	−10.59
P_2	−10.57	−10.13	−10.12	−10.27	−10.24	−10.23	−10.37	−10.35	−10.36	−10.76	−10.62
SiH_4	−13.31	−12.43	−12.40	−12.72	−12.68	−12.63	−12.96	−12.98	−12.91	−12.82	
Si_2H_6	−11.20	−10.35	−10.38	−10.62	−10.60	−10.53	−10.83	−10.82	−10.76	−10.69	−10.53
SiO	−11.98	−11.10	−11.03	−11.34	−11.24	−11.25	−11.53	−11.53	−11.53	−11.55	
ME	−0.45	0.39	0.46	0.16	0.22	0.24	−0.06	−0.07	−0.03		
MAE	0.46	0.40	0.46	0.17	0.23	0.24	0.11	0.10	0.07		

[a]The CCSD(T) total energy difference with the same basis set is given as a reference. The mean error (ME) and the mean absolute error (MAE) measure the deviation with respect to the CCSD(T) reference. The experimental vertical ionization energies are also given for comparison. The experimental values are taken from ref 43 unless otherwise stated.

experimental values are most often nonvertical energy differences, including the relaxation of the atomic positions in the excited state. Second, the experimental values include the zero-point energy of the nucleus motion for both the ground-state and the excited state, whereas the purely electronic *GW* energies do not. Finally, as the basis set is still not complete for the cc-pVQZ basis set, it induces an overestimation of the HOMO energy of 0.1−0.2 eV at most. For all these reasons, we prefer to evaluate the reference ionization energies through total energy differences obtained with a coupled-cluster method, namely CCSD(T), as obtained from Gaussian 09.[42] This allows us to employ precisely the same basis set and the same geometry for *GW* and CCSD(T).

Figure 3 summarizes the deviation of the G_0W_0 HOMO energies with respect to the reference CCSD(T) for the 34 molecule set. Whatever the choice of starting point, the G_0W_0 HOMO energy is always greatly improved with respect to that starting point. As it is well-known, all GKS schemes (with the exception of HF) perform poorly for the HOMO energy, with the MAE as large as several electronvolts. Applying *GW* on top always reduces the MAE down to 0.5 eV at most. It appears that the *GW* approximation has a slight tendency to underestimate the reference HOMO energy. Therefore, if the starting point already underestimates the HOMO energy, such as HF, the final *GW*@HF is still too low. On the contrary, when the starting point is slightly too high, such as for the hybrid functionals BH&HLYP, CAM-B3LYP, or tuned CAM-B3LYP, the final ME almost vanishes, owing to a fortunate compensation of errors. Finally, for the present benchmark, HF, LDA, and PBE perform equally well.

Not surprisingly, the most efficient GKS starting points belong to the hybrid functional family. However, the best hybrid functionals are the functionals that contains the largest fraction of exact-exchange. PBE0, B3LYP, and HSE06 that contains at most 25% of exact-exchange are not the best starting points, as far as the HOMO energy is concerned, while BH&HLYP, CAM-B3LYP, and tuned CAM-B3LYP, which

dx.doi.org/10.1021/ct300835h | J. Chem. Theory Comput. 2013, 9, 324−329

includes significantly more exact-exchange, perform impressively well, with a final MAE of 0.1 eV.

Table 1 provides the complete list of the G_0W_0 HOMO energies calculated from the different starting points and compared to the reference CCSD(T) values. We should note that our values exhibit noticeable differences with respect to some previously published results,[11,16] but much smaller differences with respect to ref 40. We believe our methodology offers easily reproducible data, since there is no technical approximation besides the basis set, and since the basis sets we use are published. We therefore hope our data can be used as a benchmark in the future.

A close look at Table 1 shows that *GW*@HF and *GW*@LDA HOMO energies always bracket the reference value, with the exception of Na_2 and P_2. The difference between *GW*@HF and *GW*@LDA, the most extreme cases, is generally of the order of 1 eV. It is then not surprising that hybrid functionals improve the G_0W_0 HOMO energy. For individual molecules, the maximum error of BH&HLYP, CAM-B3LYP, and tuned CAM-B3LYP is never higher than 0.2 eV, with the exception of P_2 once again. These functionals should therefore be used as starting points for predictive G_0W_0 calculations of ionization energies of molecules.

Our implementation does not rely on any auxiliary basis set to represent the Coulomb operator. This choice made the calculation highly accurate but also computationnally demanding. However, the nice agreement between our data and the results of Ren et al.[40] shows that the use of an auxiliary basis (or alternatively named the resolution of the identity technique) seems to affect the final results little.

