AUTHOR QUERY FORM

AIP Publishing	**Journal: J. Chem. Phys.** **Article Number: 010843JCP**	**Please provide your responses and any corrections by annotating this PDF and uploading it to AIP's eProof website as detailed in the Welcome email.**

Dear Author,

Below are the queries associated with your article. Please answer all of these queries before sending the proof back to AIP.

Article checklist: In order to ensure greater accuracy, please check the following and make all necessary corrections before returning your proof.
1. Is the title of your article accurate and spelled correctly?
2. Please check affiliations including spelling, completeness, and correct linking to authors.
3. Did you remember to include acknowledgment of funding, if required, and is it accurate?

Location in article	Query/Remark: click on the Q link to navigate to the appropriate spot in the proof. There, insert your comments as a PDF annotation.
Q1	Please check that the author names are in the proper order and spelled correctly. Also, please ensure that each author's given and surnames have been correctly identified (given names are highlighted in red and surnames appear in blue).
Q2	In the sentence beginning "To overcome these...," please specify whether "ingredients" could be changed to "criteria" or "parameters."
Q3	In the sentence beginning "Sections II C–II F are devoted...," please confirm that "the next sections" refers to Secs. II C–II F.
Q4	Please note that the labels are missing in the captions of Figs. 1–3 and 5–8 though the part figures were labeled in source files. Also note that we have stacked the part figures as supplied for Figs. 1(a), 1(b), 1(d), 2(a), 2(b), 2(c), 3(a), 5(a), 7(b), and 8(a). Kindly check all artworks for correctness and insert the labels in captions, if needed.
Q5	Please define CIPSI at first occurrence.
Q6	We have reworded the sentence beginning "By integrating over..." for clarity. Please check that your meaning is preserved.

Thank you for your assistance.

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

Emmanuel Giner,[1,a)] Barthélemy Pradines,[1,2] Anthony Ferté,[1] Roland Assaraf,[1] Andreas Savin,[1] and Julien Toulouse[1]
[1]*Laboratoire de Chimie Théorique, Sorbonne Université and CNRS, F-75005 Paris, France*
[2]*Institut des Sciences du Calcul et des Données, Sorbonne Université, F-75005 Paris, France*

(Received 21 August 2018; accepted 30 October 2018; published online XX XX XXXX)

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 on 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. *Published by AIP Publishing.* https://doi.org/10.1063/1.5052714

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 exist 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 rise to important theorems[1,2] and many robust methods, such as coupled cluster[3] or selected configuration interaction (CI).[4–10] 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 a behavior very often prohibits the reach of the so-called complete-basis-set limit which is often needed to obtain quantitative agreement with the 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.[11] To cure this problem, explicitly correlated (f_{12}) methods have emerged from the pioneering work of Hylleraas[12] 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 Hohenberg-Kohn theorems,[16] abandons the complex many-body wave function for the simple one-body density. Thanks to the so-called Kohn-Sham formalism of DFT[17] and the development of practical approximations of the exchange-correlation

[a)]Electronic mail: emmanuel.giner@lct.jussieu.fr

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[18] to many-body perturbation theory.[19] Nevertheless, the lack of a scheme to rationally and systematically improve the quality of approximate density functionals[20] 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.[22] 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,[23] coupled cluster,[24] and random-phase approximations[25,26]) or multi-reference WFT approaches (such as multi-reference CI,[27] multiconfiguration self-consistent field,[28] multi-reference perturbation theory,[29] and density-matrix renormalization group[30]). 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 μ, and there is no fully satisfying and systematic scheme to set its value, even if some interesting proposals have been made.[31–33]

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 remaining 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 on 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 Secs. II A and II B. In Sec. 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 Sec. II D (with details given in Appendixes 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 Sec. 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 Sec. II F. In Sec. 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 Sec. 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 Sec. 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 the 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 Sec. III B 3 for the oxygen atom and its first cation. Finally, we summarize the main results and conclude in Sec. 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

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

where

$$(v_{\rm ne}(\mathbf{r})|n(\mathbf{r})) = \int d\mathbf{r}\; v_{\rm ne}(\mathbf{r})\; n(\mathbf{r}) \tag{2}$$

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

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

where the minimization is over N-electron wave functions Ψ with density equal to $n(\mathbf{r})$, and $\hat{T}$ and $\hat{W}_{\rm ee}$ 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 Ψ to obtain the value $F[n(\mathbf{r})]$. Provided that the search in Eq. (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 that 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 gives a ground-state energy $E_0^{\mathcal{B}}$

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

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

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

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

$$E_0 \approx E_0^{\mathcal{B}}. \tag{6}$$

Considering the fast convergence of the density with the size of the basis set, we expect the approximation of Eq. (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 $n^{\mathcal{B}}(\mathbf{r})$:

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

where $\Psi^{\mathcal{B}}$ are the 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

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

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 Eq. (5) is good.

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

$$\begin{aligned} E_0^{\mathcal{B}} = \min_{n^{\mathcal{B}}(\mathbf{r})} \Big\{ &\min_{\Psi^{\mathcal{B}} \to n^{\mathcal{B}}(\mathbf{r})} \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi^{\mathcal{B}} \rangle \\ &+ (v_{\mathrm{ne}}(\mathbf{r})|n^{\mathcal{B}}(\mathbf{r})) + \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})] \Big\}, \end{aligned} \tag{9}$$

or, after recombining the two minimizations,

$$E_0^{\mathcal{B}} = \min_{\Psi^{\mathcal{B}}} \left\{ \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi^{\mathcal{B}} \rangle + (v_{\mathrm{ne}}(\mathbf{r})|n_{\Psi^{\mathcal{B}}}(\mathbf{r})) + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}(\mathbf{r})] \right\}, \tag{10}$$

