G2 Atomization Energies With Chemical Accuracy

Bathélemy Pradines,[1] Anthony Scemama,[1] Julien Toulouse,[2] Pierre-François Loos,[1, a)] and Emmanuel Giner[2]

[1] *Laboratoire de Chimie et Physique Quantiques (UMR 5626), Université de Toulouse, CNRS, UPS, France*

[2] *Laboratoire de Chimie Théorique, Université Pierre et Marie Curie, Sorbonne Université, CNRS, Paris, France*

I. INTRODUCTION

II. THEORY

A. The DFT basis-set correction in a nutshell

The basis-set correction investigated here proposes to use the RSDFT formalism to capture a part of the short-range correlation effects missing from the description of the WFT in a finite basis set. Here, we briefly explain the working equations and notations needed for this work, and the interested reader can find the detailed formal derivation of the theory in[?] .

B. The basic concepts

Consider an incomplete basis-set $\mathcal{B}$ for which we assume to have accurate approximations of both the FCI density $n_{\Psi_{\text{FCI}}^{\mathcal{B}}}$ and energy $E_{\text{FCI}}^{\mathcal{B}}$. According to equation (15) of[?] , one can approximate the exact ground state energy E_0 as

$$E_0 \approx E_{\text{FCI}}^{\mathcal{B}} + \bar{E}^{\mathcal{B}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}] \tag{1}$$

where $\bar{E}^{\mathcal{B}}[n]$ is the complementary density functional defined in equation (8) of[?]

$$
\begin{aligned}
\bar{E}^{\mathcal{B}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}] = &\min_{\Psi \to n_{\Psi_{\text{FCI}}^{\mathcal{B}}}} \langle \Psi | \hat{T} + \hat{W}_{\text{ee}} | \Psi \rangle \\
&- \min_{\Psi^{\mathcal{B}} \to n_{\Psi_{\text{FCI}}^{\mathcal{B}}}} \left\langle \Psi^{\mathcal{B}} \middle| \hat{T} + \hat{W}_{\text{ee}} \middle| \Psi^{\mathcal{B}} \right\rangle,
\end{aligned}
\tag{2}
$$

where Ψ is a general wave function being obtained in a complete basis. Provided that functional $\bar{E}^{\mathcal{B}}[n]$ is known exactly, the only approximation performed in (1) is that the FCI density $n_{\Psi_{\text{FCI}}^{\mathcal{B}}}$ coincides with the exact ground state density, which in general is a reasonable approximation as the density converges rapidly with the basis set.

The functional $\bar{E}^{\mathcal{B}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}]$ is not universal as it depends on the basis set $\mathcal{B}$ used. A simple analytical form for such a functional is of course not known and we approximate it in two-steps. First, we define a real-space representation of the coulomb interaction projected in $\mathcal{B}$, which is then fitted with a long-range interaction thanks to a range-separation parameter $\mu(r)$ varying in space (see **??**). Then, we choose a specific class of short-range density functionals, namely the short-range correlation functionals with multi-determinantal reference

(ECMD) introduced by Toulouse *et al*[?] , that we evaluate at the FCI density $n_{\Psi_{\text{FCI}}^{\mathcal{B}}}$ (see **??**) and with the range-separation parameter $\mu(r)$ varying in space.

C. Definition of a real-space representation of the coulomb operator truncated in a basis-set $\mathcal{B}$

One of the consequences of the use of an incomplete basis-set $\mathcal{B}$ is that the wave function does not present a cusp near the electron coalescence point, which means that all derivatives of the wave function are continuous. As the exact electronic cusp originates from the divergence of the coulomb interaction at the electron coalescence point, a cusp-free wave function could also come from a non-divergent electron-electron interaction. Therefore, the impact of the incompleteness of a finite basis-set $\mathcal{B}$ can be thought as a cutting of the divergence of the coulomb interaction at the electron coalescence point.

The present paragraph briefly describes how to obtain an effective interaction $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ which:

- is non-divergent at the electron coalescence point as long as an incomplete basis set $\mathcal{B}$ is used,

- tends to the regular $1/r_{12}$ interaction in the limit of a complete basis set $\mathcal{B}$.

