GW and Bethe-Salpeter study of small water clusters

Xavier Blase, Paul Boulanger, Fabien Bruneval, Marivi Fernandez-Serra, and Ivan Duchemin

Citation: The Journal of Chemical Physics **144**, 034109 (2016);
View online: https://doi.org/10.1063/1.4940139
View Table of Contents: http://aip.scitation.org/toc/jcp/144/3
Published by the American Institute of Physics

GW and Bethe-Salpeter study of small water clusters

Xavier Blase,[1,a)] Paul Boulanger,[1] Fabien Bruneval,[2] Marivi Fernandez-Serra,[3,4] and Ivan Duchemin[5]
[1]*CNRS, Institut NEEL, F-38042 Grenoble, France*
[2]*CEA, DEN, Service de Recherches de Métallurgie Physique, F-91191 Gif-sur-Yvette, France*
[3]*Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794-3800, USA*
[4]*Institute for Advanced Computational Sciences, Stony Brook University, Stony Brook, New York 11794-3800, USA*
[5]*INAC, SP2M/L_Sim, CEA/UJF Cedex 09, 38054 Grenoble, France*

(Received 9 November 2015; accepted 6 January 2016; published online 20 January 2016)

We study within the *GW* and Bethe-Salpeter many-body perturbation theories the electronic and optical properties of small $(H_2O)_n$ water clusters (n = 1-6). Comparison with high-level CCSD(T) Coupled-Cluster at the Single Double (Triple) levels and ADC(3) Green's function third order algebraic diagrammatic construction calculations indicates that the standard non-self-consistent G_0W_0@PBE or G_0W_0@PBE0 approaches significantly underestimate the ionization energy by about 1.1 eV and 0.5 eV, respectively. Consequently, the related Bethe-Salpeter lowest optical excitations are found to be located much too low in energy when building transitions from a non-self-consistent G_0W_0 description of the quasiparticle spectrum. Simple self-consistent schemes, with update of the eigenvalues only, are shown to provide a weak dependence on the Kohn-Sham starting point and a much better agreement with reference calculations. The present findings rationalize the theory to experiment possible discrepancies observed in previous G_0W_0 and Bethe-Salpeter studies of bulk water. The *increase* of the optical gap with increasing cluster size is consistent with the evolution from gas to dense ice or water phases and results from an enhanced screening of the electron-hole interaction. © *2016 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4940139]

I. INTRODUCTION

From biology[1] to water splitting cells for H_2 production,[2] the active role of water in processes such as electrocatalysis[3] or photocatalysis[4] calls for an exploration of its electronic and optical properties from a microscopic point of view. While a lot of work has been done on studying the atomic structure of water/metal and water/semiconductor interfaces,[5] the coupling between electronic and structural properties of the solvent at these interfaces has been much less explored. The excitation of electrons activated by an applied electric field or by an incoming photon requires an understanding of the energy offset of water occupied and excited states with respect to a metallic or semiconducting surface or to an active molecule that plays the role of a reactant in a chemical reaction.

From a theoretical standpoint, *ab initio* calculations aiming at understanding the electronic and optical properties of water have been mostly performed so far at the density functional theory (DFT) level. This is due to the need for performing expensive sampling over instantaneous configurations provided by molecular dynamic runs with large unit cells. Recently, a specific formulation of many-body perturbation theory (MBPT), namely, the so-called *GW* formalism,[6–11] has been applied to water in order to determine its band gap and band edges.[12] Further, impressive calculations within the *GW* formalism of the band offsets at water/semiconductor interfaces, for application, e.g., in water splitting technologies, are starting to appear,[13–18] benefiting from significant work at the algorithmic level to reduce the cost of such techniques.

While such a formalism was known to yield the band gap of standard extended semiconductors and insulators within 0.1-0.2 eV accuracy, the case of water may prove to be more problematic, with the calculated G_0W_0@PBE band gap and valence band maximum (VBM)[19] standing below, or on the lower side, the range of experimental values.[20–22] The estimated 8.7 eV ± 0.6 eV experimental band gap can, e.g., be compared to the G_0W_0@PBE 8.3 eV value obtained for liquid water configurations equilibrated with a standard water force field and 8.1 eV for water equilibrated with DFT within the Generalized Gradient Approximation (PBE[23]) approximation to the exchange and correlation energies. Similarly, the calculated VBM is found to be −9.0 eV and −8.8 eV for TIP4P and PBE water, to be compared to the experimental −9.3 eV to −10.1 eV energy range. The notation G_0W_0@PBE means that non-self-consistent *GW* calculations starting from PBE Kohn-Sham eigenstates were performed. More recently, a partially self-consistent GW_0 study of ice,[24] with update of the Green's function *G* only, was used to extrapolate the *GW* correction to the DFT Kohn-Sham gap of water, leading to an even smaller 7.3 eV band gap, that was compared to the 6.9 eV "adiabatic" experimental value by Coe and co-workers.[25,26] Concerning the optical properties of water, the threshold of optical absorption of water was found within the G_0W_0@LDA and Bethe-Salpeter equation (BSE) formalisms to be redshifted by about 1 eV

a)Electronic mail: xavier.blase@neel.cnrs.fr

as compared to experiments.[27,28] On the contrary, the optical absorption threshold of the water monomer was found to be in excellent agreement with experiment within the same G_0W_0/Bethe-Salpeter formalism but starting from an *ansatz* Hartree-Fock spectrum.[29–31]

These conflictive results call for an exploration of the merits of the *GW* formalism in its various implementations as compared to reference high-level quantum chemistry calculations. Along that line, very accurate studies providing valuable information on the electronic properties of water systems are available for small water clusters within the framework of CCSD(T) many-body wavefunction coupled-cluster techniques[32] or high-level ADC(3) Green's function third-order perturbation theory.[33] Even though not performed on liquid water, such calculations provide a very unique benchmark for assessing the merits of less involved theoretical frameworks to accurately reproduce the electronic and optical properties of water systems. Further, calculations on fixed common geometries allow one to separate the errors due to the formalism chosen to treat electronic correlations from geometry-dependent errors, which will always exist given that the structure of liquid water is strongly dependent on the simulation method of choice.[34]

In the present study, we explore the merits of the *GW* formalism by comparing the calculated IP of small $(H_2O)_n$ (n = 1-6) water clusters with CCSD(T) and ADC(3) data using identical geometries and basis sets. We show, in particular, that non-self-consistent G_0W_0 calculations starting from PBE[23] or PBE0[35] functionals dramatically underestimate the ionisation potential over the entire cluster range, while a simple self-consistent scheme provides very reliable results with very weak starting point dependency. As a result, the calculated Bethe-Salpeter optical absorption onset is significantly redshifted when starting from these non-self-consistent G_0W_0 quasiparticle spectra. The present results may rationalise the discrepancies with experiments observed in previous exploration of the properties of water at the G_0W_0 and Bethe-Salpeter level. The influence of the clusters geometry, comparing CCSD, PBE, PBE0 and empirical TIP4P geometries, is further analysed. Finally, we demonstrate that the optical absorption onset increases with cluster size, in great contrast with the HOMO-LUMO gap that decreases as expected.

II. METHODOLOGY

A. The quasiparticle *GW* formalism

The quasiparticle formalism, namely, the mapping of the true many-body problem onto a single (quasi)particle framework, provides a formal background for obtaining quasiparticle energies, which is the electronic energy levels associated with occupied or virtual states as measured by direct and inverse photoemission. The associated eigenvalue equation reads as

$$\left(\frac{-\nabla^2}{2} + V^{ionic}(\mathbf{r}) + V^{Hartree}(\mathbf{r})\right)\phi(\mathbf{r}) + \int d\mathbf{r}'\, \Sigma(\mathbf{r},\mathbf{r}';\varepsilon)\phi(\mathbf{r}') = \varepsilon\phi(\mathbf{r}), \quad (1)$$

where we introduce a general $\Sigma(\mathbf{r},\mathbf{r}';E)$ self-energy operator for the exchange and correlation contribution. The self-energy operator can be shown to be in general nonlocal, energy-dependent, and non-Hermitian, so that the corresponding eigenstates present an imaginary part interpreted as the lifetime of the quasiparticles with respect to electron-electron scattering.

