



**HAL**  
open science

# Curing basis-set convergence of wave-function theory using density-functional theory: a systematically improvable approach

Emmanuel Giner, Barthélémy Pradines, Anthony Ferté, Roland Assaraf, Andreas Savin, Julien Toulouse

## ► To cite this version:

Emmanuel Giner, Barthélémy Pradines, Anthony Ferté, Roland Assaraf, Andreas Savin, et al.. Curing basis-set convergence of wave-function theory using density-functional theory: a systematically improvable approach. *The Journal of Chemical Physics*, 2018, 149, pp.194301. 10.1063/1.5052714 . hal-01865504

**HAL Id: hal-01865504**

**<https://hal.sorbonne-universite.fr/hal-01865504>**

Submitted on 4 Sep 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Curing basis-set convergence of wave-function theory using density-functional theory: a systematically improvable approach

Emmanuel Giner,<sup>1, a)</sup> Barthélemy Pradines,<sup>1,2</sup> Anthony Ferté,<sup>1</sup> Roland Assaraf,<sup>1</sup> Andreas Savin,<sup>1</sup> and Julien Toulouse<sup>1</sup>

<sup>1)</sup>*Laboratoire de Chimie Théorique, Sorbonne Université and CNRS, F-75005 Paris, France*

<sup>2)</sup>*Institut des Sciences du Calcul et des Données, Sorbonne Université, F-75005 Paris, France*

(Dated: August 23, 2018)

The present work proposes to use density-functional theory (DFT) to correct for the basis-set error of wave-function theory (WFT). One of the key ideas developed here is to define a range-separation parameter which automatically adapts to a given basis set. The derivation of the exact equations are based on the Levy-Lieb formulation of DFT, which helps us to define a complementary functional which corrects uniquely for the basis-set error of WFT. The coupling of DFT and WFT is done through the definition of a real-space representation of the electron-electron Coulomb operator projected in a one-particle basis set. Such an effective interaction has the particularity to coincide with the exact electron-electron interaction in the limit of a complete basis set, and to be finite at the electron-electron coalescence point when the basis set is incomplete. The non-diverging character of the effective interaction allows one to define a mapping with the long-range interaction used in the context of range-separated DFT and to design practical approximations for the unknown complementary functional. Here, a local-density approximation is proposed for both full-configuration-interaction (FCI) and selected configuration-interaction approaches. Our theory is numerically tested to compute total energies and ionization potentials for a series of atomic systems. The results clearly show that the DFT correction drastically improves the basis-set convergence of both the total energies and the energy differences. For instance, a sub kcal/mol accuracy is obtained from the aug-cc-pVTZ basis set with the method proposed here when an aug-cc-pV5Z basis set barely reaches such a level of accuracy at the near FCI level.

## I. INTRODUCTION

The development of accurate and systematically improvable computational methods to calculate the electronic structure of molecular systems is an important research topic in theoretical chemistry as no definitive answer has been brought to that problem. The main difficulty originates from the electron-electron interaction which induces correlation between electrons, giving rise to a complexity growing exponentially with the size of the system. In this context, the two most popular approaches used nowadays, namely wave-function theory (WFT) and density-functional theory (DFT), have different advantages and limitations due to the very different mathematical formalisms they use to describe the electronic structure.

The clear advantage of WFT relies on the fact that, in a given one-electron basis set, the target accuracy is uniquely defined by the full-configuration-interaction (FCI) limit. Therefore, there exists many ways of systematically improving the accuracy by refining the wave-function ansatz, and ultimately by enlarging the basis

set. In particular, perturbation theory is a precious guide for approximating the FCI wave function and it has given birth to important theorems<sup>1,2</sup> and many robust methods, such as coupled cluster<sup>3</sup> or selected configuration interaction (CI)<sup>4-10</sup>. Despite these appealing features, the main disadvantages of WFT are certainly the slow convergence of many important physical properties with respect to the size of the one-particle basis set and the rapidly growing computational cost when one enlarges the basis set. Such behavior very often prohibits the reach of the so-called complete-basis-set limit which is often needed to obtain quantitative agreement with experiment. At the heart of the problem of slow convergence with respect to the size of the basis set lies the description of correlation effects when electrons are close, the so-called short-range correlation effects near the electron-electron cusp<sup>11</sup>. To cure this problem, explicitly correlated ( $f_{12}$ ) methods have emerged from the pioneering work of Hylleraas<sup>12</sup> and remain an active and promising field of research (for recent reviews, see Refs. 13-15). One possible drawback of the  $f_{12}$  methods is the use of a rather complex mathematical machinery together with numerically expensive quantities involving more than two-electron integrals.

An alternative formulation of the quantum many-body problem is given by DFT which, thanks to the

---

<sup>a)</sup>Electronic mail: [emmanuel.giner@lct.jussieu.fr](mailto:emmanuel.giner@lct.jussieu.fr)

Hohenberg-Kohn theorems<sup>16</sup>, abandons the complex many-body wave function for the simple one-body density. Thanks to the so-called Kohn-Sham formalism of DFT<sup>17</sup> and the development of practical approximations of the exchange-correlation density functional, DFT is nowadays the most used computational tool for the study of the molecular electronic problem. Despite its tremendous success in many areas of chemistry, Kohn-Sham DFT applied with usual semilocal density functional approximations generally fails to describe nonlocal correlation effects, such as strong correlation or dispersion forces. To overcome these problems ingredients from WFT have been introduced in DFT, starting from Hartree-Fock (HF) exchange<sup>18</sup> to many-body perturbation theory<sup>19</sup>. Nevertheless, the lack of a scheme to rationally and systematically improve the quality of approximate density functionals<sup>20</sup> remains a major limitation of DFT.

A more general formulation of DFT has emerged with the introduction of the so-called range-separated DFT (RS-DFT) (see Ref. 21 and references therein) which rigorously combines WFT and DFT. In such a formalism the electron-electron interaction is split into a long-range part which is treated using WFT and a complementary short-range part treated with DFT. The formalism is exact provided that full flexibility is given to the long-range wave function and that the exact short-range density functional is known. In practice, approximations must be used for these quantities and the splitting of the interaction has some appealing features in that regard. As the long-range wave-function part only deals with a non-diverging electron-electron interaction, the problematic cusp condition is removed and the convergence with respect to the one-particle basis set is greatly improved<sup>22</sup>. Regarding the DFT part, the approximate semilocal density functionals are better suited to describe short-range interaction effects. Therefore, a number of approximate RS-DFT schemes have been developed using either single-reference WFT approaches (such as Møller-Plesset perturbation theory<sup>23</sup>, coupled cluster<sup>24</sup>, random-phase approximations<sup>25,26</sup>) or multi-reference WFT approaches (such as multi-reference CI<sup>27</sup>, multiconfiguration self-consistent field<sup>28</sup>, multi-reference perturbation theory<sup>29</sup>, density-matrix renormalization group<sup>30</sup>). These mixed WFT/DFT schemes have shown to be able to correctly describe a quite wide spectrum of chemical situations going from weak intermolecular interactions to strong correlation effects. Nonetheless, these methods involve a range-separation parameter, often denoted by  $\mu$ , and there is no fully satisfying and systematic scheme to set its value, even if some interesting proposals have been made<sup>31-33</sup>.

The main goal of the present work is to use a DFT approach to correct for the basis-set incompleteness of WFT. The key idea developed here is to make a separation of the electron-electron interaction directly based on the one-particle basis set used and to express the re-

maining effects as a functional of the density. In practice, we propose a fit of the projected electron-electron interaction by a long-range interaction, leading to a local range-separation parameter  $\mu(\mathbf{r})$  which automatically adapts to the basis set. This is done by comparing at coalescence a real-space representation of the Coulomb electron-electron operator projected in the basis set with the long-range interaction used in RS-DFT. Thanks to this link, the theory proposed here can benefit from pre-existing short-range density functionals developed in RS-DFT.

The present paper is composed as follows. We present the general equations related to the splitting of the electron-electron interaction in a one-particle basis set in sections II A and II B. In section II C we point out the similarities and differences of this formalism with RS-DFT. A real-space representation of the electron-electron Coulomb operator developed in a one-particle basis set is proposed in section II D (with details given in Appendix A and B), which leads to the definition of a local range-separation parameter  $\mu(\mathbf{r})$  that automatically adapts to the basis set. This allows us to define in section II E a short-range local-density approximation (LDA) correcting FCI energies for the basis-set error. The formalism is then extended to the selected CI framework in section II F. In section III we test our theory on a series of atomic systems by computing both total energies and energy differences. We study the basis-set convergence of the DFT-corrected FCI total energy in the case of the helium atom in section III A. We then investigate the basis-set convergence of DFT-corrected selected CI for both total energies and ionization potentials (IPs) of the B-Ne series in section III B. In the case of the IPs, we show that chemical accuracy is systematically reached for all atomic systems already from the aug-cc-pVTZ basis set within our approach, whereas an aug-cc-pV5Z basis set is needed to reach such an accuracy at near FCI level. In order to better understand how the DFT-based correction acts for both total energies and energy differences, a detailed study is performed in section III B 3 for the oxygen atom and its first cation. Finally, we summarize the main results and conclude in section IV.

## II. THEORY

### A. Finite basis-set decomposition of the universal density functional

We begin by the standard DFT formalism for expressing the exact ground-state energy:

$$E_0 = \min_{n(\mathbf{r})} \left\{ F[n(\mathbf{r})] + (v_{\text{ne}}(\mathbf{r})|n(\mathbf{r})) \right\}, \quad (1)$$

where

$$(v_{\text{ne}}(\mathbf{r})|n(\mathbf{r})) = \int d\mathbf{r} v_{\text{ne}}(\mathbf{r}) n(\mathbf{r}) \quad (2)$$

is the nuclei-electron interaction energy, and  $F[n(\mathbf{r})]$  is the Levy-Lieb universal density functional

$$F[n(\mathbf{r})] = \min_{\Psi \rightarrow n(\mathbf{r})} \langle \Psi | \hat{T} + \hat{W}_{\text{ee}} | \Psi \rangle, \quad (3)$$

where the minimization is over  $N$ -electron wave functions  $\Psi$  with density equal to  $n(\mathbf{r})$ , and  $\hat{T}$  and  $\hat{W}_{\text{ee}}$  are the kinetic-energy and Coulomb electron-electron interaction operators, respectively. The Levy-Lieb universal functional only depends on the density  $n(\mathbf{r})$ , meaning that, given a density  $n(\mathbf{r})$ , one does not in principle need to pass through the minimization over explicit  $N$ -electron wave functions  $\Psi$  to obtain the value  $F[n(\mathbf{r})]$ . Provided that the search in equation (1) is done over  $N$ -representable densities expanded in a complete basis set, the minimizing density will be the exact ground-state density  $n_0(\mathbf{r})$ , leading to the exact ground-state energy  $E_0$ .

First, we consider the restriction on the densities over which we perform the minimization to those that can be represented within a one-electron basis set  $\mathcal{B}$ , which we denote by  $n^{\mathcal{B}}(\mathbf{r})$ . By this we mean all the densities that can be obtained from any wave function  $\Psi^{\mathcal{B}}$  expanded into  $N$ -electron Slater determinants constructed from orbitals expanded on the basis  $\mathcal{B}$ . Note that this is a sufficient but not necessary condition for characterizing these densities, as these densities can in general also be obtained from wave functions not restricted to the basis set. Therefore, the restriction on densities representable by a basis  $\mathcal{B}$  is much weaker than the restriction on wave functions representable by the same basis  $\mathcal{B}$ . With this restriction, there is a density, referred to as  $n_0^{\mathcal{B}}(\mathbf{r})$ , which minimizes the energy functional of Eq. (1) and give a ground-state energy  $E_0^{\mathcal{B}}$ :

$$\begin{aligned} E_0^{\mathcal{B}} &= \min_{n^{\mathcal{B}}(\mathbf{r})} \left\{ F[n^{\mathcal{B}}(\mathbf{r})] + (v_{\text{ne}}(\mathbf{r})|n^{\mathcal{B}}(\mathbf{r})) \right\} \\ &= F[n_0^{\mathcal{B}}(\mathbf{r})] + (v_{\text{ne}}(\mathbf{r})|n_0^{\mathcal{B}}(\mathbf{r})). \end{aligned} \quad (4)$$

Therefore, provided only that the exact ground-state density  $n_0(\mathbf{r})$  is well approximated by this density  $n_0^{\mathcal{B}}(\mathbf{r})$ ,

$$n_0(\mathbf{r}) \approx n_0^{\mathcal{B}}(\mathbf{r}), \quad (5)$$

the exact ground-state energy  $E_0$  will be well approximated by  $E_0^{\mathcal{B}}$ ,

$$E_0 \approx E_0^{\mathcal{B}}. \quad (6)$$

Considering the fast convergence of the density with the size of the basis set, we expect the approximation of equation (6) to be very good in practice for the basis sets commonly used.

