An O(N3) implementation of Hedin's GW approximation for molecules

D. Foerster, P. Koval, and D. Sánchez-Portal

Citation: J. Chem. Phys. **135**, 074105 (2011); doi: 10.1063/1.3624731

View online: http://dx.doi.org/10.1063/1.3624731

View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v135/i7

Published by the American Institute of Physics.

An $O(N^3)$ implementation of Hedin's GW approximation for molecules

D. Foerster,[1,a)] P. Koval,[2,3,b)] and D. Sánchez-Portal[2,3,c)]
[1]*CPMOH/LOMA, Université de Bordeaux 1, 351 Cours de la Liberation, 33405 Talence, France*
[2]*Centro de Física de Materiales CFM-MPC, Centro Mixto CSIC-UPV/EHU, Paseo Manuel de Lardizabal 5, E-20018 San Sebastián, Spain*
[3]*Donostia International Physics Center (DIPC), Paseo Manuel de Lardizabal 4, E-20018 San Sebastián, Spain*

(Received 10 January 2011; accepted 22 July 2011; published online 17 August 2011)

We describe an implementation of Hedin's GW approximation for molecules and clusters, the complexity of which scales as $O(N^3)$ with the number of atoms. Our method is guided by two strategies: (i) to respect the locality of the underlying electronic interactions and (ii) to avoid the singularities of Green's functions by manipulating, instead, their spectral functions using fast Fourier transform methods. To take into account the locality of the electronic interactions, we use a local basis of atomic orbitals and, also, a local basis in the space of their products. We further compress the screened Coulomb interaction into a space of lower dimensions for speed and to reduce memory requirements. The improved scaling of our method with respect to most of the published methodologies should facilitate GW calculations for large systems. Our implementation is intended as a step forward towards the goal of predicting, prior to their synthesis, the ionization energies and electron affinities of the large molecules that serve as constituents of organic semiconductors. © *2011 American Institute of Physics*. [doi:10.1063/1.3624731]

I. INTRODUCTION

Lars Hedin's GW method [1] is an approximate treatment of the propagation of electrons in condensed matter where an electron interacts with itself via a Coulomb interaction that is screened by virtual electron-hole pairs. In periodic semiconductors, the GW approximation is known to lead to surprisingly accurate gaps,[2] while for finite clusters and molecules, it provides qualitatively correct values of ionization energies and electron affinities.[3] Hedin's GW approximation is also needed, as a first step, when using the Bethe-Salpeter equation, to find the optical properties of systems in which the Coulomb interaction is only weakly screened.

The present work is motivated by the rapid progress, during the last decade, in the field of organic semiconductors, especially in organic photovoltaics and organic luminescent diodes.[4] To optimize such systems, it would be useful to know the key properties of their molecular constituents before actually synthesizing them. In order to make such predictions it is necessary to develop algorithms with a favorable complexity scaling, since many of the technologically relevant molecules are fairly large. The method presented here is a step forward along this direction. Its $O(N^3)$ scaling, with N being the number of atoms in the molecule, is an improvement over most existing methodologies.

While computational techniques for treating the GW approximation for clusters and molecules have become sophisticated enough for treating molecules of interest in photovoltaics[5] or in the physiology of vision,[6] such calculations remain computationally expensive. The scaling with the number of atoms of these recent calculations has not been published. However, in many cases it is unlikely to be better than $O(N^4)$.[7] A recently published method for computing total energies of molecules that uses the random phase approximation also has $O(N^4)$ scaling.[8] Actually, at this point it is difficult for us to envisage a scaling exponent less than three because the construction of the screened Coulomb interaction—the central element of the GW approach—requires inverting a matrix of size $O(N)$ which, in general, takes $O(N^3)$ operations.

The algorithm described in this paper is based on two main ingredients: (i) respecting the locality of the underlying interactions and, (ii) the use of spectral functions to describe the frequency/time dependence of the correlators. The latter ingredient allows the use of the fast Fourier transform (FFT) to accelerate the calculations, while the former idea of respecting locality has also been at the heart of other efficient GW methods, such as the successful "space-time approach" for periodic systems.[9]

Our method is based upon the use of spatially localized basis sets to describe electronic states within the linear combination of atomic orbitals (LCAO) technique. In particular, we have implemented our method as a post processing tool of the SIESTA code,[10] although interfaces with other LCAO codes should be simple to construct. The precision of the LCAO approach is difficult to control and to improve, but a basis of atom centered local orbitals is useful for systems that are too large to be treated by plane-wave methods.[11] In order to solve Hedin's equations we construct a basis set that gets rid of the over completeness of the orbital products while keeping the locality. In molecular computations this is frequently done through a fitting procedure (using Gaussians or other localized functions). We use an alternative mathematical procedure[12,13] that dispenses with this fitting and defines a basis of *dominant products*. The basis of dominant products

a)Electronic mail: d.foerster@cpmoh.u-bordeaux1.fr.
b)Electronic mail: koval.peter@gmail.com.
c)Electronic mail: sqbsapod@sq.ehu.es.

was instrumental to develop an efficient linear response code for molecular absorption spectra.[14,15] In the present paper we have developed an additional, non-local compression technique that further reduces the size of the product basis. The compression allows to store the whole matrix representation of the screened Coulomb interaction at all times/frequencies and needs much less memory. Moreover, the compression strongly accelerates the calculation of the screened Coulomb interaction because it involves a matrix inversion. This leads to a gain in computational efficiency which is even more important than that associated with the reduction of the needed memory.

Of course there are other methods that use a localized basis different from LCAO and, thus, equally appropriate for dealing with clusters and molecules while exploiting locality. One method uses a lattice in real space.[16] Another method uses wavelets that represent a useful compromise between localized and extended states.[17] Localized Wannier orbitals obtained from transforming plane waves[18] have also been used in GW calculations.[19,20] In this paper we use a basis of dominant functions to span the space of products of atomic orbitals[12] and we use a compression scheme to deal with the screened interaction. It is clear, however, that some of the ideas and techniques of the present paper can be combined with the alternative approaches quoted above.

The actual implementation of the algorithm that we report in this paper can be considered as a "proof of principle" only and the prefactor of our implementation leaves room for further improvement. Therefore, we validate our method with molecules of moderate size (we consider molecules of only up to three aromatic rings: benzene, naphthalene, and anthracene), leaving further improvements and applications to molecules of larger sizes for a future publication.

This paper is organized as follows. In Sec. II we recall the equations of the GW approximation. In Sec. III we rewrite the GW approximation for molecules in tensorial form. Section IV describes the instantaneous component of the self-energy, while in Sec. V we describe a spectral function technique for solving these tensorial equations. Section VI describes our GW results for benzene, a typical small molecule. Section VII describes our algorithm for the compression of the screened Coulomb interaction that is needed to treat larger molecules. Section VIII explains how $O(N^3)$ scaling can be maintained for large molecules by alternatively compressing and decompressing the Coulomb interaction. Section IX presents a summary of the entire algorithm for performing GW calculations. In Sec. X we present a convergence study with respect to the size of the orbital basis, while in Sec. XI we test our method on naphthalene and anthracene. Our conclusions are presented in Sec. XII.

II. ELEMENTARY ASPECTS OF HEDIN'S *GW* APPROXIMATION

The one-electron Green's function of a many-body system has proved to be a very useful concept in condensed matter theory. It allows to compute the total energy, the electronic density, and other quantities arising from one-particle operators. The one-electron Green's function $G(r, r', t)$ has twice as many spatial arguments as the electronic density, but it remains a far less complex object than the many-body wave function. Furthermore, Hedin has found an exact set of equations for a finite set of correlation functions of which the one-electron Green's function is the simplest element. This set of equations has not been solved so far for any system whatsoever. However, as a zeroth order starting point to his coupled equations, Hedin suggested the very successful GW approximation for the self-energy $\Sigma(r, r', t)$. This approximation describes the change of the non-interacting electron propagator $G_0(r, r', t)$ due to interactions among the electrons. With the help of a self-energy, one can find the interacting Green's function from Dyson's equation,

$$G = G_0 + G_0\Sigma G_0 + G_0\Sigma G_0\Sigma G_0 + \cdots = \frac{1}{G_0^{-1} - \Sigma}, \tag{1}$$

where products and inversions in the equation must be understood in an operator sense as required in many-body perturbation theory.

In Hedin's GW approximation, the interaction of electrons with themselves is taken into account by the following self-energy:

$$\Sigma(r, r', t) = \mathrm{i}G_0(r, r', t)W_0(r, r', t), \tag{2}$$

where $W(r, r', t)$ is a screened Coulomb interaction.

The key idea of Hedin's GW approximation[1] is to incorporate the screening of the Coulomb interaction from the very beginning in a zeroth order approximation. Let $v(r, r') = |r - r'|^{-1}$ be the bare Coulomb interaction and let $\chi_0(r, r', t - t') = \frac{\delta n(r,t)}{\delta V(r',t')}$ be the density response $\delta n(r, t)$ of non-interacting electrons with respect to a change of the external potential $\delta V(r', t')$. Hedin then replaces the original Coulomb interaction $v(r, r')$ by the screened Coulomb interaction $W(r, r', \omega)$ within the random phase approximation (RPA) (Ref. 21):

$$W_0(r, r', \omega) = \frac{1}{\delta(r - r''') - v(r, r'')\chi_0(r'', r''', \omega)}v(r''', r'). \tag{3}$$

Here and in the following we assume integration over repeated spatial coordinates (in our case r'' and r''') on the right hand side of an equation if they do not appear on its left hand side. This convention makes our equations more transparent without introducing ambiguities and it is analogous to the familiar Einstein's convention of summing over repeated indices.

We can justify expression (3) by considering an internal screening field $\delta V_{\mathrm{induced}}(r, \omega)$ that is generated by an extra external field $\delta V_{\mathrm{external}}(r, \omega)$:

$$\delta V_{\mathrm{total}}(r, \omega) = \delta V_{\mathrm{external}}(r, \omega) + \delta V_{\mathrm{induced}}(r, \omega), \tag{4}$$

where

$$\begin{aligned}\delta V_{\mathrm{induced}}(r, \omega) &= v(r, r'')\delta n_{\mathrm{induced}}(r'', \omega) \\ &= v(r, r'')\chi_0(r'', r''', \omega)\delta V_{\mathrm{total}}(r''', \omega).\end{aligned}$$

As a consequence we obtain a frequency dependent change of the total potential:

$$\delta V_{\text{total}}(r,\omega) = \frac{1}{\delta(r-r''') - v(r,r'')\chi_0(r'',r''',\omega)} \times \delta V_{\text{external}}(r''',\omega). \quad (5)$$

If we assume that large fields are screened the same way as small field changes, then we may replace $\delta V_{\text{external}}(r,\omega)$ by the singular Coulomb interaction $v(r,r')$ and we obtain the screened counterpart $W_0(r,r',\omega)$ of the original bare Coulomb interaction as in Eq. (3).

