Short range DFT combined with long-range local RPA within a range-separated hybrid DFT framework

E. Chermak and P. Reinhardt*
Laboratoire de Chimie Théorique,
Université Pierre et Marie Curie
4 place Jussieu, F — 75252 Paris, France and
CNRS, UMR 7616, 4 place Jussieu, F — 75252 Paris, France

B. Mussard
Université de Lorraine, CRM2, UMR 7036, Vandœuvre-lès-Nancy, F — 54506, France and
CNRS, CRM2, UMR 7036, Vandœuvre-lès-Nancy, F — 54506, France

J.G. Ángyán†
CNRS, CRM2, UMR 7036, Vandœuvre-lès-Nancy, F — 54506, France
Université de Lorraine, CRM2, UMR 7036, Vandœuvre-lès-Nancy, F — 54506, France and
"CEU IAS" (Central European University, Institute of Advanced studies) Nádor utca 13, Budapest 1051, Hungary

Selecting excitations in localized orbitals to calculate long-range correlation contributions to range-separated density-functional theory can reduce the overall computational effort significantly. Beyond simple selection schemes of excited determinants, the dispersion-only approximation, which avoids counterpoise-corrected monomer calculations, is shown to be particularly interesting in this context, which we apply to the random-phase approximation. The approach has been tested on dimers of formamide, water, methane and benzene.

INTRODUCTION

As it is now widely recognized, standard functionals fail dramatically when calculating the interaction energy of rare-gas dimers [1, 2]. Local-density approximations (LDA) overbind heavily [3] while hybrid functionals produce almost all possible types of results including lack of binding, like in the case of the popular B3LYP functional, to relatively reasonable van der Waals minima, and sometimes strong overbinding [4]. The imprecisions in describing London dispersion-type correlation effects have far-reaching implications, which are not limited to the flaws in modeling weakly bound van der Waals complexes. For instance, functionals that are popular in modern quantum chemistry due to their ability to describe thermochemistry data with a reasonable accuracy [5], while keeping the computational requirements on a low level as compared to wavefunction-based calculations, may lead to systematic errors in some specific cases, like the isodesmic stabilization energies of n-alkanes [6]. Therefore we could witness a rapidly increasing interest in new methodologies aiming at completing well-known density functionals by adding systematically the missing dispersion interactions. This may be achieved, for instance, by adding non-local density functional corrections based on some fundamental considerations about the LDA response function of an inhomogeneous electron gas [7, 8]. A much more pragmatic way is to add atom-atom corrections, parameterized either semi-empirically [9, 10], or by using more general considerations, like in the case of the so-called exchange-hole model [11] or in the Tkatchenko-Scheffler method [12]. A considerably more demanding computational approach is to add missing explicit correlation contributions to the total energy obtained from Kohn-Sham orbitals via the adiabatic connection fluctuation-dissipation theory (ACFDT) scheme, using the random phase approximation (RPA). The main advantage of these methods is that while the local and semi-local correlation functionals miss long-range correlation effects, the methods based on the RPA provide an accurate estimation for these. On the other hand, RPA is significantly worse in reproducing short-range correlation effects, for which standard functionals perform better.

Range-separation is one of the modern tools to overcome the deficiencies of density-functional methods in dealing with long-range dispersion interactions. It is based on a single parameter connecting pure DFT (Kohn-Sham calculations) to Hartree-Fock and post-Hartree-Fock approaches [13] by an appropriate non-linear scaling of the electron-electron interactions. The short-range interaction part is taken into account by density-functional theory, while long-range exchange and correlation contributions are left to an ab initio wavefunction-based treatment. The Hartree-Fock type description of the long-range exchange turns out to be a generalization of the concept of hybrid density functionals, sometimes evoked as range separated hybrid (RSH) method, which produces independent-particle, single-determinant wave functions. From a density functional viewpoint such an approach is a Generalized Kohn-Sham (GKS) scheme, which is an appropriate starting point for a wave function treatment of long-range electron correlation by vari-

* Peter.Reinhardt@upmc.fr
† Janos.Angyan@univ-lorraine.fr

ational or perturbational techniques.

In view of improving DFT methods for dealing with van der Waals complexes various range-separation based correlation approaches have been proposed, ranging from second-order perturbation theory [14] to most elaborate coupled-cluster approaches [15]. An intermediate ab-initio correlation method, the random-phase approximation, which sums certain perturbation diagrams to infinite order, has also been adapted to the range-separated scheme[16–20] and has lead to good results.

Since in the London dispersion interaction problem we are mainly concerned with long-range electron correlations, the selection of the most relevant excited determinants can be enormously improved by working in a localized one-electron basis, i.e. in localized orbitals. This philosophy, which follows the pioneering works by Kapuy [21] and by Pulay [22], is the basis of the family of local correlation methods, where the selection of the relevant excitation is usually done after criteria of spatial distance between the centroids of localized orbitals [23, 24]. Beyond a substantial speed-up, local correlation schemes are able to exclude a priori the excitations that contribute to the basis set superposition effects (BSSE).

The purpose of the present work is to explore the possible advantages of a local RPA method for the calculation of dispersion energies. We recall in a first section the basic ingredients of RSH scheme: the short-range DFT and the long-range RPA, as well as the two practical aspects which consist in the construction of localized orbitals and the selection of the relevant excitations. In a second part, section II, results on some selected systems (dimers of water, methane, formamide and benzene) are presented and discussed, and the significant simplifications achieved via the dispersion-only approximation will be summarized in our conclusions. Some technical details are collected in the Appendix.