Next, we consider the following decomposition of the Levy-Lieb density functional for a given density  $n^{\mathcal{B}}(\mathbf{r})$ :

$$F[n^{\mathcal{B}}(\mathbf{r})] = \min_{\Psi^{\mathcal{B}} \rightarrow n^{\mathcal{B}}(\mathbf{r})} \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\text{ee}} | \Psi^{\mathcal{B}} \rangle + \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})], \quad (7)$$

where  $\Psi^{\mathcal{B}}$  are wave functions restricted to the  $N$ -electron Hilbert space generated by the basis  $\mathcal{B}$ , and  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  is a complementary density functional

$$\begin{aligned} \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})] &= \min_{\Psi \rightarrow n^{\mathcal{B}}(\mathbf{r})} \langle \Psi | \hat{T} + \hat{W}_{\text{ee}} | \Psi \rangle \\ &\quad - \min_{\Psi^{\mathcal{B}} \rightarrow n^{\mathcal{B}}(\mathbf{r})} \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\text{ee}} | \Psi^{\mathcal{B}} \rangle. \end{aligned} \quad (8)$$

It should be pointed out that, in contrast with the density functionals used in DFT or RS-DFT, the complementary functional  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  is not universal as it depends on the basis set  $\mathcal{B}$  used to describe a specific system. As the restriction to the basis set  $\mathcal{B}$  is in general much more stringent for the  $N$ -electron wave functions  $\Psi^{\mathcal{B}}$  than for the densities  $n^{\mathcal{B}}(\mathbf{r})$ , we expect that the complementary functional  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  gives a substantial contribution, even for basis sets  $\mathcal{B}$  for which the approximation of equation (5) is good.

By using such a decomposition in equation (4), we obtain now

$$\begin{aligned} E_0^{\mathcal{B}} &= \min_{n^{\mathcal{B}}(\mathbf{r})} \left\{ \min_{\Psi^{\mathcal{B}} \rightarrow n^{\mathcal{B}}(\mathbf{r})} \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\text{ee}} | \Psi^{\mathcal{B}} \rangle \right. \\ &\quad \left. + (v_{\text{ne}}(\mathbf{r})|n^{\mathcal{B}}(\mathbf{r})) + \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})] \right\}, \end{aligned} \quad (9)$$

or, after recombining the two minimizations,

$$\begin{aligned} E_0^{\mathcal{B}} &= \min_{\Psi^{\mathcal{B}}} \left\{ \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\text{ee}} | \Psi^{\mathcal{B}} \rangle + (v_{\text{ne}}(\mathbf{r})|n_{\Psi^{\mathcal{B}}}(\mathbf{r})) \right. \\ &\quad \left. + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}(\mathbf{r})] \right\}, \end{aligned} \quad (10)$$

where  $n_{\Psi^{\mathcal{B}}}(\mathbf{r})$  is the density of  $\Psi^{\mathcal{B}}$ . By writing the Euler-Lagrange equation associated to the minimization in equation (10), we find that the minimizing wave function  $\Psi_0^{\mathcal{B}}$  satisfies the Schrödinger-like equation

$$\left( \hat{T}^{\mathcal{B}} + \hat{W}_{\text{ee}}^{\mathcal{B}} + \hat{V}_{\text{ne}}^{\mathcal{B}} + \hat{V}^{\mathcal{B}}[n_{\Psi_0^{\mathcal{B}}}(\mathbf{r})] \right) |\Psi_0^{\mathcal{B}}\rangle = \mathcal{E}_0^{\mathcal{B}} |\Psi_0^{\mathcal{B}}\rangle, \quad (11)$$

where  $\hat{T}^{\mathcal{B}}$ ,  $\hat{W}_{\text{ee}}^{\mathcal{B}}$ ,  $\hat{V}_{\text{ne}}^{\mathcal{B}}$ , and  $\hat{V}^{\mathcal{B}}[n(\mathbf{r})]$  are the restrictions to the space generated by the basis  $\mathcal{B}$  of the operators  $\hat{T}$ ,  $\hat{W}_{\text{ee}}$ ,  $\int d\mathbf{r} v_{\text{ne}}(\mathbf{r}) \hat{n}(\mathbf{r})$ , and  $\int d\mathbf{r} (\delta \bar{E}^{\mathcal{B}}[n(\mathbf{r})] / \delta n(\mathbf{r})) \hat{n}(\mathbf{r})$ , respectively, and  $\hat{n}(\mathbf{r})$  is the density operator. The potential  $\hat{V}^{\mathcal{B}}[n_{\Psi_0^{\mathcal{B}}}(\mathbf{r})]$  ensures that the minimizing wave function  $\Psi_0^{\mathcal{B}}$  gives the minimizing density  $n_0^{\mathcal{B}}(\mathbf{r})$  in equation (4). It is important to notice that the accuracy of the obtained energy  $E_0^{\mathcal{B}}$  depends only on how close the density of  $\Psi_0^{\mathcal{B}}$  is from the exact density:  $n_{\Psi_0^{\mathcal{B}}}(\mathbf{r}) = n_0(\mathbf{r}) \implies E_0^{\mathcal{B}} = E_0$ .

## B. Approximation of the FCI density in a finite basis set

In the limit where  $\mathcal{B}$  is a complete basis set, equation (10) gives the exact energy and  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})] = 0$ . When the basis set is not complete but sufficiently good,  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  can be considered as a small perturbation. Minimizing in equation (10) without  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  simply gives the FCI energy in a given basis set  $\mathcal{B}$

$$\begin{aligned} E_{\text{FCI}}^{\mathcal{B}} &= \min_{\Psi^{\mathcal{B}}} \left\{ \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\text{ee}} | \Psi^{\mathcal{B}} \rangle + (v_{\text{ne}}(\mathbf{r}) | n_{\Psi^{\mathcal{B}}}(\mathbf{r})) \right\} \\ &= \langle \Psi_{\text{FCI}}^{\mathcal{B}} | \hat{T} + \hat{W}_{\text{ee}} | \Psi_{\text{FCI}}^{\mathcal{B}} \rangle + (v_{\text{ne}}(\mathbf{r}) | n_{\Psi_{\text{FCI}}^{\mathcal{B}}}(\mathbf{r})), \end{aligned} \quad (12)$$

where we have introduced the ground-state FCI wave function  $\Psi_{\text{FCI}}^{\mathcal{B}}$  which satisfies the eigenvalue equation:

$$\left( \hat{T}^{\mathcal{B}} + \hat{W}_{\text{ee}}^{\mathcal{B}} + \hat{V}_{\text{ne}}^{\mathcal{B}} \right) | \Psi_{\text{FCI}}^{\mathcal{B}} \rangle = E_{\text{FCI}}^{\mathcal{B}} | \Psi_{\text{FCI}}^{\mathcal{B}} \rangle. \quad (13)$$

Note that the FCI energy  $E_{\text{FCI}}^{\mathcal{B}}$  is an upper bound of  $E_0^{\mathcal{B}}$  in equation (10) since  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})] \leq 0$ . By neglecting the impact of  $\hat{V}_{\text{ne}}^{\mathcal{B}}[n_{\Psi_0^{\mathcal{B}}}(\mathbf{r})]$  on the minimizing density  $n_0^{\mathcal{B}}(\mathbf{r})$ , we propose a zeroth-order approximation for the density

$$n_0^{\mathcal{B}}(\mathbf{r}) \approx n_{\Psi_{\text{FCI}}^{\mathcal{B}}}(\mathbf{r}), \quad (14)$$

which leads to a first-order-like approximation for the energy  $E_0^{\mathcal{B}}$

$$E_0^{\mathcal{B}} \approx E_{\text{FCI}}^{\mathcal{B}} + \bar{E}^{\mathcal{B}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}(\mathbf{r})]. \quad (15)$$

The term  $\bar{E}^{\mathcal{B}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}(\mathbf{r})]$  constitutes a simple DFT correction to the FCI energy which should compensate for the incompleteness of the basis set  $\mathcal{B}$ . The next sections are devoted to the analysis of the properties of  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  and to some practical approximations for this functional.

## C. Qualitative considerations for the complementary functional $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$

The definition of  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  [see equation (8)] is clear but deriving an approximation for such a functional is not straightforward. For example, defining an LDA-like approximation is not easy as the wave functions  $\Psi^{\mathcal{B}}$  used in the definition of  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  are not able to reproduce a uniform density if the basis set  $\mathcal{B}$  is not translationally invariant. Nonetheless, it is known that a finite one-electron basis set  $\mathcal{B}$  usually describes poorly the short-range correlation effects and therefore the functional  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  must recover these effects. Therefore, a natural idea is to find a mapping between this functional with the short-range functionals used in RS-DFT. Among these functionals, the multi-determinant short-range correlation functional  $\bar{E}_{\text{c,md}}^{\text{sr},\mu}[n(\mathbf{r})]$  of Toulouse *et al.*<sup>34</sup> has a definition very similar to the one of  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$ :

$$\begin{aligned} \bar{E}_{\text{c,md}}^{\text{sr},\mu}[n(\mathbf{r})] &= \min_{\Psi \rightarrow n(\mathbf{r})} \langle \Psi | \hat{T} + \hat{W}_{\text{ee}} | \Psi \rangle \\ &\quad - \langle \Psi^{\mu}[n(\mathbf{r})] | \hat{T} + \hat{W}_{\text{ee}} | \Psi^{\mu}[n(\mathbf{r})] \rangle, \end{aligned} \quad (16)$$

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

$$\Psi^{\mu}[n(\mathbf{r})] = \arg \min_{\Psi \rightarrow n(\mathbf{r})} \langle \Psi | \hat{T} + \hat{W}_{\text{ee}}^{\text{lr},\mu} | \Psi \rangle, \quad (17)$$

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

$$\hat{W}_{\text{ee}}^{\text{lr},\mu} = \frac{1}{2} \iint d\mathbf{r}_1 d\mathbf{r}_2 w^{\text{lr},\mu}(|\mathbf{r}_1 - \mathbf{r}_2|) \hat{n}^{(2)}(\mathbf{r}_1, \mathbf{r}_2), \quad (18)$$

with

$$w^{\text{lr},\mu}(|\mathbf{r}_1 - \mathbf{r}_2|) = \frac{\text{erf}(\mu|\mathbf{r}_1 - \mathbf{r}_2|)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (19)$$

and the pair-density operator  $\hat{n}^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \hat{n}(\mathbf{r}_1)\hat{n}(\mathbf{r}_2) - \delta(\mathbf{r}_1 - \mathbf{r}_2)\hat{n}(\mathbf{r}_1)$ . By comparing equation (16) to the definition of  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  in equation (8), one can see that the only difference between these two functionals relies in the wave functions used for the constrained minimization: in  $\bar{E}_{\text{c,md}}^{\text{sr},\mu}[n(\mathbf{r})]$  one uses  $\Psi^{\mu}$  whereas  $\Psi^{\mathcal{B}}$  is used in  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$ . More specifically,  $\Psi^{\mu}$  is determined by using a non-diverging long-range electron-electron interaction defined in a complete basis set [equation (18)], whereas the diverging Coulomb electron-electron interaction expanded in a finite basis set is involved in the definition of  $\Psi^{\mathcal{B}}$ . Therefore, as these two wave functions qualitatively represent the same type of physics, a possible way to link  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  and  $\bar{E}_{\text{c,md}}^{\text{sr},\mu}[n(\mathbf{r})]$  is to try to map the projection of the diverging Coulomb interaction on a finite basis set to a non-diverging long-range effective interaction.

## D. Effective Coulomb electron-electron interaction in a finite basis set

This section introduces a real-space representation of the Coulomb electron-electron operator projected in a basis set  $\mathcal{B}$ , which is needed to derive approximations for  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$ .

### 1. Expectation values over the Coulomb electron-electron operator

The Coulomb electron-electron operator restricted to a basis set  $\mathcal{B}$  is most naturally written in orbital-space second quantization:

$$\hat{W}_{\text{ee}}^{\mathcal{B}} = \frac{1}{2} \sum_{ijkl \in \mathcal{B}} V_{ij}^{kl} \hat{a}_k^\dagger \hat{a}_l^\dagger \hat{a}_j \hat{a}_i, \quad (20)$$

where the sums run over all (real-valued) orthonormal spin-orbitals  $\{\phi_i\}$  in the basis set  $\mathcal{B}$ , and  $V_{ij}^{kl}$  are the two-electron integrals. By expanding the creation and

annihilation operators in terms of real-space creation and annihilation field operators, the expectation value of  $\hat{W}_{ee}^{\mathcal{B}}$  over a wave function  $\Psi^{\mathcal{B}}$  can be written as (see Appendix A for a detailed derivation):

$$\langle \Psi^{\mathcal{B}} | \hat{W}_{ee}^{\mathcal{B}} | \Psi^{\mathcal{B}} \rangle = \frac{1}{2} \iint d\mathbf{X}_1 d\mathbf{X}_2 f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2), \quad (21)$$

where we introduced the function

$$f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \sum_{ijklmn \in \mathcal{B}} V_{ij}^{kl} \Gamma_{kl}^{mn}[\Psi^{\mathcal{B}}] \phi_n(\mathbf{X}_2) \phi_m(\mathbf{X}_1) \phi_i(\mathbf{X}_1) \phi_j(\mathbf{X}_2), \quad (22)$$

and  $\Gamma_{mn}^{pq}[\Psi^{\mathcal{B}}]$  is the two-body density matrix of  $\Psi^{\mathcal{B}}$

$$\Gamma_{mn}^{pq}[\Psi^{\mathcal{B}}] = \langle \Psi^{\mathcal{B}} | \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_n \hat{a}_m | \Psi^{\mathcal{B}} \rangle, \quad (23)$$

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

$$\mathbf{X} = (\mathbf{r}, \sigma) \quad \mathbf{r} \in \mathbb{R}^3, \quad \sigma = \pm \frac{1}{2} \quad (24)$$

$$\int d\mathbf{X} = \sum_{\sigma=\pm\frac{1}{2}} \int_{\mathbb{R}^3} d\mathbf{r}.$$

From the properties of the restriction of an operator to the space generated by the basis set  $\mathcal{B}$ , we have the following equality

$$\langle \Psi^{\mathcal{B}} | \hat{W}_{ee}^{\mathcal{B}} | \Psi^{\mathcal{B}} \rangle = \langle \Psi^{\mathcal{B}} | \hat{W}_{ee} | \Psi^{\mathcal{B}} \rangle, \quad (25)$$

which translates into

$$\frac{1}{2} \iint d\mathbf{X}_1 d\mathbf{X}_2 f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) \quad (26)$$

$$= \frac{1}{2} \iint d\mathbf{X}_1 d\mathbf{X}_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2),$$

where  $n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2)$  is the pair density of  $\Psi^{\mathcal{B}}$ . Therefore, by introducing the following function

$$W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \frac{f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)}{n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2)}, \quad (27)$$

one can rewrite equation (26) as

$$\iint d\mathbf{X}_1 d\mathbf{X}_2 W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2) \quad (28)$$

$$= \iint d\mathbf{X}_1 d\mathbf{X}_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2).$$

One can thus identify  $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  as an effective interaction, coming from the restriction to the basis set  $\mathcal{B}$ . This can be seen as a generalization of the exchange potential of Slater<sup>35</sup>. It is important to notice that all the quantities appearing in the integrals of equation (28) can be considered as functions and not operators or distributions, and therefore they can be compared pointwise. Of course, the function  $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  is not defined when  $n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2)$  vanishes, but we leave this for a future study.