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

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

where $\hat{T}^{\mathcal{B}}$, $\hat{W}_{\mathrm{ee}}^{\mathcal{B}}$, $\hat{V}_{\mathrm{ne}}^{\mathcal{B}}$, and $\hat{\bar{V}}^{\mathcal{B}}[n(\mathbf{r})]$ are the restrictions to the space generated by the basis $\mathcal{B}$ of the operators $\hat{T}$, $\hat{W}_{\mathrm{ee}}$, $\int \mathrm{d}\mathbf{r}\, v_{\mathrm{ne}}(\mathbf{r})\hat{n}(\mathbf{r})$, and $\int \mathrm{d}\mathbf{r}(\delta \bar{E}^{\mathcal{B}}[n(\mathbf{r})]/\delta n(\mathbf{r}))\hat{n}(\mathbf{r})$, respectively, and $\hat{n}(\mathbf{r})$ is the density operator. The potential $\hat{\bar{V}}^{\mathcal{B}}[n_{\Psi_0^{\mathcal{B}}}(\mathbf{r})]$ ensures that the minimizing wave function $\Psi_0^{\mathcal{B}}$ gives the minimizing density $n_0^{\mathcal{B}}(\mathbf{r})$ in Eq. (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: $n_{\Psi_0^{\mathcal{B}}}(\mathbf{r}) = n_0(\mathbf{r}) \Longrightarrow E_0^{\mathcal{B}} = E_0$.

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

In the limit where $\mathcal{B}$ is a complete basis set, Eq. (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 Eq. (10) without $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$ simply gives the FCI energy in a given basis set $\mathcal{B}$

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

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

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

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

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

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

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

The term $\bar{E}^{\mathcal{B}}[n_{\Psi_{\mathrm{FCI}}^{\mathcal{B}}}(\mathbf{r})]$ constitutes a simple DFT correction to the FCI energy which should compensate for the incompleteness of the basis set $\mathcal{B}$. Sections II C–II F 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 Eq. (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 the short-range correlation effects poorly 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 $\bar{E}^{\mathrm{sr},\mu}_{\mathrm{c,md}}[n(\mathbf{r})]$ of Toulouse *et al.*[34] has a definition very similar to the one of $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$

$$\bar{E}^{\mathrm{sr},\mu}_{\mathrm{c,md}}[n(\mathbf{r})] = \min_{\Psi\to n(\mathbf{r})} \langle\Psi|\hat{T}+\hat{W}_{\mathrm{ee}}|\Psi\rangle - \langle\Psi^{\mu}[n(\mathbf{r})]|\hat{T}+\hat{W}_{\mathrm{ee}}|\Psi^{\mu}[n(\mathbf{r})]\rangle, \tag{16}$$

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

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

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

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

with

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

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 Eq. (16) with the definition of $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$ in Eq. (8), one can see that the only difference between these two functionals relies in the wave functions used for the constrained minimization: in $\bar{E}^{\mathrm{sr},\mu}_{\mathrm{c,md}}[n(\mathbf{r})]$, one uses Ψ^{μ}, whereas $\Psi^{\mathcal{B}}$ is used in $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$. More specifically, Ψ^{μ} is determined by using a non-diverging long-range electron-electron interaction defined in a complete basis set [Eq. (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 $\bar{E}^{\mathrm{sr},\mu}_{\mathrm{c,md}}[n(\mathbf{r})]$ 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 on 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 as

$$\hat{W}^{\mathcal{B}}_{\mathrm{ee}} = \frac{1}{2}\sum_{ijkl\,\in\,\mathcal{B}} V^{kl}_{ij}\;\hat{a}^{\dagger}_k\hat{a}^{\dagger}_l\hat{a}_j\hat{a}_i, \tag{20}$$

where the sums run over all (real-valued) orthonormal spin-orbitals $\{\phi_i\}$ in the basis set $\mathcal{B}$, and V^{kl}_{ij} 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 $\hat{W}^{\mathcal{B}}_{\mathrm{ee}}$ over a wave function $\Psi^{\mathcal{B}}$ can be written as (see Appendix A for a detailed derivation)

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

where we introduced the function

$$f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = \sum_{ijklmn\,\in\,\mathcal{B}} V^{kl}_{ij}\,\Gamma^{mn}_{kl}[\Psi^{\mathcal{B}}] \times \phi_n(\mathbf{X}_2)\phi_m(\mathbf{X}_1)\phi_i(\mathbf{X}_1)\phi_j(\mathbf{X}_2), \tag{22}$$

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

$$\Gamma^{pq}_{mn}[\Psi^{\mathcal{B}}] = \langle\Psi^{\mathcal{B}}|\hat{a}^{\dagger}_p\hat{a}^{\dagger}_q\hat{a}_n\hat{a}_m|\Psi^{\mathcal{B}}\rangle, \tag{23}$$

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

$$\mathbf{X} = (\mathbf{r},\sigma) \qquad \mathbf{r}\in\mathbb{R}^3,\;\sigma=\pm\frac{1}{2}, \qquad \int \mathrm{d}\mathbf{X} = \sum_{\sigma=\pm\frac{1}{2}}\int_{\mathbb{R}^3}\mathrm{d}\mathbf{r}. \tag{24}$$

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

$$\langle\Psi^{\mathcal{B}}|\hat{W}^{\mathcal{B}}_{\mathrm{ee}}|\Psi^{\mathcal{B}}\rangle = \langle\Psi^{\mathcal{B}}|\hat{W}_{\mathrm{ee}}|\Psi^{\mathcal{B}}\rangle, \tag{25}$$

which translates into

$$\frac{1}{2}\iint \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2\; f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = \frac{1}{2}\iint \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2\;\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\;n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2), \tag{26}$$