I. THEORY

A. Range-separated density-functional theory

The electron-electron (e-e) interaction operator of the electronic Hamiltonian can be split rigorously as a combination of a long-range (lr) contribution, which dominates almost exclusively the interaction from a given e-e distance, and a complementary short-range (sr) contribution, which has a Coulomb singularity when the inter-electronic distance approaches to zero. Such a separation can be achieved in several alternative ways, e.g. by using the error function splitting:

$$\frac{1}{r_{ij}} = w_{ee}^{\rm lr} + \left(\frac{1}{r_{ij}} - w_{ee}^{\rm lr}\right) \quad \text{with} \quad w_{ee}^{\rm lr}(r_{ij}) = \frac{\mathrm{erf}(\mu\, r_{ij})}{r_{ij}} \tag{1}$$

The parameter μ (more precisely its inverse) governs the range separation, i.e. it is proportional to the distance where the sr-contribution becomes negligible besides the lr one.

The, in principle exact, ground state energy of a many-electron system can be obtained in a two-step process via:

$$E = E_{\rm RSH} + E_c^{\rm lr} \tag{2}$$

where $E_c^{\rm lr}$ is the long-range correlation energy, usually approximated by some wave-function method, and $E_{\rm RSH}$ is given by

$$E_{\rm RSH} = \min_{\Phi} \left\{ \langle \Phi | \hat{T} + \hat{V}_{\rm ne} + \hat{W}_{\rm ee}^{\rm lr} | \Phi \rangle + E_{\rm Hxc}^{\rm sr}[n_\Phi] \right\} \tag{3}$$

with the kinetic energy operator $\hat{T}$, the nuclei-electron interaction operator $\hat{V}_{ne}$ and the electron-electron interaction $\hat{W}_{\rm ee}^{\rm lr}$ written with $w_{ee}^{\rm lr}$. $E_{\rm Hxc}^{\rm sr}[n_\Phi]$ is the short-range μ-dependent Hartree-exchange-correlation functional, and Φ is a single-determinant wave function.

The minimizing single determinant is given by the Kohn-Sham-like one-electron equations with the full-range Hartree interaction of electrons $\hat{V}_{\rm H}$, the long-range part of the Hartree-Fock type exchange operator $\hat{V}_{\rm x}^{\rm lr}$, and a short-range exchange-correlation potential, $\hat{V}_{\rm xc}^{\rm sr}$ related to the sr xc functional, $E_{\rm xc}^{\rm sr}[n_\Phi]$:

$$\left(\hat{T} + \hat{V}_{\rm ne} + \hat{V}_{\rm H} + \hat{V}_{\rm x}^{\rm lr} + \hat{V}_{\rm xc}^{\rm sr}\right) |\phi^{\rm RSH}\rangle = \epsilon\, |\phi^{\rm RSH}\rangle \tag{4}$$

Since we are interested in a solution of the RSH equations (4) in localized orbitals, the iterative Singles-Configuration-Interaction scheme described in Ref. [25] will be preferred to the conventional iterative diagonalization of the RSH matrix. This procedure has the advantage of leaving the final orbitals as close as possible to the set of starting orbitals, both for occupied and virtual orbitals, and allows us to maintain the localized nature of the *a priori* constructed localized initial guess orbitals (*vide infra*). Due to the invariance of single determinants with respect to orbital rotations within the occupied subspace the resulting minimizing single-determinant wave function is equivalent with the usual canonical solution, which can be obtained from the converged RSH matrix by a single diagonalization step.

Having the localized RSH orbitals and the corresponding Fock matrix elements at hand (see Section I C), the remaining long-range correlation part in equation (2) should be calculated. In many-body perturbational approaches special attention should be paid to the nonlinear nature of the Hamiltonian, which may lead to additional contributions as compared to conventional perturbation methods Ref. [26] (see also [27]).

The separation of the correlation energy to a short-range DFT part and a long-range wave function part reduces significantly the dependence on basis sets [28], since the most strongly basis set dependent electron-electron cusps are taken into account quite well by the density functional part.

Additionally, the size of the individual integrals shows a separation into important contribution, arising from spatially close-lying orbitals, and less important ones from distant molecular orbitals. These two aspects should render the scheme in localized orbitals appealing.

B. The Random Phase Approximation

The Random Phase Approximation has become quite popular as a post-DFT correlation method [29]. It includes an infinite summation of correlation diagrams, beyond second-order perturbation theory, and it is invariant to orbital rotations[30] contrary to, for instance, Epstein-Nesbet perturbation theory. It has been shown that RPA is equivalent with a CCD (coupled cluster doubles) approximation [51]. We use RPA as a long-range correlation method in the following.

Among the different ways of expressing the RPA correlation energy [20], we have chosen the adiabatic-connection-fluctuation-dissipation-theorem (ACFDT) equation [31]. In this formulation the RPA correlation energy is written as an integral over a coupling constant α:

$$E_c = \frac{1}{2}\int_0^1 d\alpha \; tr\left\{\mathbb{F}^{lr}\mathbb{P}^{lr}_{c,\alpha}\right\} \tag{5}$$

The coupling constant α scales the electron-electron interaction, and adiabatically connects the physical system ($\alpha = 1$) to the RSH reference system ($\alpha = 0$). The correlation energy E_c is then logically written as the previous integral, between $\alpha = 0$ and $\alpha = 1$.

$\mathbb{F}^{lr}$ is a matrix involving the long-range two-electron integrals and $\mathbb{P}^{lr}_{c,\alpha}$ is the correlation part of the two-particule density matrix, obtained from the solution vectors of the long-range RPA equations. The size of all matrices is given by the product of the number of occupied and virtual orbitals ($n_{occ} \times n_{vir}$), *i.e.* the number of single excitations.