Equation (28) means that the two integrands have the same integral, but it does not mean that they are equals at each point  $(\mathbf{X}_1, \mathbf{X}_2)$ . Of course, one could argue that there exists an infinite number of functions of  $u(\mathbf{X}_1, \mathbf{X}_2)$  satisfying

$$\iint d\mathbf{X}_1 d\mathbf{X}_2 u(\mathbf{X}_1, \mathbf{X}_2) n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2) \quad (29)$$

$$= \iint d\mathbf{X}_1 d\mathbf{X}_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2),$$

which implies that the effective interaction is not uniquely defined, and that the choice of equation (27) is just one among many and might not be optimal. For instance, the definition of the effective electron-electron interaction of equation (27) implies that it can depend on the spin of the electrons, whereas the exact Coulomb electron-electron interaction does not. Nevertheless, one can show (see Appendix B) that, in the limit of a complete basis set (written as “ $\mathcal{B} \rightarrow \infty$ ”),  $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  correctly tends to the exact Coulomb interaction:

$$\lim_{\mathcal{B} \rightarrow \infty} W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad \forall (\mathbf{X}_1, \mathbf{X}_2) \text{ and } \Psi^{\mathcal{B}}. \quad (30)$$

In particular, in this limit,  $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  does not depend on  $\Psi^{\mathcal{B}}$  or on the spins of the electrons.

## 2. Effective electron-electron interaction for opposite spins $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)$ and its properties

The fact that  $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  tends to the exact Coulomb electron-electron interaction in the complete-basis-set limit supports the choice of this effective interaction. Nevertheless, it is also important to analyze a few properties of  $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  in the finite basis sets used in actual quantum chemistry calculations, and to understand how it differs from the true interaction.

We will consider the effective electron-electron interaction between electrons of opposite spins ( $\sigma$  and  $\bar{\sigma}$ )

$$W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2) = W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1\sigma, \mathbf{r}_2\bar{\sigma}), \quad (31)$$

since the interaction between same-spin electrons is normally not the limiting factor for basis convergence. The first thing to notice is that, because in practice  $\mathcal{B}$  is composed of atom-centered basis functions, the effective interaction  $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)$  is not translationally invariant nor isotropic, which means that it does not depend only on the variable  $|\mathbf{r}_1 - \mathbf{r}_2|$

$$W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2) \neq W_{\Psi^{\mathcal{B}}}(|\mathbf{r}_1 - \mathbf{r}_2|). \quad (32)$$

Thus, the quality of the representation of the Coulomb electron-electron operator (and therefore of the electron correlation effects) are not expected to be spatially uniform. Nevertheless,  $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)$  is symmetric in  $\mathbf{r}_1$  and

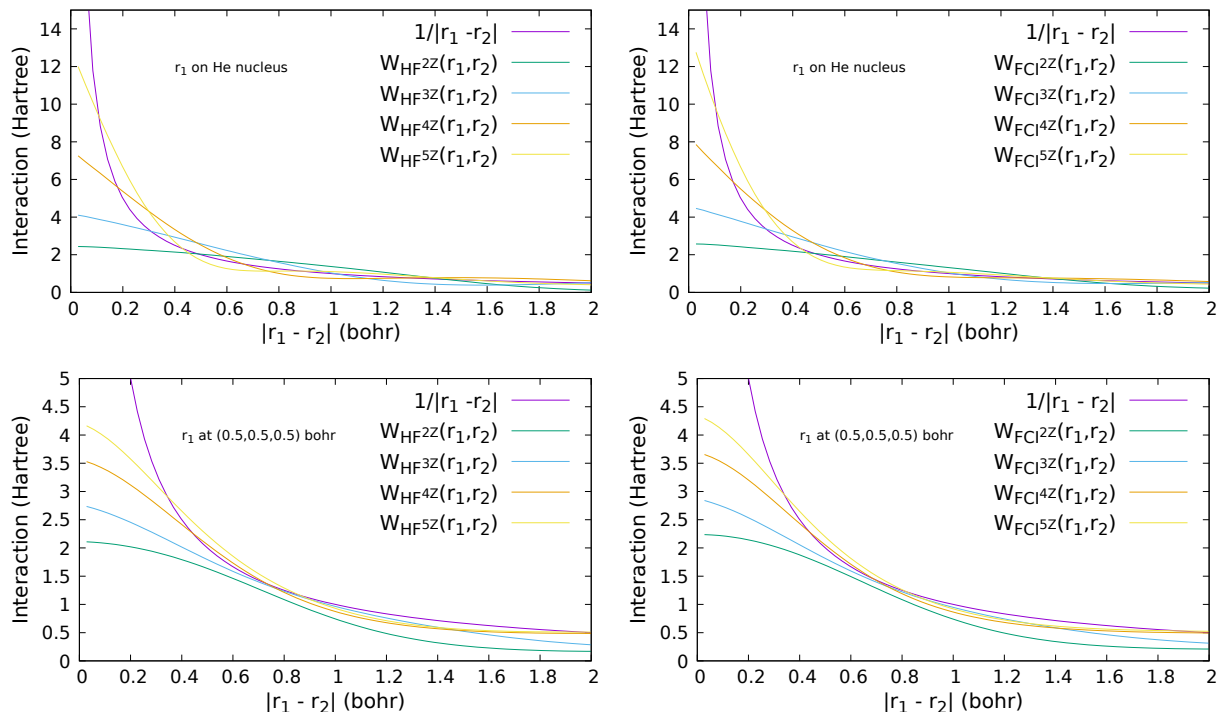


FIG. 1. Effective electron-electron interaction  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  in the helium atom for different cc-pVXZ basis sets ( $X = 2, 3, 4, 5$ ) as a function of  $|\mathbf{r}_1 - \mathbf{r}_2|$ . The two upper curves are for a reference point  $\mathbf{r}_1$  at the helium nucleus and  $\mathbf{r}_2$  moving along the diagonal of the  $xy$  plane, and the two lower curves are for a reference point  $\mathbf{r}_1$  at  $(0.5, 0.5, 0.5)$  bohr from the helium nucleus and  $\mathbf{r}_2$  moving along the diagonal of the  $xy$  plane with  $z = 0.5$  bohr. Two types of wave functions  $\Psi^B$  have been used: HF and FCI in the corresponding basis set. The exact Coulomb interaction  $1/|\mathbf{r}_1 - \mathbf{r}_2|$  is also reported for comparison.

$\mathbf{r}_2$ :

$$W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2) = W_{\Psi^B}(\mathbf{r}_2, \mathbf{r}_1). \quad (33)$$

A simple but interesting quantity is the value of the effective interaction  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  at coalescence at a given point in space  $\mathbf{r}_1$

$$W_{\Psi^B}(\mathbf{r}_1) = W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_1). \quad (34)$$

In a finite basis set,  $f_{\Psi^B}(\mathbf{X}_1, \mathbf{X}_2)$  is finite as it is obtained from a finite sum of bounded quantities [see equation (22)]. Therefore, provided that the on-top pair density does not vanish,  $n_{\Psi^B}^{(2)}(\mathbf{r}_1) = n_{\Psi^B}^{(2)}(\mathbf{r}_1\sigma, \mathbf{r}_1\sigma') \neq 0$ ,  $W_{\Psi^B}(\mathbf{r}_1)$  is necessarily finite in a finite basis set:

$$W_{\Psi^B}(\mathbf{r}_1) < \infty, \quad \forall \mathbf{r}_1 \text{ such that } n_{\Psi^B}^{(2)}(\mathbf{r}_1) \neq 0. \quad (35)$$

As mentioned above, since the effective interaction is not translationally invariant, the value  $W_{\Psi^B}(\mathbf{r}_1)$  has no reason to be independent of  $\mathbf{r}_1$ .

### 3. Illustrative examples of $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$ on the helium atom

In order to investigate how  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  behaves as a function of the basis set, the wave function  $\Psi^B$ , and the

spatial variables  $(\mathbf{r}_1, \mathbf{r}_2)$ , we performed calculations using Dunning basis sets of increasing sizes (from aug-cc-pVDZ to aug-cc-pV5Z) using a HF or a FCI wave function for  $\Psi^B$  and different reference points  $\mathbf{r}_1$ . We report these numerical results in figure 1.

From figure 1, several trends can be observed. First, for all wave functions  $\Psi^B$  and reference points  $\mathbf{r}_1$  used here, the value of  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  at coalescence is finite, which numerically illustrates equation (35). Second, the value at coalescence increases with the cardinal of the basis set, suggesting that the description of the short-range part of the interaction is improved by enlarging the basis set. Third, the global shape of the  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  is qualitatively modified by changing the reference point  $\mathbf{r}_1$ , which illustrates the lack of translational invariance of  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$ . In particular, the values of  $W_{\text{HF}^B}(\mathbf{r}_1, \mathbf{r}_2)$  and  $W_{\text{FCI}^B}(\mathbf{r}_1, \mathbf{r}_2)$  at coalescence are much larger when the reference point  $\mathbf{r}_1$  is on the He nucleus, which is a signature that the atom-centered basis set does not uniformly describe the Coulomb interaction at all points in space. Fourth, the difference between the  $W_{\text{HF}^B}(\mathbf{r}_1, \mathbf{r}_2)$  and  $W_{\text{FCI}^B}(\mathbf{r}_1, \mathbf{r}_2)$  is almost unnoticeable for all basis sets and for the two reference points  $\mathbf{r}_1$  used here.

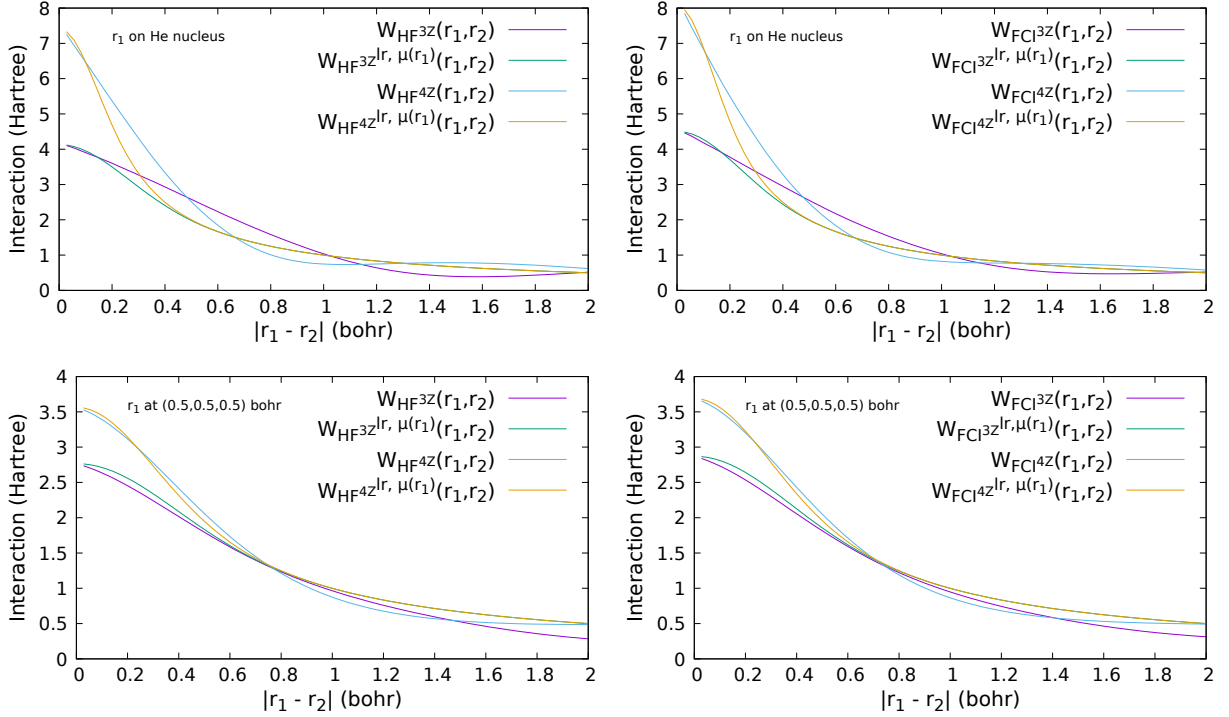


FIG. 2. Effective electron-electron interaction  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  and long-range electron-electron interaction  $W_{\Psi^B}^{\text{lr}, \mu(\mathbf{r}_1)}(\mathbf{r}_1, \mathbf{r}_2)$  for different cc-pVXZ basis sets ( $X = 3, 4$ ) as a function of  $|\mathbf{r}_1 - \mathbf{r}_2|$ . The two upper curves are for a reference point  $\mathbf{r}_1$  at the helium nucleus and  $\mathbf{r}_2$  moving along the diagonal of the  $xy$  plane, and the two lower curves are for a reference point  $\mathbf{r}_1$  at  $(0.5, 0.5, 0.5)$  bohr from the helium nucleus and  $\mathbf{r}_2$  moving along the diagonal of the  $xy$  plane with  $z = 0.5$  bohr. Two types of wave functions  $\Psi^B$  have been used: HF and FCI in the corresponding basis set.