Adopting Schwinger's functional derivative approach[6] to the many-electron problem, Hedin[7] showed that the self-energy can be given by a set of coupled equations relating self-consistently the one-body Green's function G, the screened-Coulomb potential W, and the irreducible polarizability P,

$$\begin{aligned}
G(12) &= G_0(12) + \int d(34)\; G_0(13)\Sigma(34)G(42),\\
\Sigma(12) &= i\int\int d(34)\; G(13)\Gamma(32;4)W(41),\\
W(12) &= v(12) + \int d(34)\; v(13)P(34)W(42),\\
P(12) &= -i\int d(34)\; G(13)G(41)\Gamma(34;2),\\
\Gamma(12;3) &= \delta(12)\delta(13)\\
&\quad + \int d(4567)\,\frac{\delta\Sigma(12)}{\delta G(45)}G(46)G(75)\Gamma(67;3),
\end{aligned}$$

where, e.g., $1 = (\mathbf{r}_1,t_1)$, $v(12) = v(\mathbf{r}_1,\mathbf{r}_2)\delta(t_1-t_2)$ is the bare Coulomb potential, and $\Gamma(34;2)$ is the so-called 3-body vertex correction. Such a set of equations can, in principle, be solved iteratively, starting from a zeroth-order system where the self-energy is zero, namely, the Hartree mean-field solution, yielding to first order: $\Gamma(12;3) = \delta(12)\delta(13)$. This simple approximation for the vertex correction yields the famous *GW* approximation for the self-energy,[7–11] written here in the energy representation,

$$\begin{aligned}
\Sigma(\mathbf{r},\mathbf{r}';E) &= \frac{i}{2\pi}\int d\omega e^{i\omega 0^+} G(\mathbf{r},\mathbf{r}';E+\omega)W(\mathbf{r},\mathbf{r}';\omega),\\
G(\mathbf{r},\mathbf{r}';E) &= \sum_n \frac{\phi_n(\mathbf{r})\phi_n^*(\mathbf{r}')}{E-\varepsilon_n+0^+\times sgn(\varepsilon_n-E_F)},\\
W(\mathbf{r},\mathbf{r}';\omega) &= v(\mathbf{r},\mathbf{r}') + \int d\mathbf{r}_1 d\mathbf{r}_2\; v(\mathbf{r},\mathbf{r}_1)\\
&\quad \times P_0(\mathbf{r}_1,\mathbf{r}_2;\omega)W(\mathbf{r}_2,\mathbf{r}';\omega),\\
P_0(\mathbf{r},\mathbf{r}';\omega) &= \sum_{i,j}(f_i-f_j)\frac{\phi_i^*(\mathbf{r})\phi_j(\mathbf{r})\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')}{\varepsilon_i-\varepsilon_j-\omega-i0^+},
\end{aligned}$$

where we have introduced the zeroth-order one-body (ε_n,ϕ_n) mean-field eigenstates. $P_0(\mathbf{r},\mathbf{r}';\omega)$ is the irreducible polarizability (the $f_{i/j}$ are the occupation factors). The summations over occupied and empty states lead to an $O(N^4)$ scaling for *GW* calculations with respect to system size, a scaling larger than the standard $O(N^3)$ scaling for DFT calculations with (semi)local functionals.

In practice, the mean-field starting point is never the Hartree solution, but more traditionally DFT Kohn-Sham eigenstates which represent, in general, the "best available" mean-field starting point. This leads to the standard "single-shot perturbative" G_0W_0 treatment where the exchange-correlation contribution to the DFT Kohn-Sham eigenvalues is replaced by the *GW* self-energy operator expectation value

onto the "frozen" Kohn-Sham DFT eigenstates, namely,

$$E_n^{QP} = \varepsilon_n^{DFT} + \langle \phi_n^{DFT} | \Sigma^{GW}(E_n^{QP}) - v^{XC,DFT} | \phi_n^{DFT} \rangle.$$

Taking, e.g., the local density approximation (LDA) to the exchange-correlation potential $v^{XC,DFT}$ leads to the so-called G_0W_0@LDA scheme, one of the most common approaches for GW calculations in solids. In the following, we will explore, in particular, the merits of the non-self-consistent G_0W_0@PBE and G_0W_0@PBE0 schemes, namely, single-shot G_0W_0 calculations aiming at correcting the Kohn-Sham PBE and PBE0 electronic energy levels.

As shown here below, and as demonstrated in several recent studies, the standard G_0W_0@LDA or G_0W_0@PBE non-self-consistent schemes do not offer sufficient accuracy in the case of isolated molecular systems, leading to underestimated ionization potential (IP) and HOMO-LUMO gaps.[29,36–50] This can be ascribed to the fact that the starting point (zeroth-order) Kohn-Sham spectrum used to build the one-body Green's function and screening potential W is too far off the experimental reference when using the PBE, or even the PBE0, functionals. A solution to that problem may consist in finding the "best" (optimized) DFT starting point, such as hybrid functionals with tuned amount of exact exchange[39–42,45,49] or starting with generalized Kohn-Sham formulations designed to provide accurate frontier orbital energies.[51,52] The important issue to be addressed within this approach is the fact that such a best starting point for single-shot G_0W_0 calculation is system dependent, since Hartree-Fock appears to be a good starting point for very small molecules,[29,36,37,53] while PBE0 with its reduced 25% of exact exchange would be much better for intermediate size systems,[41,42,45,49,54] raising the standard problem of the proper functional for a given system and physical observable to be calculated. This will be demonstrated here below in the case of water clusters with increasing size.

Another approach to improve on the calculated quasiparticle energies is a self-consistent GW calculation that removes the starting point dependency. By self-consistency, we mean that the corrected eigenvalues, and potentially eigenfunctions as well, are reinjected into the calculation of G, W, and Σ. Such an approach, in its various formulations, has been shown in the case of extended solids to significantly improve the accuracy of the GW formalism in many situations where single-shot G_0W_0 calculations would not be as accurate as desired.[55–59] Concerning molecular systems, for which much less data are available, full self-consistency of both eigenvalues and eigenstates has been demonstrated to improve on the standard non-self-consistent G_0W_0@PBE approach, but it is not clear whether it improves systematically on a single-shot G_0W_0 calculations starting from a hybrid functional including a proper amount of exact exchange,[36,47,48] questioning its practical interest given its very high cost.

To improve on the accuracy of the GW formalism and reduce the starting-point dependency problem, while preserving a scheme computationally efficient that allows one to tackle systems comprising several hundred atoms, we will explore here below a simple self-consistent strategy where only the corrected eigenvalues are reinjected in the construction of the polarizability P and the Green's function G. Such a scheme was shown by several groups to lead to much improved ionization potentials and HOMO-LUMO gaps[37,40,44,50] as compared to G_0W_0@LDA or G_0W_0@PBE values and subsequently to improved Bethe-Salpeter excitation energies[37,44,60,61] and even electron-phonon coupling strengths.[62,63] The partial nature of the self-consistency will be justified here below by showing that the impact of "freezing" (not updating) the starting-point Kohn-Sham wavefunction has limited impact on the final result. Such a simple partially self-consistent scheme will be labeled in the following evGW@PBE or evGW@PBE0 when starting from input PBE or PBE0 Kohn-Sham eigenstates.

As an additional approximation, we only calculate the GW correction to the highest occupied (HOMO) and lowest unoccupied (LUMO) levels, shifting rigidly the occupied (virtual) manifold according to the HOMO (LUMO) level self-energy correction. In this simplified evGW scheme, the GW HOMO-LUMO gap is thus reinjected self-consistently in the construction of G and W, but the original Kohn-Sham energy spacing within the occupied (virtual) manifold is frozen. As shown in the supplementary material (Table S2),[64] this restriction only affects the calculated ionization potential and optical absorption onset by less than 20 meV, as compared to an approach in which a large number of energy levels are corrected, while reducing dramatically the computing time. Such a marginal effect originates in the relative stability of the GW correction for occupied (virtual) energy levels (see Fig. S2 of the supplementary material).[64]

B. The Bethe-Salpeter formalism

While the GW formalism aims at obtaining accurate quasiparticle (occupied and virtual) energy levels, the study of optical properties (neutral excitations) requires the introduction of the (screened) electron-hole interaction. This is the goal of the BSE[65–71] that requires the quasiparticle energy and screened-Coulomb potential as generated by the preceding GW calculation.

The BSE formalism can be straightforwardly compared to time-dependent density functional theory (TDDFT)[72–74] in the so-called Casida's formulation[74] which recasts the TDDFT problem as a similar eigenvalue problem in the electron-hole two-body basis, namely,

$$\begin{pmatrix} R & C \\ -C^* & -R^* \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \Omega^{BSE} \begin{pmatrix} A \\ B \end{pmatrix}, \tag{2}$$

where the vector (A,B) represents the coefficients of the (ψ^{eh}) excitations with energy Ω^{BSE} on the occupied/virtual molecular orbitals product basis,

$$\psi^{eh}(\mathbf{r}_e, \mathbf{r}_h) = \sum_{ai} \{ A_{ai} \phi_a(\mathbf{r}_e) \phi_i(\mathbf{r}_h) + B_{ai} \phi_a(\mathbf{r}_h) \phi_i(\mathbf{r}_e) \}.$$

The indexes (i,j) and (a,b) indicate the occupied and virtual orbitals, and $(\mathbf{r}_e, \mathbf{r}_h)$ indicate the electron and hole positions, respectively. With these notations, the $\phi_a(\mathbf{r}_e)\phi_i(\mathbf{r}_h)$ components represent all excitations (note, e.g., that $\phi_a(\mathbf{r}_e)$ means that an electron is put into a virtual orbital), while the components $\phi_i(\mathbf{r}_e)\phi_a(\mathbf{r}_h)$ represent all disexcitations. As such, R (R^*) describes the resonant coupling between electron-hole

excitations (disexcitations), while the off-diagonal blocks C and C^* account for non-resonant coupling between excitations and disexcitations. The BSE resonant Hamiltonian matrix elements read, in particular,

$$R^{BSE}_{ai,bj} = \delta_{a,b}\delta_{i,j}\left(\varepsilon^{QP}_a - \varepsilon^{QP}_i\right) \quad (3)$$

$$- \left\langle \phi_a(\mathbf{r})\phi_i(\mathbf{r}')W(\mathbf{r},\mathbf{r}')\phi_b(\mathbf{r})\phi_j(\mathbf{r}')\right\rangle \quad (4)$$

$$+ 2\left\langle \phi_a(\mathbf{r})\phi_i(\mathbf{r})v(|\mathbf{r}-\mathbf{r}'|)\phi_b(\mathbf{r}')\phi_j(\mathbf{r}')\right\rangle . \quad (5)$$

We use the notation $\langle\cdots\rangle$ for the $\int\int d\mathbf{r}d\mathbf{r}'$ double integral. The middle (last) line gathers terms with occupied and virtual orbitals taken at different (identical) integration variables.[75] For isolated systems, the wavefunctions can be taken to be real. The resonant terms have been written here above for singlet excitations, and we have considered spin-unpolarized systems.