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

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

one can rewrite Eq. (26) as

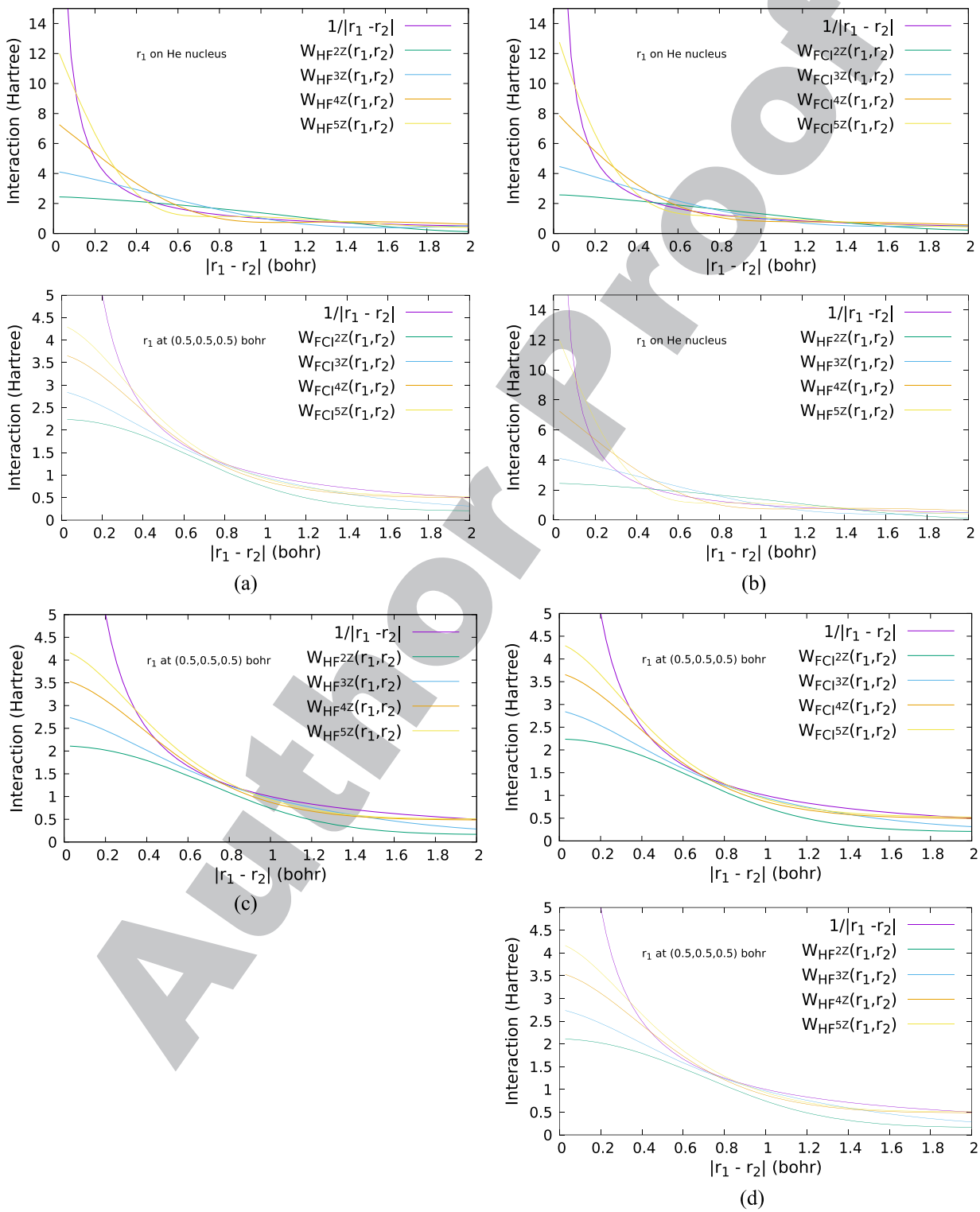
$$\iint \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2\;W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2)\;n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = \iint \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2\;\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\;n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2). \tag{28}$$

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.[35] It is important to notice that all the quantities appearing in the integrals of Eq. (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 $n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2)$ 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 equal at each point $(\mathbf{X}_1,\mathbf{X}_2)$. Of course, one could argue that there exist an infinite number of functions of $u(\mathbf{X}_1,\mathbf{X}_2)$ satisfying

$$\iint \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2\;u(\mathbf{X}_1,\mathbf{X}_2)\;n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = \iint \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2\;\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\;n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2), \tag{29}$$

which implies that the effective interaction is not uniquely defined, and that the choice of Eq. (27) is just one among the many and might not be optimal. For instance, the definition of the effective electron-electron interaction of Eq. (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 "$\mathcal{B} \to \infty$"), $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ correctly tends



Q4

FIG. 1. Effective electron-electron interaction $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)$ 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 $\Psi^{\mathcal{B}}$ 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.

to the exact Coulomb interaction

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

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 (σ 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 the 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 $\mathbf{r}_2$

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

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

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

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

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

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

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

In order to investigate how $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ behaves as a function of the basis set, the wave function $\Psi^{\mathcal{B}}$, 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 $\Psi^{\mathcal{B}}$ and different reference points $\mathbf{r}_1$. We report these numerical results in Fig. 1.

From Fig. 1, several trends can be observed. First, for all wave functions $\Psi^{\mathcal{B}}$ and reference points $\mathbf{r}_1$ used here, the value of $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ at coalescence is finite, which numerically illustrates Eq. (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 $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ is qualitatively modified by changing the reference point $\mathbf{r}_1$, which illustrates the lack of transitional invariance of $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$. In particular, the values of $W_{\mathrm{HF}^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ and $W_{\mathrm{FCI}^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ 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 $W_{\mathrm{HF}^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ and $W_{\mathrm{FCI}^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ is almost unnoticeable for all basis sets and for the two reference points $\mathbf{r}_1$ used here.