Several versions of long-range RPA co-exist in the literature [19, 20], depending whether exchange is included in the kernel of the long-range RPA equations used to compute $\mathbb{P}^{lr}_{c,\alpha}$ and whether anti-symmetrized integrals are used in the definition of $\mathbb{F}^{lr}$. In the direct-RPA (dRPA) version of long-range RPA one neglects exchange, while in the RPA-exchange (RPAx) variant the exchange is included in the kernel. Further variants of these two versions of long-range RPA can be sought depending on using or dropping antisymmetrized integrals when forming the ACFDT integrand. The variants without antisymmetrization are labelled dRPA-I and RPAx-I (*-I* for single-bar), whereas variants with antisymmetrized integrals are named dRPA-II and RPAx-II (*-II* for double-bar). For a detailed overview on these energy expressions, see Ref. [31].

For the present calculations, the version RPAx-I is employed throughout, as this flavor of long-range RPA seems to yield the most reliable results in a range-separated context [18]. In this version, we have :

$$\mathbb{F}^{lr} = \begin{pmatrix} \mathbf{K} & \mathbf{K} \\ \mathbf{K} & \mathbf{K} \end{pmatrix} \tag{6}$$

with $K_{ia,jb} = \langle ab|ij\rangle^{lr}$, and $\mathbb{P}^{lr}_{c,\alpha}$ is constructed from the solution vectors of :

$$\begin{pmatrix} \epsilon + \alpha\mathbf{A}' & \alpha\mathbf{B} \\ \alpha\mathbf{B} & \epsilon + \alpha\mathbf{A}' \end{pmatrix} \begin{pmatrix} \mathbf{X}_{n,\alpha} \\ \mathbf{Y}_{n,\alpha} \end{pmatrix} = \omega_{n,\alpha} \begin{pmatrix} \mathbf{X}_{n,\alpha} \\ -\mathbf{Y}_{n,\alpha} \end{pmatrix} \tag{7}$$

with $\mathbf{A}'_{ia,jb} = \langle ib||aj\rangle^{lr}$, $\mathbf{B}_{ia,jb} = \langle ab||ij\rangle^{lr}$, and $\epsilon_{ia,jb} = F^{lr}_{ab}\delta_{ij} - F^{lr}_{ij}\delta_{ab}$, where F^{lr} is the corresponding RSH operator.

C. Localized orbitals

Many methods are available to generate localized *occupied* orbitals, by projection or maximization of functionals like the common Foster-Boys, Pipek-Mezey, von Niessen, or Edmiston-Ruedenberg methods. All these can be applied equally well to Hartree-Fock or Kohn-Sham orbitals. *Virtual* localized orbitals may be generated as non-orthogonal projected atomic orbitals [22], as pair natural orbitals [32], as Optimized Virtual Orbital Space, OVOS [33], as complementary Boys-localized orbitals, or simply Pipek-Mezey localized ones, occupying the least number of expansion centers per orbital.

Having well-localized orbitals within the monomers is crucial to render the selection of the excitations efficient, and to allow us to separate inter- and intra-molecular excitations in the context of intermolecular interactions. To obtain this separation automatically, without the need for a beforehand construction of canonical, delocalized molecular dimer orbitals, we choose a slightly different route than the previously cited ones: localized occupied and virtual orbitals will be constructed on the same footing. First the RSH equations are solved in canonical orbitals for the monomers in the monomer basis, and for the monomers in the dimer basis. Starting from the first set, the orbitals are localized through a projection procedure, described in reference [34], and recalled in the appendix. This method constructs a minimal set of molecular occupied and virtual orbitals, and adds the remaining virtual space as projected, orthogonalized atomic orbitals.

With these localized monomer orbitals in the *monomer* basis and a converged Kohn-Sham operator of the monomers in the *dimer* basis, a Singles-CI procedure is used to add the ghost basis set as supplementary virtual orbitals.

From the calculation of the monomers in the monomer basis the virtual orbitals are taken, and from the monomers in the dimer basis the occupied ones, in order to generate starting orbitals for the Singles-CI orbital optimization of the real dimer system [25]. In this

way the orbitals and the corresponding excitations can be clearly identified in the respective monomer and the dimer systems.

At the end, three orbital sets are available (monomer orbitals in the monomer basis, counterpoise-corrected monomer orbitals and dimer orbitals), generated from one single guess and with overlaps in the vicinity to one.

D. Selection of single excitations

As it has been mentioned, the localized-orbitals framework can reduce the computational effort by reducing the number of significant long-range two-electron integrals to be taken into account. In order to reduce the dimensions of the matrices to be constructed and treated for the long-range RPA calculations, further considerations are needed to optimize the selection of determinants. A possibility is to use an energy criterion of perturbative nature, by considering the second-order approximation to long-range RPA, which happens to be the standard MP2 energy expression in the RPAx-I variant. It will be supposed that only those single excited determinants Φ_i^a (*i.e.* a single excitation), which satisfy the following condition:

$$\frac{\langle\Phi_0|\hat{W}_{\rm ee}^{\rm lr}|\Phi_{i\bar{i}}^{a\bar{a}}\rangle^2}{E(\Phi_{i\bar{i}}^{a\bar{a}})-E_0} = \frac{(\langle ii|aa\rangle^{lr})^2}{2(\epsilon_a-\epsilon_i)} > \tau \qquad (8)$$

are going to provide a significant contribution to the long-range RPA energy. This evaluation is very rapid, since it implies only long-range two-electron integrals and diagonal elements of the RSH matrix. $|\Phi_0\rangle$ and E_0 being respectively the RSH wavefunction and energy, ϵ_a and ϵ_i repectively the RSH virtual and occupied orbital energies, $\hat{W}_{\rm ee}^{\rm lr}$ the range separation operator from the equation 1, $\Phi_{i\bar{i}}^{a\bar{a}}$ a determinant of two single excitations, and $E(\Phi_{i\bar{i}}^{a\bar{a}})$ its corresponding energy. The dimension of the long-range RPA matrix is exactly the number of singly excited determinants thus selected, in contrast to Configuration-Interaction-based methods, where the matrix dimension is roughly proportional to the square of the selected single excitations.