4. DISCUSSION: FROZEN CORE APPROXIMATION

As mentioned before, the data presented in the previous section exhibit non-negligible discrepancies with earlier works.[11,16]

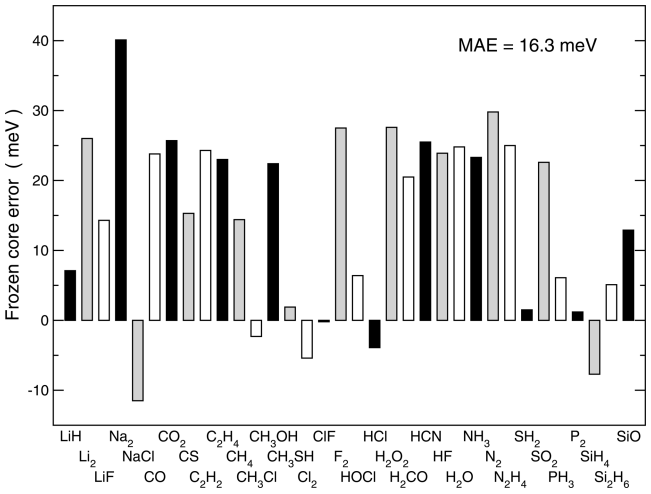


Figure 4. Error due to the frozen core approximation in the G_0W_0 correlation employing the *GW*@BH&HLYP approximation and the cc-pVQZ basis set.

Here, we would like to discuss the possible origins for the differences. In particular, the frozen core approximation is sometimes mentioned as a possible cause of error.

Freezing the core means skipping the core states in the calculation of the correlation part of *GW*. The technique allows one to save time both in the calculation of the screening (the RPA equation does not include transitions from or to the core states) and in the calculation of the *GW* self-energy (the Green's function does not have poles at the core state energies). The frozen-core approximation for *GW* is completely analogous to the frozen-core approximation for MP2 calculations.

The consequences of the frozen-core approximation can be evaluated from Figure 4 that shows the difference between frozen-core and all-electron *GW* calculations. The approximation is completely innocuous, with a maximum error of 40 meV. This confirms the good quality of the frozen core approximation that was observed in solids.[20,21] Moreover, the computational effort is noticeably reduced by the frozen-core approximation. For instance, for Na_2 the diagonalized RPA matrix is reduced from 3014 × 3014 to 274 × 274, with a change of only 40 meV in the HOMO energy. We therefore advocate the use of the frozen-core approximation in *GW* calculations of molecules.

Considering the very slow convergence of the G_0W_0 HOMO energy with respect to the basis set size, we would tentatively attribute the discrepancy between the published data to the difference in the basis sets. When the same basis set is employed, the differences are completely under control. Comparing with ref 14, where cc-pVTZ basis sets were used together with B3LYP relaxed geometries, our ionization energy of C_2H_2 is 11.59 eV against their value of 11.55 eV, and our ionization energy of C_2H_4 is 10.75 eV against 10.69 eV. This comparison gives a compelling hint that the basis set choice is indeed a crucial parameter.

5. CONCLUSIONS

Our implementation of the *GW* approximation in Gaussian basis sets has allowed us to provide highly accurate estimates of the performance of the *GW* approximation for molecules. We assessed the performance of different GKS schemes as starting points for a subsequent perturbative *GW* calculation. We used the benchmark of 34 small closed-shell molecules proposed in ref 11; however, we examined the error with respect to the higher level quantum-chemistry method, namely CCSD(T), rather than with respect to experiment.

For all the benchmarked molecules but two, the reference ionization energy lies between the *GW*@HF and *GW*@LDA HOMO energies. The hybrid functional outperforms largely the simpler local or semilocal approximations of DFT. Among the different hybrid functionals, the most efficient ones, as a starting points for the G_0W_0 procedure, are definitely the functionals that include the most significant contribution of exact-exchange, such as BH&HLYP, CAM-B3LYP, and tuned CAM-B3LYP.

We also discussed the basis set convergence issue in the G_0W_0 calculations for small molecules. G_0W_0 HOMO energies converge very slowly indeed, with a convergence rate similar to MP2. Therefore, when comparing two different published G_0W_0 results, the basis sets should be compared with great care.

Finally, we proved the reliability of the frozen-core approximation for the correlation part of the *GW* self-energy. This approximation can induce a noticeable acceleration of the calculations, at the expense of a very small error in the *GW* energies ($\lesssim$40 meV).

AUTHOR INFORMATION

Corresponding Author

*E-mail: fabien.bruneval@cea.fr.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was performed using HPC resources from GENCI-CCRT (Grant 2012-gen6018). M.A.L.M. acknowledges support from the French ANR (ANR-08-CEXC8-008-01).