1. General definition of an effective interaction for the basis set $\mathcal{B}$

Consider the coulomb operator projected in the basis-set $\mathcal{B}$

$$\hat{W}_{\text{ee}}^{\mathcal{B}} = \frac{1}{2} \sum_{ijkl \,\in\, \mathcal{B}} V_{ij}^{kl} \,\hat{a}_k^{\dagger}\hat{a}_l^{\dagger}\hat{a}_j\hat{a}_i, \tag{3}$$

where the indices run over all orthonormal spin-orbitals in $\mathcal{B}$ and V_{ij}^{kl} are the usual coulomb two-electron integrals. Consider now the expectation value of $\hat{W}_{\text{ee}}^{\mathcal{B}}$ over a general wave function $\Psi^{\mathcal{B}}$ belonging to the $N-$electron Hilbert space spanned by the basis set $\mathcal{B}$. After a few mathematical work (see appendix A of[?] for a detailed derivation), such an expectation value can be rewritten as an integral over $\mathbb{R}^6$:

$$\left\langle \Psi^{\mathcal{B}} \middle| \hat{W}_{\text{ee}}^{\mathcal{B}} \middle| \Psi^{\mathcal{B}} \right\rangle = \frac{1}{2} \iint d\mathbf{X}_1 \, d\mathbf{X}_2 \, f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2), \tag{4}$$

where the function $f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ is

$$
\begin{aligned}
f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = &\sum_{ijklmn \,\in\, \mathcal{B}} V_{ij}^{kl} \; \Gamma_{kl}^{mn}[\Psi^{\mathcal{B}}] \\
&\phi_n(\mathbf{X}_2)\phi_m(\mathbf{X}_1)\phi_i(\mathbf{X}_1)\phi_j(\mathbf{X}_2),
\end{aligned}
\tag{5}
$$

a)Corresponding author: loos@irsamc.ups-tlse.fr

$\Gamma^{pq}_{mn}[\Psi^{\mathcal{B}}]$ is the two-body density tensor of $\Psi^{\mathcal{B}}$

$$\Gamma^{pq}_{mn}[\Psi^{\mathcal{B}}] = \left\langle \Psi^{\mathcal{B}} \middle| \hat{a}^{\dagger}_p \hat{a}^{\dagger}_q \hat{a}_n \hat{a}_m \middle| \Psi^{\mathcal{B}} \right\rangle, \qquad (6)$$

and $\mathbf{X}$ collects the space and spin variables,

$$\mathbf{X} = (\mathbf{r}, \sigma) \qquad \mathbf{r} \in \mathbb{R}^3, \; \sigma = \pm\frac{1}{2}$$
$$\int \mathrm{d}\mathbf{X} = \sum_{\sigma = \pm\frac{1}{2}} \int_{\mathbb{R}^3} \mathrm{d}\mathbf{r}. \qquad (7)$$

Then, consider the expectation value of the exact coulomb operator over $\Psi^{\mathcal{B}}$

$$\left\langle \Psi^{\mathcal{B}} \middle| \hat{W}_{\mathrm{ee}} \middle| \Psi^{\mathcal{B}} \right\rangle = \frac{1}{2} \iint \mathrm{d}\mathbf{X}_1 \, \mathrm{d}\mathbf{X}_2 \, \frac{1}{r_{12}} n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) \quad (8)$$

where $n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ is the two-body density associated to $\Psi^{\mathcal{B}}$. Because $\Psi^{\mathcal{B}}$ belongs to $\mathcal{B}$, such an expectation value coincides with the expectation value of $\hat{W}^{\mathcal{B}}_{\mathrm{ee}}$

$$\left\langle \Psi^{\mathcal{B}} \middle| \hat{W}^{\mathcal{B}}_{\mathrm{ee}} \middle| \Psi^{\mathcal{B}} \right\rangle = \left\langle \Psi^{\mathcal{B}} \middle| \hat{W}_{\mathrm{ee}} \middle| \Psi^{\mathcal{B}} \right\rangle, \qquad (9)$$

which can be rewritten as:

$$\iint \mathrm{d}\mathbf{X}_1 \, \mathrm{d}\mathbf{X}_2 \, W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) \, n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$$
$$= \iint \mathrm{d}\mathbf{X}_1 \, \mathrm{d}\mathbf{X}_2 \, \frac{1}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \, n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2). \qquad (10)$$

where we introduced $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$

$$W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \frac{f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)}{n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)}, \qquad (11)$$

which is the effective interaction in the basis set $\mathcal{B}$.