Because of the relation

$$\mathrm{i}\chi_0(r,r',t) = 2G_0(r,r',t)G_0(r',r,-t), \quad (6)$$

the screening in Eq. (3) may be interpreted as being due to the creation of virtual electron-hole pairs. The screening by virtual electron-hole pairs is the quantum analogue of classical Debye screening in polarizable media.[22] The factor of 2 in Eq. (6) takes into account the summation over spins.

Many body theory uses Feynman-Dyson perturbation theory[23] and the latter is formulated in terms of time ordered correlators. For instance, a Green's function is represented as a time ordered correlator of electron creation $\psi^+(r,t)$ and annihilation $\psi(r,t)$ operators:

$$\mathrm{i}G(r,r',t-t') = \theta(t-t')\langle 0|\psi(r,t)\psi^+(r',t')|0\rangle - \theta(t'-t)\langle 0|\psi^+(r',t')\psi(r,t)|0\rangle, \quad (7)$$

where the minus sign is due to Fermi statistics, $|0\rangle$ denotes the electronic ground state, and $\theta(t)$ denotes Heaviside's step function. This completes our formal description of Hedin's GW approximation.

In practice, Hedin's equations are solved "on top" of a density functional or Hartree-Fock calculation. The framework of density functional theory (DFT) (Refs. 24, 25) already includes electron correlations at the mean-field level via the exchange correlation energy $E_{\text{xc}}[n(r)]$, where $[n]$ denotes the functional dependence of E_{xc} on the electron density. DFT calculations are usually performed using the Kohn-Sham scheme,[25] in which electrons move as independent particles in an effective potential. The Kohn-Sham Hamiltonian H_{KS} reads

$$H_{\text{KS}} = -\frac{1}{2}\nabla^2 + V_{\text{KS}}, \quad (8)$$

$$V_{\text{KS}} = V_{\text{ext}} + V_{\text{Hartree}} + V_{\text{xc}}, \text{ where } V_{\text{xc}}(r) = \frac{\delta E_{\text{xc}}}{\delta n(r)}.$$

To avoid including the interaction twice, the exchange correlation potential $V_{\text{xc}}(r)$ must be subtracted from $\Sigma(r,r',t)$ in Eq. (2) when using the output of a DFT calculation as an input for a GW calculation. This is done by making the replacement

$$\Sigma(r,r',t) \to \Sigma(r,r',t) - \delta(r-r')\delta(t)V_{\text{xc}}(r) \quad (9)$$

in Dyson's equation (1).

Our aim is to compute the electronic density of states (DOS) that is defined as the trace of the imaginary part of the electron propagator:

$$\rho(\omega+\mathrm{i}\varepsilon) = -\frac{1}{\pi}\mathrm{Im}\int G(\omega+\mathrm{i}\varepsilon,r,r)d^3r. \quad (10)$$

The electronic DOS $\rho(\omega+\mathrm{i}\varepsilon)$ can be compared with experimental data from direct and inverse photo-emission.[26] From it, we can read off the energy position of the highest occupied and the lowest unoccupied molecular orbitals (HOMO and LUMO) or, alternatively, the ionization energy and the electron affinity.

Finally, let us list the equations that define the GW approximation:

$$\begin{aligned}
\mathrm{i}\chi_0(r,r',t) &= 2G_0(r,r',t)G_0(r',r,-t); && \text{non-interacting response,}\\
W_0(r,r',\omega) &= [\delta(r-r''') - v(r,r'')\chi_0(r'',r''',\omega)]^{-1}v(r''',r'); && \text{RPA screening,}\\
\Sigma(r,r',t) &= \mathrm{i}G_0(r,r',t)W_0(r,r',t); && GW \text{ self-energy,}\\
G^{-1}(r,r',\omega+\mathrm{i}\varepsilon) &= G_0^{-1}(r,r',\omega+\mathrm{i}\varepsilon) - \Sigma(r,r',\omega+\mathrm{i}\varepsilon). && \text{Dyson equation.}
\end{aligned} \quad (11)$$

Sections III–VI describe the tensor form of Eq. (11) as well as the main ingredients of our implementation of the GW approximation as embodied in Eq. (11) for the case of small molecules. Sections VII and VIII will describe the compression/decompression of the Coulomb interaction that is needed for treating large molecules without over flooding the computer memory.

III. TENSOR FORM OF HEDIN'S EQUATIONS

In order to compute the non-interacting Green's function (7), we will use the LCAO method where one expresses the electron operator in terms of a set of fermions $c_a(t)$ that belong to localized atomic orbitals:[27]

$$\psi(r,t) \sim \sum_a f^a(r)c_a(t). \quad (12)$$

Such a parametrization is parsimonious in the number of degrees of freedom, although its quality is difficult to control and to improve in a systematic way.

The output of a DFT calculation (that serves as an input for the GW calculation) is the Kohn-Sham Hamiltonian H^{ab} and the overlap matrix S^{ab} of the LCAO basis functions $f^a(r)$.[28] One may use the eigenvectors $\{X_a^E\}$ of the

Kohn-Sham Hamiltonian,

$$H^{ab}X_b^E = ES^{ab}X_b^E \tag{13}$$

to express the (time-ordered) propagation of electrons between localized atomic orbitals:

$$G_{ab}^0(\omega \pm \mathrm{i}\varepsilon) = \sum_E \frac{X_a^E X_b^E}{\omega \pm \mathrm{i}\varepsilon - E}. \tag{14}$$

In this paper we measure energies relative to a Fermi energy, so that $E < 0$ ($E > 0$) refers to occupied (empty) states, respectively, and the infinitesimal constant ε shifts the poles of the Greens function away from the real axis. Moreover, to avoid cluttering up the notation, we will often use Einstein's convention of summing over repeated indices, as in Eq. (13).

The set of Eq. (11) contains correlation functions, such as the density response function $\chi(r, r', t)$ that must be represented in a basis of *products of atomic orbitals:* [23]

$$\mathrm{i}\chi(r, r', t - t') = \theta(t - t')\langle 0|n(r, t)n(r', t')|0\rangle + \theta(t' - t)\langle 0|n(r', t')n(r, t)|0\rangle. \tag{15}$$

Indeed, by virtue of Eq. (12), the electronic density $n(r, t) = \psi^+(r, t)\psi(r, t)$ involves products of atomic orbitals:

$$n(r, t) = \psi^+(r, t)\psi(r, t) = \sum_{a,b} f^a(r)f^b(r)c_a^+(t)c_b(t).$$

The set of products $\{f^a(r)f^b(r)\}$ is well known to be strongly linearly dependent.[29] As an improved solution of this very old technical difficulty, we previously developed an algorithm to construct a local basis of "dominant products" $F^\mu(r)$ that (i) spans the space of orbital products with exponential accuracy and which (ii) respects the locality of the original atomic orbitals.[12] Moreover, the products of atomic orbitals $f^a(r)f^b(r)$ relate to dominant products $F^\mu(r)$ via a *product vertex* V_μ^{ab}:

$$f^a(r)f^b(r) = \sum_\mu V_\mu^{ab}F^\mu(r). \tag{16}$$

Because the dominant products $F^\mu(r)$ are themselves special linear combinations of the original products, arbitrary extra fitting functions do not enter into this scheme. In order to respect the principle of locality, the above decomposition is carried out separately for each pair of atoms, the orbitals of which overlap. By their construction, the set of coefficients V_μ^{ab} is sparse in the sense that $V_\mu^{ab} \neq 0$ only if a, b, μ all reside on the same atom pair.[12] In the construction of the dominant product basis, we made use of Talman's algorithms and computer codes for the expansion of products of orbitals about an arbitrary center and we also used his fast Bessel transform.[30]

To rewrite the defining equations of GW approximation (11) in our basis, we expand both $G(r, r', t - t')$ and $\Sigma(r, r', t - t')$ in atomic orbitals $f^a(r)$:[14]

$$G(r, r', t - t') = G_{ab}(t - t')f^a(r)f^b(r');$$
$$\Sigma(r, r', t - t') = \Sigma_{ab}(t - t')f^a(r)f^b(r'). \tag{17}$$

We also develop the screened Coulomb interaction in dominant products:

$$W_0^{\mu\nu}(t - t') = \int d^3r d^3r' F^\mu(r)W_0(r, r', t - t')F^\nu(r'). \tag{18}$$

Using Eqs. (16)–(18) it is easy to show[14] that Hedin's Eq. (11) takes the following tensorial form in our basis:

$$\mathrm{i}\chi_{\mu\nu}^0(t) = 2V_\mu^{aa'}G_{ab}^0(t)V_\nu^{bb'}G_{a'b'}^0(-t); \quad \text{non-interacting response,} \tag{19}$$

$$W_0^{\mu\nu}(\omega) = \frac{1}{\delta_\alpha^\mu - v^{\mu\beta}\chi_{\beta\alpha}^0(\omega)}v^{\alpha\nu}; \quad \text{RPA screening,} \tag{20}$$

$$\Sigma^{ab}(t) = \mathrm{i}V_\mu^{aa'}G_{a'b'}^0(t)V_\nu^{b'b}W_0^{\mu\nu}(t); \quad GW \text{ approximation,} \tag{21}$$

$$G_{ab}^{-1}(\omega + \mathrm{i}\varepsilon) = G_{0ab}^{-1}(\omega + \mathrm{i}\varepsilon) - \Sigma_{ab}(\omega + \mathrm{i}\varepsilon). \quad \text{Dyson's equation.} \tag{22}$$

Here, $v^{\mu\nu}$ denotes the Coulomb interaction $v^{\mu\nu} = \int d^3r d^3r' F^\mu(r)\frac{1}{|r-r'|}F^\nu(r')$ which, due to its positivity and symmetry, we also refer to as a "Coulomb metric." Indices are raised or lowered using either the overlaps of the dominant functions $F^\mu(r)$ or the overlaps of the atomic orbitals $f^a(r)$ and which are defined as follows:

$$O^{\mu\nu} = \int d^3r\, F^\mu(r)F^\nu(r), \; S^{ab} = \int d^3r\, f^a(r)f^b(r). \tag{23}$$

Figure 1 shows the Feynman diagram corresponding to Eq. (21). The local character of the product vertex $V_\mu^{aa'}$ is emphasized in this figure.