#### 4. Link with range-separated DFT: Introduction of a local range-separated parameter $\mu(\mathbf{r})$

From the numerical illustration of the properties of  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  given in section IID3, it appears that the development of approximations for the density functional  $\bar{E}^B[n^B(\mathbf{r})]$  seems rather complicated since the effective interaction  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  is system- and basis-dependent, non translationally invariant, and non isotropic. Nevertheless, as it was numerically illustrated, the effective interaction  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  typically describes a long-range interaction which is finite at coalescence. Therefore, a possible way to approximate  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  is to locally fit  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$  by the long-range interaction  $w^{\text{lr}, \mu}(|\mathbf{r}_1 - \mathbf{r}_2|)$  of equation (19) used in RS-DFT. To do so, we propose here to determine a local value of the range-separation parameter  $\mu$  such that the value of the long-range interaction at coalescence is identical to the value of the effective interaction  $W_{\Psi^B}(\mathbf{r}_1)$  at coalescence at point  $\mathbf{r}_1$ . More specifically, the range-separation parameter  $\mu(\mathbf{r}_1; \Psi^B)$  is thus determined for each  $\mathbf{r}_1$  and  $\Psi^B$  by the condition:

$$W_{\Psi^B}(\mathbf{r}_1) = w^{\text{lr}, \mu(\mathbf{r}_1; \Psi^B)}(0), \quad (36)$$

with  $W_{\Psi^B}(\mathbf{r}_1)$  given by equations (34) which, since  $w^{\text{lr}, \mu}(0) = 2\mu/\sqrt{\pi}$ , simply gives

$$\mu(\mathbf{r}_1; \Psi^B) = \frac{\sqrt{\pi}}{2} W_{\Psi^B}(\mathbf{r}_1). \quad (37)$$

Therefore, defining the function  $W_{\Psi^B}^{\text{lr}, \mu(\mathbf{r}_1)}(\mathbf{r}_1, \mathbf{r}_2)$  as

$$W_{\Psi^B}^{\text{lr}, \mu(\mathbf{r}_1)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\text{erf}(\mu(\mathbf{r}_1; \Psi^B)|\mathbf{r}_1 - \mathbf{r}_2|)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (38)$$

we make the following approximation:

$$W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2) \approx W_{\Psi^B}^{\text{lr}, \mu(\mathbf{r}_1)}(\mathbf{r}_1, \mathbf{r}_2), \quad \forall (\mathbf{r}_1, \mathbf{r}_2). \quad (39)$$

One can notice that the definition of  $\mu(\mathbf{r}_1; \Psi^B)$  in equation (37) depends on the choice of  $\Psi^B$ , and therefore the approximation of equation (39) depends also on  $\Psi^B$ . Nevertheless, in the limit of a complete basis set the dependence on  $\Psi^B$  vanishes.

In order to illustrate how  $W_{\Psi^B}^{\text{lr}, \mu(\mathbf{r}_1)}(\mathbf{r}_1, \mathbf{r}_2)$  compares to  $W_{\Psi^B}(\mathbf{r}_1, \mathbf{r}_2)$ , we report in figure 2 these two functions for several basis sets, for different reference points  $\mathbf{r}_1$ , and for two different wave functions  $\Psi^B$ . From these plots it appears that the approximation of equation (39) is reasonably accurate when the reference point  $\mathbf{r}_1$  is on the helium nucleus and becomes even more accurate when



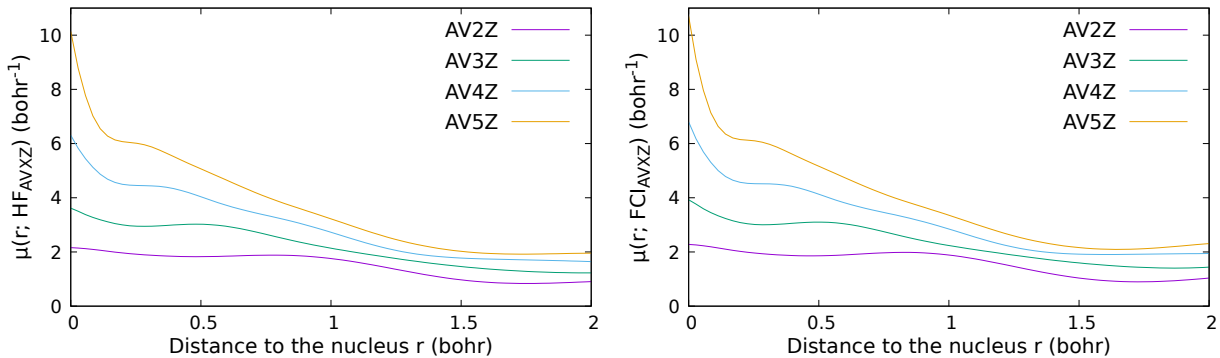


FIG. 3. Local range-separated parameter  $\mu(\mathbf{r}; \Psi^{\mathcal{B}})$  for the helium atom for different aug-cc-pVXZ basis sets ( $X = 2, 3, 4, 5$ ) as a function of the position  $\mathbf{r}$  along the diagonal of the  $xy$  plane. The curve on the left is when using the HF wave function for  $\Psi^{\mathcal{B}}$ , and the curve on the right is when using the FCI wave function for  $\Psi^{\mathcal{B}}$ .

the reference point  $\mathbf{r}_1$  is farther away from the helium nucleus.

In figure 3, we report the local range-separation parameter  $\mu(\mathbf{r}; \Psi^{\mathcal{B}})$ , as determined by equation (37), for different basis sets and when  $\Psi^{\mathcal{B}}$  is the HF or FCI wave function. It clearly appears that the magnitude of  $\mu(\mathbf{r}; \Psi^{\mathcal{B}})$  increases when the size of the basis set increases, which translates the fact that the electron-electron interaction is better described by enlarging the basis set. Also, for all basis sets, the maximal value of  $\mu(\mathbf{r}; \Psi^{\mathcal{B}})$  is reached when  $\mathbf{r}$  is at the nucleus, which demonstrates the non-homogeneity of the description of the electron-electron interaction with atom-centered basis functions. Finally, one can notice that the values of  $\mu(\mathbf{r}; \Psi^{\mathcal{B}})$  are very similar when using the HF or FCI wave function for  $\Psi^{\mathcal{B}}$ , but nevertheless slightly larger for the FCI wave function which reflects the fact that the corresponding effective interaction is slightly stronger.

### E. Practical approximations for the complementary functional $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$ : a short-range LDA-like functional with a local $\mu(\mathbf{r})$

A proper way to define an LDA-like approximation for the complementary density functional  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  would be to perform a uniform-electron gas calculation with the function  $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)$  as the electron-electron interaction. However, such a task would be rather difficult and ambiguous as  $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)$  is not translationally invariant nor isotropic, which thus questions how a uniform density could be obtained from such an interaction. Instead, by making the approximation of equation (39), one can define for each point  $\mathbf{r}_1$  an effective interaction which only depends on  $|\mathbf{r}_1 - \mathbf{r}_2|$ . For a given point in space  $\mathbf{r}_1$  one can therefore use the multi-determinant short-range correlation density functional of equation (16) with the range-separation parameter value  $\mu(\mathbf{r}_1; \Psi^{\mathcal{B}})$  corresponding to a local effective interaction at  $\mathbf{r}_1$  [see equation

(37)]. Therefore, we define an LDA-like functional for  $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$  as

$$\bar{E}_{\text{LDA}}^{\mathcal{B}, \Psi^{\mathcal{B}}}[n^{\mathcal{B}}(\mathbf{r})] = \int d\mathbf{r} n^{\mathcal{B}}(\mathbf{r}) \bar{\varepsilon}_{\text{c,md}}^{\text{sr,unif}}(n^{\mathcal{B}}(\mathbf{r}); \mu(\mathbf{r}; \Psi^{\mathcal{B}})), \quad (40)$$

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. 36. 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. One can interpret equation (40) as follows: the total correction to the energy in a given basis set is approximated by the sum of local LDA corrections obtained, at each point, from a uniform electron gas with a specific electron-electron interaction which approximatively coincides with the local effective interaction obtained in the basis set. Within the LDA approximation, the final working equation for our basis-correction scheme is thus

$$E_{\text{FCI+LDA}}^{\mathcal{B}, \Psi^{\mathcal{B}}} = E_{\text{FCI}}^{\mathcal{B}} + \bar{E}_{\text{LDA}}^{\mathcal{B}, \Psi^{\mathcal{B}}}[n_{\Psi_{\text{FCI}}^{\mathcal{B}}}], \quad (41)$$

We will refer to this approach as FCI+LDA $_{\Psi^{\mathcal{B}}}$  where  $\Psi^{\mathcal{B}}$  indicates the wave function used to define the effective interaction within the basis set  $\mathcal{B}$  employed in the calculation.

### F. Basis-set-corrected CIPSI: the CIPSI+LDA $_{\Psi^{\mathcal{B}}}$ approach

Equation (41) requires the calculation of the FCI energy and density whose computational cost can be very rapidly prohibitive. In order to remove this bottleneck, we propose here a similar approximation to correct the so-called CIPSI energy which can be used to approximate the FCI energy in systems where the latter is out of reach.

## 1. The CIPSI algorithm in a nutshell

The CIPSI algorithm approximates the FCI wave function through an iterative selected CI procedure, and the FCI energy through a second-order multi-reference perturbation theory. The CIPSI algorithm belongs to the general class of methods build upon selected CI<sup>4–10</sup> which have been successfully used to converge to FCI correlation energies, one-body properties, and nodal surfaces.<sup>8,37–44</sup> The CIPSI algorithm used in this work uses iteratively enlarged selected CI spaces and Epstein–Nesbet<sup>45,46</sup> multi-reference perturbation theory. The CIPSI energy is

$$E_{\text{CIPSI}} = E_{\text{v}} + E^{(2)}, \quad (42)$$

where  $E_{\text{v}}$  is the variational energy

$$E_{\text{v}} = \min_{\{c_{\text{I}}\}} \frac{\langle \Psi^{(0)} | \hat{H} | \Psi^{(0)} \rangle}{\langle \Psi^{(0)} | \Psi^{(0)} \rangle}, \quad (43)$$

where the reference wave function  $|\Psi^{(0)}\rangle = \sum_{\text{I} \in \mathcal{R}} c_{\text{I}} |\text{I}\rangle$  is expanded in Slater determinants  $\text{I}$  within the CI reference space  $\mathcal{R}$ , and  $E^{(2)}$  is the second-order energy correction

$$E^{(2)} = \sum_{\kappa} \frac{|\langle \Psi^{(0)} | \hat{H} | \kappa \rangle|^2}{E_{\text{v}} - \langle \kappa | H | \kappa \rangle} = \sum_{\kappa} e_{\kappa}^{(2)}, \quad (44)$$

where  $\kappa$  denotes a determinant outside  $\mathcal{R}$ . To reduce the cost of the evaluation of the second-order energy correction, the semi-stochastic multi-reference approach of Garniron *et al.*<sup>47</sup> was used, adopting the technical specifications recommended in that work. The CIPSI energy is systematically refined by doubling the size of the CI reference space at each iteration, selecting the determinants  $\kappa$  with the largest  $|e_{\kappa}^{(2)}|$ . The calculations are stopped when a target value of  $E^{(2)}$  is reached.

## 2. Working equations for the CIPSI+LDA <sub>$\Psi^{\mathcal{B}}$</sub> approach

The CIPSI algorithm being an approximation to FCI, one can straightforwardly apply the DFT correction developed in this work to correct the CIPSI energy error due to the basis set. For a given basis set  $\mathcal{B}$  and a given reference wave function  $\Psi^{(0)}$ , one can estimate the FCI energy and density by the following approximations:

$$E_{\text{FCI}}^{\mathcal{B}} \approx E_{\text{CIPSI}}^{\mathcal{B}}, \quad (45)$$

$$n_{\Psi_{\text{FCI}}^{\mathcal{B}}}(\mathbf{r}) \approx n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r}), \quad (46)$$

with

$$n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r}) = \langle \Psi^{(0)} | \hat{n}(\mathbf{r}) | \Psi^{(0)} \rangle. \quad (47)$$

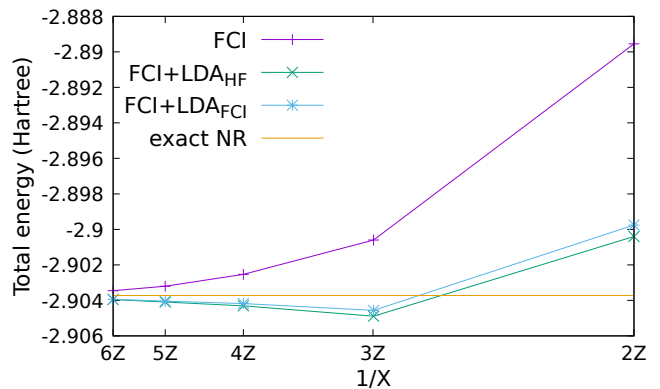


FIG. 4. Convergence of the total energy of the helium atom for FCI and FCI+LDA <sub>$\Psi^{\mathcal{B}}$</sub> , where  $\Psi^{\mathcal{B}}$  is either the HF or FCI wave function, as a function of the inverse of the cardinal number  $X$  of the AVXZ basis sets ( $X = 2, 3, 4, 5, 6$ ). The exact non-relativistic (NR) energy is also reported.