As already discussed in[?], such an effective interaction is symmetric, *a priori* non translational nor rotational invariant if the basis set $\mathcal{B}$ does not have such symmetries and is necessary *finite* at the electron coalescence point for an incomplete basis set $\mathcal{B}$. Also, as demonstrated in the appendix B of[?], $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ tends to the regular coulomb interaction $1/r_{12}$ for all points in $(\mathbf{X}_1, \mathbf{X}_2)$ in the limit of a complete basis set $\mathcal{B}$.

2. Definition of a valence effective interaction

As the average inter electronic distances are very different between the valence electrons and the core electrons, it can be advantageous to define an effective interaction taking into account only for the valence electrons.

According to (15) and (4), the effective interaction is defined by the expectation value of the coulomb operator over a wave function $\Psi^{\mathcal{B}}$. Therefore, to define an effective interaction accounting only for the valence electrons, one needs to define a function $f^{\mathrm{val}}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ satisfying

$$\left\langle \Psi^{\mathcal{B}} \middle| \hat{W}^{\mathcal{B}_{\mathrm{val}}}_{\mathrm{ee}} \middle| \Psi^{\mathcal{B}} \right\rangle = \frac{1}{2} \iint \mathrm{d}\mathbf{X}_1 \, \mathrm{d}\mathbf{X}_2 \, f^{\mathrm{val}}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2), \quad (12)$$

where $\hat{W}^{\mathcal{B}_{\mathrm{val}}}_{\mathrm{ee}}$ is the valence coulomb operator defined as

$$\hat{W}^{\mathcal{B}_{\mathrm{val}}}_{\mathrm{ee}} = \frac{1}{2} \sum_{ijkl \, \in \, \mathcal{B}_{\mathrm{val}}} V^{kl}_{ij} \, \hat{a}^{\dagger}_k \hat{a}^{\dagger}_l \hat{a}_j \hat{a}_i \; , \qquad (13)$$

and $\mathcal{B}_{\mathrm{val}}$ is a given set of molecular orbitals associated to the valence space which will be defined later on. Following the spirit of (5), the function $f^{\mathrm{val}}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ can be defined as

$$f^{\mathrm{val}}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \sum_{ij \, \in \, \mathcal{B}} \sum_{klmn \, \in \, \mathcal{B}_{\mathrm{val}}} V^{kl}_{ij} \, \Gamma^{mn}_{kl}[\Psi^{\mathcal{B}}]$$
$$\phi_n(\mathbf{X}_2)\phi_m(\mathbf{X}_1)\phi_i(\mathbf{X}_1)\phi_j(\mathbf{X}_2). \qquad (14)$$

Then, the effective interaction associated to the valence $W^{\mathrm{val}}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ is simply defines as

$$W^{\mathrm{val}}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \frac{f^{\mathrm{val}}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)}{n^{(2)}_{\Psi^{\mathcal{B}}, \, \mathrm{val}}(\mathbf{X}_1, \mathbf{X}_2)}, \qquad (15)$$

where $n^{(2)}_{\Psi^{\mathcal{B}}, \, \mathrm{val}}(\mathbf{X}_1, \mathbf{X}_2)$ is the two body density associated to the valence electrons:

$$n^{(2)}_{\Psi^{\mathcal{B}}, \, \mathrm{val}}(\mathbf{X}_1, \mathbf{X}_2) = \sum_{klmn \, \in \, \mathcal{B}_{\mathrm{val}}} \Gamma^{kl}_{mn}[\Psi^{\mathcal{B}}] \, \phi_m(\mathbf{X}_1)\phi_n(\mathbf{X}_2)\phi_k(\mathbf{X}_1)\phi_l(\mathbf{X}_2). \qquad (16)$$

It is important to notice in (14) the difference between the set of orbitals for the indices (i, j), which span the full set of MOs within $\mathcal{B}$, and the (k, l, m, n), which span only the valence space $\mathcal{B}_{\mathrm{val}}$. With such a definition, one can show (see annex) that $f^{\mathrm{val}}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ fulfills (12) and tends to the exact interaction $1/r_{12}$ in the limit of a complete basis set $\mathcal{B}$, whatever the choice of subset $\mathcal{B}_{\mathrm{val}}$.