While the literature on the Bethe-Salpeter formalism is too large to be overviewed, we can focus here on isolated molecular systems for which it was demonstrated that not only "standard" Frenkel excitations[44,76–83] but also problematic charge-transfer[60,84–88] or cyanine[89,90] excitations could be very well described with such an approach that scales exactly like Casida's eigenvalue problem within TDDFT, once the screened-Coulomb potential W is built from the preceding GW calculations. Recently, extensive benchmark calculations on a standard molecular set[61,91] concluded that the present formalism, namely, Bethe-Salpeter calculations based on partially self-consistent evGW calculations would lead to a mean-absolute error (MAE) of the order of 0.2 eV as compared to reference high-level quantum chemistry calculation (e.g., CC3 calculations) with a very reduced dependence on the starting DFT functional.[61] Alternatively, non-self-consistent BSE/G_0W_0 calculation starting from Kohn-Sham states generated with optimally tuned DFT functionals would lead to similar accuracy.[92]

C. Technical details

Our GW calculations are performed at the all-electron level with standard *aug*-cc-pVXZ correlation consistent basis sets[93] using the Fiesta code,[37,85] implementing "Coulomb-fitting" resolution-of-the-identity (RI-V) techniques. Additionally, and for sake of verification, we also use the Molgw code[45,53] for the water monomer and dimer that explicitly calculates the required four-center two-electron Coulomb integrals. As shown here below, with well-converged auxiliary basis sets, both codes give results agreeing at the very few meV level. For the Fiesta code, the input Kohn-Sham eigenstates are generated with the NWChem package.[94]

The implemented resolution-of-the-identity, or Coulomb fitting technique, expresses four-center integrals in terms of three-center integrals combined with an auxiliary basis. We will show that the so-called "PAuto" auxiliary basis,[95] obtained from the onsite product of the DFT basis used to generate the Kohn-Sham eigenstates, and the *aug*-cc-pVXZ-RI basis by Weigend and co-workers[96] reproduce very accurately the "exact" (no-RI) calculations. All virtual states are included in the construction of the polarizability P and of the self-energy Σ. The energy integration required to calculate the correlation self-energy is achieved by contour deformation techniques in the Fiesta code and with a direct analytic integration in the Molgw code thanks to a spectral decomposition of the screened Coulomb potential W. In particular, both codes do not involve any plasmon-pole approximation.

Concerning the calculation of the correlation contribution to the self-energy $\Sigma^{GW}(E^{QP})$ at the targeted quasiparticle energy, we do not use the linear interpolation technique exploiting the value of the self-energy $\Sigma^{GW}(\varepsilon^{DFT})$ at the input ε^{DFT} Kohn-Sham energy and its energy derivative, since this may lead to substantial errors in the present case where Kohn-Sham and GW energies differ by several eVs. We explicitly calculate $\Sigma^{GW}(E)$ on a large energy grid to explicitly solve $E^{QP}_n = \varepsilon^{DFT}_n + \langle\phi^{DFT}_n|\Sigma^{GW}(E^{QP}_n) - v^{XC,DFT}|\phi^{DFT}_n\rangle$.

For the Bethe-Salpeter calculations, performed with the Fiesta code, all valence and virtual states are included in the electron-hole product space within which the electronic excitations are built. We go beyond the Tamm-Dancoff approximation, namely, we mix resonant and non-resonant contributions.

III. COMPARING *GW* AND CCSD(T) IONIZATION ENERGIES FOR SMALL WATER CLUSTERS

In order to compare the present GW formalism with the high-level quantum chemistry CCSD(T) calculations of Segarra-Martí and co-workers in Ref. 32 for small $(H_2O)_n$ (n = 2,6) clusters, we first perform calculations using the same *aug*-cc-pVTZ basis and the same CCSD/*aug*-cc-pVDZ geometries. For the water monomer, we adopt the MP2/*aug*-cc-pVDZ geometry used by Müller and Cederbaum in their ADC(3) study of the $(H_2O)_n$ (n = 1–4) clusters.[33] The ADC(3) calculations were also performed at the *aug*-cc-pVTZ level, except for the $(H_2O)_4$ cluster where a smaller *aug*-cc-pVDZ basis was adopted.[33] The auxiliary basis is the large PAuto basis generated by Gaussian09[95] for the *aug*-cc-pVTZ basis which yields for the water monomer and dimer IPs within less than 5 meV as compared to a full calculation where two-electrons four-center Coulomb integrals are explicitly calculated without any RI-V approximation (see the supplementary material[64] and numbers in parenthesis in Table I).

Our results are summarized in Table I and in Fig. 1 where we plot the reference CCSD(T) ionization potential values for small $(H_2O)_n$ (n = 2,6) clusters (black squares) together with the Green's function ADC(3) *aug*-cc-pVTZ data (white squares). As emphasized in many studies of molecular systems, the standard "single-shot" G_0W_0@PBE approach (blue circles) significantly underestimates the IP by more than 1 eV. As shown here below, the quality of the Kohn-Sham wavefunctions is not responsible for such a deficiency. Following the discussion in the Introduction concerning the comparison between G_0W_0@PBE and experimental data for liquid water, this is a strong indication that the G_0W_0@PBE approach may lead to too small ionization energies over the entire range of $(H_2O)_n$ clusters, from the (n = 1) monomer limit to the (n → ∞) water limit.

In several studies, the use of a hybrid functional such as PBE0[35] as starting point for GW calculations was advocated

TABLE I. Calculated ionization potential (IP) in eV. CCSD(T) and ADC(3) calculations are from Refs. 32 and 33. All data correspond to calculations performed at the *aug*-cc-pVTZ level. Geometries are the CCSD/*aug*-cc-pVDZ of Ref. 32, except for the monomer where the available MP2/*aug*-cc-pVDZ from Ref. 33 is chosen. Numbers in parenthesis for the water monomer and dimer are obtained without any RI-V approximation (we provide energies at the meV level for sake of comparison). The mean absolute errors (MAEs) take the CCSD(T) results as a reference.

	G_0W_0			ev-GW					
	PBE	PBE0	HF	PBE	PBE0	HF	QSGW	ADC(3)[a]	CCSD(T)
(n = 1)	11.615 (11.611)	12.139 (12.138)	12.861 (12.864)	12.88	12.77	12.76	12.94	12.8	12.653
(n = 2)	10.659 (10.662)	11.268 (11.268)	12.049 (12.049)	12.03	11.93	11.93	12.13	11.9	11.79[b]
(n = 3)	11.15	11.74	12.55	12.45	12.37	12.42		12.4	12.27[b]
(n = 4)	11.20	11.79	12.60	12.49	12.41	12.47			12.27[b]
(n = 5)	11.01	11.61	12.44	12.32	12.24	12.30			12.10[b]
6-ring	11.08	11.68	12.50	12.40	12.32	12.37			12.14[b]
6-cage	10.73	11.33	12.19	12.04	11.95	12.04			11.99[b]
6-book	10.60	11.22	12.08	11.91	11.85	11.94			11.69[b]
6-prism	10.57	11.18	12.04	11.88	11.81	11.90			11.65[b]
MAE	1.1	0.51	0.31	0.20	0.13	0.17			

[a]Reference 33.
[b]Reference 32.

for small molecular systems.[41,42,45,54] We confirm indeed that the G_0W_0@PBE0 IP values (red circles) are in much better agreement with the corresponding ADC(3) or CCSD(T) references. However, the obtained binding energies are still ~0.5 eV too small over the entire range of cluster sizes. As a matter of fact, as shown first by Hahn and co-workers in their study of small molecules[29] such as the water monomer and silane and as confirmed later by several groups,[36,37,45,53]