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

From the numerical illustration of the properties of $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ given in Sec. II D 3, it appears that the development of approximations for the density functional $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}(\mathbf{r})]$ seems rather complicated since the effective interaction $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ is system- and basis-dependent, non translationally invariant, and non-isotropic. Nevertheless, as it was numerically illustrated, the effective interaction $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ typically describes a long-range interaction which is finite at coalescence. Therefore, a possible way to approximate $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ is to locally fit $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ by the long-range interaction $w^{\mathrm{lr},\mu}(|\mathbf{r}_1-\mathbf{r}_2|)$ of Eq. (19) used in RS-DFT. To do so, here we propose to determine a local value of the range-separation parameter μ such that the value of the long-range interaction at coalescence is identical to the value of the effective interaction $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1)$ at coalescence at point $\mathbf{r}_1$. More specifically, the range-separation parameter $\mu(\mathbf{r}_1;\Psi^{\mathcal{B}})$ is thus determined for each $\mathbf{r}_1$ and $\Psi^{\mathcal{B}}$ by the condition

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

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

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

Therefore, defining the function $W^{\mathrm{lr},\mu(\mathbf{r}_1)}_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ as

$$W^{\mathrm{lr},\mu(\mathbf{r}_1)}_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2) = \frac{\mathrm{erf}\big(\mu(\mathbf{r}_1;\Psi^{\mathcal{B}})|\mathbf{r}_1-\mathbf{r}_2|\big)}{|\mathbf{r}_1-\mathbf{r}_2|}, \quad (38)$$

we make the following approximation:

$$W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2) \approx W^{\mathrm{lr},\mu(\mathbf{r}_1)}_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2), \qquad \forall\ (\mathbf{r}_1,\mathbf{r}_2). \quad (39)$$

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

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

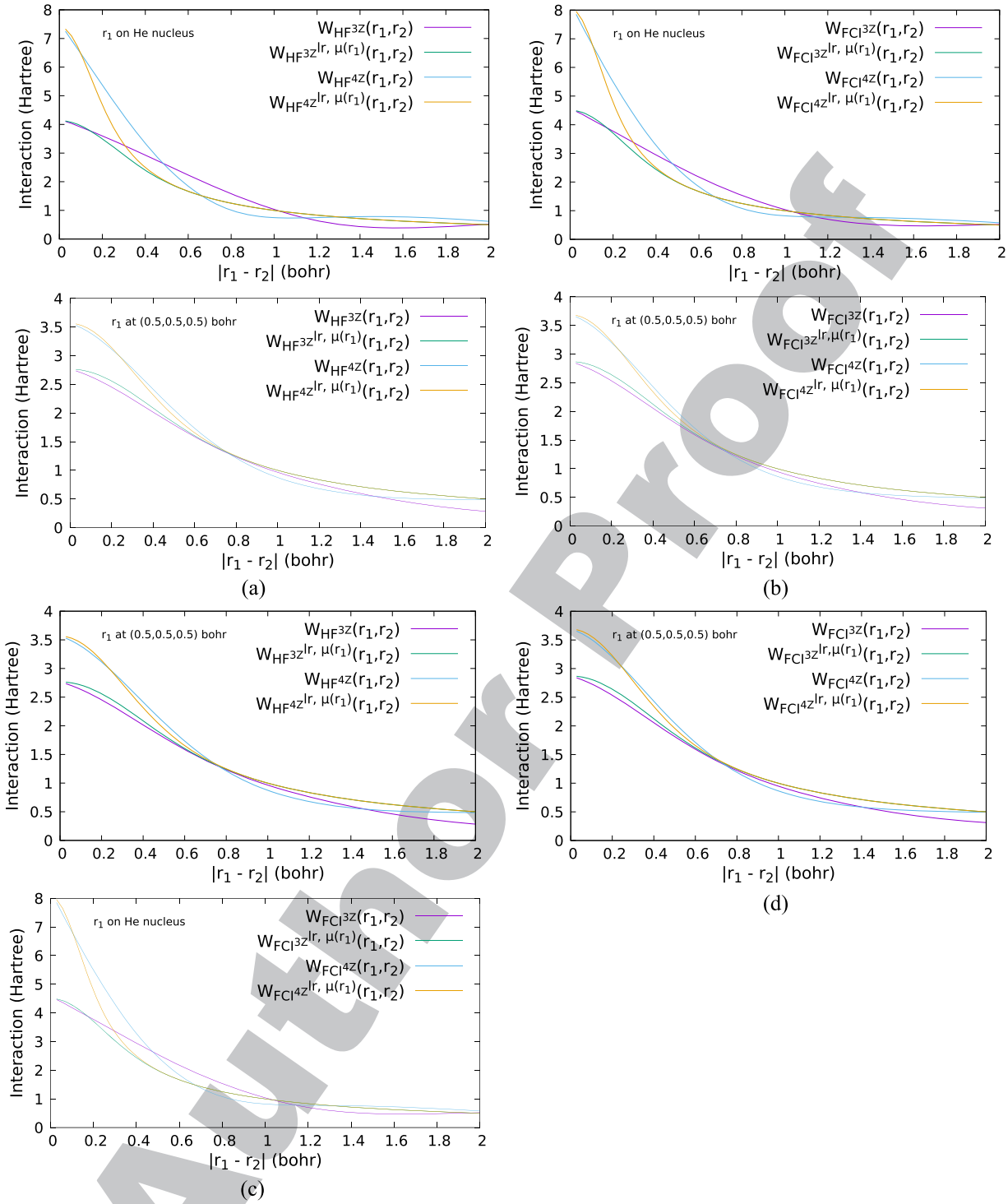


FIG. 2. Effective electron-electron interaction $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1,\mathbf{r}_2)$ and long-range electron-electron interaction $W_{\Psi^{\mathcal{B}}}^{\mathrm{lr},\mu(\mathbf{r}_1)}(\mathbf{r}_1,\mathbf{r}_2)$ for different cc-pVXZ basis sets (X = 3, 4) as a function of $|\mathbf{r}_1 - \mathbf{r}_2|$. 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 $\Psi^{\mathcal{B}}$ have been used: HF and FCI in the corresponding basis set.