3. Definition of a range-separation parameter varying in space

To be able to approximate the complementary functional $\bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}_{\mathrm{FCI}}}]$ thanks to functionals developed in the field of RS-DFT, we fit the effective interaction with a long-range interaction having a range-separation parameter *varying in space*. More precisely, if we define the value of the interaction at coalescence as

$$W_{\Psi^{\mathcal{B}}}(\mathbf{r}) = W_{\Psi^{\mathcal{B}}}(\mathbf{X}, \bar{\mathbf{X}}). \qquad (17)$$

where $(\mathbf{X}, \bar{\mathbf{X}})$ means a couple of anti-parallel spins at the same point in $\mathbf{r}$, we propose a fit for each point in $\mathbb{R}^3$ of $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1)$ with a long-range-like interaction:

$$W_{\Psi^{\mathcal{B}}}(\mathbf{r}) = w^{\mathrm{lr}, \mu(\mathbf{r}; \Psi^{\mathcal{B}})}(\mathbf{r}, \mathbf{r}) \qquad (18)$$

where the long-range-like interaction is defined as:

$$w^{\mathrm{lr}, \mu(\mathbf{r})}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2}\left(\frac{\mathrm{erf}(\mu(\mathbf{r}_1) \, r_{12})}{r_{12}} + \frac{\mathrm{erf}(\mu(\mathbf{r}_2) \, r_{12})}{r_{12}} \right). \qquad (19)$$

The equation (17) is equivalent to the following condition for $\mu(\mathbf{r}; \Psi^{\mathcal{B}})$:

$$\mu(\mathbf{r}; \Psi^{\mathcal{B}}) = \frac{\sqrt{\pi}}{2} W_{\Psi^{\mathcal{B}}}(\mathbf{r}). \qquad (20)$$

As we defined an effective interaction for the valence electrons, we also introduce a valence range-separation parameter as

$$\mu_{\text{val}}(\mathbf{r}; \Psi^{\mathcal{B}}) = \frac{\sqrt{\pi}}{2} W_{\Psi^{\mathcal{B}}}^{\text{val}}(\mathbf{r}). \qquad (21)$$

D. Approximations for the complementary functional

1. General scheme

In[?] the authors have proposed to approximate the complementary functional $\bar{E}^{\mathcal{B}}[n]$ by using a specific class of SRDFT energy functionals, namely the ECMD whose general definition is:

$$\bar{E}_{\text{c,md}}^{\text{sr}}[n(\mathbf{r}); \mu] = \min_{\Psi \to n(\mathbf{r})} \langle \Psi | \hat{T} + \hat{W}_{\text{ee}} | \Psi \rangle \\ - \langle \Psi^{\mu}[n(\mathbf{r})] | \hat{T} + \hat{W}_{\text{ee}} | \Psi^{\mu}[n(\mathbf{r})] \rangle, \qquad (22)$$

where the wave function $\Psi^{\mu}[n(\mathbf{r})]$ is defined by the constrained minimization

$$\Psi^{\mu}[n(\mathbf{r})] = \arg \min_{\Psi \to n(\mathbf{r})} \langle \Psi | \hat{T} + \hat{W}_{\text{ee}}^{\text{lr},\mu} | \Psi \rangle, \qquad (23)$$

where $\hat{W}_{\text{ee}}^{\text{lr},\mu}$ is the long-range electron-electron interaction operator

$$\hat{W}_{\text{ee}}^{\text{lr},\mu} = \frac{1}{2} \iint d\mathbf{r}_1 d\mathbf{r}_2 \, w^{\text{lr},\mu}(|\mathbf{r}_1 - \mathbf{r}_2|) \hat{n}^{(2)}(\mathbf{r}_1, \mathbf{r}_2), \qquad (24)$$

with

$$w^{\text{lr},\mu}(|\mathbf{r}_1 - \mathbf{r}_2|) = \frac{\text{erf}(\mu |\mathbf{r}_1 - \mathbf{r}_2|)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \qquad (25)$$

and the pair-density operator $\hat{n}^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \hat{n}(\mathbf{r}_1)\hat{n}(\mathbf{r}_2) - \delta(\mathbf{r}_1 - \mathbf{r}_2)\hat{n}(\mathbf{r}_1)$.