Beyond the selection through the energetic importance of a determinant and in the context of inter-molecular interactions single excitations can be classified into intra-molecular and inter-molecular ones. The former ones have complete analogs in the individual monomer system, as ideally orbitals are little changed during the construction of the dimer wave function. The latter ones should be of limited importance as the correlation part of the interaction between monomers is mainly governed by dispersion-type interactions, which imply two coupled intra-monomer excitations.

A significant saving for the calculation of the correlation contribution to the interaction energy is obtained when (1) considering only single excitations within the same monomers, assuming that inter-monomer excitations have only small amplitudes with our choice of the construction of the orbitals perfectly centered on the individual monomers, and (2) considering for the calculation of the long-range RPA correlation contribution to the interaction energy only the dispersion-type combination of the intra-monomer single excitations (see the schema of figure 1). This assumes that the intra-monomer correlation contributions are about the same for the dimer and the isolated monomers, and should not contribute significantly to the binding. The partial summation over the long-range RPA amplitudes exploits the property that the correlation energy is a linear functional of the amplitudes.

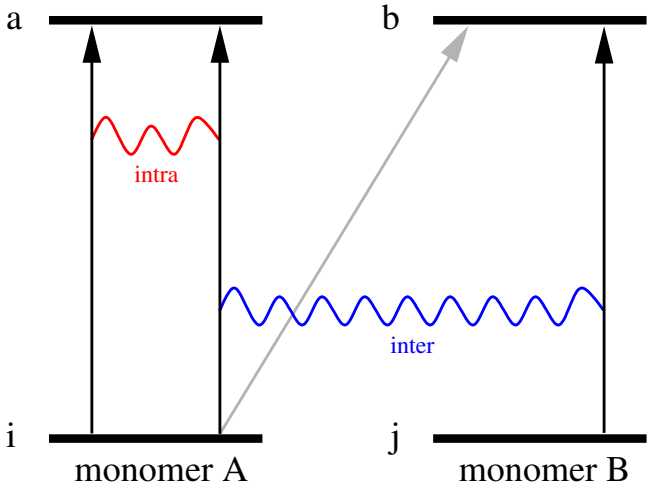

FIG. 1. Single excitations (arrows) between an occupied (i or j) and a virtual (a or b) orbital in two monomers. These single excitations can either be intra-molecular (black arrows) or inter-molecular (gray arrow). One can then consider the coupling of two intra-molecular single excitations either within the same monomer (red double wavy line) or between two monomers (blue wavy line). These last couplings we call 'dispersion-type combination of intra-monomer excitations'.

II. RESULTS

In srDFT+lrRPA, one can tweak the functional, the version of long-range RPA and the range separation parameter to obtain a kind of best combination. The present study aims, however, at providing insights in the general context of the srDFT+lrRPA. We will fix the functional as the srPBE [35] and the long-range RPA version as the RPAx-I (see previous sections), and vary only the range-separation parameter. Conclusions should not be affected by the specific choices made before.

We will study a few selected dispersive and hydrogen-bonded systems: the dimers of water (one hydrogen bond), methane (dispersion-only), formamide (two hydrogen bonds) and, finally, the larger benzene dimer,

bound by dispersion interactions. Geometries are taken from the S22 test set [36], and as basis set we use (excepted the benzene dimer) the one given by Voisin [37], which has been designed specifically for intermolecular interactions (see Appendix for more details). For benzene we employed the standard aug-cc-pvdz basis [38]. All correlation energies are evaluated with frozen core orbitals.

If we look at the overlaps produced by the proposed Singles-CI procedure (see section I C), we have more than 99.9 % throughout for the *occupied* dimer orbitals and the monomer orbitals, and 0.96±0.06, 0.97±0.05, 0.96±0.07 and 0.99±0.03, repectively, for the virtual dimer and monomer orbitals for the four complexes of water, methane, formamide and benzene. The procedure of perturbing the monomer orbitals through single excitations toward the dimer orbitals deforms indeed only very slightly the occupied orbitals, and little the virtual ones.

A. Correlation energy differences

In this first part we employ the selection of single excitations individually to the dimer and the separate monomers. Applying selection criteria of 10^{-8}, 10^{-9}, and 10^{-10} Hartrees, for a range separation parameter $\mu = 0.5$ a.u. the results converge to the result obtained without any selection, as displayed in Table I, together with reference data of explicitely correlated coupled-cluster (F12-CCSD(T)) results of Ref. [39]. Larger thresholds than 10^{-8} Hartree lead eventually to repulsive lr-correlation contributions.

ΔE(mH)/dimer	water	methane	formamide	benzene
srPBE	−7.423	+0.393	−22.98	+0.898
+lrRPAx-I (10^{-8})	−7.548	+0.307	−24.21	−0.420
+lrRPAx-I (10^{-9})	−8.178	−0.140	−25.44	−1.742
+lrRPAx-I (10^{-10})	−8.396	−0.500	−25.91	−2.880
+lrRPAx-I (all)	−8.628	−0.664	−27.27	−3.759
F12-CCSD(T)	−7.865[a]	−0.818[a]	−25.38[a]	−4.318[a]

[a] from Ref. [50], aug-cc-pvtz basis set

TABLE I. srPBE and srPBE+lrRPAx-I (for a range separation parameter of μ=0.5 a.u.) interaction energies for the water, methane and formamide dimers in the Voisin-ANO basis and the benzene dimer in the aug-cc-pvdz basis, for different selection thresholds. All core orbitals are frozen, and results are BSSE corrected. For comparison we give as well published explicitly correlated coupled-cluster results.