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

In Fig. 3, we report the local range-separation parameter $\mu(\mathbf{r};\Psi^{\mathcal{B}})$, as determined by Eq. (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

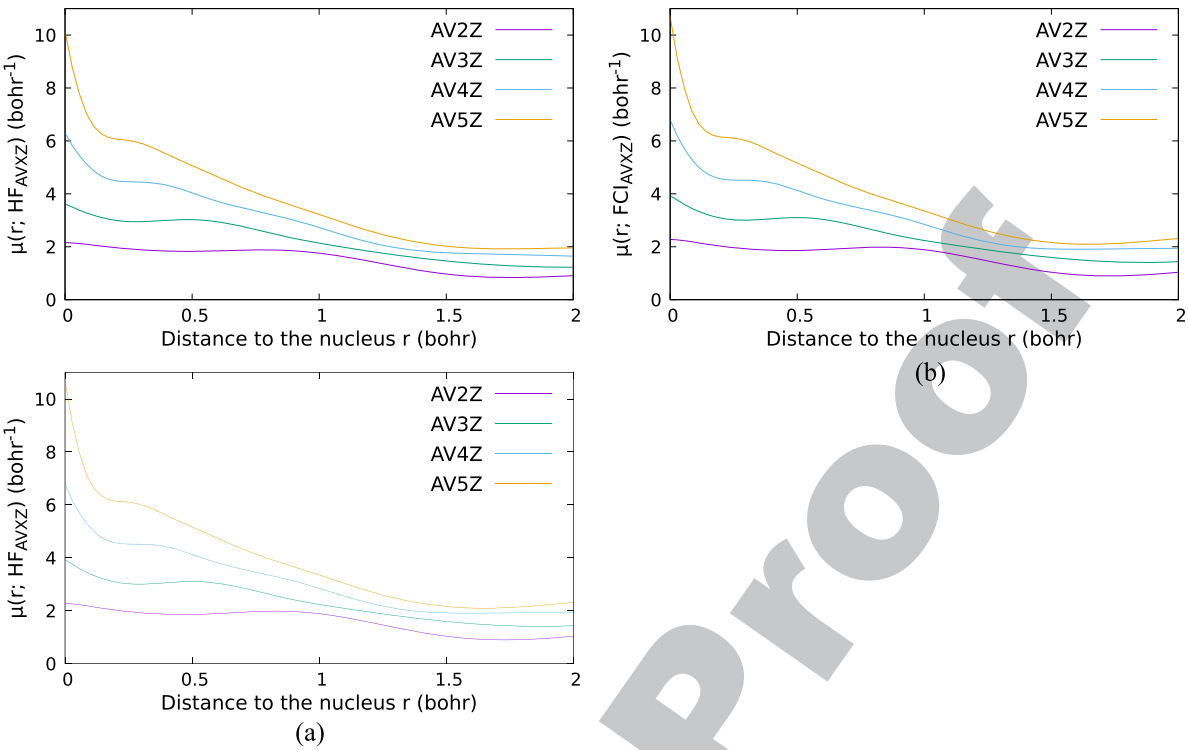


FIG. 3. Local range-separated parameter $\mu(\mathbf{r};\Psi^{\mathcal{B}})$ for the helium atom for different aug-cc-pVXZ basis sets ($X = 2, 3, 4, 5$) as a function of the position $\mathbf{r}$ along the diagonal of the xy plane. The curve on the left is when using the HF wave function for $\Psi^{\mathcal{B}}$,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

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

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

where $\bar{\varepsilon}^{\mathrm{sr,unif}}_{\mathrm{c,md}}(n,\mu)$ 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 Eq. (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 an uniform electron gas with a specific electron-electron interaction which approximately 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

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

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 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 built upon selected CI[4–10] which have been successfully used to converge to FCI correlation energies, one-body properties, and nodal surfaces.[8,37–44] The CIPSI algorithm used

in this work uses iteratively enlarged selected CI spaces and Epstein–Nesbet[45,46] multi-reference perturbation theory. The CIPSI energy is

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

where E_{v} is the variational energy

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

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 I within the CI reference space $\mathcal{R}$, and $E^{(2)}$ is the second-order energy correction

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

where κ denotes the 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.*[47] 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 κ 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$_{\Psi^{\mathcal{B}}}$ 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:

$$E_{\text{FCI}}^{\mathcal{B}} \approx E_{\text{CIPSI}}^{\mathcal{B}}, \tag{45}$$

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

with

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

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

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

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 Eq. (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$_{\text{HF}}$.

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 Eq. (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[48] using the Dunning aug-cc-pVXZ basis sets which are referred here as AVXZ.

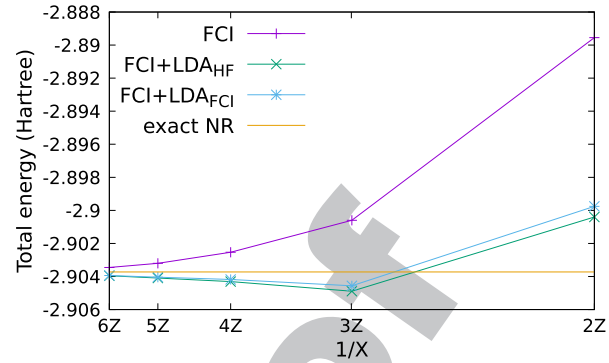


FIG. 4. Convergence of the total energy of the helium atom for FCI and FCI+LDA$_{\Psi^{\mathcal{B}}}$, where $\Psi^{\mathcal{B}}$ is either the HF or FCI wave function, as a function of the inverse of the cardinal number X of the AVXZ basis sets (X = 2, 3, 4, 5, 6). The exact non-relativistic (NR) energy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

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

We report in Fig. 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$_{\Psi^{\mathcal{B}}}$ where $\Psi^{\mathcal{B}}$ is either the HF or FCI wave function. The first striking observation from these data is that the FCI+LDA$_{\Psi^{\mathcal{B}}}$ energies rapidly converge to the exact energy as one increases the size of the basis set and that FCI+LDA$_{\Psi^{\mathcal{B}}}$ 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 more 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$_{\text{HF}}$ and FCI+LDA$_{\text{FCI}}$ 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 Eq. (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 Eq. (22) reduces

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$_{\text{HF}}$, and FCI+LDA$_{\text{FCI}}$ with the AVXZ basis sets (X = 2, 3, 4, 5, 6).