Assuming these approximations, for a given choice of  $\Psi^{\mathcal{B}}$  to define the effective interaction and within the LDA approximation of equation (40), one can define the corrected CIPSI energy as

$$E_{\text{CIPSI+LDA}_{\Psi^{\mathcal{B}}}}^{\mathcal{B}, \Psi^{\mathcal{B}}} = E_{\text{CIPSI}}^{\mathcal{B}} + \bar{E}_{\text{LDA}}^{\mathcal{B}, \Psi^{\mathcal{B}}} [n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})]. \quad (48)$$

Note that the reference wave function  $\Psi^{(0)}$  can be used for the definition of the effective interaction through its two-body density matrix [see equation (22)], but we leave that for further investigation and for the rest of the calculations we use the HF wave function for  $\Psi^{\mathcal{B}}$  in the definition of the effective interaction, and we denote the method by CIPSI+LDA<sub>HF</sub>.

## III. NUMERICAL TESTS: TOTAL ENERGY OF HE AND IONIZATION POTENTIALS FOR THE B-NE ATOMIC SERIES

For the present study, we use the LDA approximation of equation (40) and investigate the convergence of the total energies and energy differences as a function of the basis set. All calculations were performed with Quantum Package<sup>48</sup> using the Dunning aug-cc-pVXZ basis sets which are referred here as AVXZ.

### A. FCI+DFT: Total energy of the helium atom

We report in figure 4 and table I the convergence of the total energies computed for the helium atom in the Dunning basis sets AVXZ ( $X = 2, 3, 4, 5, 6$ ) using FCI and FCI+LDA <sub>$\Psi^{\mathcal{B}}$</sub>  where  $\Psi^{\mathcal{B}}$  is either the HF or FCI wave function. The first striking observation from these data

TABLE I. Total energies (in Hartree) of the helium atom and errors (in mH) with respect to the exact non-relativistic energy for FCI, FCI+LDA<sub>HF</sub>, and FCI+LDA<sub>FCI</sub> with the AVXZ basis sets ( $X = 2, 3, 4, 5, 6$ ).

	FCI		FCI+LDA <sub>HF</sub>		FCI+LDA <sub>FCI</sub>	
	Total energy	Error	Total energy	Error	Total energy	Error
AV2Z	-2.88955	14.17	-2.90040	3.3187	-2.89976	3.962
AV3Z	-2.90060	03.12	-2.90489	-1.1698	-2.90456	-0.840
AV4Z	-2.90253	01.18	-2.90430	-0.5849	-2.90418	-0.460
AV5Z	-2.90320	00.52	-2.90409	-0.3710	-2.90404	-0.321
AV6Z	-2.90346	00.26	-2.90396	-0.2367	-2.90394	-0.217
Exact non-relativistic total energy						
-2.90372						

is that the FCI+LDA <sub>$\Psi^{\mathcal{B}}$</sub>  energies rapidly converge to the exact energy as one increases the size of the basis set and that FCI+LDA <sub>$\Psi^{\mathcal{B}}$</sub>  is systematically closer to the exact energy than the FCI energy. Also, one can observe that  $\bar{E}_{\text{LDA}}^{\mathcal{B},\Psi^{\mathcal{B}}}[n^{\mathcal{B}}(\mathbf{r})]$  overestimates the correlation energy (in absolute value) for the AV3Z basis and the larger ones, which is consistent with the fact that LDA is known to give too negative correlation energies in regular Kohn-Sham DFT or in RS-DFT. Interestingly,  $\bar{E}_{\text{LDA}}^{\mathcal{B},\Psi^{\mathcal{B}}}[n^{\mathcal{B}}(\mathbf{r})]$  is almost independent of the choice of the wave function  $\Psi^{\mathcal{B}}$  used for the definition of the effective interaction within  $\mathcal{B}$ , as the FCI+LDA<sub>HF</sub> and FCI+LDA<sub>FCI</sub> energies are overall very close and get closer as one increases the size of the basis set. This last point is the numerical illustration that, in the limit of a complete basis set, the effective interaction is independent of the wave function  $\Psi^{\mathcal{B}}$  [see equation (30)]. Nevertheless, one observes that the correction obtained using the FCI wave function for  $\Psi^{\mathcal{B}}$  is systematically smaller in absolute value than the one obtained with the HF wave function for  $\Psi^{\mathcal{B}}$ . This result can be qualitatively understood by noticing that the introduction of the HF two-body density matrix in equation (22) reduces the number of two-electron integrals involved in the definition of  $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  [see equation (27)]. This reduction implies that the effective interaction  $W_{\text{HF}}(\mathbf{X}_1, \mathbf{X}_2)$  misses a part of the interaction within the basis set, namely the repulsion between electrons in virtual orbitals. However, the fact that  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n(\mathbf{r})]$  and  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{FCI}}[n(\mathbf{r})]$  are close suggests that  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n(\mathbf{r})]$  misses only a small part of the interaction. This statement can be intuitively understood by noticing that some two-electron integrals involved in the definition of  $W_{\text{HF}}(\mathbf{X}_1, \mathbf{X}_2)$  are of the type  $V_{ij}^{ab}$  (where  $i, j$  and  $a, b$  run over the occupied and virtual orbitals, respectively) which are the ones giving rise to the dominant part of the MP2 correlation energy in a given basis set.

## B. CIPSI+LDA: Total energies and energy differences for atomic systems

### 1. Convergence of the CIPSI+LDA<sub>HF</sub> total energy with the number of determinants

We report in figure 5, in the case of the oxygen ground state using the AV4Z basis set, the convergence of the variational energy  $E_{\text{v}}$ , the CIPSI energy, the CIPSI+LDA<sub>HF</sub> energy, and the LDA correction  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})]$  as a function of the number of Slater determinants in the reference wave function. The behavior of  $E_{\text{v}}$  and  $E_{\text{CIPSI}}$  reported in figure 5 are typical of a CIPSI calculation: a rapid convergence of the variational energy and an even faster convergence of the CIPSI energy. In this case,  $E_{\text{CIPSI}}$  with a reference wave function including  $2 \times 10^3$  and  $5 \times 10^5$  determinants provides an estimation of the FCI energy with an error smaller than 1 mH and 0.1 mH, respectively, whereas the size of the FCI space of this system for this basis set is approximately of  $10^{11}$  determinants. Regarding  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})]$ , it varies by about 0.08 mH between 100 and  $4 \times 10^6$  determinants. The very small variation of  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})]$  can be qualitatively understood by noticing that, within the LDA approximation of equation (40) and choosing a HF wave function for  $\Psi^{\mathcal{B}}$  to define the effective interaction,  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})]$  only depends on the one-body density which is known to converge rapidly with the level of correlation treatment, especially for atomic systems. To conclude this part of the study, it can be stated that the convergence of the CIPSI+LDA<sub>HF</sub> energy is only limited by the convergence of the CIPSI algorithm itself as  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})]$  converges very rapidly with the quality of the wave function.

### 2. The ionization potentials of the B-Ne series using CIPSI+LDA<sub>HF</sub>

In order to investigate how the correction  $\bar{E}_{\text{LDA}}^{\mathcal{B},\Psi^{\mathcal{B}}}[n^{\mathcal{B}}(\mathbf{r})]$  performs for energy differences, we report calculations

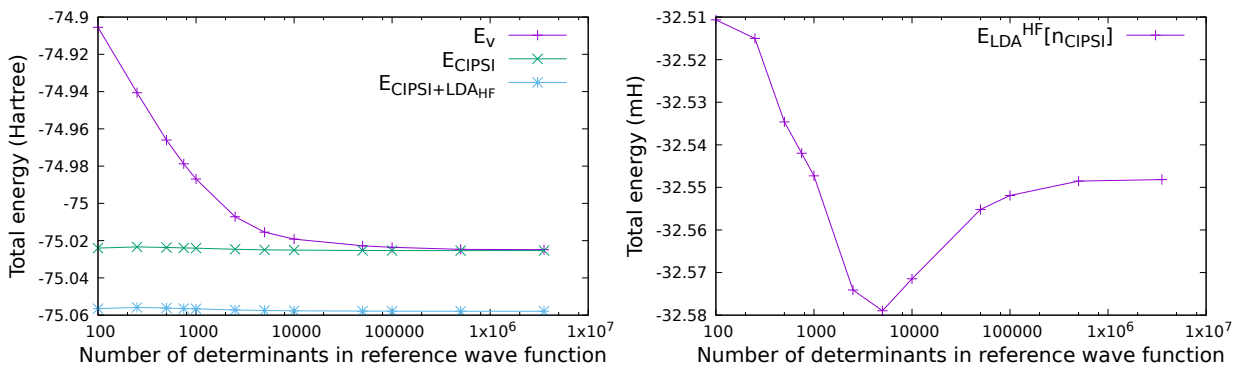


FIG. 5. Convergence of the variational total energy  $E_V$ , the CIPSI total energy, and the CIPSI+LDA<sub>HF</sub> total energy (left plot), and of the LDA correction  $\bar{E}_{\text{LDA}}^{\text{B, HF}}[n_{\text{CIPSI}}^{\text{B}}(\mathbf{r})]$  (right plot) of the oxygen atom as a function of the number of Slater determinants in the reference wave function using the AV4Z basis set.

of IPs for the B-Ne series using Dunning AVXZ basis sets ( $X = 2, 3, 4, 5$ ). These quantities have already been investigated at the initiator FCI quantum Monte Carlo (*i*-FCIQMC) level by Alavi and coworkers<sup>49</sup> and the authors have shown that obtaining errors of the IPs of the order of 1 mH for these simple atomic systems having at most ten electrons requires the use of large basis sets. As FCI in large basis sets is very rapidly out of reach for these systems, we use here the CIPSI+LDA<sub>HF</sub> method. The total energies are reported in table II and the IPs in table III. A graphical representation of the errors with respect to the estimated exact non-relativistic IPs at the CIPSI and CIPSI+LDA<sub>HF</sub> levels is also reported in figure 6. All electrons were correlated in the CIPSI calculations and the calculations were stopped when  $|E^{(2)}| < 10^{-3}$  Hartree, except for the Ne atom with the AV5Z basis set for which the calculation was stopped at  $|E^{(2)}| = 1.3 \times 10^{-3}$  Hartree.

From table II it clearly appears that all available *i*-FCIQMC total energies values are perfectly reproduced by the CIPSI total energies, which can thus be considered as good approximations of the FCI energies. Also, considering the small threshold on  $|E^{(2)}|$  and that the LDA correction  $\bar{E}_{\text{LDA}}^{\text{B, HF}}[n_{\text{CIPSI}}^{\text{B}}(\mathbf{r})]$  converges very rapidly with respect to the number of Slater determinants (see figure 5), the approximation of equation (46) can be considered as valid and therefore our CIPSI+LDA<sub>HF</sub> results can be considered as virtually identical to the ones that would be obtained with FCI+LDA<sub>HF</sub>. Finally, the CIPSI+LDA<sub>HF</sub> total energies obtained with the AV5Z basis set are remarkably close to the estimated exact total energies for the whole series, with an error ranging from 3.8 mH for the B<sup>+</sup> cation to 7.2 mH for the Ne atom.

Regarding the quality of the IPs (table III and figure 6), at near FCI level (either *i*-FCIQMC or CIPSI) the typical chemical accuracy of 1 kcal/mol ( $\approx 1.6$  mH) is reached with the AV4Z basis set for the B, C, and N atoms, whereas such a level of accuracy is barely reached with the AV5Z basis set for the O, F, and Ne atoms.

This illustrates how demanding the accurate computation of energy differences on these simple atomic systems is. Also one can notice that the IPs computed at the CIPSI level are systematically too small compared to the estimated exact values, showing that the cations are systematically better described than the neutral atoms in a given basis set. This result can be intuitively understood by the fact that the neutral atom has necessarily more correlated electron pairs than the cation and therefore, in the same basis, the cation is favored.

Considering now the convergence of the results obtained at the CIPSI+LDA<sub>HF</sub> level with respect to the basis set, it is striking to observe how the addition of the DFT correction improves the accuracy of the energy differences, with a sub kcal/mol accuracy being obtained for all atoms from the AV3Z to the AV5Z basis sets. With the AV2Z basis set, the error is overall strongly reduced, the average error being about 3 mH at the CIPSI+LDA<sub>HF</sub>, whereas it is of about 9 mH at the CIPSI level. From the AV3Z and larger basis sets, the maximum error occurs for the IP of the oxygen atom, which is overestimated by only 1.1 mH with the AV4Z basis set and by 0.9 mH with the AV5Z basis set, showing the accuracy of the approach. One can nevertheless observe a global trend of CIPSI+LDA<sub>HF</sub> to overestimate the IP, which is due to an over-correlation of the neutral species.