The possible savings while applying the selection criteria can be seen in Figure 2, where we display the the evolution of the individual correlation energies with the number of selected determinants. We see that with roughly half of the determinants more than 90 % of the complete long-range RPA energy is recovered. Moreover, the selection criterion becomes more and more efficient the larger the dimer systems is, as the overlaps between occupied and virtual orbitals become less and less important. The figure gives as well an impression of the orders of magnitude between the actual correlation energies and the contribution to the interaction, leading to the slow convergence of the selection scheme as applied here.

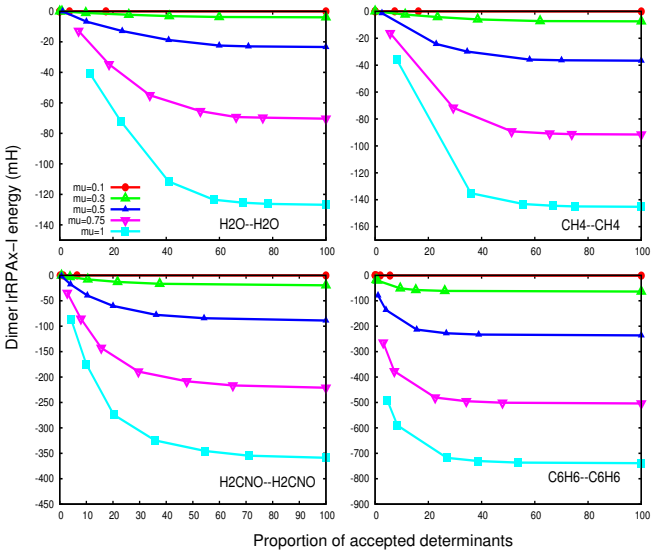

FIG. 2. Evolution of the lrRPAx-I dimer energies (in mH) with the number of selected determinants (in % of the possible determinants) for thresholds of 10^{-5}, 10^{-6}, 10^{-7}, 10^{-8}, 10^{-9}, 10^{-10} Hartrees (from left to right, as tighter thresholds select more determinants) for the water (top left), the methane (top right), the formamide (bottom left) and the benzene dimer (bottom right).

B. Dispersion-only contribution to the interaction energy

As we see the poor convergence of the correlation contributions to the interaction energy with tighter selection criteria, a scheme selecting determinants homogeneously within the monomers and the dimer may be more adequate, making use of the clear identifiability of the molecular orbitals of monomers and the dimer. We propose here to regard only dispersion-type correlation diagrams (see figure 1) of all possible classes of diexcitations in the dimer system, without considering the monomers explicitly, but as furnishing starting orbitals for constructing the monomer-localized dimer orbitals.

As a supplementary simplification in terms of computational effort, we may construct and solve the long-range RPA equations only within that space of single excitations, neglecting all excitations having the occupied and virtual orbital on different monomers. In addition, as in the previous section, the single excitations can be selected through the previously employed scheme.

We present in Figure 3 the result of this selection of dispersion-only correlation contributions, for the commonly employed range-separation parameter $\mu = 0.5$ a.u. As the correlation energy is displayed at a whole, all the inter-monomer excitations other than dispersion are missing in the right-most bar of the graph. Nevertheless, the dispersion part (topmost part of the bars), is nearly unchanged between the two dimer calculations. As the attribution to the different classes of diexcitations follows the attribution of orbitals to the monomers, the explicit form of the starting monomer orbitals (canonical, localized, approximated ...) has no impact on the final result.

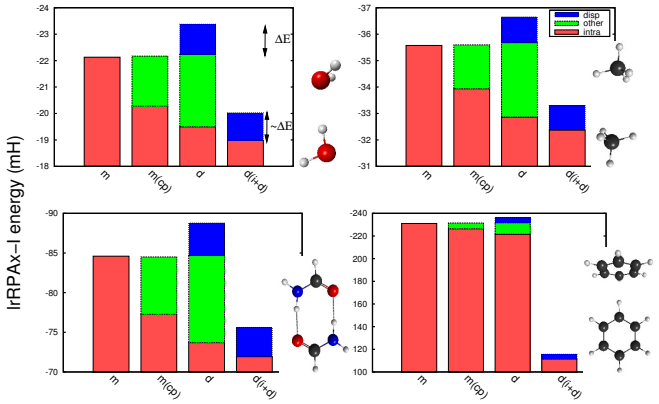

FIG. 3. Decomposition of the lrRPAx-I correlation energies into different classes of excitations for the water(top left), methane (top right), formamide (bottom left) and benzene (bottom right) dimer, for a range-separation parameter of $\mu = 0.5$ a.u.. The left-most bar in each part shows the monomer in the monomer basis (by definition only intra-molecular correlation energy), the next the monomer in the whole dimer basis (separation into intra-molecular and other contributions like BSSE), then the dimer correlation energy with the monomer part, the dispersion part and the remaining contributions, and finally, the right-most bar, the dimer correlation energy when using only the intra-molecular and dispersion-type excitations in the RPA equations. Note that the zero of the energy scale is not included in the diagram.