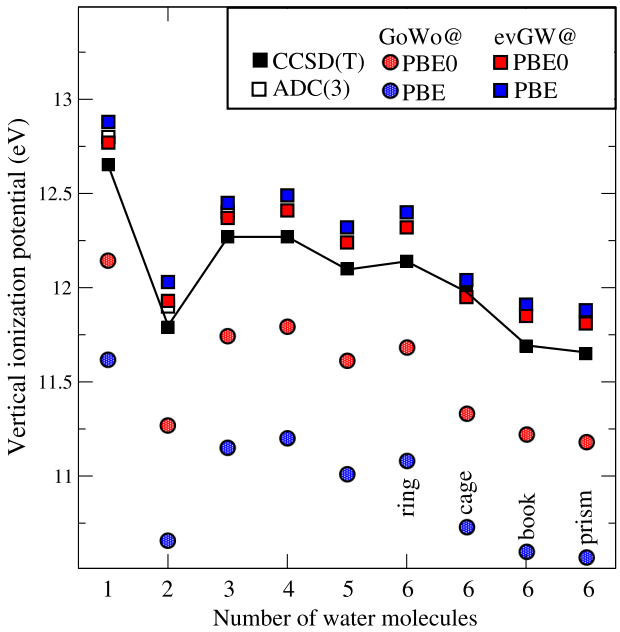

FIG. 1. Theoretical ionization potential (IP) in electronvolts. Calculations are performed at the aug-cc-pVTZ level. Reference CCSD(T) (black squares) and ADC(3) (white squares) calculations are from Refs. 32 and 33, respectively. Single-shot (G_0W_0) calculations starting from PBE (blue circles) and PBE0 (red circles) are compared to partially self-consistent (evGW) calculations starting from PBE (blue squares) and PBE0 (red squares). The black line serves as a guide to the eyes for the reference CCSD(T) data.

better agreement with experiment at the non-self-consistent G_0W_0 level can be obtained provided that one starts with Hartree-Fock eigenvalues which, for small systems, are much closer to the experimental values than the Kohn-Sham spectrum with PBE or PBE0 functionals.

In the original G_0W_0 study by Hahn and co-workers of the water monomer, the starting spectrum was a Hartree-Fock "like" *ansatz*, namely, the authors started their G_0W_0 calculations from a corrected DFT (PBE) spectrum where the exchange-correlation contribution to the Kohn-Sham eigenvalues was replaced by the expectation value of the exchange-only operator (Σ^X) on the Kohn-Sham DFT eigenfunctions (ψ_n^{DFT}), namely,

$$\varepsilon_n^{HF-like} = \varepsilon_n^{DFT} + \langle \psi_n^{DFT} | \Sigma^X - V_{xc}^{DFT} | \psi_n^{DFT} \rangle.$$

In such a scheme, the DFT Kohn-Sham wavefunctions (ψ_n^{DFT}) are kept frozen.

In the present study, we start from "full" Hartree-Fock calculations, namely, we build our G_0W_0@HF approach starting from Hartree-Fock eigenvalues and eigenfunctions. The results reported in Fig. 2 and Table I clearly demonstrate that indeed for the water monomer, the Hartree-Fock starting point is much better than PBE or PBE0. Further, for the monomer and dimer, G_0W_0@HF data are in better agreement with the CCSD(T) reference than fully self-consistent GW calculations with update of eigenstates and eigenvalues (see Table I) as calculated by the MOLGW code.[53] The G_0W_0@HF monomer IP value falls 0.07 eV and 0.12 eV above the ADC(3) and CCSD(T) results. Since the starting point Hartree-Fock ionization potential and HOMO-LUMO gap are larger than the target reference values,[97] the single-shot perturbative G_0W_0@HF data land above the CCSD(T) reference. However, with increasing size, this discrepancy becomes larger with an overall MAE of 0.3 eV and an ~0.4 eV discrepancy for the largest hexamers.

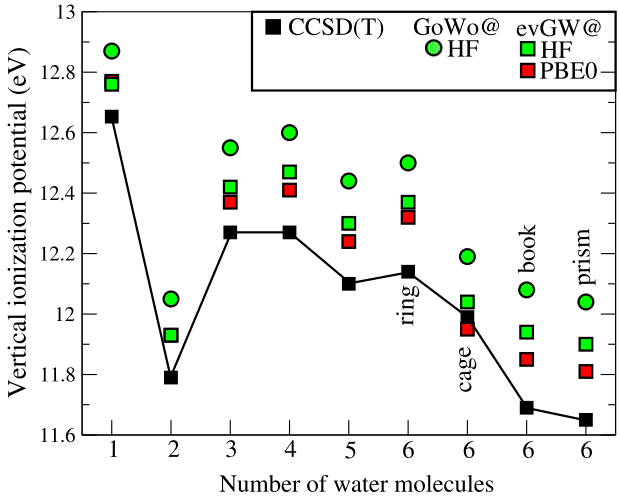

FIG. 2. Theoretical ionization potential (IP) in electronvolts. Calculations are performed at the aug-cc-pVTZ level. Reference CCSD(T) (black squares) calculations are from Ref. 32. Single-shot (G_0W_0) calculations starting from Hartree-Fock (green circles) are compared to partially self-consistent (evGW) calculations starting from Hartree-Fock (green squares) and PBE0 (red squares). The black line serves as a guide to the eyes for the reference CCSD(T) data.

While the Hartree-Fock starting point may be excellent for very small molecules, it is expected to lead to underscreening in the bulk limit due to the much too large Hartree-Fock band gap for solids. This is what we start to see here above in the case of our largest clusters. As such, one observes that the best starting point for a non-self-consistent G_0W_0 calculation may be system dependent. In the present case of water clusters, it is cluster size dependent. In order to minimize this problem of the starting-point dependency, and come-up with an approach providing reliable results from small to large extended systems, we apply the self-consistent evGW scheme discussed above with update of the eigenvalues only. Namely, we assume that the starting Kohn-Sham or Hartree-Fock wavefunctions are in good agreement with the fully self-consistent quasiparticle ones, an assumption that is validated here below. Such a simple scheme will be labeled evGW in the following.

We provide in Fig. 1 the evGW@PBE (blue squares) and evGW@PBE0 (red squares) IPs for the $(H_2O)_n$ (n = 1,6) clusters. Clearly, the corresponding IPs are in much better agreement with the reference CCSD(T) and ADC(3) values, with a nearly systematic ~0.1-0.2 eV overestimation for the evGW@PBE0 values and a MAE of 0.13 eV. Concerning the evGW@HF data, they are within 0.1 eV of the evGW@PBE and evGW@PBE0 data, demonstrating the large removal of starting-point dependence despite the partially self-consistent scheme adopted in the present study. This close agreement between all evGW approaches indicates, in particular, that the (frozen) Hartree-Fock or Kohn-Sham PBE or PBE0 wavefunctions are similar enough so that their differences hardly affect the quasiparticle energies, confirming that the crucial issue is the update of the occupied and virtual energy levels, as a validation of the simple evGW scheme.[98]

To conclude this section related to the importance of self-consistency, we now comment on an intermediate scheme, labeled $(evG)W_0$, where only the Green's function G is updated, keeping the screened Coulomb potential W frozen to its W_0 value obtained from a polarizability built with starting Kohn-Sham eigenstates. Such a scheme was advocated for solids[56,99] and proposed recently for molecular systems.[48] Consistently with our simplified evGW scheme, only the HOMO-LUMO gap is updated self-consistently in the present $(evG)W_0$ approach; the other occupied (virtual) energy levels being rigidly shifted according to the HOMO (LUMO) self-energy correction. Our findings, reported in Table S2 of the supplementary material,[64] indicate that such a scheme provides certainly much better results than a non-self-consistent G_0W_0 approach, but with a rather large residual dependency on the starting point Kohn-Sham spectrum. For sake of illustration, the water 6-prism hexamer *aug*-cc-pVTZ $(evG)W_0$ IP values read 11.3 eV, 11.5 eV, and 12.0 eV, starting from PBE, PBE0, and HF, respectively. Overall, the $(evG)W_0$@PBE approach underestimates the CCSD(T) value with a MAE of 0.32 eV and a maximum error of about 0.5 eV for the 6-cage hexamer structure.