IV. THE INSTANTANEOUS PART OF THE SELF-ENERGY

When the screened Coulomb interaction $W_0^{\mu\nu}$ in Eq. (20) is expanded as a function of $v\chi^0$, its first term is the bare Coulomb interaction $v^{\mu\nu}\delta(t - t')$ and the corresponding self-

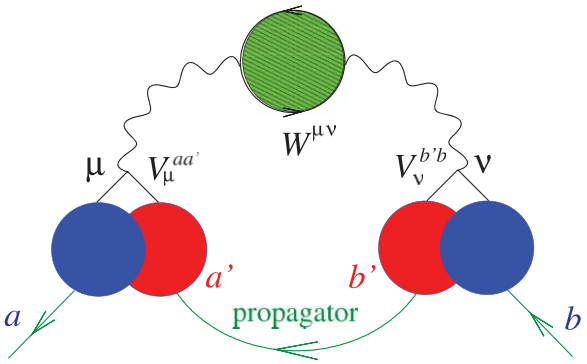

FIG. 1. Feynman diagram for the *GW* self-energy expressed in our local LCAO and dominant products basis.

energy in Eq. (21) is frequency independent. In textbook treatments of the theory of the electron gas, it is explained[23] that the Green's function $G_{ab}(t-t')$ of the electron gas must be defined, at $t=t'$, by setting $t-t'=0^-$ or by first annihilating and then creating electrons. Using this prescription and Eqs. (20) and (21), one finds the following result for the frequency-independent self-energy that corresponds to the exchange operator:

$$\Sigma_{\mathrm{x}}^{ab}(t-t') = \mathrm{i}V_\mu^{aa'}G_{a'b'}^0(0^-)V_\nu^{b'b}\delta(t-t')v^{\mu\nu}.$$

In the frequency domain, the last operator becomes a frequency independent matrix

$$\Sigma_{\mathrm{x}}^{ab} = V_\mu^{aa'}\sum_{E<0} X_{a'}^E X_{b'}^E V_\nu^{b'b} v^{\mu\nu}, \tag{24}$$

which can be computed in $O(N^3)$ operations by using the sparsity of the product vertex $V_\mu^{aa'}$.

For small molecules and clusters, the instantaneous self-energy that incorporates the effects of electron exchange may dominate over the remaining frequency dependent self-energy. If this is the case, we may substitute Σ_{x}^{ab} into Eq. (21) and finish the calculation by computing the DOS,

$$\rho(\omega+\mathrm{i}\varepsilon) = -\frac{1}{\pi}S^{ab}\,\mathrm{Im}G_{ba}(\omega+\mathrm{i}\varepsilon),$$

where we have emphasized the non-orthogonality of the basis orbitals by the explicit inclusion of the overlap S^{ab}.

However, the frequency dependent part of the self-energy contains correlation effects that significantly improve the calculation quantitatively and qualitatively as we demonstrate in Secs. VI and XI. Therefore, we shall present our approach for the frequency dependent part of the self-energy in Sec. V.

V. USING SPECTRAL FUNCTIONS TO COMPUTE THE SELF-ENERGY

One might want to solve Eqs. (19) and (21) directly as matrix valued equations in time t and to use FFTs (Ref. 31) to shuttle back and forth between the time and frequency domains. Unfortunately, however, this direct approach is doomed to fail—the behaviour of the functions $\{G_{ab}(t), \Sigma^{ab}(t), W_0^{\mu\nu}(t)\}$ at $t=0$ leads to oscillations in their Fourier transforms that die out very slowly at high frequencies.[13] We will now show how spectral functions come to the rescue and allow us to (i) respect locality in our calculations and to (ii) accelerate our calculation by means of FFT.

Let us consider the energy dependent density matrix

$$\rho_{ab}(\omega) = \sum_E X_a^E X_b^E \delta(\omega - E), \tag{25}$$

and rewrite the electronic propagator equation (14) with the help of it:

$$G_{ab}^0(\omega \pm i\varepsilon) = \int_{-\infty}^{\infty} ds\, \frac{\rho_{ab}(s)}{\omega \pm i\varepsilon - s}. \tag{26}$$

Integral representations such as these are very useful, even in finite systems where the spectral weight is concentrated at isolated frequencies.[13,14,32] Because a spectral function is broadened by the experimental resolution ε, it can be represented on a discrete mesh of frequencies, with the distance between mesh points somewhat smaller than ε. All the response functions considered in the present paper have a spectral representation because their retarded and advanced parts taken together define a single analytic function in the complex frequency plane with a cut on the real axis.

A spectral representation is merely a rather thinly disguised Cauchy integral as we can see by considering the Cauchy integral representation of the electronic Greens function:

$$G_{ab}(z) = \frac{1}{2\pi\mathrm{i}}\oint_C \frac{G_{ab}(\xi)d\xi}{\xi - z}, \tag{27}$$

where C is a path surrounding the point z with $\mathrm{Im}z > 0$ in an anti-clockwise direction. If the point $z=\infty$ is regular, the complex plane may be treated like a sphere and we may deform the contour on this sphere in such a way that it wraps around the cut on the real axis in a clockwise direction. Finally, because Green's functions take mutually hermitian conjugate values $G_{ab}(z^*) = G_{ba}^*(z)$ across the cut on the real axis, the above integral can the rewritten as

$$G_{ab}(z) = \int ds \frac{\rho_{ab}(s)}{z-s},\ \rho_{ab}(z) = -\frac{1}{\pi}\mathrm{Im}G_{ab}^0(z) \tag{28}$$

with z on the upper branch of the cut. In writing the preceding equation we have used the simplifying feature that the electronic Green's function is a symmetric matrix in our real representation of angular momenta (the same is true for the screened Coulomb interaction). In the following, we will always reconstruct correlation functions such as $\{G_{ab}(\omega),\ \Sigma^{ab}(\omega),\ W_0^{\mu\nu}(\omega)\}$ from their imaginary part or from their spectral functions. The time ordered version of such correlators is determined above (below) the real axis for positive (negative) frequencies, respectively.

A. The spectral function of a product of two correlators

The well known convolution theorem[31] tells us that the spectral content of a product of two signals is the convolution of the spectral contents of its factors. The situation is entirely analogous for Green's functions and the other correlators considered here and their spectral functions. To see this, we use the spectral representations of the time ordered factors $G_{ab}(t)$,

$\Sigma^{ab}(t)$, $W_0^{\mu\nu}(t)$ (the quantities in Eqs. (19)–(22) are time ordered):

$$\begin{aligned} G_{ab}(t) &= -\mathrm{i}\theta(t)\int_0^\infty ds\, \rho_{ab}^+(s)e^{-\mathrm{i}st} + \mathrm{i}\theta(-t) \\ &\quad\times\int_{-\infty}^0 ds\, \rho_{ab}^-(s)e^{-\mathrm{i}st}; \\ \Sigma^{ab}(t) &= -\mathrm{i}\theta(t)\int_0^\infty ds\, \sigma_+^{ab}(s)e^{-\mathrm{i}st} + \mathrm{i}\theta(-t) \\ &\quad\times\int_{-\infty}^0 ds\sigma_-^{ab}(s)e^{-\mathrm{i}st}; \\ W_0^{\mu\nu}(t) &= -\mathrm{i}\theta(t)\int_0^\infty ds\, \gamma_+^{\mu\nu}(s)e^{-\mathrm{i}st} + \mathrm{i}\theta(-t) \\ &\quad\times\int_{-\infty}^0 ds\gamma_-^{\mu\nu}(s)e^{-\mathrm{i}st}, \end{aligned} \tag{29}$$

where "positive" and "negative" spectral functions define the whole spectral function by means of Heaviside functions. For instance, the spectral function of the electronic Green's function reads

$$\rho_{ab}(s) = \theta(s)\rho_{ab}^+(s) + \theta(-s)\rho_{ab}^-(s).$$

These representations can be checked by transforming (for example) the representation of $G_{ab}(t)$ into the frequency domain and by comparing with the known expression (26). We then compute $\Sigma^{ab}(t)$ from Eq. (21) and compare the result with the second equation of Eqs. (29). The spectral function of the self-energy is seen to have the expected convolution form:

$$\begin{aligned} \sigma_+^{ab}(s) &= \int_0^\infty\int_0^\infty \delta(s_1+s_2-s)\, V_\mu^{aa'} \\ &\quad\times \rho_{a'b'}^+(s_1)V_\nu^{b'b}\gamma_+^{\mu\nu}(s_2)ds_1ds_2, \\ \sigma_-^{ab}(s) &= -\int_{-\infty}^0\int_{-\infty}^0 \delta(s_1+s_2-s)V_\mu^{aa'} \\ &\quad\times \rho_{a'b'}^-(s_1)V_\nu^{b'b}\gamma_-^{\mu\nu}(s_2)ds_1ds_2. \end{aligned} \tag{30}$$

Note that, as commented above, the $V_\mu^{aa'}$ matrices are sparse and respect spatial locality. Finally, we can easily construct the full self-energy from its spectral functions $\sigma_\pm^{ab}(s)$ by a Cauchy type integral:

$$\Sigma^{ab}(\omega \pm i\varepsilon) = \int_{-\infty}^\infty \frac{\sigma^{ab}(s)ds}{\omega \pm i\varepsilon - s}. \tag{31}$$

By entirely analogous arguments, we can find the spectral function of the non-interacting response $\chi_{\mu\nu}^0$ from Eq. (19):

$$\begin{aligned} a_{\mu\nu}(s) &= \int_0^\infty\int_0^\infty V_\mu^{ab}\rho_{bc}^+(s_1)V_\nu^{cd}\rho_{da}^-(-s_2) \\ &\quad\times\delta(s_1+s_2-s)ds_1ds_2, && \text{for } s>0; \\ a_{\mu\nu}(-s) &= -a_{\mu\nu}(s), && \text{for all } s; \\ \chi_{\mu\nu}^0(\omega+\mathrm{i}\varepsilon) &= \int_{-\infty}^\infty ds\,\frac{a_{\mu\nu}(s)}{\omega+\mathrm{i}\varepsilon-s}, && \text{for } \omega>0. \end{aligned} \tag{32}$$

We implemented the convolutions in Eqs. (30) and (32) conveniently by FFT without encountering any singularities. Please observe that analytic continuations are not needed in our approach.

We have seen in this subsection that the locality of the expressions for $\Sigma^{ab}(t)$ and $\chi_{\mu\nu}^0(t)$ in Eqs. (19) and (21) can be taken into account without multiplying singular Green's functions and by focusing instead on the spectral functions of their products.

B. The second window technique

Although we only need results in a suitable low energy window $-\lambda \le \omega \le \lambda$ of a few electronvolts, Eqs. (31) and (32) show that high energy processes at $|\omega| > |\lambda|$ influence quantities at low energies, for example, the self-energy. Therefore, these high energy processes cannot be ignored and we need the imaginary part of the screened Coulomb interaction $W_0^{\mu\nu}$ not only for small $|\omega| \le \lambda$ but also for larger frequencies. To find the imaginary part of $W_0^{\mu\nu}$, we also need, in view of Eq. (20), the non-interacting response $\chi_{\mu\nu}^0$ both at small and at large frequencies.