These functionals differ from the standard RSDFT correlation functional by the fact that the reference is not the Konh-Sham determinant but a multi determinant wave function, which makes them much more adapted in the present context where one aims at correcting the FCI energy.

The general scheme for estimating $\bar{E}^{\mathcal{B}}[n]$ is the following. Consider a given approximated ECMD functional $\bar{E}_{\text{c,md-}\mathcal{X}}^{\text{sr}}[n; \mu]$ labelled by ECMD-$\mathcal{X}$. Such a functional which might depend on the density $n(\mathbf{r})$ (and potentially its derivatives $\nabla n(\mathbf{r})$) is defined for any value of the range-separation parameter μ. A general scheme to approximate $\bar{E}^{\mathcal{B}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}]$ is to use $\bar{E}_{\text{c,md-}\mathcal{X}}^{\text{sr}}[n; \mu(\mathbf{r})]$ with the $\mu(\mathbf{r})$ defined in (15) and the FCI density $n_{\Psi_{\text{FCI}}^{\mathcal{B}}}$

$$\bar{E}^{\mathcal{B}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}] \approx \bar{E}_{\text{c,md-}\mathcal{X}}^{\text{sr}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}; \mu(\mathbf{r})] \qquad (26)$$

Therefore, any approximated ECMD can be used to estimate $\bar{E}^{\mathcal{B}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}]$.

2. LDA approximation for the complementary functional

Therefore, one can define an LDA-like functional for $\bar{E}^{\mathcal{B}}[n]$ as

$$\bar{E}_{\text{LDA}}^{\mathcal{B},\Psi^{\mathcal{B}}}[n] = \int d\mathbf{r} \, n(\mathbf{r}) \, \bar{\varepsilon}_{\text{c,md}}^{\text{sr,unif}} \left(n(\mathbf{r}); \mu(\mathbf{r}; \Psi^{\mathcal{B}}) \right), \qquad (27)$$

where $\bar{\varepsilon}_{\text{c,md}}^{\text{sr,unif}}(n, \mu)$ is the multi-determinant short-range correlation energy per particle of the uniform electron gas for which a parametrization can be found in Ref. ? . In practice, for open-shell systems, we use the spin-polarized version of this functional (i.e., depending on the spin densities) but for simplicity we will continue to use only the notation of the spin-unpolarized case.

3. New PBE interpolated ECMD functional

The LDA-like functional defined in (38) relies only on the transferability of the physics of UEG which is certainly valid for large values of μ but which is known to over correlate for small values of μ. In order to correct such a defect, we propose here a new ECMD functional inspired by the recently proposed functional of some of the present authors[?] which interpolates between the usual PBE correlation functional when $\mu \to 0$ and the exact behaviour which is known when $\mu \to \infty$.

The exact behaviour of the $\bar{E}_{\text{c,md}}^{\text{sr}}[n(\mathbf{r}); \mu]$ is known in the large μ limit[?] :

$$\bar{E}_{\text{c,md}}^{\text{sr}}[n(\mathbf{r}); \mu] = \frac{2\sqrt{\pi}\left(1 - \sqrt{2}\right)}{3\mu^3} \int d\mathbf{r} \, n^{(2)}(r) \qquad (28)$$

where $n^{(2)}(\mathbf{r})$ is the *exact* on-top pair density for the ground state of the system. As the exact ground state on-top pair density $n^{(2)}(r)$ is not known, we propose here to approximate it by that of the UEG at the density of the system:

$$n^{(2)}(r) \approx n_{\text{UEG}}^{(2)}(n_\uparrow(r), n_\downarrow(r)) \qquad (29)$$

where $n_\uparrow(r)$ and $n_\downarrow(r)$ are, respectively, the up and down spin densities of the physical system at r, $n_{\text{UEG}}^{(2)}(n_\uparrow n_\downarrow)$ is the UEG on-top pair density

$$n_{\text{UEG}}^{(2)}(n_\uparrow n_\downarrow) = 4 n_\uparrow n_\downarrow g_0(n_\uparrow(, n_\downarrow) \qquad (30)$$

and $g_0(n_\uparrow, n_\downarrow)$ is the correlation factor of the UEG whose parametrization can be found in[?] .