### 3. A case study: The oxygen atom and cation

In order to better understand how  $\bar{E}_{\text{LDA}}^{\text{B, HF}}[n_{\text{CIPSI}}^{\text{B}}(\mathbf{r})]$  corrects for the basis-set incompleteness and its impact on the energy differences, we perform a detailed study of the behavior of two quantities related to  $\bar{E}_{\text{LDA}}^{\text{B, HF}}[n_{\text{CIPSI}}^{\text{B}}(\mathbf{r})]$  for the oxygen atom and its first cation.

We first define the spherically averaged local basis-set

TABLE II. Total energies (in Hartree) of the neutral atoms and first cations for the B-Ne series with the AVXZ basis sets ( $X = 2, 3, 4, 5$ ) using CIPSI and CIPSI+LDA<sub>HF</sub>. The  $i$ -FCIQMC values from Ref. 49 are also reported for comparison with CIPSI.

	Method	AV2Z	AV3Z	AV4Z	AV5Z	Exact NR <sup>a</sup>
B	$i$ -FCIQMC <sup>b</sup>	-24.59242(1)	-24.60665(2)	-24.62407(11)	-24.63023(2)	
	CIPSI	-24.592418	-24.606654	-24.624109	-24.630233	-24.65390
	CIPSI+LDA <sub>HF</sub>	-24.641525	-24.641706	-24.648135	-24.650243	
B <sup>+</sup>	$i$ -FCIQMC <sup>b</sup>	-24.29450(1)	-24.30366(2)	-24.32005(2)	-24.32553(9)	
	CIPSI	-24.294496	-24.303660	-24.320044	-24.325531	-24.34889
	CIPSI+LDA <sub>HF</sub>	-24.338930	-24.336580	-24.343043	-24.345024	
C	$i$ -FCIQMC <sup>b</sup>	-37.76656(1)	-37.79163(2)	-37.81301(2)	-37.82001(4)	
	CIPSI	-37.766573	-37.791623	-37.813025	-37.820016	-37.8450
	CIPSI+LDA <sub>HF</sub>	-37.824730	-37.830667	-37.838253	-37.840544	
C <sup>+</sup>	$i$ -FCIQMC <sup>b</sup>	-37.35960(1)	-37.37967(2)	-37.39991(1)	-37.40605(1)	
	CIPSI	-37.359602	-37.379703	-37.399932	-37.406342	-37.43095
	CIPSI+LDA <sub>HF</sub>	-37.413086	-37.416631	-37.424109	-37.426321	
N	$i$ -FCIQMC <sup>b</sup>	-54.48881(2)	-54.52797(1)	-54.55423(3)	-54.56303(2)	
	CIPSI	-54.488814	-54.527941	-54.554235	-54.563027	-54.5893
	CIPSI+LDA <sub>HF</sub>	-54.556940	-54.571576	-54.581128	-54.584048	
N <sup>+</sup>	$i$ -FCIQMC <sup>b</sup>	-53.96106(10)	-53.99535(1)	-54.01838(1)	-54.02865(2)	
	CIPSI	-53.961062	-53.995355	-54.020414	-54.028633	-54.0546
	CIPSI+LDA <sub>HF</sub>	-54.024314	-54.036820	-54.046204	-54.049068	
O	$i$ -FCIQMC <sup>b</sup>	-74.92772(2)	-74.99077(4)	-75.02534(4)	-75.03749(6)	
	CIPSI	-74.927696	-74.990750	-75.025340	-75.037527	-75.0674
	CIPSI+LDA <sub>HF</sub>	-75.014946	-75.044685	-75.057889	-75.061639	
O <sup>+</sup>	$i$ -FCIQMC <sup>b</sup>	-74.444194(6)	-74.49701(1)	-74.52799(4)	-74.53869(6)	
	CIPSI	-74.444191	-74.497018	-74.527968	-74.538630	-74.5669
	CIPSI+LDA <sub>HF</sub>	-74.517650	-74.543804	-74.556296	-74.560233	
F	$i$ -FCIQMC <sup>b</sup>	-99.55223(1)	-99.64036(2)	-99.68460(10)	-99.70029(5)	
	CIPSI	-99.552228	-99.640295	-99.684561	-99.700258	-99.7341
	CIPSI+LDA <sub>HF</sub>	-99.658315	-99.704195	-99.722750	-99.727639	
F <sup>+</sup>	$i$ -FCIQMC <sup>b</sup>	-98.923015(6)	-99.00542(1)	-99.04599(3)	-99.06082(4)	
	CIPSI	-98.923000	-99.005441	-99.046481	-99.060808	-99.0930
	CIPSI+LDA <sub>HF</sub>	-99.016909	-99.062981	-99.080847	-99.085872	
Ne	$i$ -FCIQMC <sup>b</sup>	-128.71145(3)	-128.82577(5)	-128.88065(6)	-	
	CIPSI	-128.711476	-128.825813	-128.880658	-128.900438	-128.9383
	CIPSI+LDA <sub>HF</sub>	-128.835474	-128.898894	-128.924219	-128.931038	
Ne <sup>+</sup>	$i$ -FCIQMC <sup>b</sup>	-127.92411(2)	-128.03691(2)	-128.08816(11)	-	
	CIPSI	-127.924068	-128.036898	-128.088901	-128.107479	-128.1437
	CIPSI+LDA <sub>HF</sub>	-128.037019	-128.104203	-128.128973	-128.135914	

<sup>a</sup> Estimated exact non-relativistic (NR) values from Ref. 50.

<sup>b</sup> From Ref. 49. The statistical errors are given in parenthesis.

correction as

$$\bar{E}_{\text{LDA}}^{\mathcal{B}}(r) = \iint d\Omega r^2 n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r}) \bar{\epsilon}_{\text{c,md}}^{\text{sr,unif}}(n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r}); \mu(\mathbf{r}; \text{HF})),$$

such that

$$\int dr \bar{E}_{\text{LDA}}^{\mathcal{B}}(r) = \bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})], \quad (49)$$

where we use the largest CIPSI wave function to obtain the density  $n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})$ . With  $\bar{E}_{\text{LDA}}^{\mathcal{B}}(r)$  one can analyze in real space how  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})]$  corrects for the incompleteness of the basis set in near FCI calculations.

We report in figure 7 the plot of  $\bar{E}_{\text{LDA}}^{\mathcal{B}}(r)$  for the oxygen atom and its cation for different basis sets. One can observe that, with all basis sets used here, the LDA correction for the neutral atom is overall larger in absolute value than for the cation, which confirms that the cation is better described in a given basis set than the neutral atom. Also, it clearly explains why  $\bar{E}_{\text{LDA}}^{\mathcal{B},\text{HF}}[n_{\text{CIPSI}}^{\mathcal{B}}(\mathbf{r})]$  has a differential effect on the IPs. Regarding the behavior as a function of the distance to the nucleus, all the curves show that the dominant contributions, in absolute value, are in the region of high density. As expected,  $\bar{E}_{\text{LDA}}^{\mathcal{B}}(r)$  gets smaller as the size of the basis set is increased. With the largest basis set,  $\bar{E}_{\text{LDA}}^{\mathcal{B}}(r)$  is small in the valence shell ( $r > 0.5$  bohr), but remains substantial in the core re-

TABLE III. IPs (in mH) calculated by CIPSI and CIPSI+LDA<sub>HF</sub> for the B-Ne series with the AVXZ basis sets ( $X = 2, 3, 4, 5$ ). The errors with respect to the estimated exact non-relativistic values are given in parenthesis.

Method	AV2Z	AV3Z	AV4Z	AV5Z	Exact NR <sup>a</sup>
B CIPSI	297.92 (7.05)	302.99 (1.98)	304.06 (0.91)	304.70 (0.27)	304.98
CIPSI+LDA <sub>HF</sub>	302.59 (2.38)	305.12 (-0.14)	305.09 (-0.11)	305.21 (-0.23)	
C CIPSI	406.97 (7.10)	411.92 (2.15)	413.09 (0.98)	413.67 (0.40)	414.08
CIPSI+LDA <sub>HF</sub>	411.64 (2.43)	414.03 (0.04)	414.14 (-0.06)	414.22 (-0.14)	
N CIPSI	527.75 (7.13)	532.58 (2.30)	533.82 (1.06)	534.39 (0.49)	534.89
CIPSI+LDA <sub>HF</sub>	532.62 (2.26)	534.75 (0.13)	534.92 (-0.03)	534.97 (-0.08)	
O CIPSI	483.50 (16.90)	493.73 (6.67)	497.37 (3.03)	498.89 (1.51)	500.41
CIPSI+LDA <sub>HF</sub>	497.29 (3.11)	500.88 (-0.47)	501.59 (-1.18)	501.40 (-0.99)	
F CIPSI	629.22 (11.90)	634.85 (6.27)	638.07 (3.05)	639.45 (1.67)	641.13
CIPSI+LDA <sub>HF</sub>	641.40 (-0.27)	641.21 (-0.08)	641.90 (-0.77)	641.76 (-0.63)	
Ne CIPSI	787.40 (7.23)	788.91 (5.72)	791.75 (2.88)	792.95 (1.68)	794.64
CIPSI+LDA <sub>HF</sub>	798.45 (-3.81)	794.69 (-0.05)	795.24 (-0.60)	795.12 (-0.48)	

<sup>a</sup> Estimated exact non-relativistic (NR) values from Ref. 50.

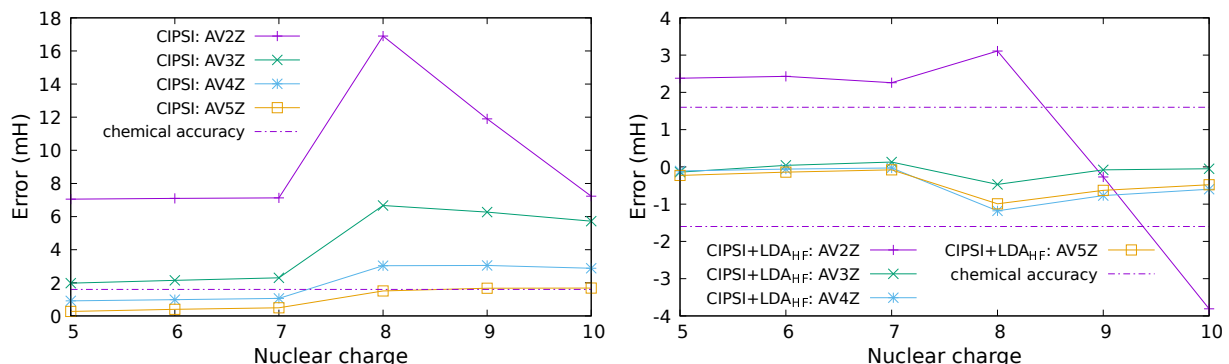


FIG. 6. Errors on the IPs calculated at the CIPSI (left plot) and CIPSI+LDA<sub>HF</sub> (right plot) levels for the B-Ne series with the AVXZ basis sets ( $X = 2, 3, 4, 5$ ). Note the different scales of the two plots.

gion. The fact that the basis sets used here do not contain functions optimized for core correlation explains why the LDA correction remains important in the core region, even with the AV5Z basis set.

In order to investigate the differential impact of the DFT correction on O and O<sup>+</sup>, we also define the following function:

$$\Delta \bar{E}_{\text{LDA}}^{\text{B}}(r) = \bar{E}_{\text{LDA},\text{O}}^{\text{B}}(r) - \bar{E}_{\text{LDA},\text{O}^+}^{\text{B}}(r). \quad (50)$$

We report in figure 8 the values of  $\Delta \bar{E}_{\text{LDA}}^{\text{B}}(r)$  for different basis sets. It clearly appears that the differential effects are mainly located in the valence region, which is what is expected since the electron can be qualitatively considered to be removed from the valence region. Also, except for the inner core region,  $\Delta \bar{E}_{\text{LDA}}^{\text{B}}(r)$  is always negative which means that  $\bar{E}_{\text{LDA}}^{\text{B,HF}}[n_{\text{CIPSI}}^{\text{B}}(\mathbf{r})]$  corrects more the neutral atom than the cation for the basis-set incompleteness. The fact that  $\Delta \bar{E}_{\text{LDA}}^{\text{B}}(r)$  is positive near 0.1 bohr means that the cation is more correlated in this region, which could be a sign that the two 1s electrons are closer to each other in the cation than in the neutral atom.

## IV. CONCLUSION

In the present work we have proposed a theory based on DFT to correct for the basis-set incompleteness of WFT. The key point here is the definition of a local range-separation parameter  $\mu(\mathbf{r})$  which automatically adapts to the basis set.