Such a starting functional residual dependence in the $(evG)W_0$ scheme clearly indicates that part of the problem in the G_0W_0 scheme lies in the construction of the screened Coulomb potential W and not only in the proper description of the poles of G. As analyzed in several recent papers,[29,36,37,39] the too small Kohn-Sham gap at the PBE level leads to overscreening in the construction of W, resulting in a too small energy gap and ionization potential. As discussed above, the same argument led Hahn and co-workers to argue that Hartree-Fock-like eigenvalues were preferable over the Kohn-Sham LDA ones to build the screened Coulomb potential W in their original study of the water monomer.[29]

IV. BENCHMARK *GW* CALCULATIONS FOR $(H_2O)_n$: INFLUENCE OF BASIS AND GEOMETRY

While an *aug*-cc-pVTZ basis was adopted for sake of comparison with CCSD(T) calculations, our convergency tests presented in the supplementary material (Table S1)[64] indicate that the use of a larger *aug*-cc-pVQZ basis increases the IPs by about 0.2 eV. Explicit comparison with the *aug*-cc-pV5Z basis in the monomer and dimer cases indicates further that calculations at the *aug*-cc-pVQZ level are converged within less than 50 meV. Further, we adopt the more compact *aug*-cc-pVQZ-RI auxiliary basis approach by Weigend and co-workers[96] which reproduces the data obtained with the corresponding PAuto basis within a few meVs (see the supplementary material). For sake of example, taking the monomer experimental geometry from the NIST database,[100] our G_0W_0@PBE0 ionization potential for the water monomer amounts to 12.35 eV at the *aug*-cc-pVQZ level with the *aug*-cc-pVQZ-RI auxiliary basis, to be compared to the latest 12.37 eV value obtained with a planewave implementation of the GW formalism for the same geometry.[12]

We provide in Table II the IP of water clusters adopting the *aug*-cc-pVQZ basis and several geometries, namely, the same CCSD geometry as above, the PBE0 geometries, and the empirical TIP4P potential geometries as obtained from the Cambridge database.[101] For the monomer, we use as a

TABLE II. Calculated ionization potential (IP) in eV. The effect of geometry and self-consistency is explored. For the *GW* and BSE data, all calculations are performed at the *aug*-cc-pVQZ level with the *aug*-cc-pVQZ-RI auxiliary basis by Weigend and co-workers. Geometries are indicated in the second column. In the case of the monomer, we take the experimental geometry as a reference (see Ref. 100). Note that the TIP4P geometry matches the experimental one by construction.

	Geometry	PBE0	G_0W_0@PBE0	evGW@PBE0
(n = 1)	Expt.	9.09	12.35	13.02
	PBE0	9.08	12.34	13.00
	PBE	9.06	12.30	12.97
	TIP4P	9.09	12.35	13.02
(n = 2)	CCSD	8.37	11.45	12.15
	PBE0	8.34	11.42	12.11
	PBE	8.33	11.38	12.08
	TIP4P	8.21	11.27	11.97
(n = 3)	CCSD	8.98	11.93	12.60
	PBE0	8.98	11.91	12.57
	PBE	8.98	11.87	12.53
	TIP4P	9.00	11.99	12.65
(n = 4)	CCSD	9.08	11.97	12.63
	PBE0	9.06	11.93	12.58
	PBE	9.03	11.86	12.51
	TIP4P	9.10	12.02	12.68
(n = 5)	CCSD	8.92	11.80	12.47
	PBE0	8.89	11.74	12.40
	PBE	8.89	11.69	12.35
	TIP4P	8.84	11.77	12.44
6-ring	CCSD	8.98	11.87	12.54
	PBE0	8.97	11.78	12.42
	PBE	8.91	11.63	12.25
6-cage	CCSD	8.76	11.52	12.18
	PBE0	8.78	11.53	12.18
	PBE	8.79	11.53	12.19
6-book	CCSD	8.64	11.40	12.07
	PBE0	8.64	11.38	12.02
	PBE	8.64	11.33	11.97
6-prism	CCSD	8.60	11.37	12.04
	PBE0	8.60	11.35	12.00
	PBE	8.61	11.32	11.96

reference the experimental geometry.[100] For the hexamer, the lowest TIP4P geometry does not seem to correspond to the ring, cage, book, and prism geometries studied at the CCSD level, so that we do not include the TIP4P hexamer case. As shown in the recent G_0W_0@PBE study of water,[12] differences in liquid structure resulting from equilibrating water with different force fields (*ab initio* or empirical) can result in significant variations of the electronic properties, calling for an exploration of such issues on the much simpler case of static clusters.

The present results suggest that the ionization potential associated with high-level CCSD, DFT, or empirical TIP4P potential structures is rather similar. The maximum discrepancy between the CCSD and PBE0 geometries evGW@PBE0 IP values is found to be 0.11 eV for the 6-ring structure, with a mean absolute error of 47 meV for all clusters. As shown by Zhang and co-workers,[102] the PBE0 functional provides an excellent description of the structural properties of water and solvated ions, and the present findings corroborate the fact that evGW@PBE0 quasiparticle energies obtained for the CCSD or PBE0 geometries are in good agreement. The MAE associated with the PBE geometries is found to be 0.1 eV as compared to CCSD geometries, with a maximum value of 0.3 eV, indicating that PBE geometries may not be as accurate as the PBE0 ones. Comparing finally the available TIP4P and corresponding CCSD geometries, the maximum discrepancy amounts to 0.18 eV for the small dimer case, with a MAE of 76 meVs, indicating on this reduced sampling set the quality of the TIP4P empirical potential.

An interesting issue is related to the influence of the monomer geometry as compared to intermolecular orientations and distances. The two effects cannot be fully separated since the monomer geometry affects the structuration of water. In general, DFT calculations with nonhybrid functionals tend to have a poor water monomer geometry description. In particular, PBE predicts too long OH bond lengths. This results in shorter H-bonds and a more structured liquid.[34,103] However, and as shown in Table II for the monomer (n = 1), the influence of the monomer geometry on its ionisation potential is very marginal, much smaller than the differences observed between various cluster geometries. This indicates that it is not so much the monomer geometry that matters, but rather the influence of the monomer geometry on the relative water molecule structuration that governs the electronic properties of water clusters.

Overall, and despite the differences observed in Table II, one can conclude that the possible variations in ionisation potential induced by the underlying water geometry remain much smaller than the impact of self-consistency at the *GW* level, or the choice of the starting functional in the case of non-self-consistent calculations.

V. OPTICAL PROPERTIES

We finally address the problem of the optical properties of these small water clusters within the BSE formalism. As reported in the Introduction, a previous BSE study of water, starting from a quasiparticle spectrum calculated at the G_0W_0@PBE level, was shown to lead to an absorption onset redshifted by ∼1 eV as compared to the experimental values.[27] Since the non-self-consistent G_0W_0@PBE scheme was shown above to lead to erroneous electronic energy levels, we now explore the influence of self-consistency at the *GW* level on the subsequent Bethe-Salpeter lowest (singlet) excitation energies for the same family of water clusters. Our results are reported in Table III and Fig. 3.

From the analysis of the optical data, we observe that BSE calculations based on non-self-consistent G_0W_0@PBE and G_0W_0@PBE0 lead to an ∼1.2 eV and ∼0.6 eV underestimation of the absorption threshold as compared to BSE calculations starting from self-consistent evGW@PBE or evGW@PBE0. We observe, in particular, that the 1 eV redshift observed at the BSE@G_0W_0@PBE level is very similar to the redshift, as compared to experiment, observed in the BSE@G_0W_0@PBE study of bulk water by Garbuio and co-workers.[27]

TABLE III. Lowest calculated Bethe-Salpeter singlet excitation (S_1) energies in eV. The associated oscillator strengths are given in parenthesis. Various schemes are explored depending on the starting GW approach adopted. All calculations are performed at the *aug*-cc-pVQZ level with the corresponding *aug*-cc-pVQZ auxiliary RI basis by Weigend and co-workers.[96] Geometries are the CCSD geometries, except the (n = 1) monomer for which the MP2 geometry is adopted (see text). In the 6-ring case, we also provide the lowest singlet excitation with non-zero oscillator strength.

	BSE@			
	G_0W_0@PBE	G_0W_0@PBE0	evGW@PBE	evGW@PBE0
(n = 1)	5.80 (0.017)	6.31 (0.018)	6.97 (0.020)	6.97 (0.020)
(n = 2)	6.04 (0.020)	6.63 (0.021)	7.31 (0.022)	7.24 (0.022)
(n = 3)	6.29 (0.020)	6.93 (0.020)	7.48 (0.021)	7.50 (0.021)
(n = 4)	6.49 (0.032)	7.10 (0.034)	7.65 (0.035)	7.66 (0.036)
(n = 5)	6.34 (0.011)	6.97 (0.011)	7.53 (0.012)	7.54 (0.012)
6-ring	6.46 (0.000)	7.10 (0.000)	7.65 (0.000)	7.67 (0.000)
	6.47 (0.043)	7.11 (0.035)	7.65 (0.048)	7.68 (0.037)
6-cage	6.26 (0.018)	6.89 (0.019)	7.46 (0.020)	7.46 (0.020)
6-book	6.24 (0.019)	6.90 (0.020)	7.44 (0.021)	7.47 (0.021)
6-prism	6.53 (0.014)	7.16 (0.019)	7.72 (0.017)	7.73 (0.020)

A second observation is the remarkable stability of the BSE@evGW data which remain nearly constant when starting from PBE or PBE0 Kohn-Sham eigenstates, with a ~20 meV MAE discrepancy between the two approaches. This again indicates that the starting Kohn-Sham wavefunctions, which are kept frozen in the present partially self-consistent scheme, are close enough so that the final excitation energies are not significantly affected.