As such a form diverges for small values of μ as $1/\mu^3$, we follow the work proposed in[?] and interpolate with the Kohn-Sham correlation functional at $\mu = 0$. More precisely, we propose the following expression for the

$$\bar{E}_{\text{c,md}}^{\text{sr}}[n(\mathbf{r}); \mu] = \int d\mathbf{r} \, \bar{e}_{\text{c,md}}^{\text{PBE}}(n(r), \nabla n(r); \mu) \qquad (31)$$

with

$$\bar{e}_{\text{c,md}}^{\text{PBE}}(n, \nabla n; \mu) = \frac{e_c^{PBE}(n, \nabla n)}{1 + \beta_{\text{c,md PBE}}(n, \nabla n; \mu)\mu^3} \quad (32)$$

$$\beta(n, \nabla n; \mu) = \frac{3e_c^{PBE}(n, \nabla n)}{2\sqrt{\pi}\left(1 - \sqrt{2}\right)n_{\text{UEG}}^{(2)}(n_\uparrow n_\downarrow)}. \quad (33)$$

Therefore, we propose this approximation for the complementary functional $\bar{E}^{\mathcal{B}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}]$:

$$\bar{E}_{\text{PBE}}^{\mathcal{B},\Psi^{\mathcal{B}}}[n] = \int d\mathbf{r}\, \bar{e}_{\text{c,md}}^{\text{PBE}}(n(\mathbf{r}), \nabla n(\mathbf{r}); \mu(\mathbf{r})) \quad (34)$$

E. Valence-only approximation for the complementary functional

We now introduce a valence-only approximation for the complementary functional, which, as we shall see, performs much better than the usual approximations in the context of atomization energies. Defining the valence one-body spin density matrix as

$$\rho_{ij,\sigma}^{\text{val}}[\Psi^{\mathcal{B}}] = \left\langle \Psi^{\mathcal{B}}\middle| a_{i,\sigma}^\dagger a_{j,\sigma}\middle|\Psi^{\mathcal{B}}\right\rangle \quad \text{if } (i,j) \in \mathcal{B}_{\text{val}}$$
$$= 0 \quad \text{in other cases} \quad (35)$$

then one can define the valence density as:

$$n_\sigma^{\text{val}}(\mathbf{r}) = \sum_{i,j} \rho_{ij,\sigma}^{\text{val}}[\Psi^{\mathcal{B}}]\phi_i(\mathbf{r})\phi_j(\mathbf{r}) \quad (36)$$

Therefore, we propose the following valence-only approximations for the complementary functional

$$\bar{E}_{\text{LDA, val}}^{\mathcal{B},\Psi^{\mathcal{B}}}[n] = \int d\mathbf{r}\, n^{\text{val}}(\mathbf{r})\, \bar{\varepsilon}_{\text{c,md}}^{\text{sr,unif}}\left(n^{\text{val}}(\mathbf{r}); \mu_{\text{val}}(\mathbf{r}); \Psi^{\mathcal{B}}\right), \quad (37)$$

$$\bar{E}_{\text{PBE, val}}^{\mathcal{B},\Psi^{\mathcal{B}}}[n] = \int d\mathbf{r}\, \bar{e}_{\text{c,md}}^{\text{PBE}}(n^{\text{val}}(\mathbf{r}), \nabla n^{\text{val}}(\mathbf{r}); \mu_{\text{val}}(\mathbf{r})) \quad (38)$$

III. RESULTS

A. The case of C_2, N_2, O_2, F_2 and the impact of the lack of basis functions adapted to core correlation

We begin the investigation of the behavior of the basis-set correction by the study of the atomization energies of the C_2, N_2, O_2, F_2 homo-nuclear diatomic molecules in the Dunning cc-pVXZ and cc-pCVXZ (X=D,T,Q,5) using the CIPSI algorithm to obtain reliable estimate of $E_{\text{FCI}}^{\mathcal{B}}$ and $n_{\Psi_{\text{FCI}}^{\mathcal{B}}}$.