Both the exact theory (see section II A) and a series of approximations (see sections II B, II E, and II F) were derived for FCI and selected CI wave functions. Our theory combines WFT with a complementary density functional, as in RS-DFT. Unlike the latter theory, the electron-electron interaction is split directly in the one-electron basis set (see section II A). Here, the part of the electron-electron interaction expanded in the basis set is treated by WFT and the remaining interaction by the density functional. Thanks to a definition of the real-space representation of the basis-set-projected electron-electron interaction (see section II D), we show that the effect of the incompleteness of a given basis set can be mapped into a non-diverging effective electron-electron interaction. We

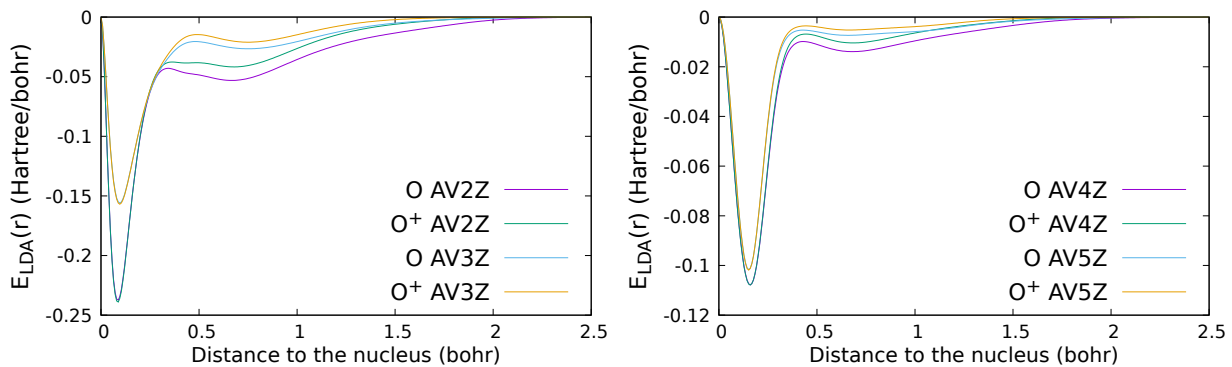


FIG. 7. Behavior of  $\bar{E}_{\text{LDA}}^{\text{B}}(r)$  for the AV2Z and AV3Z basis sets (left plot) and AV4Z and AV5Z basis sets (right plot) for the oxygen atom and its first cation.

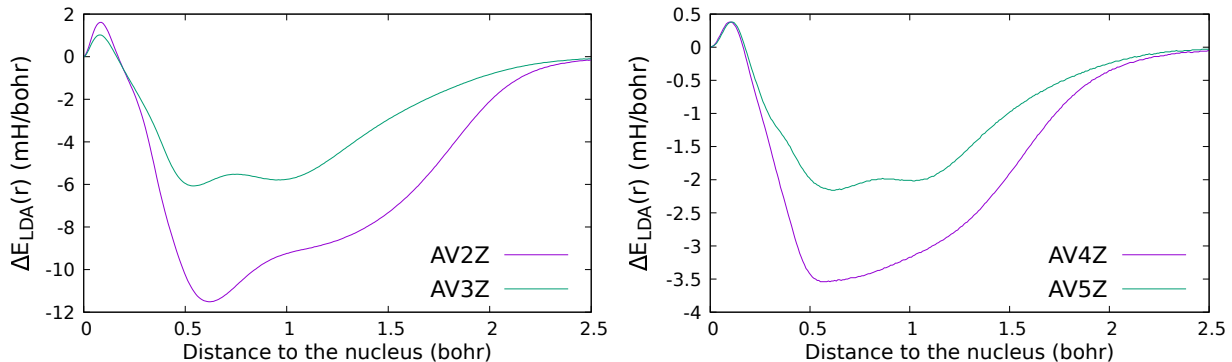


FIG. 8. Behavior of  $\Delta \bar{E}_{\text{LDA}}^{\text{B}}(r)$  for the AV2Z and AV3Z basis sets (left plot) and AV4Z and AV5Z basis sets (right plot) for the oxygen atom.

derive some of the important exact properties of the effective electron-electron interaction (see Appendix B and section IID 2), which helps us to physically motivate such a choice for an effective electron-electron interaction. A mapping between RS-DFT and our theory is proposed through the non-diverging behavior of the interactions in both theories (see section IIC), and such a mapping is done in practice through a comparison at coalescence of the effective electron-electron interaction with the long-range interaction used in the RS-DFT framework (see section IID 4). More specifically, this link between the basis-set splitting and range separation of the electron-electron interaction is done through the definition of a range-separation parameter  $\mu(\mathbf{r})$  which now depends on the spatial coordinate in  $\mathbb{R}^3$ . The computation of  $\mu(\mathbf{r})$  nonetheless requires the computation of bi-electronic integrals. This allows us to benefit from all pre-existing methodologies developed in the RS-DFT framework and therefore to produce numerically tractable approximations for our theory (see section IIE for the definition of an LDA-like functional in the present context). As the local range-separation parameter  $\mu(\mathbf{r})$  is automatically defined for a given physical system in a given basis set, we completely remove the choice of the parameter  $\mu$  which is inherent in the RS-DFT framework. Also this

local range-separation parameter  $\mu(\mathbf{r})$  can be seen as a measure of the incompleteness of a given basis set together with its non uniformity in the description of the correlation effects in  $\mathbb{R}^3$ . Finally, our theory produces a DFT-based correction for a given basis set which is added to the approximation of the FCI energy obtained in the same basis set.

We performed numerical tests both for total energies and energy differences for atomic systems (see section III). Using FCI wave functions (see section III A), we demonstrated that our approach is able to accelerate the basis convergence toward the exact non-relativistic total energy for the helium atom, which numerically illustrates its systematically improvable character. Then, we investigated the accuracy of our basis-set corrected CIPSI approach to describe the IPs of the B-Ne series (see section III B) which are known to be challenging for WFT methods, even at near FCI level. The main result of this study is that the level of accuracy of the energy differences is drastically improved even using the small aug-cc-pVDZ basis set, and that a sub-kcal/mol error is reached for all atoms from the aug-cc-pVTZ up to the aug-cc-pV5Z basis sets. Such results have to be compared with near FCI results for which a comparable error is barely reached only using the aug-cc-pV5Z basis set. In order to have

a better understanding of the origin of the systematic improvement of IPs brought by the DFT correction, we performed a detailed study of the oxygen atom and its first cation (see section III B 3). By introducing spherical averaged quantities, we show that the major differential contribution brought by the DFT correction comes from the valence region, which is physically meaningful and therefore tends to confirm that the good results obtained with our approach do not come from fortuitous error cancellations. Finally, it is important to stress here that the computational cost of the DFT corrections used here represents a negligible percentage of the computational cost of the CIPSI calculations.

### Appendix A: Derivation of the real-space representation of the effective interaction projected in a basis set

The exact Coulomb electron-electron operator can be expressed in real-space second quantization as

$$\hat{W}_{ee} = \frac{1}{2} \iiint d\mathbf{X}_1 d\mathbf{X}_2 d\mathbf{X}_3 d\mathbf{X}_4 \delta(\mathbf{X}_1 - \mathbf{X}_4) \delta(\mathbf{X}_2 - \mathbf{X}_3) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \hat{\Psi}^\dagger(\mathbf{X}_4) \hat{\Psi}^\dagger(\mathbf{X}_3) \hat{\Psi}(\mathbf{X}_2) \hat{\Psi}(\mathbf{X}_1), \quad (\text{A1})$$

where  $\hat{\Psi}(\mathbf{X})$  and  $\hat{\Psi}^\dagger(\mathbf{X})$  are annihilation and creation field operators, and  $\mathbf{X} = (\mathbf{r}, \sigma)$  collects the space and spin variables. The Coulomb electron-electron operator restricted to a basis set  $\mathcal{B}$  can be written in orbital-space second quantization:

$$\hat{W}_{ee}^{\mathcal{B}} = \frac{1}{2} \sum_{ijkl \in \mathcal{B}} V_{ij}^{kl} \hat{a}_k^\dagger \hat{a}_l^\dagger \hat{a}_j \hat{a}_i, \quad (\text{A2})$$

where the summations run over all (real-valued) orthonormal spin-orbitals  $\{\phi_i(\mathbf{X})\}$  in the basis set  $\mathcal{B}$ ,  $V_{ij}^{kl}$  are the two-electron integrals, the annihilation and creation operators can be written in terms of the field operators as

$$\hat{a}_i = \int d\mathbf{X} \phi_i(\mathbf{X}) \hat{\Psi}(\mathbf{X}), \quad (\text{A3})$$

$$\hat{a}_i^\dagger = \int d\mathbf{X} \phi_i(\mathbf{X}) \hat{\Psi}^\dagger(\mathbf{X}). \quad (\text{A4})$$

Therefore, by defining

$$w^{\mathcal{B}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) = \sum_{ijkl \in \mathcal{B}} V_{ij}^{kl} \phi_k(\mathbf{X}_4) \phi_l(\mathbf{X}_3) \phi_j(\mathbf{X}_2) \phi_i(\mathbf{X}_1), \quad (\text{A5})$$

we can rewrite  $\hat{W}_{ee}^{\mathcal{B}}$  in real-space second quantization as

$$\hat{W}_{ee}^{\mathcal{B}} = \frac{1}{2} \iiint d\mathbf{X}_1 d\mathbf{X}_2 d\mathbf{X}_3 d\mathbf{X}_4 w^{\mathcal{B}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) \hat{\Psi}^\dagger(\mathbf{X}_4) \hat{\Psi}^\dagger(\mathbf{X}_3) \hat{\Psi}(\mathbf{X}_2) \hat{\Psi}(\mathbf{X}_1). \quad (\text{A6})$$

In the limit of a complete basis set (written as “ $\mathcal{B} \rightarrow \infty$ ”),  $\hat{W}_{ee}^{\mathcal{B}}$  coincides with  $\hat{W}_{ee}$ :

$$\lim_{\mathcal{B} \rightarrow \infty} \hat{W}_{ee}^{\mathcal{B}} = \hat{W}_{ee}, \quad (\text{A7})$$

which implies that

$$\lim_{\mathcal{B} \rightarrow \infty} w^{\mathcal{B}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) = \delta(\mathbf{X}_1 - \mathbf{X}_4) \delta(\mathbf{X}_2 - \mathbf{X}_3) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (\text{A8})$$

It is important here to stress that the definition  $w^{\mathcal{B}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4)$  tends to a distribution in the limit of a complete basis set, and therefore such an object must really be considered as a distribution acting on test functions and not as a function to be evaluated pointwise. This is why we need to use an expectation value in order to make sense out of  $w^{\mathcal{B}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4)$ .

From equation (A1), the expectation value of the Coulomb electron-electron operator over a wave function  $\Psi$  is, after integration over  $\mathbf{X}_3$  and  $\mathbf{X}_4$ ,

$$\langle \Psi | \hat{W}_{ee} | \Psi \rangle = \frac{1}{2} \iint d\mathbf{X}_1 d\mathbf{X}_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \langle \Psi | \hat{\Psi}^\dagger(\mathbf{X}_1) \hat{\Psi}^\dagger(\mathbf{X}_2) \hat{\Psi}(\mathbf{X}_2) \hat{\Psi}(\mathbf{X}_1) | \Psi \rangle, \quad (\text{A9})$$

which, by introducing the two-body density matrix,

$$n_{\Psi}^{(2)}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) = \langle \Psi | \hat{\Psi}^\dagger(\mathbf{X}_4) \hat{\Psi}^\dagger(\mathbf{X}_3) \hat{\Psi}(\mathbf{X}_2) \hat{\Psi}(\mathbf{X}_1) | \Psi \rangle, \quad (\text{A10})$$

turns into

$$\langle \Psi | \hat{W}_{ee} | \Psi \rangle = \frac{1}{2} \iint d\mathbf{X}_1 d\mathbf{X}_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} n_{\Psi}^{(2)}(\mathbf{X}_1, \mathbf{X}_2), \quad (\text{A11})$$

where  $n_{\Psi}^{(2)}(\mathbf{X}_1, \mathbf{X}_2) = n_{\Psi}^{(2)}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_2, \mathbf{X}_1)$  is the pair density of  $\Psi$ . Equation (A11) holds for any wave function  $\Psi$ . Consider now the expectation value of  $\hat{W}_{ee}^{\mathcal{B}}$  over a wave function  $\Psi^{\mathcal{B}}$ . From equation (A6), we get

$$\langle \Psi^{\mathcal{B}} | \hat{W}_{ee}^{\mathcal{B}} | \Psi^{\mathcal{B}} \rangle = \frac{1}{2} \iiint d\mathbf{X}_1 d\mathbf{X}_2 d\mathbf{X}_3 d\mathbf{X}_4 w^{\mathcal{B}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4), \quad (\text{A12})$$

where  $n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4)$  is expressed as

$$n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) = \sum_{mnpq \in \mathcal{B}} \phi_p(\mathbf{X}_4) \phi_q(\mathbf{X}_3) \phi_n(\mathbf{X}_2) \phi_m(\mathbf{X}_1) \Gamma_{mn}^{pq}[\Psi^{\mathcal{B}}], \quad (\text{A13})$$



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

$$\Gamma_{mn}^{pq}[\Psi^{\mathcal{B}}] = \langle \Psi^{\mathcal{B}} | \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_n \hat{a}_m | \Psi^{\mathcal{B}} \rangle. \quad (\text{A14})$$

By integrating over  $\mathbf{X}_3$  and  $\mathbf{X}_4$  in equation (A12), it comes:

$$\langle \Psi^{\mathcal{B}} | \hat{W}_{\text{ee}}^{\mathcal{B}} | \Psi^{\mathcal{B}} \rangle = \frac{1}{2} \iint d\mathbf{X}_1 d\mathbf{X}_2 f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2), \quad (\text{A15})$$

where we introduced the function

$$f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \sum_{ijklmn \in \mathcal{B}} V_{ij}^{kl} \Gamma_{kl}^{mn}[\Psi^{\mathcal{B}}] \phi_n(\mathbf{X}_2) \phi_m(\mathbf{X}_1) \phi_i(\mathbf{X}_1) \phi_j(\mathbf{X}_2). \quad (\text{A16})$$