	FCI		FCI+LDA$_{\text{HF}}$		FCI+LDA$_{\text{FCI}}$	
	Total energy	Error	Total energy	Error	Total energy	Error
AV2Z	−2.88 955	14.17	−2.90 040	3.3187	−2.89 976	3.962
AV3Z	−2.90 060	03.12	−2.90 489	−1.1698	−2.90 456	−0.840
AV4Z	−2.90 253	01.18	−2.90 430	−0.5849	−2.90 418	−0.460
AV5Z	−2.90 320	00.52	−2.90 409	−0.3710	−2.90 404	−0.321
AV6Z	−2.90 346	00.26	−2.90 396	−0.2367	−2.90 394	−0.217
Exact non-relativistic total energy −2.90 372						

the number of two-electron integrals involved in the definition of $W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2)$ [see Eq. (27)]. This reduction implies that the effective interaction $W_{\mathrm{HF}}(\mathbf{X}_1, \mathbf{X}_2)$ misses a part of the interaction within the basis set, namely, the repulsion between electrons in virtual orbitals. However, the fact that $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n(\mathbf{r})]$ and $\bar{E}^{\mathcal{B},\mathrm{FCI}}_{\mathrm{LDA}}[n(\mathbf{r})]$ are close suggests that $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n(\mathbf{r})]$ 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 $W_{\mathrm{HF}}(\mathbf{X}_1, \mathbf{X}_2)$ are of the type V^{ab}_{ij} (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$_{HF}$ total energy with the number of determinants

We report in Fig. 5, in the case of the oxygen ground state using the AV4Z basis set, the convergence of the variational energy E_{v}, the CIPSI energy, the CIPSI+LDA$_{\mathrm{HF}}$ energy, and the LDA correction $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})]$ as a function of the number of Slater determinants in the reference wave function. The behavior of E_{v} and E_{CIPSI} reported in Fig. 5 is 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_{CIPSI} with a reference wave function including 2×10^3 and 5×10^5 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 $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})]$, it varies by about 0.08 mH between 100 and 4×10^6 determinants. The very small variation of $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})]$ can be qualitatively understood by noticing that, within the LDA approximation of Eq. (40) and choosing a HF wave function for $\Psi^{\mathcal{B}}$ to define the effective interaction, $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})]$ 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$_{\mathrm{HF}}$ energy is only limited by the convergence of the CIPSI algorithm itself as $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})]$ converges rapidly with the quality of the wave function.

2. The ionization potentials of the B-Ne series using CIPSI+LDA$_{HF}$

In order to investigate how the correction $\bar{E}^{\mathcal{B},\Psi^{\mathcal{B}}}_{\mathrm{LDA}}[n^{\mathcal{B}}(\mathbf{r})]$ performs for energy differences, we report calculations 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 co-workers,[49] 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 rapidly out of reach for these systems, here we use the CIPSI+LDA$_{\mathrm{HF}}$