1. CIPSI calculations

All CIPSI calculations were performed in two steps. First, a CIPSI calculation was performed until the zeroth-order wave function reaches 10^6 Slater determinants, from which we extracted the natural orbitals. From this set of natural orbitals, we performed CIPSI calculations until the $E_{\text{exFCI}}^{\mathcal{B}}$ reaches about 0.1 mH convergence for each systems. Such convergence criterion is more than sufficient for the CIPSI densities $n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})$. Therefore, from now on, we assume that

$$E_{\text{FCI}}^{\mathcal{B}} \approx E_{\text{exFCI}}^{\mathcal{B}} \quad (39)$$

and that

$$n(\mathbf{r})_{\Psi_{\text{FCI}}^{\mathcal{B}}} \approx n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r}). \quad (40)$$

Regarding the wave function chosen to define the local range-separation parameter $\mu(\mathbf{r})$, we take a single Slater determinant built with the natural orbitals of the first CIPSI calculation.

2. Treating the valence electrons

TABLE I. Dissociation energy (D_e) in kcal/mol of the C_2, O_2, N_2 and F_2 molecules computed with various methods and basis sets.

Molecule	Method	Dunning's basis set				
		cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	E^∞_{QZ5Z}
C_2	(FC)FCIQMC	130.0(1)	139.9(3)	143.3(2)		144.9
	(FC)FCIQMC+F12	142.3	145.3			
	ex (FC)FCI	132.0	140.3	143.6	144.3	
	ex (FC)FCI+LDA-val	143.0	145.4	146.4	146.0	
	ex (FC)FCI+PBE-val	147.4	146.1	146.4	145.9	
	exFCI+PBE-on-top-val	143.3	144.7	145.7	145.6	
		cc-pCVDZ	cc-pCVTZ	cc-pCVQZ	cc-pCV5Z	E^∞_{CQZC5Z}
	ex FCI	131.0	141.5	145.1	146.1	147.1
	ex FCI+LDA	141.4	146.7	147.8	147.6	
	ex FCI+PBE	145.1	147.0	147.7	147.5	

Molecule	Method	Dunning's basis set				
		cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	E^∞_{QZ5Z}
N_2	ex (FC)FCI	201.1	217.1	223.5	225.7	227.8
	ex (FC)FCI+LDA-val	217.9	225.9	228.0	228.6	
	ex (FC)FCI+PBE-val	227.7	227.8	228.3	228.5	
	exFCI+PBE-on-top-val	224.8	226.7	228.3	228.3	
	(FC)CCSD(T)	199.9	216.3	− − − − −	− − − − −	227.4
	ex (FC)CCSD(T)+LDA-val	216.3	224.8	− − − − −	− − − − −	
	ex (FC)CCSD(T)+PBE-val	225.9	226.7	− − − − −	− − − − −	
		cc-pCVDZ	cc-pCVTZ	cc-pCVQZ	cc-pCV5Z	E^∞_{QZ5Z}
	ex FCI	202.2	218.5	224.4	226.6	228.8
	ex FCI+LDA	218.0	226.8	229.1	229.4	
	ex FCI+PBE	226.4	228.2	229.1	229.2	

[a] Results from Ref. ? .
[b] Results from Ref. ? .

TABLE II. Dissociation energy (D_e) in kcal/mol of the C_2, O_2, N_2 and F_2 molecules computed with various methods and basis sets.

Molecule	Method	cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	E^∞_{QZ5Z}
O_2	exFCI	105.2	114.5	118.0	119.1	120.0
	exFCI+LDA-val	112.4	118.4	120.2	120.4	
	exFCI+PBE-val	117.2	119.4	120.3	120.4	
	(FC)CCSD(T)	103.9	113.6	− − − − −	− − − − −	120.5
	ex (FC)CCSD(T)+LDA-val	110.6	117.2	− − − − −	− − − − −	
	ex (FC)CCSD(T)+PBE-val	115.1	118.0	− − − − −	− − − − −	
Molecule	Method	cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	E^∞_{QZ5Z}
F_2	exFCI	26.7	35.1	37.1	38.0	39.0
	exFCI+LDA-val	30.4	37.2	38.4	38.9	
	exFCI+PBE -val	33.1	37.9	38.5	38.9	
	(FC)CCSD(T)	25.7	34.4	− − − − −	− − − − −	38.8
	ex (FC)CCSD(T)+LDA-val	29.2	36.5	− − − − −	− − − − −	
	ex (FC)CCSD(T)+PBE-val	31.5	37.1	− − − − −	− − − − −	

[a] Results from Ref. ? .
[b] Results from Ref. ? .