From the definition of the restriction of an operator to the space generated by the basis set  $\mathcal{B}$ , we have the following equality

$$\langle \Psi^{\mathcal{B}} | \hat{W}_{\text{ee}}^{\mathcal{B}} | \Psi^{\mathcal{B}} \rangle = \langle \Psi^{\mathcal{B}} | \hat{W}_{\text{ee}} | \Psi^{\mathcal{B}} \rangle, \quad (\text{A17})$$

which translates into

$$\begin{aligned} & \frac{1}{2} \iint d\mathbf{X}_1 d\mathbf{X}_2 f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) \\ &= \frac{1}{2} \iint d\mathbf{X}_1 d\mathbf{X}_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2), \end{aligned} \quad (\text{A18})$$

and holds for any  $\Psi^{\mathcal{B}}$ . Therefore, by introducing the following function

$$W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \frac{f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)}{n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2)}, \quad (\text{A19})$$

one can rewrite equation (A18) as

$$\begin{aligned} & \iint d\mathbf{X}_1 d\mathbf{X}_2 W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2) \\ &= \iint d\mathbf{X}_1 d\mathbf{X}_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2). \end{aligned} \quad (\text{A20})$$

## Appendix B: Behavior of the effective electron-electron interaction $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ in the limit of a complete basis set

To study how  $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  behaves in the limit of a complete basis set, one only needs to study  $f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ . By expliciting the two-electron integrals,  $f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  can be written as

$$\begin{aligned} f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) &= \\ & \sum_{ijklmn \in \mathcal{B}} \Gamma_{kl}^{mn}[\Psi^{\mathcal{B}}] \phi_n(\mathbf{X}_2) \phi_m(\mathbf{X}_1) \phi_i(\mathbf{X}_1) \phi_j(\mathbf{X}_2) \\ & \iint d\mathbf{X} d\mathbf{X}' \phi_k(\mathbf{X}) \phi_l(\mathbf{X}') \phi_i(\mathbf{X}) \phi_j(\mathbf{X}') \frac{1}{|\mathbf{r} - \mathbf{r}'|}, \end{aligned} \quad (\text{B1})$$

which, after regrouping the summations over the indices  $i$  and  $j$ , becomes:

$$\begin{aligned} f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) &= \sum_{mnkl \in \mathcal{B}} \Gamma_{kl}^{mn}[\Psi^{\mathcal{B}}] \phi_n(\mathbf{X}_2) \phi_m(\mathbf{X}_1) \\ & \int d\mathbf{X} \left( \sum_{i \in \mathcal{B}} \phi_i(\mathbf{X}_1) \phi_i(\mathbf{X}) \right) \phi_k(\mathbf{X}) \\ & \int d\mathbf{X}' \left( \sum_{j \in \mathcal{B}} \phi_j(\mathbf{X}_2) \phi_j(\mathbf{X}') \right) \phi_l(\mathbf{X}') \frac{1}{|\mathbf{r} - \mathbf{r}'|}. \end{aligned} \quad (\text{B2})$$

One can recognize in equation (B2) the expression of the restriction of a Dirac distribution to the basis set  $\mathcal{B}$ :

$$\delta^{\mathcal{B}}(\mathbf{Y} - \mathbf{Y}') = \sum_{i \in \mathcal{B}} \phi_i(\mathbf{Y}) \phi_i(\mathbf{Y}'). \quad (\text{B3})$$

Such a distribution  $\delta^{\mathcal{B}}(\mathbf{Y} - \mathbf{Y}')$  maintains the standard Dirac distribution properties only when applied to functions which are exactly representable in  $\mathcal{B}$ . More precisely, if  $g$  is a test function from  $\mathbb{R}^3$  to  $\mathbb{R}$ ,  $g^{\mathcal{B}}$  its component in  $\mathcal{B}$  and  $g^{\perp}$  the orthogonal component:

$$g = g^{\mathcal{B}} + g^{\perp} \quad \text{with} \quad \int d\mathbf{r} g^{\mathcal{B}}(\mathbf{r}) g^{\perp}(\mathbf{r}) = 0, \quad (\text{B4})$$

then

$$\int d\mathbf{r} \delta^{\mathcal{B}}(\mathbf{r} - \mathbf{r}') g(\mathbf{r}) = g(\mathbf{r}') \quad \text{iff} \quad g^{\perp}(\mathbf{r}') = 0 \quad \forall \mathbf{r}'. \quad (\text{B5})$$

In the limit of a complete basis set, the function  $\phi_l(\mathbf{X}') \frac{1}{|\mathbf{r} - \mathbf{r}'|}$  is necessarily within  $\mathcal{B}$ , and thus one has:

$$\begin{aligned} & \lim_{\mathcal{B} \rightarrow \infty} \int d\mathbf{X}' \delta^{\mathcal{B}}(\mathbf{X}_2 - \mathbf{X}') \phi_l(\mathbf{X}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \\ &= \phi_l(\mathbf{X}_2) \frac{1}{|\mathbf{r} - \mathbf{r}_2|}, \end{aligned} \quad (\text{B6})$$

and

$$\begin{aligned} & \lim_{\mathcal{B} \rightarrow \infty} \int d\mathbf{X} \delta^{\mathcal{B}}(\mathbf{X}_1 - \mathbf{X}) \phi_k(\mathbf{X}) \int d\mathbf{X}' \frac{\delta^{\mathcal{B}}(\mathbf{X}_2 - \mathbf{X}') \phi_l(\mathbf{X}')}{|\mathbf{r} - \mathbf{r}'|} \\ &= \phi_k(\mathbf{X}_1) \phi_l(\mathbf{X}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}. \end{aligned} \quad (\text{B7})$$

Inserting this expression in  $f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  leads to:

$$\begin{aligned} \lim_{\mathcal{B} \rightarrow \infty} f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) &= \sum_{klmn \in \mathcal{B}} \Gamma_{kl}^{mn}[\Psi^{\mathcal{B}}] \\ & \phi_m(\mathbf{X}_1) \phi_n(\mathbf{X}_2) \phi_l(\mathbf{X}_2) \phi_k(\mathbf{X}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}, \end{aligned} \quad (\text{B8})$$

which is nothing but

$$\lim_{\mathcal{B} \rightarrow \infty} f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = n_{\Psi^{\mathcal{B}}}^{(2)}(\mathbf{X}_1, \mathbf{X}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (\text{B9})$$

Therefore, in the limit of a complete basis set, the effective electron-electron interaction  $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$  correctly

reduces to the true Coulomb interaction interaction for all points  $(\mathbf{X}_1, \mathbf{X}_2)$ .

$$\lim_{\beta \rightarrow \infty} W_{\Psi^\beta}(\mathbf{X}_1, \mathbf{X}_2) = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad \forall (\mathbf{X}_1, \mathbf{X}_2) \text{ and } \Psi^\beta. \quad (\text{B10})$$

- <sup>1</sup>J. Goldstone, *Proc. R. Soc. A* **239**, 267 (1957).
- <sup>2</sup>I. Lindgren, *Phys. Scr.* **32**, 291 (1985).
- <sup>3</sup>R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007).
- <sup>4</sup>C. F. Bender and E. R. Davidson, *Phys. Rev.* **183**, 23 (1969).
- <sup>5</sup>B. Huron, J. Malrieu, and P. Rancurel, *J. Chem. Phys.* **58**, 5745 (1973).
- <sup>6</sup>R. J. Buenker and S. D. Peyerimhoff, *Theor. Chim. Acta* **35**, 33 (1974).
- <sup>7</sup>R. J. Buenker, S. D. Peyerimhoff, and P. J. Bruna, *Comp. Theor. Org. Chem.* (Reidel, Dordrecht, 1981) p. 55.
- <sup>8</sup>S. Evangelisti, J.-P. Daudey, and J.-P. Malrieu, *Chem. Phys.* **75**, 91 (1983).
- <sup>9</sup>R. J. Harrison, *J. Chem. Phys.* **94**, 5021 (1991).
- <sup>10</sup>A. A. Holmes, N. M. Tubman, and C. J. Umrigar, *J. Chem. Theory Comput.* **12**, 3674 (2016).
- <sup>11</sup>T. Kato, *Comm. Pure Appl. Math.* **10**, 151 (1957).
- <sup>12</sup>E. Hylleraas, *Z. Physik* **54**, 347 (1929).
- <sup>13</sup>C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, *Chem. Rev.* **112**, 4 (2012).
- <sup>14</sup>L. Kong, F. A. Bischoff, and E. F. Valeev, *Chem. Rev.* **112**, 75 (2012).
- <sup>15</sup>A. Grüneis, S. Hirata, Y.-Y. Ohnishi, and S. Ten-no, *J. Chem. Phys.* **146**, 080901 (2017).
- <sup>16</sup>P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- <sup>17</sup>W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- <sup>18</sup>A. D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).
- <sup>19</sup>L. Goerigk and S. Grimme, *WIREs Comput. Mol. Sci.* **4**, 576 (2014).
- <sup>20</sup>R. O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61**, 689 (1989).
- <sup>21</sup>J. Toulouse, F. Colonna, and A. Savin, *Phys. Rev. A* **70**, 062505 (2004).
- <sup>22</sup>O. Franck, B. Mussard, E. Luppi, and J. Toulouse, *J. Chem. Phys.* **142**, 074107 (2015).
- <sup>23</sup>J. G. Ángyán, I. C. Gerber, A. Savin, and J. Toulouse, *Phys. Rev. A* **72**, 012510 (2005).
- <sup>24</sup>E. Goll, H.-J. Werner, and H. Stoll, *Phys. Chem. Chem. Phys.* **7**, 3917 (2005).
- <sup>25</sup>J. Toulouse, I. C. Gerber, G. Jansen, A. Savin, and J. G. Ángyán, *Phys. Rev. Lett.* **102**, 096404 (2009).
- <sup>26</sup>B. G. Janesko, T. M. Henderson, and G. E. Scuseria, *J. Chem. Phys.* **130**, 081105 (2009).
- <sup>27</sup>T. Leininger, H. Stoll, H.-J. Werner, and A. Savin, *Chem. Phys. Lett.* **275**, 151 (1997).
- <sup>28</sup>E. Fromager, J. Toulouse, and H. J. A. Jensen, *J. Chem. Phys.* **126**, 074111 (2007).
- <sup>29</sup>E. Fromager, R. Cimraglia, and H. J. A. Jensen, *Phys. Rev. A* **81**, 024502 (2010).
- <sup>30</sup>E. D. Hedegård, S. Knecht, J. S. Kielberg, H. J. A. Jensen, and M. Reiher, *J. Chem. Phys.* **142**, 224108 (2015).
- <sup>31</sup>L. Kronik, T. Stein, S. Refaely-Abramson, and R. Baer, *J. Chem. Theory Comput.* **8**, 1515 (2012).
- <sup>32</sup>A. V. Krukau, G. E. Scuseria, J. P. Perdew, and A. Savin, *J. Chem. Phys.* **129**, 124103 (2008).
- <sup>33</sup>T. M. Henderson, B. G. Janesko, G. E. Scuseria, and A. Savin, *Int. J. Quantum Chem.* **109**, 2023 (2009).
- <sup>34</sup>J. Toulouse, P. Gori-Giorgi, and A. Savin, *Theor. Chem. Acc.* **114**, 305 (2005).
- <sup>35</sup>J. C. Slater, *Phys. Rev.* **81**, 385 (1951).
- <sup>36</sup>S. Pazziani, S. Moroni, P. Gori-Giorgi, and G. B. Bachelet, *Phys. Rev. B* **73**, 155111 (2006).
- <sup>37</sup>J. Rubio, J. Novoa, and F. Illas, *Chem. Phys. Lett.* **126**, 98 (1986).
- <sup>38</sup>R. Cimraglia and M. Persico, *J. Comp. Chem.* **8**, 39 (1987).
- <sup>39</sup>C. Angeli and M. Persico, *Theor. Chem. Acc.* **98**, 117 (1997).
- <sup>40</sup>C. Angeli, R. Cimraglia, and J.-P. Malrieu, *Chem. Phys. Lett.* **317**, 472 (2000).
- <sup>41</sup>E. Giner, A. Scemama, and M. Caffarel, *Can. J. Chem.* **91**, 879 (2013).
- <sup>42</sup>A. Scemama, T. Applencourt, E. Giner, and M. Caffarel, *J. Chem. Phys.* **141**, 244110 (2014).
- <sup>43</sup>E. Giner, A. Scemama, and M. Caffarel, *J. Chem. Phys.* **142**, 044115 (2015).
- <sup>44</sup>E. Giner, R. Assaraf, and J. Toulouse, *Mol. Phys.* **114**, 910 (2016).
- <sup>45</sup>P. S. Epstein, *Phys. Rev.* **28**, 695 (1926).
- <sup>46</sup>R. K. Nesbet, *Proc. R. Soc. A* **230**, 312 (1955).
- <sup>47</sup>Y. Garniron, A. Scemama, P.-F. Loos, and M. Caffarel, *J. Chem. Phys.* **147**, 034101 (2017).
- <sup>48</sup>A. Scemama, T. Applencourt, Y. Garniron, E. Giner, G. David, and M. Caffarel, "Quantum package v1.0," (2016).
- <sup>49</sup>G. H. Booth and A. Alavi, *J. Chem. Phys.* **132**, 174104 (2010).
- <sup>50</sup>S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia, and C. Froese Fischer, *Phys. Rev. A* **47**, 3649 (1993).