Concerning the comparison with experimental or reference calculations, we observe that our best Bethe-Salpeter singlet excitation for the water monomer, starting from our evGW@PBE0 or evGW@PBE quasiparticle spectrum, amounts to ~7 eV, which is smaller than the available experimental gas phase ~7.4 eV reference[104] and the 7.54 eV value reported at the *aug*-cc-pVTZ EOM-CCSD(T) level.[105] Our calculated value is closer to the 7.12 eV *aug*-cc-pVTZ CIS(D) result,[105] and within the range of the 6.95 eV, 7.09 eV, and 7.15 eV TD-DFT values obtained at the B3LYP, CAM-B3LYP, and PBE0 level, respectively, for the same geometry and *aug*-cc-pVQZ basis (NWChem calculations).

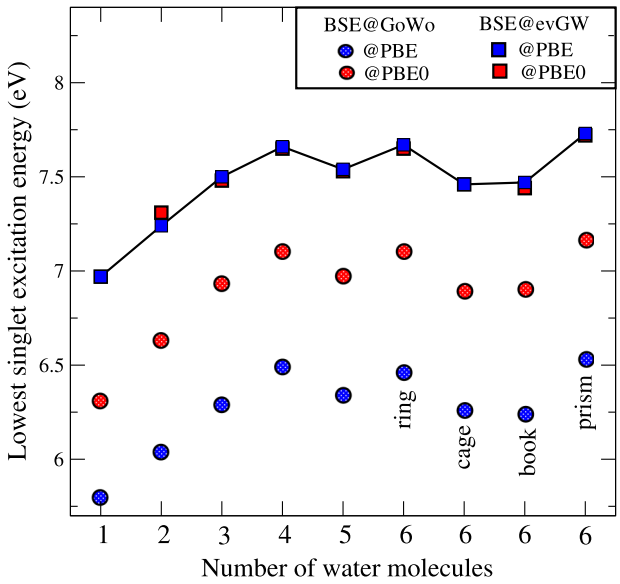

FIG. 3. Lowest calculated Bethe-Salpeter singlet excitation (S_1) energies in eV. Various schemes are explored depending on the starting GW approach adopted. Blue/red circles correspond to Bethe-Salpeter calculations performed on top of non-self-consistent G_0W_0@PBE and G_0W_0@PBE0, respectively. Blue/red squares correspond to Bethe-Salpeter calculations performed on top of self-consistent evGW@PBE and evGW@PBE0, respectively. The black line serves as a guide to the eyes for BSE@evGW@PBE0 data.

Recently, two benchmark studies of the so-called Thiels' set of medium sized molecules were conducted at the GW/BSE level.[61,92] It was shown that indeed BSE calculations starting from non-self-consistent G_0W_0@PBE data would lead to a dramatic underestimation of the reference values by about 1 eV, while BSE calculations starting from G_0W_0@PBE0 quasiparticle energies still resulted in redshifted optical absorption lines with a mean absolute error of about 0.6 eV. Performing, as in the present study, BSE calculations on top of evGW@PBE0 calculations led to a mean signed error of −0.14 eV and a mean absolute error of 0.25 eV.[61] These results are consistent with the present findings concerning the water molecule. Such an underestimation of the excitation energies in small molecular systems remains to be understood. We keep as a final message that the standard non-self-consistent G_0W_0@PBE scheme leads to dramatically too small Bethe-Salpeter excitation energies.

We now conclude this section by commenting on the evolution of the optical gap (lowest singlet excitation energy) as a function of cluster size. The most striking result is the *increase* of the absorption threshold with increasing cluster size, in great contrast with the standard result that due to confinement, the optical gap in aggregates or clusters increases usually with decreasing size. Our findings are consistent with the experimental observation that the optical gap of water is shifted *upwards* by about 1.3 eV from the monomer isolated limit to the condensed water liquid or ice phases.[30] The present results already confirm this trend in the small cluster limit, with a more than 0.7 eV increase between the monomer and the 6-prism hexamer structures.

An analysis of the DFT (Kohn-Sham) or GW energy levels (see Table S4 of the supplementary material)[64] indicates that as expected, the HOMO-LUMO gap tends to decrease with increasing cluster size, even though large fluctuations can be noticed in this small size limit. Overall, from the monomer to the 6-prism hexamer, the evGW@PBE0 HOMO-LUMO gap decreases by 1.25 eV, while the lowest singlet (S_1) excitation energy increases by 0.76 eV. As such, the electron-hole binding energy must decrease much faster than the HOMO-LUMO gap so as to lead to an increased optical gap. Such a behaviour was attributed to a delocalization of the electron-hole pair on neighbouring molecules, thus reducing the exciton binding energy.[30] In another study, the blueshift of the optical absorption onset in water under pressure was attributed to an increase of the electrostatic reaction field of the (polarizable) surrounding water environment.[106] In the present case of small water clusters, a plot of the hole-averaged (electron-averaged) electron (hole) charge distribution, namely—for the electron

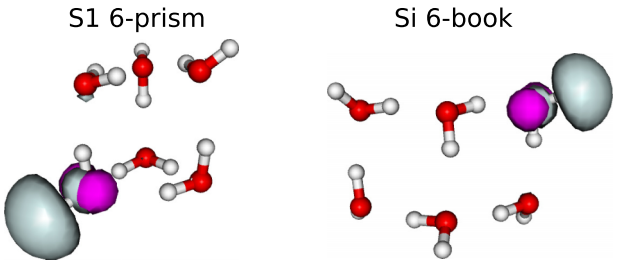

FIG. 4. Calculated hole-averaged electron (grey) and electron-averaged hole (pink) distribution for the lowest lying singlet excitation in the 6-prim and 6-book clusters. Isocontours have been taken at the 5% value of the maximum corresponding hole or electron density.

density—the expectation value of the $\delta(\mathbf{r}-\mathbf{r}_e)$ electron density operator over the 2-body Bethe-Salpeter $\psi(\mathbf{r}_e,\mathbf{r}_h)$ lowest lying (S_1) eigenstates, does not seem to favour a delocalization of the exciton mechanism (see Fig. 4). The rather delocalized nature of HOMO and LUMO eigenstates (see Fig. S4 of the supplementary material for the 6-prism cluster case)[64] indicates that it is the Coulomb interaction that relocalizes the interacting electron-hole pair on the same water molecule. As a result, the lowest S_1 excitation must mix the HOMO-LUMO component with higher-lying transitions.

To characterize the origin of the increase of the S_1 energy onset, we further decompose the BSE lowest eigenvalue in terms of its diagonal $\langle\varepsilon_a-\varepsilon_i\rangle$ contribution, its direct screened Coulomb potential $\langle W_D\rangle$, and "exchange" bare Coulomb $\langle V_X\rangle$ contributions, with the relation: $S_1=\langle\varepsilon_a-\varepsilon_i\rangle+2\langle V_X\rangle-\langle W_D\rangle$, following Eqs. (3)-(5). Namely, we take the expectation value of the corresponding operators over the lowest BSE electron-hole $\psi(\mathbf{r}_e,\mathbf{r}_h)$ eigenstate. While the resulting energies are provided in Table S5 of the supplementary material,[64] we plot in Fig. 5 the changes, as compared to the monomer, of the S_1 energy and various contributions for the (n = 1-, 2-, 3-, 4-, 6-prism) $(H_2O)_n$ clusters characterized by a steady increase of the absorption onset energy. To gather additional information, we also provide the changes in the direct bare Coulomb potential $\langle V_D\rangle$. The difference between $\langle W_D\rangle$ and $\langle V_D\rangle$ gives a direct information on the screening efficiency of the electron-hole interaction. Further, the inverse of $\langle V_D\rangle$ is a relevant measure of the average electron-hole distance. For sake of simplicity in the analysis, the data in Fig. 5 and in Table S5 of the supplementary material are performed within the Tamm-Dancoff approximation that yields S_1 transition energies blue-shifted by ~30 meV as compared to the "full" calculations discussed here above.

An important observation is that the direct bare Coulomb energy $\langle V_D\rangle$ hardly changes (see stars in Fig. 5). In particular, the related average electron-hole distance varies between 3.075 a.u. and 3.088 a.u. for all the considered structures, a negligible variation inconsistent with the delocalization scenario. On the contrary, the screened Coulomb $\langle W_D\rangle$ matrix elements decrease steadily with increasing size, indicating that it is the enhancement of the screening that can explain the reduction of the electron-hole binding energy. While the increase of the diagonal $\langle\varepsilon_a-\varepsilon_i\rangle$ energy, resulting from the participation of transitions with energy larger than the HOMO-LUMO gap, contributes as well to the S_1 increase from the dimer to the quadrimer, its decrease between the quadrimer and the hexamer leaves the enhancement of screening as the main ingredient in the optical onset blueshift with increasing sizes.