Let us see in the case of the density response, how the necessary spectral information can be obtained from two separate calculations in two distinct frequency windows.[13] In the large spectral window $-\Lambda \le \omega \le \Lambda$, a low resolution calculation with a large broadening (and, therefore, a coarse grid of frequencies) is sufficient to find $\chi_{\mu\nu}^0$ at large energies $|\omega| > |\lambda|$:

$$\chi_{\mu\nu}^0(\omega+\mathrm{i}\varepsilon_{\text{large}}) = \int_{-\Lambda}^\Lambda ds\,\frac{a_{\mu\nu}(s)}{\omega+\mathrm{i}\varepsilon_{\text{large}}-s}. \tag{33}$$

To get correct results in the low energy window $-\lambda \le \omega \le \lambda$, we must take into account the spectral weight in this window:

$$\begin{aligned} &\chi_{\mu\nu}^0(\omega+\mathrm{i}\varepsilon_{\text{small}}) \\ &= \int_{-\lambda}^\lambda ds\,\frac{a_{\mu\nu}(s)}{\omega+\mathrm{i}\varepsilon_{\text{small}}-s} + \left(\int_{-\Lambda}^{-\lambda} + \int_\lambda^\Lambda\right) ds\,\frac{a_{\mu\nu}(s)}{\omega+\mathrm{i}\varepsilon_{\text{large}}-s} \\ &= \chi_{\mu\nu}^{\text{small window}}(\omega+\mathrm{i}\varepsilon_{\text{small}}) \\ &\quad + \left[\chi_{\mu\nu}^{\text{large window}}(\omega+\mathrm{i}\varepsilon_{\text{large}})\right]_{\text{truncated spectral function}}. \end{aligned} \tag{34}$$

Instead of doing the second Cauchy integral directly, we construct $\chi_{\mu\nu}^{\text{large window}}$ from the large spectral window, using spectral data that are truncated for $|s| < \lambda$ to avoid counting the spectral weight in $-\lambda \le \omega \le \lambda$ twice. Moreover, the broadening constant ε is set differently in the first and second windows and corresponds to the spectral resolution in these windows. The broadening constant ε is chosen automatically in each frequency window, by setting $\varepsilon = 1.5\Delta\omega$, where $\Delta\omega$ is the distance between two points on the corresponding frequency grid. We use the second window technique (presented in Eq. (34) for the case of the density response) again in the calculation of the self-energy $\Sigma^{ab}(\omega)$, where we also need the screened interaction in two windows. We combine the spectral functions of the self-energy in exactly the same way as for the density response. For the cases considered here, computations using two spectral windows were up to one order of magnitude faster than computations using a single spectral window. An alternative way of efficiently taking into account

high energy features is by eliminating empty states altogether as recently explored by several authors.[33]

VI. TESTING OUR IMPLEMENTATION OF *GW* ON A SMALL MOLECULE

The methods presented above are sufficient to compute the self-energy equation (2) of small molecules.[34] As a test, we will compute the interacting Green's function by solving Dyson's equation (1). From this Green's function we can obtain the DOS and estimate the positions of the HOMO and LUMO levels. Here, we illustrate this procedure in the case of benzene. This molecule has been chosen because extensive theoretical results and experimental data are available for it. Our calculations show a considerable improvement using the *GW* approximation as compared to the results obtained with plain DFT calculations using local or semi-local functionals. In general, for small molecules we find a reasonable agreement with experimental data and previous *GW* calculations of the ionization potentials and electron affinities.

The input for our *GW* method has been obtained from calculations using the local density approximation (LDA) and the SIESTA package.[10] SIESTA uses a basis set of strictly confined numerical atomic orbitals. The extension of these orbitals is consistently determined by an *energy shift* parameter. In general, the smaller the energy shift, the larger the extension of the orbitals, although the procedure results in different cutoff radii for each multiplet of orbitals.[35] In the present calculations we have used the default double-ζ polarized (DZP) basis, along with the Perdew-Zunger LDA exchange-correlation functional[36] and pseudo-potentials of the Troullier-Martins type.[37] Our calculations indicate (see Table I) that it is necessary to use rather extended orbitals to obtain converged results for the HOMO and LUMO levels. For the most extended basis used here (determined from an energy shift of 3 meV), all the orbitals in benzene have a non-zero overlap and, in principle, the number of products of orbitals is $108(108+1)/2 = 5886$. This number is reduced using the algorithm described in Ref. 12, and the dominant product basis (see Eq. (16)) only contains 2325 functions. The spectral functions have been discretized using a grid with $N_\omega = 1024$ points in the range from -80 eV to 80 eV. The broadening constant has been set automatically to $\varepsilon = 1.5\Delta\omega = 0.23$ eV.

TABLE I. Ionization potentials and electron affinities for benzene versus the extension of the basis functions. The extension of the atomic orbitals is determined using the energy shift parameter of the SIESTA method (see Ref. 35). Note that rather extended orbitals are necessary to achieve converged results. Differences associated with the use of the second window technique introduced in Sec. V B are of the order of 0.1 eV (see also Table II). The experimental ionization potential is taken from the NIST server (see Ref. 40). The electron affinity of benzene is taken from Ref. 41.

Energy-shift,	One window		Two windows	
meV	IP, eV	EA, eV	IP, eV	EA, eV
150	8.48	−1.89	8.48	−2.01
30	8.71	−1.45	8.72	−1.57
3	8.76	−1.29	8.78	−1.41
Experiment	9.25	−1.12	9.25	−1.12

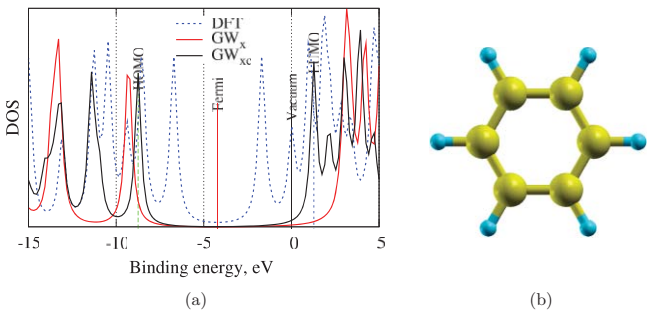

FIG. 2. (a) Density of states of benzene computed from different Green's functions using as an input the results of a DFT-LDA calculation performed with the SIESTA package. A DZP basis set, with orbital radii determined using a value of the energy shift parameter of 3 meV, has been used. The results shown in this figure are obtained with a single energy window. GW_x refers to the results obtained with only the instantaneous part of the self-energy (only exchange), while GW_{xc} labels the results obtained with the whole self-energy (incorporating additional correlation effects). (b) Ball and stick model of benzene produced with the XCRYSDEN package (see Ref. 39).

The frequency range was chosen manually by inspecting the non-interacting absorption spectrum. The results of the calculation depend only weakly on the frequency range.

Figure 2 shows the DOS calculated with different Green's functions. As one can see, the input Green's function G_0 from a DFT-LDA calculation has a very small HOMO-LUMO gap. The Green's function G obtained with the instantaneous part of the self-energy (see Eq. (24)) opens the HOMO-LUMO gap. This part of the self-energy Σ_x incorporates the effect of exchange and is very important for small molecules. However, the gap is over-estimated as one can already anticipate from typical mean-field Hartree-Fock calculations. Correlation effects are partially taken into account by the dynamical part of the *GW* self-energy. This brings the HOMO-LUMO gap closer to the experimental value. Our results stay also in agreement with other works using similar approximations (G_0W_0 on top of DFT-LDA).[19,38]

Apart from the *GW* approximations to the self-energy equation (2), our numerical method is controlled by precision parameters of a more technical nature. Table I presents the results for the ionization (IP) and electron affinity (EA) as a function of the extension of the atomic orbitals in the original LDA calculation as determined from the energy shift parameter.[35] An energy shift of 150 meV is usually sufficient to have an appropriate description of the ground-state properties of the molecules.[10] However, we can see that our *GW* calculation requires more extended (smaller energy shift) orbitals. The slow convergence of the ionization potential of benzene with respect to the quality/completeness of the basis set has also been observed in the plane-wave calculations (using Wannier functions) of Ref. 19.

Table I also shows the results of calculations using one and two energy windows. The former calculation is more straightforward but requires the same density of frequency points at higher energies as in the region of interest near the HOMO and LUMO levels. The latter calculation uses two separate frequency grids: as described in Subsection V B, a lower resolution and a larger imaginary part of the energy are used for the whole spectral range, while a high resolution and a small width are used in the low energy range

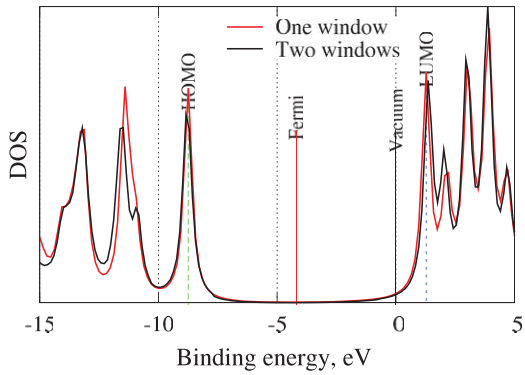

FIG. 3. The density of states of benzene computed with a uniformly discretized spectral function and using the second window technique. The peak positions are very weakly perturbed by using the two windows technique. The parameters of the calculation are identical with those of Fig. 2. The two windows technique allowed to reduce the number of frequency points from $N_\omega = 1024$ to $N_\omega = 192$.

to resolve the HOMO and LUMO levels. Thus, the second window technique requires the computation of both the response function and the screened Coulomb interaction at far fewer frequencies than the one-window calculation. For instance, the one-window results presented above have been obtained with $N_\omega = 1024$ frequencies, while the two-window results used only $N_\omega = 192$ frequencies in both windows, implying a gain of a factor of 2.7 in speed and in memory. The first and second window are defined by $\lambda = 12.58$ eV and $\Lambda = 80$ eV, respectively. The first window is chosen as $2.5(E_{\text{DFT LUMO}} - E_{\text{DFT HOMO}})$. The computational result depends only weakly on the extension of the first window. The second window is chosen manually as in the one-window calculation above. The broadening constant ε has been set separately for each spectral window, using the default value $\varepsilon = 1.5\Delta\omega$. The lower number of frequencies obviously accelerates the calculation and saves the memory, while introducing very small inaccuracies in the low frequency region. According to Table I the positions of the HOMO and LUMO calculated with one and two energy windows agree within 0.1 eV. Figure 3 shows that the second window leads to changes that are small, both in the HOMO and LUMO positions, and in the density of states at low energies. Moreover, Table II shows the dependence of the calculated IP and EA on the extension of the first frequency window (λ in Eq. (34)) when using the two windows technique. The calculations have been done using a first window of different sizes ($\lambda = 40$, 20, and 10 eV). The fineness of the frequency discretization $\Delta\omega$ was maintained; thus, a different number of frequencies is used in each case (N_ω= 512, 256, 128). The second window has been

TABLE II. The ionization potentials and electron affinities computed with different first frequency windows in the second window technique. The second energy window extends over a range of ±80 eV. Calculations were performed using atomic orbitals whose extension is defined by a 3 meV *energy shift*.

λ, eV	IP, eV	EA, eV
40	8.75	−1.29
20	8.78	−1.33
10	8.91	−1.40

set as previously with $\Lambda = 80$ eV. One can clearly see the increase of the error in the position of the level as we decrease the size of the first window. However, it is interesting to note that the results remain far more accurate than the uncertainties associated with the use of a coarse frequency grid over the whole energy interval, as would be the case in one-window calculations using a similar number of frequency points.[48] It is also important to see that using a first window defined by $\lambda = 40$ eV we recover the results obtained with a single window extended over the ±80 eV range. This clearly indicates that the use of two energy windows is an appropriate way to reduce the computational load without substantially affecting the quality of the results.