One should note the small difference of the monomer correlation energies when including or not the ghost basis sets of the other monomer, showing again the weak basis-set dependence of the RSH+lrRPA calculations. When including the ghost basis sets, the excitations can be grouped into those within the occupied and virtual orbitals of the monomer, and those from the occupied orbitals toward the orbitals originating from the ghost basis set. This part seems to be relatively large (green parts of the second monomer bar in each panel), as the virtual space of the starting orbitals in this calculation is the Löwdin (or $S^{-1/2}$) orthogonalized ensemble of the virtual orbitals of the monomer calculation in the monomer orbitals and the additional ghost-basis atomic orbitals, without a hierarchy of virtual and ghost-orbital space. As RPA is invariant to rotations of the orbital spaces,

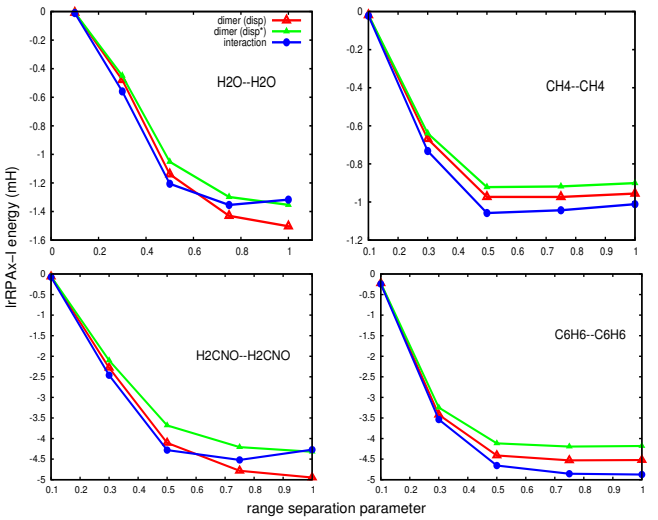

FIG. 4. lrRPAx-I interaction energies in mH (in full circles) for the water(top left) , methane (top right), formamide (bottom left) and benzene (bottom right) dimer, versus the range separation parameter μ. "disp" (in hollow triangles) and "disp*" (in full triangles) represent the dispersion contribution to the dimer energy computed respectively from the whole RPA matrices and the intramolecular plus dispersion based RPA matrices.

the overall correlation energy is independent of this technical detail here, delocalizing slightly orbitals due to the orthogonality constraints.

We remark two coincidences that should not be fortuitous: the coincidence of a) the dispersion contribution with the contribution to the interaction energy on the one hand, and of b) the solution of the long-range RPA equation for all excitations with the solution of the long-range RPA equation only for the intra- and dispersion-type excitations on the other hand. Other correlation contributions to the interaction energy like induction, due to the deformation of the monomers in the dimer system and the corresponding change in correlation energy, are taken into account by the short-range density functional. Only the dispersion should be accounted for by the lr-RPA equations.

Changing μ from very small values (toward pure DFT calculations) and large values (regular RPA) shows clearly this aspect (see Figure 4): the long-range RPA contributions to the interaction energies deviates from the dispersion-only approximation in the case of the two hydrogen-bound systems for larger values of μ. This qualitative difference in the weight of the dispersion contributions to the long-range RPA interaction energies should be still more pronounced in the case of charged monomers. This will be studied elsewhere.

As a final step we apply the selection of the single excitations to the calculation of the dispersion-only correlation contribution to the interaction energy. Table II shows the data for $\mu = 0.5$ a.u., as a function of the se-

lection threshold.

E (mH)/dimer	water	methane	formamide	benzene
ΔE_{srPBE}	−7.423	+0.393	−22.984	−0.564
$+E_{lrRPAx-I}$	−7.433	+0.364	−23.02	−0.074
(disp, 10^{-5})	*−7.433*	*+0.364*	*−23.02*	*−0.074*
$+E_{lrRPAx-I}$	−7.758	+0.033	−23.94	−1.101
(disp, 10^{-6})	*−7.758*	*+0.033*	*−23.94*	*−1.101*
$+E_{lrRPAx-I}$	−8.021	−0.294	−24.90	−2.301
(disp, 10^{-7})	*−8.031*	*−0.296*	*−24.98*	*−2.230*
$+E_{lrRPAx-I}$	−8.299	−0.398	−25.74	−2.943
(disp, 10^{-8})	*−8.331*	*−0.407*	*−25.89*	*−2.806*
$+E_{lrRPAx-I}$	−8.452	−0.522	−26.37	−3.300
(disp, 10^{-9})	*−8.510*	*−0.554*	*−26.62*	*−3.088*
$+E_{lrRPAx-I}$	−8.474	−0.528	−26.57	−3.429
(disp, 10^{-10})	*−8.547*	*−0.568*	*−26.89*	*−3.180*
$+E_{lrRPAx-I}$	−8.475	−0.528	−26.67	−3.513
(disp)	*−8.562*	*−0.579*	*−27.09*	*−3.218*
$+\Delta E_{lrRPAx-I}$ (all)	−8.628	−0.664	−27.27	−3.759
$\Delta E_{F12-CCSD(T)}$	−7.865	−0.818	−25.38	−4.318

TABLE II. Interaction energies (in mH) for the water, methane, formamide and benzene dimer, for a range separation parameter of 0.5 a.u., and different selection thresholds. The correlation contribution is evaluated from the dispersion-type part of the RPA energies, as part of the all dimer RPA components, and, in *italics, from the solution of the RPA equations in intra- and dispersion parts only*. All core orbitals are frozen.

The convergence is much more rapid than for the energy differences, however not really smooth, as for instance the energy difference between selection thresholds of 10^{-7} and 10^{-8} Hartree are in the same order of magnitude as for the difference of 10^{-9} and 10^{-10} Hartree. We could not yet determine a consistent extrapolation scheme to the final value, taking into account all determinants without selection. Nevertheless, we see from the table that taking into account all of the long-range RPA equations or just those for intra- and dispersion-type diexcitations does not make a large difference. On the contrary: the latter values are slightly closer to the complete long-range RPA interaction without selection than the result of the evaluation of all of the long-range RPA equations.