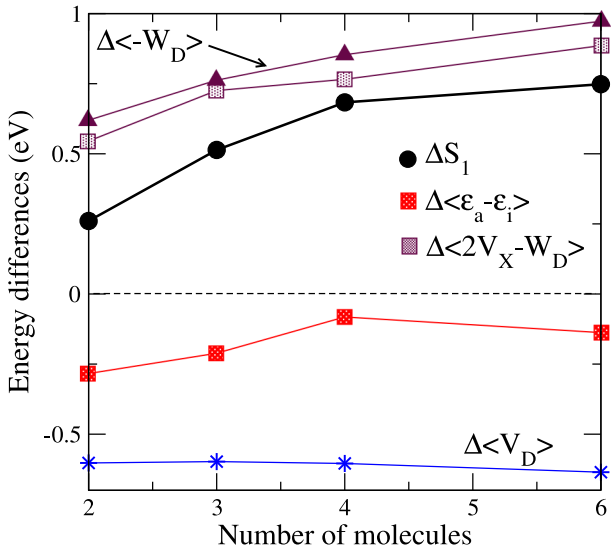

FIG. 5. Absorption energy differences ΔS_1 for the (n = 1-, 2-, 3-, 4-, 6-prism) $(H_2O)_n$ clusters as compared to the water monomer. The corresponding energy differences for the diagonal $\langle\varepsilon_a-\varepsilon_i\rangle$ contribution, where (a,i) index virtual and occupied levels, respectively, the direct screened Coulomb potential $\langle W_D\rangle$, and exchange bare Coulomb potential $\langle V_X\rangle$ contributions, are provided. For sake of analysis, the evolution of the direct bare Coulomb potential $\langle V_D\rangle$ matrix elements is also considered (blue stars). Calculations performed at the Tamm-Dancoff BSE@evGW@PBE0 *aug*-cc-pVTZ level. Energies are in eV.

VI. CONCLUSIONS

We have studied within the many-body Green's function GW and Bethe-Salpeter perturbation theories the electronic and optical properties of small $(H_2O)_n$ water clusters (n = 1-6). Comparison with high-level CCSD(T) and ADC(3) calculation for the corresponding ionization potentials clearly indicates that standard non-self-consistent G_0W_0 calculation starting from either PBE or PBE0 Kohn-Sham eigenstates significantly underestimates the reference values. We show that a simple self-consistent scheme, with update of the quasiparticle energies, leads to a much better agreement with CCSD(T) data with a mean absolute error of 0.13 eV when starting from PBE0. Further, self-consistency on the eigenvalues dramatically reduces the starting-point dependency, with variations smaller than 0.1 eV between calculations started from PBE, PBE0, or even Hartree-Fock eigenstates. The Bethe-Salpeter optical onsets follow the same trend, namely, much too low excitation energies are obtained when starting from single-shot G_0W_0@PBE quasiparticle energies. The increase of the optical gap with increasing cluster sizes is consistent with the same trend observed earlier going from the isolated water molecule to the dense ice or water phases. This blueshift as a function of size in the onset of absorption results from an increased screening of the electron-hole interaction and not from the delocalization of the electron-hole pair.

ACKNOWLEDGMENTS

X.B. acknowledges support from the French National Funding Agency (ANR) under contract "PANELS" No. ANR-12-BS04. Computing time has been provided by the National Supercomputing Centers (CCRT/Curie) with Project Nos. 2015096655 and 2015096018. F.B. acknowledges the Enhanced Eurotalent program and the France Berkeley Fund. M.V.F-S. acknowledges DOE Award No. DE-FG02-09ER16052.

[1]G. Hummer and A. Tokmakoff, J. Chem. Phys. **141**, 22D101 (2014), this issue of J. Chem. Phys. contains a section dedicated to the role of water in biological processes.
[2]M. G. Walter, E. L. Warren, J. R. McKone, S. W. Boettcher, Q. Mi, E. A. Santori, and N. S. Lewis, Chem. Rev. **110**, 6446-6473 (2010).
[3]Y. Jiao, Y. Zheng, M. Jaroniec, and S. Z. Qiao, Chem. Soc. Rev. **44**, 2060 (2015).
[4]T. S. Teets and D. G. Nocera, Chem. Commun. **47**, 9268-9274 (2011).
[5]A. J. Nozik and R. Memming, J. Phys. Chem. **100**, 13061-13078 (1996).
[6]P. C. Martin and J. Schwinger, Phys. Rev. **115**, 1342 (1959).
[7]L. Hedin, Phys. Rev. **139**, A796 (1965).
[8]M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[9]R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).
[10]G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).
[11]F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).
[12]T. A. Pham, C. Zhang, E. Schwegler, and G. Galli, Phys. Rev. B **89**, 060202(R) (2014).
[13]N. Kharche, J. T. Muckerman, and M. S. Hybertsen, Phys. Rev. Lett. **113**, 176802 (2014).
[14]V. Stevanovi, S. Lany, D. S. Ginley, W. Tumas, and A. Zunger, Phys. Chem. Chem. Phys. **16**, 3706-3714 (2014).
[15]T. A. Pham, D. Lee, E. Schwegler, and G. Galli, J. Am. Chem. Soc. **136**, 17071-17077 (2014).
[16]M. Govoni and G. Galli, J. Chem. Theory Comput. **11**, 2680-2696 (2015).
[17]A. Migani, D. J. Mowbray, J. Zhao, and H. Petek, J. Chem. Theory Comput. **11**, 239 (2015).
[18]H. Sun, D. J. Mowbray, A. Migani, J. Zhao, H. Petek, and A. Rubio, ACS Catal. **5**, 4242-4254 (2015).
[19]The uncertainty on the experimental valence band maximum value was discussed in Ref. 12 (see Table 1).
[20]P. Delahay and K. Von Burg, Chem. Phys. Lett. **83**, 250 (1981).
[21]P. Delahay, Acc. Chem. Res. **15**, 40 (1982).
[22]B. Winter, R. Weber, W. Widdra, M. Dittmar, M. Faubel, and I. Hertel, J. Phys. Chem. A **108**, 2625 (2004).
[23]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865-3868 (1996).
[24]C. Fang, W.-F. Li, R. S. Koster, J. Klimeš, A. van Blaaderen, and M. A. van Huis, Phys. Chem. Chem. Phys. **17**, 365 (2015).
[25]J. V. Coe, A. D. Earhart, M. H. Cohen, G. J. Hoffman, H. W. Sarkas, and K. H. Bowen, J. Chem. Phys. **107**, 6023 (1997).
[26]We note that the *GW* calculations aim at reproducing the vertical excitation energies, with, e.g., polarization effects governed by the optical refractive index and the related $\epsilon_\infty \sim 1.75$ dielectric constant, to be compared to the much larger $\epsilon \sim 78$ dielectric constant when slow molecular reorientation degrees of freedom are accounted for.
[27]V. Garbuio, M. Cascella, L. Reining, R. Del Sole, and O. Pulci, Phys. Rev. Lett. **97**, 137402 (2006).
[28]V. Garbuio, M. Cascella, and O. Pulci, J. Phys.: Condens. Matter **21**, 033101 (2009).
[29]P. H. Hahn, W. G. Schmidt, and F. Bechstedt, Phys. Rev. B **72**, 245425 (2005).
[30]P. H. Hahn, W. G. Schmidt, K. Seino, M. Preuss, F. Bechstedt, and J. Bernholc, Phys. Rev. Lett. **94**, 037404 (2005).
[31]A. Hermann, W. G. Schmidt, and P. Schwerdtfeger, Phys. Rev. Lett. **100**, 207403 (2008).
[32]J. Segarra-Martí, M. Merchán, and D. Roca-Sanjuán, J. Chem. Phys. **136**, 244306 (2012).
[33]I. B. Müller and L. S. Cederbaum, J. Chem. Phys. **125**, 204305 (2006).
[34]F. Corsetti, E. Artacho, J.-M. Soler, S. S. Alexandre, and M.-V. Fernández-Serra, J. Chem. Phys. **139**, 194502 (2013).
[35]C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158-6170 (1999); M. Ernzerhof and G. E. Scuseria, *ibid.* **110**, 5029-5036 (1999).
[36]C. Rostgaard, K. W. Jacobsen, and K. S. Thygesen, Phys. Rev. B **81**, 085103 (2010).
[37]X. Blase, C. Attaccalite, and V. Olevano, Phys. Rev. B **83**, 115103 (2011).
[38]C. Faber, C. Attaccalite, V. Olevano, E. Runge, and X. Blase, Phys. Rev. B **83**, 115123 (2011).
[39]N. Marom, X. Ren, J. E. Moussa, J. R. Chelikowsky, and L. Kronik, Phys. Rev. B **84**, 195143 (2011).
[40]S. Sharifzadeh, A. Biller, L. Kronik, and J. B. Neaton, Phys. Rev. B **85**, 125307 (2012).
[41]T. Körzdörfer and N. Marom, Phys. Rev. B **86**, 041110(R) (2012).
[42]N. Marom, F. Caruso, X. Ren, O. T. Hofmann, T. Körzdörfer, J. R. Chelikowsky, A. Rubio, M. Scheffler, and P. Rinke, Phys. Rev. B **86**, 245127 (2012).
[43]M. J. van Setten, F. Weigend, and F. Evers, J. Chem. Theory Comput. **9**, 232 (2013).
[44]C. Hogan, M. Palummo, J. Gierschner, and A. Rubio, J. Chem. Phys. **138**, 024312 (2013).
[45]F. Bruneval and M. A. L. Marques, J. Chem. Theory Comput. **9**, 324-329 (2013).
[46]T. A. Pham, H.-V. Nguyen, D. Rocca, and G. Galli, Phys. Rev. B **87**, 155148 (2013).
[47]F. Caruso, P. Rinke, X. Ren, A. Rubio, and M. Scheffler, Phys. Rev. B **88**, 075105 (2013).
[48]J. Lischner, S. Sharifzadeh, J. Deslippe, J. B. Neaton, and S. G. Louie, Phys. Rev. B **90**, 115130 (2014).
[49]M. Pinheiro, Jr., M. J. Caldas, P. Rinke, V. Blum, and M. Scheffler, Phys. Rev. B **92**, 195134 (2015).
[50]F. Kaplan, F. Weigend, F. Evers, and M. J. van Setten, J. Chem. Theory Comput. **11**, 5152 (2015).
[51]S. Kümmel and L. Kronik, Rev. Mod. Phys. **80**, 3 (2008).
[52]G. Borghi, A. Ferretti, N. L. Nguyen, I. Dabo, and N. Marzari, Phys. Rev. B **90**, 075135 (2014).
[53]F. Bruneval, J. Chem. Phys. **136**, 194107 (2012).
[54]S. Körbel, P. Boulanger, I. Duchemin, X. Blase, M. A. L. Marques, and S. Botti, J. Chem. Theory Comput. **10**, 3934-3943 (2014).
[55]F. Bruneval, N. Vast, and L. Reining, Phys. Rev. B **74**, 045102 (2006).
[56]M. Shishkin and G. Kresse, Phys. Rev. B **75**, 235102 (2007).
[57]M. Gatti, F. Bruneval, V. Olevano, and L. Reining, Phys. Rev. Lett. **99**, 266402 (2007).
[58]F. Trani, J. Vidal, S. Botti, and M. A. L. Marques, Phys. Rev. B **82**, 085115 (2010).
[59]T. Rangel, D. Kecik, P. E. Trevisanutto, G.-M. Rignanese, H. Van Swygenhoven, and V. Olevano, Phys. Rev. B **86**, 125125 (2012).
[60]B. Baumeier, D. Andrienko, and M. Rohlfing, J. Chem. Theory Comput. **8**, 2790-2795 (2012).
[61]D. Jacquemin, I. Duchemin, and X. Blase, J. Chem. Theory Comput. **11**, 3290-3304 (2015).
[62]C. Faber, J. Laflamme Janssen, M. Côté, E. Runge, and X. Blase, Phys. Rev. B **84**, 155104 (2011).
[63]S. Ciuchi, R. C. Hatch, H. Höchst, C. Faber, X. Blase, and S. Fratini, Phys. Rev. Lett. **108**, 256401 (2012).
[64]See supplementary material at http://dx.doi.org/10.1063/1.4940139 for (I) basis set sizes convergency tests, (II) an analysis of the accuracy of the "scissor approximation," (III) an analysis of the partially self-consistent $(evG)W_0$ scheme, (IV) a table containing the HOMO-LUMO gaps, and (V) the decomposition of the S_1 transition energies in terms of its diagonal and various Coulomb energy contributions.
[65]L. J. Sham and T. M. Rice, Phys. Rev. **144**, 708-714 (1966).
[66]W. Hanke and L. J. Sham, Phys. Rev. Lett. **43**, 387-390 (1979).
[67]G. Strinati, Phys. Rev. Lett. **49**, 1519 (1982).
[68]G. Strinati, Riv. Nuovo Cimento **11**, 1 (1988).
[69]M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **80**, 3320-3323 (1998).
[70]L. X. Benedict, E. Shirley, and R. B. Bohn, Phys. Rev. Lett. **80**, 4514-4517 (1998).
[71]S. Albrecht, L. Reining, R. Del Sole, and G. Onida, Phys. Rev. Lett. **80**, 4510-4513 (1998).
[72]E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997-1000 (1984).
[73]*Time-Dependent Density Functional Theory*, edited by L. Marques, C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. K. U. Gross (Springer, New York, NY, 2006).
[74]M. E. Casida, J. Mol. Struct.: THEOCHEM **914**, 3-18 (2009).
[75]The structure of the screened Coulomb integrals explains, in particular, the success of the BSE formalism in treating charge-transfer excitations.
[76]M. L. Tiago and J. R. Chelikowsky, Solid State Commun. **136**, 333-337 (2005).