The calculations presented in this section needed a fairly large amount of random access memory (RAM). The amount of RAM increases as N^2 with N being the number of dominant products, which prohibits the treatment of larger molecules using the methods described above in a straightforward manner. However, as we will see in Sec. VII, we can use a compression method that dramatically reduces the required memory.

VII. COMPRESSION OF THE COULOMB INTERACTION

As shown in Ref. 14 it is possible to solve the Petersilka-Gossmann-Gross equations[42] for time-dependent density functional theory (TDDFT) using a Lanczos type approach if, for example, we are only interested in the polarizability tensor of the system. In this way, we avoid keeping the entire linear response matrix $\chi^0_{\mu\nu}(\omega)$ in the computer memory. Unfortunately, we were unable to find an analogous Lanczos type procedure for the self-energy matrix. However, we have found an alternative solution to this problem. It consists of taking into account the electron dynamics and keeping preferentially those dominant products that are necessary to describe $\chi^0_{\mu\nu}$ in the relevant range of frequencies.

A. Defining a subspace within the space of products

Consider the following closed form expression of the non-interacting response $\chi^0_{\mu\nu}(\omega)$ of Eq. (19):

$$\chi^0_{\mu\nu}(\omega + i\varepsilon) = 2\sum_{E,F} V^{EF}_\mu \frac{n_F - n_E}{\omega + i\varepsilon - (E - F)} V^{EF}_\nu,$$

$$\text{where } V^{EF}_\mu = X^E_a V^{ab}_\mu X^F_b. \tag{35}$$

This is a well known expression, but rewritten in the basis of dominant products.[13] It must be emphasized that we do not use this equation to compute $\chi^0_{\mu\nu}(\omega)$ (it would require $O(N^4)$ operations), but this explicit representation of the exact non-interacting response is nonetheless crucial for motivating our method of compression.

Clearly, $\chi^0_{\mu\nu}(\omega)$ is built up from $O(N^2)$ vectors V^{EF}_μ. On the other hand, the entire space of products is, by construction, of only $O(N)$ dimensions. Therefore, there must be a significant amount of collinearity in the set of vectors V^{EF}_μ and a much smaller subset of such vectors should span the space where $\chi^0_{\mu\nu}(\omega)$ acts. As candidates for the generators of this subspace, we sort the vectors $\{V^{EF}_\mu\}$ according to $|E - F|$

up to a certain *rank* N_{rank}:

$$\{X_\mu^n\} \equiv \text{subset of } \left\{V_\mu^{EF}\right\} \text{ limited according to } \times|E-F| < E_{\text{threshold}}, n = 1 \ldots N_{\text{rank}}. \quad (36)$$

Here, we treat $\{E, F\}$ as electron-hole pairs, i.e., $E < 0$ and $F > 0$.

As a first test of whether the subspace carries enough information, we define a projector onto it:

$$g^{mn} = X_\mu^m v^{\mu\nu} X_\nu^n;$$
$$P_{\mu\nu} = X_\mu^m g_{mn} X_\nu^n, \text{ where } g_{mn} = (g^{mn})^{-1};$$
$$P_\nu^\mu = v^{\mu\mu'} P_{\mu'\nu}. \quad (37)$$

It can be shown without difficulty that P_ν^μ is indeed a projector in the sense of $P^2 = P$. We can use it to project the screened Coulomb interaction onto the subspace generated by the set $\{X_\mu^n\}_{n=1\ldots N_{\text{rank}}}$:

$$W_{\text{projected}}^{\mu\nu}(\omega) = P_{\mu'}^\mu P_{\nu'}^\nu W_0^{\mu'\nu'}(\omega). \quad (38)$$

We must choose N_{rank} large enough so that the trace of the projected spectral density $W_{\text{projected}}^{\mu\nu}(\omega)$ is sufficiently close to the original one. We checked that this works even for N_{rank} considerably smaller than the original dimension of the space of products. We can go further and reduce the dimension of the subspace by eliminating collinear vectors from it. We do this by diagonalizing the matrix g^{mn} in Eq. (37) and by defining new vectors Z_μ^κ.[43]

$$g^{mn}\xi_n^\kappa = \kappa\xi_m^\kappa,$$
$$Z_\mu^\kappa \equiv X_\mu^m \xi_m^\kappa/\sqrt{\kappa}. \quad (39)$$

To define the vector space $\{Z_\mu^\kappa\}$, we first discard the eigenvectors ξ_m^κ that correspond to eigenvalues κ smaller than a threshold κ_{min} with respect to the Coulomb metric $v^{\mu\nu}$ and we then normalize the remaining vectors, for simplicity. As a result of this procedure, we obtain a smaller set of vectors that we denote again by $\{Z_{n\mu}\}$ with $n = 1 \ldots N_{\text{subrank}}$, with $N_{\text{subrank}} \leq N_{\text{rank}}$, and the additional property that they are orthonormal with respect to the Coulomb metric $v^{\mu\nu}$:

$$Z_m^\mu Z_{n\mu} = \delta_{mn}, \text{ where}$$
$$Z_m^\mu = v^{\mu\nu} Z_{m\nu}, \quad \text{for } m, n = 1 \ldots N_{\text{subrank}}. \quad (40)$$

B. Construction of the screened interaction from the action of the response function in the subspace

From the preceding discussion we know that $\chi_{\mu\nu}^0$ can be adequately represented in the previously constructed subspace $\{Z_\mu^n\}$ in the sense of

$$\chi_{\mu\nu}^0 = Z_{m\mu}\chi_{mn}^0 Z_{n\nu}, \text{ with } \chi_{mn}^0 = Z_m^\mu \chi_{\mu\nu}^0 Z_n^\nu. \quad (41)$$

To see which form the screened Coulomb interaction (3) takes for such a density response $\chi_{\mu\nu}^0$, we write it as a series:

$$W_0^{\mu\nu} = \left(\frac{1}{v^{-1}-\chi^0}\right)^{\mu\nu} = v^{\mu\nu} + v^{\mu\mu'}\chi_{\mu'\nu'}^0 v^{\nu'\nu} + v^{\mu\mu'}\chi_{\mu'\nu'}^0 v^{\nu'\mu''}\chi_{\mu''\nu''}^0 v^{\nu''\nu} + \cdots. \quad (42)$$

Because χ^0 acts—by hypothesis—only in the subspace, the series may be simplified. Lets insert the representation of the response function $\chi_{\mu\nu}^0$ of Eq. (41) into the series (42):

$$W_0^{\mu\nu} = v^{\mu\nu} + v^{\mu\mu'}\left[Z_{m\mu'}\chi_{mn}^0 Z_{n\nu'}\right]v^{\nu'\nu} + v^{\mu\mu'}\left[Z_{m\mu'}\chi_{mn}^0 Z_{n\nu'}\right] \times v^{\nu'\mu''}\left[Z_{m\mu''}\chi_{mn}^0 Z_{n\nu''}\right]v^{\nu''\nu} + \cdots,$$

then use the orthogonality property of the basis vectors $Z_{m\mu}$, Eq. (40), and find

$$W_0^{\mu\nu} = v^{\mu\nu} + Z_m^\mu \chi_{mr}^0\left[\delta_{rn} + \chi_{rn}^0 + \chi_{rs}^0\chi_{sn}^0 + \cdots\right]Z_n^\nu = v^{\mu\nu} + Z_m^\mu \chi_{mn}^{\text{RPA}} Z_n^\nu. \quad (43)$$

Here, we introduced the new response function $\chi_{mn}^{\text{RPA}} \equiv (\delta_{mk} - \chi_{mk}^0)^{-1}\chi_{kn}^0$. From the preceding arguments we conclude that the dynamically screened Coulomb interaction $W_0^{\mu\nu}$ can be computed in terms of the response function χ_{mn}^{RPA} within the previously constructed subspace, and the matrix inversion in this smaller space is of course much cheaper than in the original space. This is a welcome feature—the number of operations for matrix inversion scales with the cube of the dimension and a compression by a factor of 10 will lead to a 1000 fold acceleration of this part of the computation.

It is important to note that, although an energy cut-off $E_{\text{threshold}}$ is used to choose the relevant $\{V_\mu^{EF}\}$ vectors, high frequency components of the response and the screened interaction are explicitly calculated. $E_{\text{threshold}}$ only serves to construct the frequency independent basis vectors $\{Z_n^\nu\}$ according to Eq. (39). This basis is later used to calculate the response function $\chi_{\mu\nu}^0(\omega)$ in the whole frequency range (see Eq. (41)). Of course, we can expect that, if $E_{\text{threshold}}$ is chosen too small, the ability of the compressed basis to represent the high energy components of the response will eventually deteriorate. However, we are interested in the low energy excitations of the system and, as we will show in Subsection VII C, those can be accurately described using values of $E_{\text{threshold}}$ which allow for a large reduction in the size of the product basis. Furthermore, it is also important to note that the instantaneous self-energy Σ_x, for which a compression criterion based on our definition of $E_{\text{threshold}}$ is dubious, is calculated within the original dominant product basis, i.e., before this non-local compression is performed.

C. The compression in the case of benzene

The non-local compression depends on two parameters: (i) the maximum energy $E_{\text{threshold}}$ of the Kohn-Sham electron-hole pairs $\{V_\mu^{EF}\}$ in Eq. (36) , and (ii) the eigenvalue threshold κ for identifying the important basis vectors $\{Z_n^\nu\}$ in Eq. (39).

Table III shows the electron affinity of benzene as a function of $E_{\text{threshold}}$ and κ_{min}. The computational parameters have been chosen as described in Sec. VI and the energy shift to define the extension of the orbitals is 3 meV. Table III illustrates a general feature that we have found in many tests for several systems: N_{rank} can be chosen of the order of the number of atomic orbitals N_{orb}. We have found that $N_{\text{rank}} \approx 5N_{\text{orb}}$ usually guarantees a converged result for the HOMO and LUMO levels. In any case, the number of relevant linear combinations N_{subrank} was always much smaller than the number of

TABLE III. Electron affinities for benzene versus the compression parameters $E_{\text{threshold}}$ and eigenvalues cutoff κ_{min} in Eq. (39). In brackets the dimension of compressed subspace is given, N_{subrank}. The dimension N_{rank} is governed by $E_{\text{threshold}}$ and was 39, 297, and 765 for $E_{\text{threshold}} = 10$ eV, $E_{\text{threshold}} = 20$ eV, and $E_{\text{threshold}} = 40$ eV, accordingly. The number of atomic orbitals is $N_{\text{orb}} = 108$, while the number of dominant products is $N_{\text{prod}} = 2325$.

	$\kappa_{min} = 10^{-2}$	$\kappa_{min} = 10^{-3}$	$\kappa_{min} = 10^{-4}$
$E_{\text{threshold}} = 10$eV	2.48 (33)	2.46 (37)	2.45 (39)
$E_{\text{threshold}} = 20$eV	1.38 (96)	1.39 (133)	1.40 (171)
$E_{\text{threshold}} = 40$eV	1.41 (132)	1.41 (192)	1.41 (279)

dominant functions, with a typical compression ratio of ten or more.