In terms of considered determinants, and thus the computational effort, we can go back to Figure 2, from which we see that for a threshold of 10^{-8} Hartree only 40 % of the total number of determinants are involved (for benzene only 20 %). This savings enter quadratically in the number of matrix elements.

III. CONCLUSIONS

In this article we propose to select excitations within long-range RPA correlation corrections to range-separated hybrid density-functional theory.

From the presented data we conclude that the chosen selection of excitations via an energy criterion leads to imprecise results when applied separately to the dimer system and the individual monomers, due to the relative smallness of the calculated interaction energy (several percent of the individual contributions only). In the context of range-separated density-functional theory, however, we observe that the direct calculation of the long-range RPA contribution to the interaction energy via the dispersion-type excitations in the long-range RPA equations leads to about the same result as the complete long-range RPA calculations without selections, with a much more rapid convergence towards the complete long-range RPA interaction energy. The computational effort is significantly reduced as well, as a result of a much smaller number of excitations involved and the unnecessary evaluation of the counterpoise corrected monomers energy. This seems consistently satisfied for the dispersion-type interactions of the benzene and the methane dimer, and as well for the single and double hydrogen bonds in the water and formamide dimer. For the latter, the good coincidence of the dispersion-only and the complete long-range RPA calculation is lost for large values of the range-separation parameter, showing that in these cases other contributions than dispersion are not any more taken into account by the short-range DFT functional.

The construction of the orbital space (occupied and virtual), assigned to the monomers in a fragment-oriented approach is essential for this decomposition of the correlation energy.

ACKNOWLEDGEMENTS

All calculations have been carried out in the Laboratory of Theoretical Chemistry in Paris. Support from the ANR project WADEMECOM was very helpful, including the local version of the Molpro program package, developed within this collaboration. Discussions with J. Toulouse, K. Sharkas and A. Savin (Paris), and N. Ben Amor and D. Maynau (Toulouse) are gratefully acknowledged. J. G. A. is grateful to the CEU IAS for the senior fellowship.

TECHNICAL DETAILS

A. The Voisin-ANO basis

The "Voisin basis" is a 7s4p(O, N, C)/3s(H) type Van Duijneveldt [40] basis, contracted from a 12s7p/6s primitive basis, and augmented by Voisin [41] with diffuse and

polarization functions, leading to a 13s8p3d/10s2p primitive basis, contracted once for each angular momentum to lead to an overall basis set described as 8s5p3d/4s2p [41]. This basis set has been used in our group for previous studies [42, 43] on similar molecules.

The localization method [34] we use in this study constructs in a first step local guess orbitals from linear combinations of atom-like, step-wise (core, valence, ...) orthogonalized orbitals from a minimal set of basis functions through the chemical intuition of the bonding in a molecule (σ or π bonding/antibonding orbitals, lone pairs). Occupied orbitals from this construction should represent already closely the electronic SCF density of the molecule, hence the use of an Atomic Natural Orbital (ANO [44, 45]) basis set when possible. The guess orbitals are projected onto the occupied space, then hierarchially orthogonalized, i.e. occupied orbitals among themselves, then the occupied with the virtual orbitals and finally the virtual and diffuse orbitals among themselves.

From the completely decontracted Voisin basis we constructed thus $1s$, $2s$ and $2p$ atomic orbitals with a modified atomic Hartree-Fock program [46], and we made sure that both the original and the partial-ANO Voisin basis yield the same RHF interaction energy for small complexes.

B. Employed computer codes

For the calculation of short-range DFT energies and the construction of the RSH matrix in a given set of orbitals a development version of the Molpro package based on the 2010.1 release [47] was used. The SCF procedure itself and the hierarchical generation of orbital sets (see section I C) was carried out with the local-orbital code of Paris [48]. Long-range integrals for the long-range RPA part are calculated using an intermediate version of the Dalton 2011 package [49], transformed on the molecular-orbital basis by a local program [48], that was used as well as for the selection of determinants and the generation of the input lists for the evaluation of the long-range RPA energy. For that latter task routines have been written (B. M.) as part of a development version of Molpro and compiled as a stand-alone tool reading the generated lists of integrals and excitations. In canonical orbitals, results are identical to corresponding calculations employing Molpro.

The explicit construction of starting orbitals relies on the code of the university of Toulouse [34].