[77]M. L. Tiago, P. R. C. Kent, R. Q. Hood, and F. A. Reboredo, J. Chem. Phys. **129**, 084311 (2008).
[78]Y. Ma, M. Rohlfing, and C. Molteni, Phys. Rev. B **80**, 241405 (2009).
[79]M. Palummo, C. Hogan, F. Sottile, P. Bagalá, and A. Rubio, J. Chem. Phys. **131**, 084102 (2009).
[80]Y. Ma, M. Rohlfing, and C. Molteni, J. Chem. Theory Comput. **6**, 257-265 (2010).
[81]M. S. Kaczmarski, Y. Ma, and M. Rohlfing, Phys. Rev. B **81**, 115433 (2010).
[82]D. Varsano, E. Coccia, O. Pulci, A. Mosca Conte, and L. Guidoni, Comput. Theor. Chem. **1040-1041**, 338-346 (2014).
[83]E. Coccia, D. Varsano, and L. Guidoni, J. Chem. Theory Comput. **10**, 501-506 (2014).
[84]D. Rocca, D. Lu, and G. Galli, J. Chem. Phys. **133**, 164109 (2010).
[85]X. Blase and C. Attaccalite, Appl. Phys. Lett. **99**, 171909 (2011).
[86]I. Duchemin, T. Deutsch, and X. Blase, Phys. Rev. Lett. **109**, 167801 (2012).
[87]C. Faber, I. Duchemin, T. Deutsch, and X. Blase, Phys. Rev. B **86**, 155315 (2012).
[88]C. Faber, P. Boulanger, I. Duchemin, C. Attaccalite, and X. Blase, J. Chem. Phys. **139**, 194308 (2013).
[89]P. Boulanger, S. Chibani, B. Le Guennic, I. Duchemin, X. Blase, and D. Jacquemin, J. Chem. Theory Comput. **10**, 4548-4556 (2014).
[90]P. Boulanger, D. Jacquemin, I. Duchemin, and X. Blase, J. Chem. Theory Comput. **10**, 1212-1218 (2014).
[91]M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. **128**, 134110 (2008).
[92]F. Bruneval, S. M. Hamed, and J. B. Neaton, J. Chem. Phys. **142**, 244101 (2015).
[93]T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
[94]M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, and W. A. de Jong, Comput. Phys. Commun. **181**, 1477 (2010).
[95]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, GAUSSIAN 09, Revision B.01, Gaussian, Inc., Wallingford, CT, 2010.
[96]F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[97]Taking as an example the monomer and 6prism cases, the PBE0, Hartree-Fock, and evGW@PBE0 HOMO-LUMO gaps are found to be (8.7;14.7;13.4) eV and (7.9;13.8;12.4) eV, respectively (aug-cc-pVTZ values), showing that the Hartree-Fock gap is closer to the final self-consistent evGW@PBE0 value, but overestimates as expected the reference gap.
[98]We note that the present scheme, where all the occupied (virtual) energy levels are shifted rigidly according to the HOMO (LUMO) energy correction, preserves the original Kohn-Sham energy spacing between occupied (virtual) levels, representing formally another source, even though small (see Table S2 of the supplementary material), of starting point dependency.
[99]M. P. Surh, J. E. Northrup, and S. G. Louie, Phys. Rev. B **38**, 5976 (1988).
[100]*NIST Computational Chemistry Comparison and Benchmark Database*, edited by R. D. Johnson III, NIST Standard Reference Database Number Vol. 101 (NIST, Gaithersburg, MD, 2013), see http://cccbdb.nist.gov/.
[101]See http://www-wales.ch.cam.ac.uk/~wales/CCD/TIP4P-water.html for atomic coordinates.
[102]C. Zhang, T. A. Pham, F. Gygi, and G. Galli, J. Chem. Phys. **138**, 181102 (2013).
[103]B. Pamuk, J.-M. Soler, R. Ramírez, C. P. Herrero, P. W. Stephens, P. B. Allen, and M.-V. Fernández-Serra, Phys. Rev. Lett. **108**, 193003 (2012).
[104]W. F. Chan, G. Cooper, and C. E. Brion, Chem. Phys. **178**, 387 (1993).
[105]D. M. Chipman, J. Chem. Phys. **122**, 044111 (2005).
[106]A. Hermann and P. Schwerdtfeger, Phys. Rev. Lett. **106**, 187403 (2011).