To further illustrate the quality of the basis, we will compare the trace of the original screened interaction with the trace of the projected screened interaction for the benzene molecule. The result of the comparison can be seen in Fig. 4. In this test calculation, the dominant product basis consists of 921 functions, while the compressed basis contains only 248 functions. Examples of compression for larger molecules will be presented in Sec. XI.

The examples presented in this section show that the screened Coulomb interaction can be effectively compressed. A practical algorithm that uses the non-local compression and maintains the $O(N^3)$ complexity scaling of the calculation will be presented in Sec. VIII.

VIII. MAINTAINING $O(N^3)$ COMPLEXITY SCALING BY COMPRESSING/DECOMPRESSING

The favorable $O(N^2)$ scaling of the construction of the uncompressed non-interacting response $\chi^0_{\mu\nu}$ is due to its locality. On the other hand, we need compression for $\chi^0_{\mu\nu}$ to fit into the computer memory and the compressed χ^0_{mn} is no longer local. To satisfy the two mutually antagonistic criteria of (i) locality (for computational speed) and (ii) small dimension (to fit into the computer memory), we shuttle back and forth as needed between the uncompressed/local and the compressed/non-local representations of the response χ^0 and of the screened interaction W_0. Both compression and decompression are matrix operations that scale as $O(N^3)$ and this, along with the matrix inversion in Eq. (43) in the computation of the screened Coulomb interaction, and the computation of the spectral densities $\rho^{\pm}_{ab}(s)$ is the reason why our implementation of GW scales as $O(N^3)$.

A. A construction of the subspace response in $O(N^3)$ operations

Let us describe an efficient construction of the response $\chi^0_{\mu\nu}$ and its compressed counterpart χ^0_{mn} that, besides, gives us an opportunity to describe our use of frequency and time domains during the calculation. Consider Eq. (32) that involves convolutions of the spectral functions $\rho^+_{bc}(\omega)$ and $\overline{\rho^-_{ad}}(\omega) \equiv \rho^-_{ad}(-\omega)$. To make use of the convolution theorem, we will first compute the spectral function of the non-interacting response $a_{\mu\nu}(s)$ in the time domain:

$$\begin{aligned} a_{\mu\nu}(t) &= \int \frac{ds}{2\pi} a_{\mu\nu}(s) e^{\mathrm{i}st} = 2\pi \int_0^\infty V^{ab}_\mu \rho^+_{bc}(s_1) e^{\mathrm{i}s_1 t} \frac{ds_1}{2\pi} \\ &\quad \cdot \int_0^\infty V^{cd}_\nu \rho^-_{ad}(-s_2) e^{\mathrm{i}s_2 t} \frac{ds_2}{2\pi} \\ &= 2\pi V^{ab}_\mu \rho^+_{bc}(t) V^{cd}_\nu \overline{\rho}^-_{ad}(t). \end{aligned} \tag{44}$$

In other words, we prepare the use of the FFT driven convolution by first computing the Fourier transforms of the electronic spectral densities $\rho^{\pm}$ and once $a_{\mu\nu}(t)$ is computed, we return to $a_{\mu\nu}(s)$ by an inverse Fourier transformation. This is nothing else but the fast convolution method with the Fourier transform of the spectral densities $\rho^+_{bc}(\omega)$ and $\overline{\rho}^-_{ad}(\omega)$ carried out prior to the tensor operations in Eq. (44).

Above we saw how to compute the spectral function of the non-interacting density response. However, as we mentioned before, we cannot easily store this information in the memory of the computer and we must therefore compress this quantity as soon as it is found to avoid over flooding the computer memory. An efficient way to do this is to compute $a_{\mu\nu}(t)$, the spectral function of the non-interacting response in the time domain, in a "time by time" fashion, with the

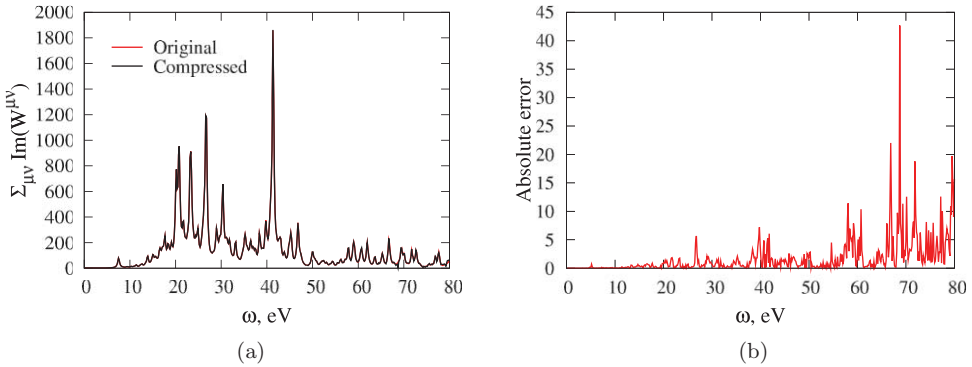

FIG. 4. (a) Comparison of the screened interaction calculated for benzene using our original dominant product basis and the screened interaction projected to a compressed product basis (see Eq. (38)). We plot the sum of all the matrix elements of the imaginary part of the screened interaction. (b) A plot of the difference of the functions represented in panel (a). The change in spectral weight of the screened interaction due to compressing the space of dominant products is seen to be small. Please notice the different scales of the y-axis in both panels.

time variable t in the outer loop. Fortunately, the compression of the response $\chi^0_{\mu\nu}$ according to Eq. (41) can be done on the level of its spectral function $a_{\mu\nu}(t)$ separately for each time t.

B. A construction of the self-energy in $O(N^3)$ operations

Although we use the spectral function given by Eq. (30) to compute the self-energy in the second equation of Eqs. (29), it is useful to think also of Eq. (21) that corresponds to the Feynman diagram of Fig. 1 and which has the same locality properties. Please recall that the product vertex in Eq. (16) is sparse and local and that the indices $\{a, a', \mu\}$ and $\{b, b', \nu\}$ must each reside on a single pair of overlapping atoms. Once the indices a, b of the self-energy are specified, there are only $O(N^0)$ possibilities of choosing the remaining indices. Therefore, the calculation of $\Sigma^{ab}(t)$ requires asymptotically $O(N^2)$ operations provided that the screened Coulomb interaction $W_0^{\mu\nu}$ in a basis of localized functions is known. However, the local screened Coulomb interaction $W_0^{\mu\nu}$ in the original space of dominant products does not fit into the computer memory as opposed to the compressed, but non-local, response χ^{RPA}_{mn} that we store (see Eq. (43)). We may, however, regain locality by decompressing χ^{RPA}_{mn} at the cost of $O(N^3)$ operations, using the identity $W^{\mu\nu}_{\mathrm{dynamical}} = Z^\mu_m \chi^{\mathrm{RPA}}_{mn} Z^\nu_n$ in Eq. (43). As we cannot keep $W^{\mu\nu}_{\mathrm{dynamical}}$ in the computer memory, we must try to "decompress χ^{RPA}_{mn} on the fly." To do this, let us transform the first equation of Eqs. (30) into the time domain. For instance, for the positive part of the spectral density $\sigma^{ab}_+(t)$ of the self-energy, we find

$$\sigma^{ab}_+(t) = 2\pi V^{aa'}_\mu \rho^+_{a'b'}(t) V^{b'b}_\nu \gamma^{\mu\nu}_+(t).$$

Again, the representation in time of $\rho^+_{a'b'}(t)$ is prepared only once. However, the transform $\gamma^{\mu\nu}_+(t) = -\frac{1}{\pi} Z^\mu_m \mathrm{Im}\chi^{\mathrm{RPA}}_{mn}(t) Z^\nu_n$ for all times does not fit into the computer memory. Therefore, we also decompress $\gamma^{\mu\nu}_+(t)$ time by time by letting the time t run in the outer loop, by computing $\gamma^{\mu\nu}_+(t)$ via decompression for a single time, and by storing only the result $\sigma^{ab}_\pm(t)$ for each time. Once we have computed $\sigma^{ab}_\pm(t)$ for all times, we can find $\sigma^{ab}_\pm(s)$ from it.

IX. A SUMMARY OF THE COMPLETE ALGORITHM

At this point, it is useful to briefly recapitulate the different steps of our implementation of Hedin's *GW* approximation. It consists of the following steps:

1. Export the results of a DFT code that uses numerical local atomic orbitals as a basis set. Here, we use the SIESTA code,[10] but other codes like the FHI-AIMS code[44] could also be used.
2. Set up a basis of dominant products in $O(N)$ operations. Here, we use the method of Ref. 12.
3. Set up a space of reduced dimension where the screened Coulomb interaction will act and exploiting the low effective rank of this set. Such a subspace is determined by a set of N_{rank} vectors V^{EF}_μ that correspond to electron-hole pairs with a predetermined maximum value of $|E - F|$. Further compression is obtained by diagonalizing the Coulomb metric projected onto this subspace. This step requires $O(N^3)$ operations.
4. Choose low and high energy spectral windows and a frequency grid. Prepare the electronic spectral density $\rho_{ab}(s)$ in these two windows from the output of the DFT calculation.
5. Find χ^{RPA}_{mn} by constructing and compressing the local $\chi^0_{\mu\nu}$ on the fly in $O(N^3)$ operations and by solving for χ^{RPA}_{mn} for all frequencies in $O(N^3)$ operations. The construction must be done in two frequency windows. Truncate the spectral data where needed in order to avoid double counting and store χ^{RPA}_{mn} in the two windows.
6. Find the spectral function of the self-energy by decompressing $\chi^{RPA}_{mn} \rightarrow W^{\mu\nu}_0$ on the fly in $O(N^3)$ operations and by convolving it with the electronic spectral function. Again this must be done in two frequency windows and the results must be combined consistently.
7. Construct the self-energy from its spectral representation.
8. Solve Dyson's equation and find the density of states from the interacting Green's function. Obtain the desired spectroscopic information from the density of states.

Results obtained with the above algorithm will be discussed in Secs. X and XI.

X. TESTS FOR DIFFERENT BASIS SET SIZES

The concept of locality is a key ingredient in the algorithm described in the present paper. To take into account the locality of the electronic interactions we use a basis set of atomic orbitals to describe the electronic states. In particular, the results presented here have been obtained starting from DFT Kohn-Sham calculations performed with the SIESTA code.[10] SIESTA uses a basis set numerical atomic orbitals of finite support which has been specially designed to improve the efficiency of the calculations by reducing the interaction range, and thus the number of non-zero Hamiltonian and overlap matrix elements.[10] However, it is important to stress that the algorithm presented here does not depend critically on this choice. Our *GW* can also be coupled to other codes using numerical atomic orbitals as a basis set.