[1] D. C. Patton, M. R. Pederson, Phys. Rev. A 56 (1997) 2495–2498.
[2] S. Kristyán, P. Pulay, Chem. Phys. Lett. 229 (1994) 175–180.
[3] J. Harris, Phys. Rev. B 31 (1985) 1770–1779.
[4] I. C. Gerber, J. G. Ángyán, J. Chem. Phys. 126 (2007) 044103.
[5] W. Koch, M. C. Holthausen, A Chemist's Guide to Density Functional Theory, Wiley-VCH Verlag GmbH, Weinheim, pp. i–xiii.
[6] M. D. Wodrich, C. Corminboeuf, P. v. R. Schleyer, Organic Letters 8 (2006) 3631–3634.
[7] Y. Andersson, D. Langreth, B. Lundqvist, Phys. Rev. Lett. 76 (1996) 102–105.
[8] K. Lee, E. Murray, L. Kong, B. Lundqvist, D. Langreth, Phys. Rev. B 82 (2010) 081101.
[9] S. Grimme, J. Comp. Chem. 25 (2004) 1463–1473.
[10] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, J. Chem. Phys. 132 (2010) 154104.
[11] A. D. Becke, E. R. Johnson, J. Chem. Phys. 122 (2005) 154104.
[12] A. Tkatchenko, M. Scheffler, Phys. Rev. Lett. 102 (2009) 073005.
[13] A. Savin, in: J. M. Seminario (Ed.), Recent developments and Applications of Modern Density Functional Theory, Elsevier, Amsterdam, 1996, pp. 327–257.
[14] J. G. Ángyán, I. C. Gerber, A. Savin, J. Toulouse, Phys. Rev. A 72 (2005) 012510.
[15] E. Goll, H. J. Werner, H. Stoll, T. Leininger, P. Gori-Giorgi, A. Savin, Chem. Phys. 329 (2006) 276–282.
[16] J. Toulouse, I. C. Gerber, G. Jansen, A. Savin, J. G. Ángyán, Phys. Rev. Lett. 102 (2009) 096404.
[17] B. Janesko, T. Henderson, G. Scuseria, J. Chem. Phys. 131 (2009) 034110.
[18] W. Zhu, J. Toulouse, A. Savin, J. Ángyán, J. Chem. Phys. 132 (2010) 244108.
[19] J. Toulouse, W. Zhu, A. Savin, G. Jansen, J. Ángyán, J. Chem. Phys. 135 (2011) 084119.
[20] J. G. Ángyán, R.-F. Liu, J. Toulouse, G. Jansen, J. Chem. Theory Comput. 7 (2011) 3116–3130.
[21] E. Kapuy, C. Kozmutza, J. Chem. Phys. 94 (1991) 5565–5573.
[22] P. Pulay, Chem. Phys. Lett. 100 (1983) 151–154.
[23] M. Schütz, G. Rauhut, H. Werner, J. Phys. Chem. A 102 (1998) 5997–6003.
[24] N. Ben Amor, F. Bessac, S. Hoyau, D. Maynau, J. Chem. Phys. 135 (2011) 014101.
[25] P. Reinhardt, J. P. Piquemal, A. Savin, J. Chem. Theo. Comp. 4 (2008) 2020–2029.
[26] J. G. Ángyán, Phys. Rev. A 78 (2008) 022510.
[27] E. Fromager, H. J. A. Jensen, Phys Rev A 78 (2008) 022504.
[28] R. H. Hertwig, W. Koch, J. Comp. Chem. 16 (1995) 576–585.
[29] H. Eshuis, J. E. Bates, F. Furche, Theor. Chem. Acc. 131 (2012) 1084.
[30] T. Bouman, B. Voigt, A. Hansen, J. Am. Chem. Soc. 101 (1979) 550–558.
[31] J. G. Ángyán, R.-F. Liu, J. Toulouse, G. Jansen, J. Chem. Theo. Comp. 7 (2011) 3116–3130.
[32] F. Neese, F. Wennmohs, A. Hansen, J. Chem. Phys. 130 (2009) 114108.
[33] M. Pitonak, F. Holka, P. Neogrady, M. Urban, J. Mol.

Struc.: THEOCHEM 768 (2006) 79–89.
[34] D. Maynau, S. Evangelisti, N. Guihéry, C. Calzado, J. Malrieu, J. Chem. Phys. 116 (2002) 10060.
[35] E. Goll, H. J. Werner, H. Stoll, Phys. Chem. Chem. Phys. 7 (2005) 3917–3923.
[36] P. Jurecka, J. Sponer, J. Cerny, P. Hobza, Phys. Chem. Chem. Phys. 8 (2006) 1985–1993.
[37] C. Voisin, Contributions to the computation of the induction term in intermolecular potentials for the modelisation of polypeptids, Ph.D. thesis, Nancy University, France, 1991. In french.
[38] T. H. Dunning, J. Chem. Phys. 90 (1989) 1007–1023.
[39] O. Marchetti, H. Werner, J. Phys. Chem. A 113 (2009) 11580–11585.
[40] F. B. Van Duijneveldt, Gaussian Basis Sets for the Atoms H–Ne for Use in Molecular Calculations, Technical Report RJ 945, IBM, San José, 1971.
[41] C. Voisin, A. Cartier, J. L. Rivail, J. Phys. Chem. 96 (1992).
[42] J. Langlet, J. Caillet, M. Caffarel, J. Chem. Phys. 103 (1995) 8043.
[43] J. Langlet, J. Bergès, P. Reinhardt, J. Mol. Struc.: THEOCHEM 685 (2004) 43–56.
[44] J. P. Foster, F. Weinhold, J. Am. Chem. Soc. 102 (1980) 7211–7218.
[45] P. Widmark, P. Malmqvist, B. Roos, Theor. Chem. Acc. 77 (1990) 291–306.
[46] B. Roos, B. Salez, A. Veillard, E. Clementi, Atomic Hartree-Fock program, Technical Report RJ 815, IBM, San José, 1968. Modified by L. Gianola (1978), later by B.A. Heß (1986) and P. Reinhardt (1995).
[47] H.-J. Werner, P. Knowles, G. Knizia, F. Manby, M. Schütz, et al., Molpro, version 2010.1, a package of ab initio programs, 2010. See `http://www.molpro.net`.
[48] P. Reinhardt, Ortho, series of ab-initio programs in localized orbitals, 1996 – . . .. Unpublished.
[49] K. Ruud, T. Helgaker, J. Olsen, P. Jørgensen, H. J. A. Jensen, et al., Dalton2011, a molecular electronic structure program, see `http://www.daltonprogram.org`, 2011.
[50] O. Marchetti, H. Werner, Phys. Chem. Chem. Phys. 10 (2008) 3400–3409.
[51] G. E. Scuseria, T. M. Henderson, D. C. Sorensen, J. Chem. Phys. 129 (2008) 231101.