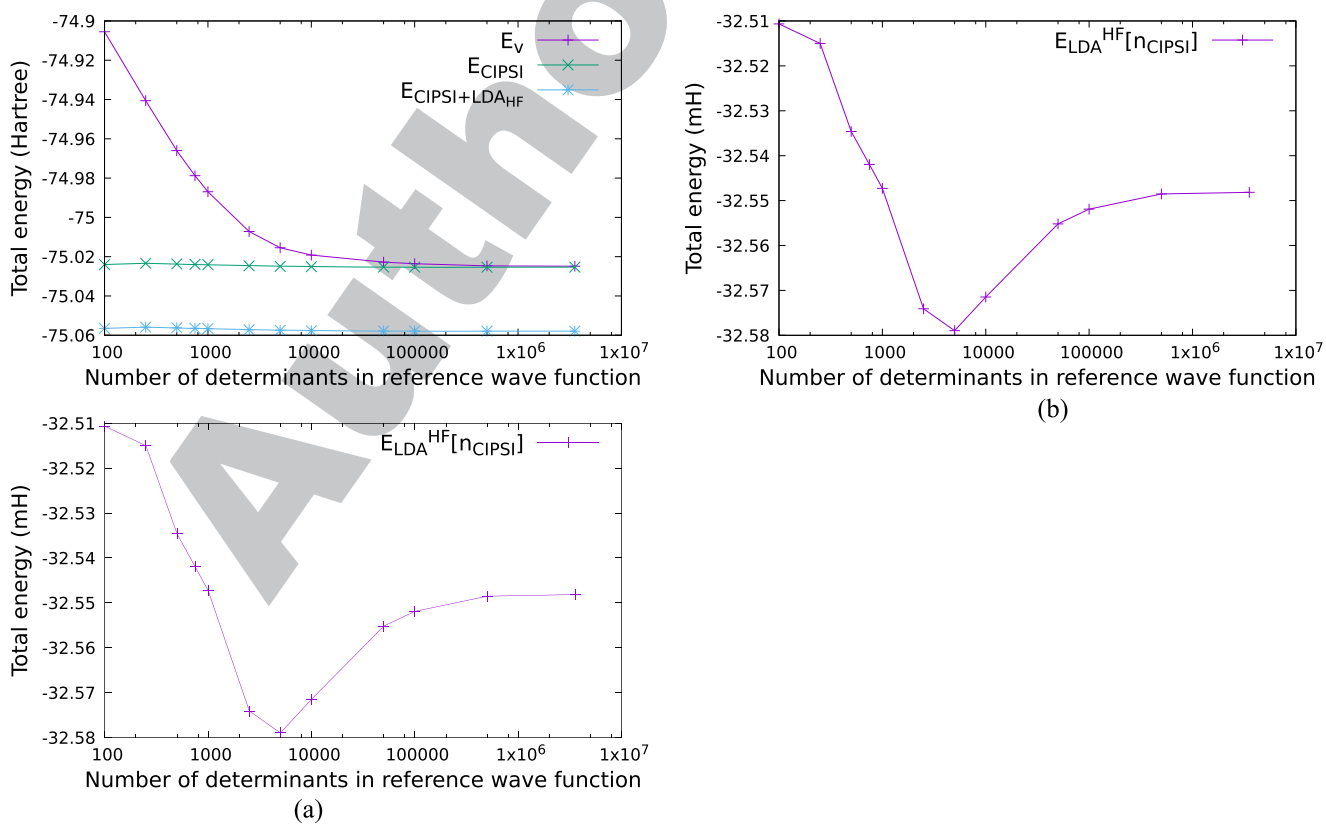


FIG. 5. Convergence of the variational total energy E_{v}, the CIPSI total energy, and the CIPSI+LDA$_{\mathrm{HF}}$ total energy (left plot), and of the LDA correction $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})]$ (right plot) of the oxygen atom as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

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$_{\mathrm{HF}}$ levels is also reported in Fig. 6. All electrons were correlated with 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 $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})]$ converges rapidly with respect to the number of Slater determinants (see Fig. 5), the approximation of Eq. (46) can be considered as valid, and therefore, our CIPSI+LDA$_{\mathrm{HF}}$ results can be considered as virtually identical to the ones that would be obtained with FCI+LDA$_{\mathrm{HF}}$. Finally, the CIPSI+LDA$_{\mathrm{HF}}$ 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^+ cation to 7.2 mH for the Ne atom.

TABLE II. Total energies (in Hartree) of the neutral atoms and first cations for the B-Ne series with the AV*X*Z basis sets (X = 2, 3, 4, 5) using CIPSI and CIPSI+LDA$_{\mathrm{HF}}$. The *i*-FCIQMC values from Ref. 49 are also reported for comparison with CIPSI.

	Method	AV2Z	AV3Z	AV4Z	AV5Z	Exact NR[a]
B	*i*-FCIQMC[b]	−24.59 242(1)	−24.60 665(2)	−24.62 407(11)	−24.63 023(2)	−24.65 390
	CIPSI	−24.592 418	−24.606 654	−24.624 109	−24.630 233	
	CIPSI+LDA$_{\mathrm{HF}}$	−24.641 525	−24.641 706	−24.648 135	−24.650 243	
B^+	*i*-FCIQMC[b]	−24.29 450(1)	−24.30 366(2)	−24.32 005(2)	−24.32 553(9)	
	CIPSI	−24.294 496	−24.303 660	−24.320 044	−24.325 531	−24.34 889
	CIPSI+LDA$_{\mathrm{HF}}$	−24.338 930	−24.336 580	−24.343 043	−24.345 024	
C	*i*-FCIQMC[b]	−37.76 656(1)	−37.79 163(2)	−37.81 301(2)	−37.82 001(4)	
	CIPSI	−37.766 573	−37.791 623	−37.813 025	−37.820 016	−37.8 450
	CIPSI+LDA$_{\mathrm{HF}}$	−37.824 730	−37.830 667	−37.838 253	−37.840 544	
C^+	*i*-FCIQMC[b]	−37.35 960(1)	−37.37 967(2)	−37.39 991(1)	−37.40 605(1)	
	CIPSI	−37.359 602	−37.379 703	−37.399 932	−37.406 342	−37.43 095
	CIPSI+LDA$_{\mathrm{HF}}$	−37.413 086	−37.416 631	−37.424 109	−37.426 321	
N	*i*-FCIQMC[b]	−54.48 881(2)	−54.52 797(1)	−54.55 423(3)	−54.56 303(2)	
	CIPSI	−54.488 814	−54.527 941	−54.554 235	−54.563 027	−54.5 893
	CIPSI+LDA$_{\mathrm{HF}}$	−54.556 940	−54.571 576	−54.581 128	−54.584 048	
N^+	*i*-FCIQMC [b]	−53.96 106(10)	−53.99 535(1)	−54.01 838(1)	−54.02 865(2)	
	CIPSI	−53.961 062	−53.995 355	−54.020 414	−54.028 633	−54.0 546
	CIPSI+LDA$_{\mathrm{HF}}$	−54.024 314	−54.036 820	−54.046 204	−54.049 068	
O	*i*-FCIQMC[b]	−74.92 772(2)	−74.99 077(4)	−75.02 534(4)	−75.03 749(6)	
	CIPSI	−74.927 696	−74.990 750	−75.025 340	−75.037 527	−75.0 674
	CIPSI+LDA$_{\mathrm{HF}}$	−75.014 946	−75.044 685	−75.057 889	−75.061 639	
O^+	*i*-FCIQMC[b]	−74.444 194(6)	−74.49 701(1)	−74.52 799(4)	−74.53 869(6)	
	CIPSI	−74.444 191	−74.497 018	−74.527 968	−74.538 630	−74.5 669
	CIPSI+LDA$_{\mathrm{HF}}$	−74.517 650	−74.543 804	−74.556 296	−74.560 233	
F	*i*-FCIQMC[b]	−99.55 223(1)	−99.64 036(2)	−99.68 460(10)	−99.70 029(5)	
	CIPSI	−99.552 228	−99.640 295	−99.684 561	−99.700 258	−99.7 341
	CIPSI+LDA$_{\mathrm{HF}}$	−99.658 315	−99.704 195	−99.722 750	−99.727 639	
F^+	*i*-FCIQMC[b]	−98.923 015(6)	−99.00 542(1)	−99.04 599(3)	−99.06 082(4)	
	CIPSI	−98.923 000	−99.005 441	−99.046 481	−99.060 808	−99.0 930
	CIPSI+LDA$_{\mathrm{HF}}$	−99.016 909	−99.062 981	−99.080 847	−99.085 872	
Ne	*i*-FCIQMC[b]	−128.71 145(3)	−128.82 577(5)	−128.88 065(6)	...	
	CIPSI	−128.711 476	−128.825 813	−128.880 658	−128.900 438	−128.9 383
	CIPSI+LDA$_{\mathrm{HF}}$	−128.835 474	−128.898 894	−128.924 219	−128.931 038	
Ne^+	*i*-FCIQMC[b]	−127.92 411(2)	−128.03 691(2)	−128.08 816(11)	...	
	CIPSI	−127.924 068	−128.036 898	−128.088 901	−128.107 479	−128.1 437
	CIPSI+LDA$_{\mathrm{HF}}$	−128.037 019	−128.104 203	−128.128 973	−128.135 914	

[a]Estimated exact non-relativistic (NR) values from Ref. 50.
[b]From Ref. 49. The statistical errors are given in parentheses.