■ REFERENCES

(1) Hedin, L. *Phys. Rev.* **1965**, *139*, A796.
(2) Strinati, G.; Mattausch, H. J.; Hanke, W. *Phys. Rev. B* **1982**, *25*, 2867−2888.
(3) Hybertsen, M. S.; Louie, S. G. *Phys. Rev. B* **1986**, *34*, 5390−5413.
(4) Aryasetiawan, F.; Gunnarsson, O. *Rep. Prog. Phys.* **1998**, *61*, 237−312.
(5) Shirley, E. L.; Martin, R. M. *Phys. Rev. B* **1993**, *47*, 15404.
(6) Rohlfing, M. *Int. J. Quantum Chem.* **2000**, *80*, 807.
(7) Grossman, J. C.; Rohlfing, M.; Mitas, L.; Louie, S. G.; Cohen, M. L. *Phys. Rev. Lett.* **2001**, *86*, 472−475.
(8) Tiago, M. L.; Chelikowsky, J. R. *Phys. Rev. B* **2006**, *73*, 205334.
(9) Morris, A. J.; Stankovski, M.; Delaney, K. T.; Rinke, P.; García-González, P.; Godby, R. W. *Phys. Rev. B* **2007**, *76*, 155106.
(10) Bruneval, F. *Phys. Rev. Lett.* **2009**, *103*, 176403.
(11) Rostgaard, C.; Jacobsen, K. W.; Thygesen, K. S. *Phys. Rev. B* **2010**, *81*, 085103.
(12) Blase, X.; Attaccalite, C.; Olevano, V. *Phys. Rev. B* **2011**, *83*, 115103.
(13) Foerster, D.; Koval, P.; Sanchez-Portal, D. *J. Chem. Phys.* **2011**, *135*, 074105.
(14) Ke, S.-H. *Phys. Rev. B* **2011**, *84*, 205415.
(15) Bruneval, F. *J. Chem. Phys.* **2012**, *136*, 194107.
(16) Caruso, F.; Rinke, P.; Ren, X.; Scheffler, M.; Rubio, A. *Phys. Rev. B* **2012**, *86*, 081102.
(17) Körzdörfer, T.; Marom, N. *Phys. Rev. B* **2012**, *86*, 041110.
(18) Bruneval, F.; Gonze, X. *Phys. Rev. B* **2008**, *78*, 085125.
(19) Ku, W.; Eguiluz, A. G. *Phys. Rev. Lett.* **2002**, *89*, 126401.
(20) Gómez-Abal, R.; Li, X.; Scheffler, M.; Ambrosch-Draxl, C. *Phys. Rev. Lett.* **2008**, *101*, 106404.
(21) Li, X.-Z.; Gómez-Abal, R.; Jiang, H.; Ambrosch-Draxl, C.; Scheffler, M. *New J. Phys.* **2012**, *14*, 023006.
(22) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.
(23) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
(24) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029.
(25) Stephens, P.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
(26) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 224106.
(27) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
(28) Yanai, T.; Tew, D. P.; Handy, N. C. *J. Chem. Phys.* **2004**, *393*, 51.
(29) Okuno, K.; Shigeta, Y.; Kishi, R.; Miyasaka, H.; Nakano, M. *J. Photochem. Photobiol.* **2012**, *A235*, 29.
(30) Seidl, A.; Görling, A.; Vogl, P.; Majewski, J. A.; Levy, M. *Phys. Rev. B* **1996**, *53*, 3764.
(31) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
(32) Marques, M. A. L.; Oliveira, M. J. T.; Burnus, T. *Chem. Phys. Chem.* **2012**, *183*, 2272.
(33) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
(34) Woon, D. E.; D., T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.
(35) Feller, D. *J. Comput. Chem.* **1996**, *17*, 1571.
(36) Schuchardt, K.; Didier, B.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. *J. Chem. Inf. Model.* **2007**, *47*, 1045.
(37) http://sourceforge.net/p/libint/ (accessed Nov 22, 2012).
(38) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618−622.
(39) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.
(40) Ren, X.; Rinke, P.; Blum, V.; Wieferink, J.; Tkatchenko, A.; Sanfilippo, A.; Reuter, K.; Scheffler, M. *New J. Phys.* **2012**, *14*, 053020.
(41) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 42.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision C.01. 2009; Gaussian Inc.: Wallingford CT, 2009.
(43) http://cccbdb.nist.gov (accessed Nov 22, 2012).
(44) Bieri, G.; Åsbrink, L.; von Niessen, W. *J. Electron Spectrosc. Relat. Phenom.* **1982**, *27*, 129.