So far we have only used a DZP basis in our calculations. DZP basis are known to provide results in reasonably good agreement with those obtained with plane-wave calculations for a large variety of systems[35] and, for this reason, they have become the standard choice for ground-state calculations with SIESTA.[10] However, the performance of DZP basis for the calculation of the IP and EA is more uncertain. For this reason we present here a study of the effect of the size of the basis on the results of our *GW* calculations for benzene. We will consider the cases of single-ζ polarized (SZP), DZP, triple-ζ polarized (TZP), and triple-ζ doubly-polarized (TZDP) basis sets. SZP basis set contains only one radial shape to describe each shell of valence electrons occupied in the atomic ground-state configuration plus a polarization shell. Therefore, for carbon it has one $2s$ and three $2p$ orbitals, plus five

TABLE IV. Convergence of the calculated ionization potential and affinity level as a function of the number N_{orb} of numerical atomic orbitals (see Ref. 10) in the basis set describing a benzene molecule. The convergence of the total energy with respect to the TZDP result $\Delta E_{\text{total}} = E_{\text{total,TZDP}} - E_{\text{total,Basis}}$ and that of the HOMO and LUMO levels in the DFT-LDA calculations is also displayed. This allows to check the improvement of the description of the ground state of the molecule as we increase the size of the basis. The calculations are performed for a fixed geometry obtained using the TZDP basis. DZP is the standard basis used in ground-state SIESTA (see Ref. 10) calculations. See the text for the definition of the acronyms describing the different basis. N_{prod} is the number of dominant products, i.e., before performing a non-local compression, for different basis sizes.

Basis type	N_{orb}	ΔE_{total}, eV	$E^{\text{LDA}}_{\text{HOMO}}$, eV	$E^{\text{LDA}}_{\text{LUMO}}$, eV	IP, eV	EA, eV	N_{prod}
SZP	78	7.680	−7.90	−2.98	9.85	−0.32	2319
DZP	108	0.506	−6.72	−1.67	8.78	−1.31	4155
TZP	138	0.329	−6.71	−1.66	8.78	−1.27	5493
TZDP	186	0.000	−6.67	−1.62	8.78	−1.24	8487

polarization orbitals with *d* symmetry. For hydrogen, we have one 1*s* orbital and three polarization orbitals with *p* symmetry. In DZP and TZP basis the number of valence orbitals is doubled and tripled, respectively, while the number of polarization orbitals is also doubled in the TZDP basis. We only consider automatic basis sets generated following the algorithms presented in Ref. 10. The radii of the orbitals is chosen using an *energy shift* of 3 meV. We restrict here to this type of basis because currently it is the most widely used with the SIESTA code. It is possible to further improve the accuracy of the ground-state calculations using smoothly confined variationally optimized orbitals as described in Ref. 35. The geometry of the molecule is optimized with the TZDP basis and then used in other calculations.

The second column in Table IV shows the convergence of the total energy of the benzene molecule with respect to the basis size. The energy increases as the size of the basis decreases, reflecting the limitations of the small basis to accurately represent the molecular ground state. The main change takes place when moving from the SZP to the DZP basis. This strong effect of the doubling of the basis is well known. It is related to limitations in the generation of optimal basis functions for molecules from isolated atoms.[10,35,47] In comparison, the effect of including a third radial shape or an additional shell of polarization orbitals is much smaller. As shown in Table IV, the situation is similar for the Kohn-Sham eigenvalues corresponding to the HOMO and LUMO levels. They are relatively insensitive to the increase of the size of the basis beyond DZP when using this type of basis.[10]

The IP and EA computed with different basis sets are also collected in Table IV. The *GW* calculations have been carried out with a single frequency window defined by $\Lambda = 80$ eV and a large number of frequency points $N_\omega = 2048$. The convergence is strikingly fast for the IP, for which the DZP basis provides the same result as the larger basis. For the EA the convergence is slower, although the DZP basis already provides a result which is only 0.07 eV lower than the value obtained with our largest basis. Our converged results are close to those obtained in other G_0W_0 calculations on top of DFT-LDA results.[19,38] However, their very fast convergence is somewhat surprising if we consider, for example, the results by Umari *et al.*[19] This reference reports a very slow convergence of the IP of the benzene molecule as a function of the number of unoccupied bands included in the calculation. Although our value of 8.78 eV is close to the values reported in that work using a large number of bands (energy cutoff close to ~100 eV), it is still far from the extrapolated result of 9.1 eV for an infinite energy cutoff. This clearly indicates that the convergence shown in Table IV does not reflect a saturation of the basis, but rather the fact that the additional orbitals do not provide the necessary additional degrees of freedom for full convergence. Thus, more work should be carried out to investigate how this kind of numerical basis can be improved for the description of one-electron excitations. A possible route that we plan to investigate is a projection technique similar to that described in Ref. 47, which can be applied to improve the description of unoccupied states. Still the results obtained with our DZP basis are quite satisfactory, particularly if we take into account the relatively small number of orbitals in such basis, and we will use it for larger molecules in Sec. XI.

XI. TESTS FOR MOLECULES OF INTERMEDIATE SIZE

The compression technique has been carefully tested in the case of the benzene molecule. The tests show excellent agreement of the density of states computed with and without compression. In this section, we will consider larger molecules, such as the hydrocarbons naphthalene and anthracene.[34] These molecules are well known to differ in their character as electron acceptors: naphthalene, like

TABLE V. Ionization potentials and electron affinities for naphthalene and anthracene and their dependence on the extension of the atomic orbitals. For naphthalene we compared the results obtained with spectral functions discretized in one or two windows. The experimental data has been taken from the NIST server (see Ref. 40). For naphthalene and anthracene, vertical ionization potentials are not available at the NIST database. Therefore, we give experimental ionization energies including effects of geometry relaxation.

	Naphthalene				Anthracene	
Energy-shift	One window		Two windows		Two windows	
meV	IP, eV	EA, eV	IP, eV	EA, eV	IP, eV	EA, eV
200	7.24	−0.68	7.27	−0.79	6.44	0.20
20	7.61	−0.083	7.67	−0.18	6.89	0.77
Experiment	8.14	−0.191	8.14	−0.191	7.439	0.530

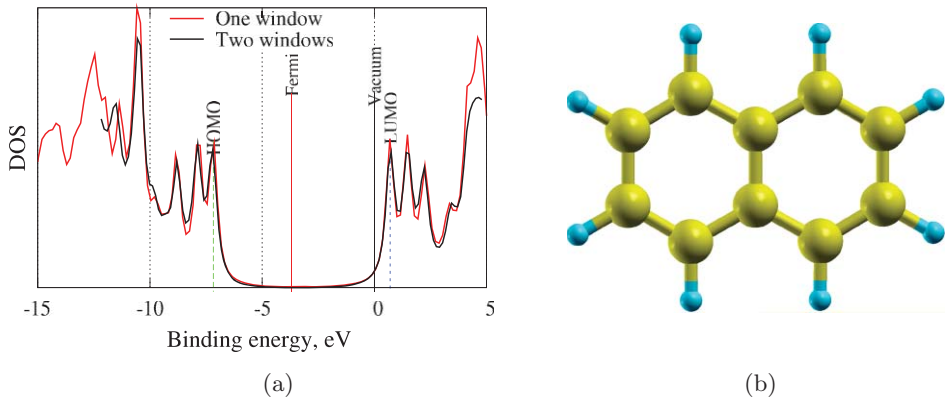

FIG. 5. (a) Density of states for naphthalene. The results have been obtained with our most extended basis orbitals (corresponding to an energy shift of 20 meV (see Ref. 35)). We can appreciate the accuracy of the second window technique. (b) Ball and stick model of naphthalene produced with the XCRYSDEN package (see Ref. 39).

benzene, has a negative electron affinity, while anthracene is an electron acceptor with positive electron affinity.

A compression of the dominant product basis is necessary to treat the molecules considered in this section. These molecules are too large for a calculation without compression on ordinary desktop machines because of memory requirements. For naphthalene the dominant product basis contained 4003 functions, which were reduced to 433 functions after compression. In the case of anthracene, the dominant product basis contained 5796 functions, while the compressed basis had only 598 functions.

Table V shows our results for naphthalene and anthracene. The computational details were similar to those used for benzene and were already described in Sec. VI. The two-window results were obtained with frequency grids of only 128 points for naphthalene (in the ranges ± 8.32 eV and ± 80 eV) and yet it provides an accuracy on the 0.1 eV level, while the one-window calculation is done again with 1024 frequencies (in the range of ± 80 eV). In the case of anthracene, results using frequency grids of 256 points (in the ranges ± 16 eV and ± 0 eV) are presented.

Again we find a large improvement over the position of the Kohn-Sham levels in a DFT-LDA calculation. The agreement with the experimental data is certainly improved, although there are still significant deviations, particularly with respect to the reported ionization potentials. Interestingly, however, our calculations recover the important qualitative feature of anthracene being an electron acceptor. In the case of anthracene, the results obtained with our most extended basis orbitals (energy shift of 20 meV) are in excellent agreement with the recent calculations by Blase *et al.*[5] In the case of naphthalene we can see that the two frequency windows technique introduces only tiny differences (below 0.1 eV) in the positions of the HOMO and LUMO levels.

The corresponding DOS is shown in Figs. 5 and 6. One can see in Fig. 6 that it is the dynamical part of the self-energy, including correlation effects, that turns our theoretical anthracene into an acceptor, while including only the instantaneous self-energy predicts anthracene to be a donor.

These results for molecules of modest size are just a first application of our algorithm. With its favorable scaling, our method aims at *GW* calculations for larger molecules of the type used in organic semiconductors. However, before carrying out such studies, we should reduce the initial number of dominant products, i.e., before any compression is applied to it.

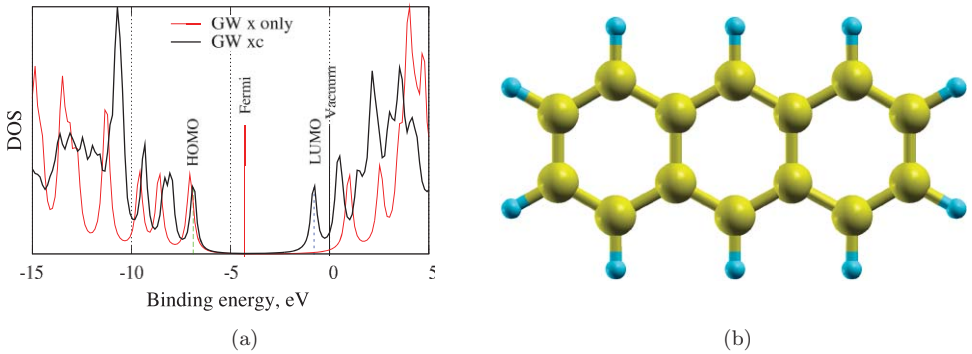

FIG. 6. (a) Density of states for anthracene. The results have been obtained using the extended basis orbitals corresponding to an energy shift of 20 meV (see Ref. 35). Here, we compare calculations using the instantaneous (exchange-only) self-energy and the full self-energy (including correlation effects). The correlation component of the self-energy is crucial to reproduce the experimental observation that anthracene is an acceptor. In contrast, the exchange-only calculation locates the LUMO level above the vacuum level. (b) Ball and stick model of anthracene produced with the XCRYSDEN package (see Ref. 39).

XII. CONCLUSIONS AND OUTLOOK

In the present paper we have described our approach to Hedin's GW approximation for finite systems. This approach provides results for densities of states and gaps that are in reasonable agreement with experiment and it requires only modest computer resources[34] for the systems presented here. The complexity of our algorithm scales asymptotically as the third power of the number of atoms, while the needed memory grows with the second power of the number of atoms. We hope that these features, along with a further reduction of the size of the basis describing the products of localized orbitals, will allow to apply our method to describe the electronic structure of large molecules and contribute to an *ab initio* design of organic semiconductors for technological applications.

The algorithm described here is built upon the LCAO technique[10] and uses a previously constructed basis in the space of orbital products that preserves their locality and avoids fitting procedures.[12] Moreover, a (non-local) compression technique has been developed to reduce the size of this basis. This allows to store the whole matrix representation of the screened Coulomb interaction at all time/frequencies in random access memory while significantly reducing the computational time. The time (and frequency) dependence of observables is treated with the help of spectral functions. This avoids analytical continuations and allows for operations to be accelerated by the use of FFTs. As a useful byproduct of our focus on spectral functions we obtain, as primary result, an electronic spectral function of the type observed in photoemission and from which we then read off the HOMO and LUMO levels.

We have applied our method to benzene, naphthalene, and anthracene. As expected, we find that our estimations of the HOMO and LUMO positions and the corresponding gaps are significantly improved over the results obtained from the Kohn-Sham eigenvalues in a plain DFT-LDA calculation. Our results approach the experimental data but, as observed by other authors,[5] these "single-shot" GW-LDA calculations (or G_0W_0-LDA using a more standard terminology) still present sizeable deviations from the measured ionization potentials and electron affinities. In general, our results are in good agreement with previous G_0W_0-LDA calculations for similar systems.[5,19,38] Thus, we expect further improvements by iterating our procedure until self consistency or, as suggested by other authors in the case relatively small molecules,[5,20,45] by using Hartree-Fock results as an input for our G_0W_0 calculations. For periodic systems it is well known that G_0W_0-LDA systematically underestimates the size of the gaps of semiconductors. The best results so far were found using the so-called "improved quasi-particle method."[2,46] A realization of this method in our framework should also improve the precision of our results.

The method presented in this paper depends crucially on the quality and size of the original LCAO basis. A possible limitation is that the typical LCAO basis used in electronic structure calculations is constructed and optimized in order to describe ground-state properties.[35] However, it is possible to optimize an LCAO basis, for example, using a technique similar to that described in Ref. 47, to represent electronically excited states. This will increase the accuracy and applicability of the method and could even allow to reduce the size of the original LCAO basis used to represent the electronic states. Moreover, by comparing our basis with that of other authors, there are indications that the (local) basis of dominant products used in this paper can be reduced in dimension without changing the physical results.[5] Such a reduction should lead to an important improvement of the prefactor in our implementation of GW, but also, as a side effect, introduce a similar acceleration in our published TDDFT algorithm[14] that is already competitive, in its present form, with other TDDFT codes.

The quantities calculated in the presented algorithm can be useful in other branches of many-body perturbation theory. For instance, the screened Coulomb interaction is a crucial ingredient of the Bethe-Salpeter technique that is needed to study excitons and the optical response of excitonic systems. In this context it is interesting to note [49] that the solution of the Bethe-Salpeter equation scales as $O(N^3)$ for clusters of size N, at least when suppressing the dynamic part of the fermion self-energy and the dynamic part of the screening of the Coulomb interaction. Calculations of the transport properties of molecular junctions[50] are another possible application of the GW approach described here.

ACKNOWLEDGMENTS

We thank Olivier Coulaud for useful advice on computing, and Isabelle Baraille and Ross Brown for discussions on chemistry, both in the context of the ANR project "NOSSI." James Talman has kindly provided essential computer algorithms and codes, and we thank him, furthermore, for inspiring discussions and for correspondence. We are indebted to the organizers of the ETSF2010 meeting at Berlin for feedback and perspective on the ideas of this paper. Arno Schindlmayr, Xavier Blase, and Michael Rohlfing helped with extensive correspondence on various aspects of the GW method. D.S.P. and P.K. acknowledge financial support from the Consejo Superior de Investigaciones Científicas (CSIC), the Basque Departamento de Educación, UPV/EHU (Grant No. IT-366-07), the Spanish Ministerio de Ciencia e Innovación (Grant No. FIS2010-19609-C02-02), and the ETORTEK program funded by the Basque Departamento de Industria and the Diputación Foral de Guipuzcoa.

[1] L. Hedin, Phys. Rev. **139**, A796 (1965); For a review see F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998); C. Friedrich and A. Schindlmayr, *Many-Body Perturbation Theory: The GW Approximation*, NIC Series Vol. 31 (Lecture Notes, 2006 Winter School, Forschungszentrum Julich, Germany), p. 335, available online at http://www2.fz-juelich.de/nic-series/volume31/friedrich.pdf.
[2] M. van Schilfgaarde, T. Kotani, and S. Faleev, Phys. Rev. Lett. **96**, 226402 (2006); especially Fig. 1 of this paper.
[3] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).
[4] H. Hoppe and N. S. Sariciftci, J. Mater. Res. **19**, 1924 (2004).
[5] X. Blase, C. Attaccalite, and V. Olevano, Phys. Rev. B **83**, 115103 (2011).
[6] M. S. Kaczmarski, Y. Ma, and M. Rohlfing, Phys. Rev. B **81**, 115433 (2010).
[7] X. Blase, private communication (15 December 2010); J. Chelikowsky and M. Tiago, private communication (15 December 2010).
[8] H. Eshuis, J. Yarkony, and F. Furche, J. Chem. Phys. **132**, 234114 (2010).

[9]M. M. Rieger, L. Steinbeck, I. D. White, H. N. Rojas, and R. W. Godby, Comput. Phys. Commun. **117**, 211 (1999).
[10]J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, J. Phys.: Condens. Matter **14**, 2745 (2002); E. Artacho, E. Anglada, O. Dieguez, J. D. Gale, A. García, J. Junquera, R. M. Martin, P. Ordejón, J. M. Pruneda, D. Sánchez-Portal, and J. M. Soler, J. Phys.: Condens. Matter **20**, 064208 (2008).
[11]M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[12]D. Foerster, J. Chem. Phys. **128**, 34108 (2008).
[13]D. Foerster and P. Koval, J. Chem. Phys. **131**, 044103 (2009).
[14]P. Koval, D. Foerster, and O. Coulaud, J. Chem. Theory Comput. **6**, 2654 (2010).
[15]The code solves the Petersilka-Gossmann-Gross equations of TDDFT linear response and will become available under an appropriate non-commercial license.
[16]A. Natan, A. Benjamini, D. Naveh, L. Kronik, M. L. Tiago, S. P. Beckman, and J. R. Chelikowsky, Phys. Rev. B **78**, 75109 (2008).
[17]L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman, and R. Schneider, J. Chem. Phys. **129**, 014109 (2008).
[18]N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).
[19]P. Umari, G. Stenuit, and S. Baroni, Phys. Rev. B **79**, 201104R (2009).
[20]C. Rostgaard, K. W. Jacobsen, and K. S. Thygesen, Phys. Rev. B **81**, 085103 (2010).
[21]D. Pines, *Elementary Excitations in Solids* (Wiley, New York, 1964).
[22]P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, England, 1995).
[23]A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (Dover, New York, 2003).
[24]P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[25]W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[26]F. Reinert and S. Hüfner, New J. Phys. **7**, 97 (2005).
[27]P. Fulde, *Electron Correlations in Molecules and Solids*, Springer Series in Solid-State Sciences Vol. 100 (Springer, Berlin, 1991).
[28]R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, England, 2004).
[29]J. E. Harriman, Phys. Rev. A **34**, 29 (1986).
[30]J. D. Talman, J. Chem. Phys. **80**, 2000 (1984); J. Comput. Phys. **29**, 35 (1978); Comput. Phys. Commun. **30**, 93 (1983); **180**, 332 (2009).
[31]W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes* (Cambridge University Press, Cambridge, England, 2007).
[32]The technique of discretizing spectral functions was also used by M. Shishkin and G. Kresse, Phys. Rev. B **74**, 035101 (2006).
[33]P. Umari, G. Stenuit, and S. Baroni, Phys. Rev. B **81**, 115104 (2010); J. A. Berger, L. Reining, and F. Sottile, Phys. Rev. B **82**, 41103R (2010).
[34]All the calculations presented in this paper have been performed on one core of an Intel Core Quad CPU Q9400 2.66 GHz using the MKL BLAS library.
[35]J. Junquera, Ó. Paz, D. Sánchez-Portal, and E. Artacho, Phys. Rev. B **64**, 235111 (2001).
[36]J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[37]N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[38]M. L. Tiago and J. R. Chelikowsky, Phys. Rev. B **73**, 205334 (2006); T. A. Niehaus, M. Rohlfing, F. Della Sala, A. Di Carlo, and Th. Frauenheim, Phys. Rev. A **71**, 022508 (2005).
[39]A. Kokalj, Comput. Mater. Sci. **28**, 155 (2003).
[40]Experimental data obtained from the NIST Computational Chemistry Comparison and Benchmark DataBase at http://cccbdb.nist.gov.
[41]J. C. Rienstra-Kiracofe, Ch. J. Barden, Sh. T. Brown, and H. F. Schaefer, J. Phys. Chem. A **105**, 524 (2001).
[42]M. Petersilka, U. J. Gossmann, and E. K.U. Gross, Phys. Rev. Lett. **76**, 1212 (1996); M. E. Casida, in *Recent Advances in Density Functional Theory*, edited by D. P. Chong, (World Scientific, Singapore, 1995), p. 155.
[43]Including $O(N^2)$ electron-hole pairs EF into the diagonalization procedure would scale as $O(N^6)$, but this is unnecessary, as the total dimension of the product space is only $O(N)$. Furthermore, our computations show explicitly that a small number of low energy pairs is already sufficient to get correct results (see Subsection VIIC).
[44]V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comput. Phys. Commun. **180**, 2175 (2009).
[45]P. H. Hahn, W. G. Schmidt, and F. Bechstedt, Phys. Rev. B **72**, 245425 (2005).
[46]San-Huang Ke, eprint arXiv: 1012.1084.
[47]D. Sánchez-Portal, E. Artacho, and J. M. Soler, J. Phys.: Condens. Matter **8**, 3859 (1996).
[48]Notice that the use of 512, 256, and 128 points over a range of 160 eV would lead to $\Delta\omega = 0.31$, 0.63, and 1.25 eV, respectively. These values are much larger than $\Delta\omega = 0.16$ eV used for the first window in our calculations using the two-windows technique.
[49]L. X. Benedict and E. L. Shirley, Phys. Rev. B **59**, 5441 (1991); W. G. Schmidt, S. Glutsch, P. H. Hahn, and F. Bechstedt, Phys. Rev. B **67**, 085307 (2003).
[50]M. Brandbyge, J. L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, Phys. Rev. B **65**, 165401 (2002).