TABLE III. IPs (in mH) calculated by CIPSI and CIPSI+LDA$_{\mathrm{HF}}$ 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 parentheses.

	Method	AV2Z	AV3Z	AV4Z	AV5Z	Exact NR[a]
B	CIPSI	297.92 (7.05)	302.99 (1.98)	304.06 (0.91)	304.70 (0.27)	304.98
	CIPSI+LDA$_{\mathrm{HF}}$	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$_{\mathrm{HF}}$	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$_{\mathrm{HF}}$	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$_{\mathrm{HF}}$	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$_{\mathrm{HF}}$	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$_{\mathrm{HF}}$	798.45 (−3.81)	794.69 (−0.05)	795.24 (−0.60)	795.12 (−0.48)	

[a]Estimated exact non-relativistic (NR) values from Ref. 50.

Regarding the quality of the IPs (Table III and Fig. 6), at the near FCI level (either i-FCIQMC or CIPSI), the typical chemical accuracy of 1 kcal/mol (≈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 very 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$_{\mathrm{HF}}$ 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$_{\mathrm{HF}}$, 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$_{\mathrm{HF}}$ 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 $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})]$ 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 $\bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})]$ for the oxygen atom and its first cation.

We first define the spherically averaged local basis-set correction as

$$\bar{E}^{\mathcal{B}}_{\mathrm{LDA}}(r) = \iint \mathrm{d}\Omega\; r^2 n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})\; \bar{\varepsilon}^{\mathrm{sr,unif}}_{\mathrm{c,md}}\left(n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r}); \mu(\mathbf{r};\mathrm{HF})\right)$$

such that

$$\int \mathrm{d}r\; \bar{E}^{\mathcal{B}}_{\mathrm{LDA}}(r) = \bar{E}^{\mathcal{B},\mathrm{HF}}_{\mathrm{LDA}}[n^{\mathcal{B}}_{\mathrm{CIPSI}}(\mathbf{r})], \tag{49}$$

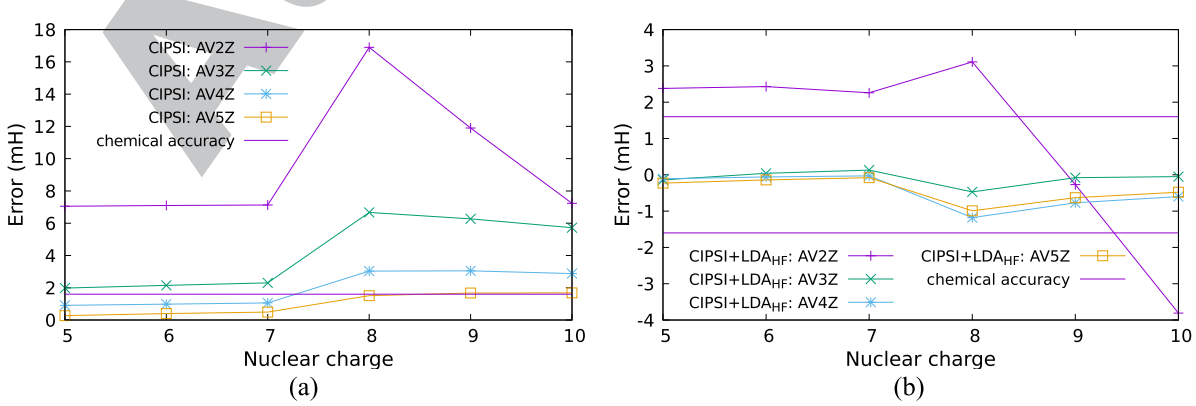


FIG. 6. Errors on the IPs calculated at the CIPSI (left plot) and CIPSI+LDA$_{\mathrm{HF}}$ (right plot) levels for the B-Ne series with the AVXZ basis sets (X = 2, 3, 4, 5). Note the different scales of the two plots.

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

We report in Fig. 7 the plot of $\bar{E}^{\mathcal{B}}_{\text{LDA}}(r)$ 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 $\bar{E}^{\mathcal{B},\text{HF}}_{\text{LDA}}[n^{\mathcal{B}}_{\text{CIPSI}}(\mathbf{r})]$ 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, $\bar{E}^{\mathcal{B}}_{\text{LDA}}(r)$ gets smaller as the size of the basis set is increased. With the largest basis set, $\bar{E}^{\mathcal{B}}_{\text{LDA}}(r)$ is small in the valence shell ($r > 0.5$ bohr) but remains substantial in the core region. 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^+, we also define the following function:

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

We report in Fig. 8 the values of $\Delta\bar{E}^{\mathcal{B}}_{\text{LDA}}(r)$ 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, $\Delta\bar{E}^{\mathcal{B}}_{\text{LDA}}(r)$ is always negative which means that $\bar{E}^{\mathcal{B},\text{HF}}_{\text{LDA}}[n^{\mathcal{B}}_{\text{CIPSI}}(\mathbf{r})]$ corrects more the neutral atom than the cation for the basis-set incompleteness. The fact that $\Delta\bar{E}^{\mathcal{B}}_{\text{LDA}}(r)$ 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.

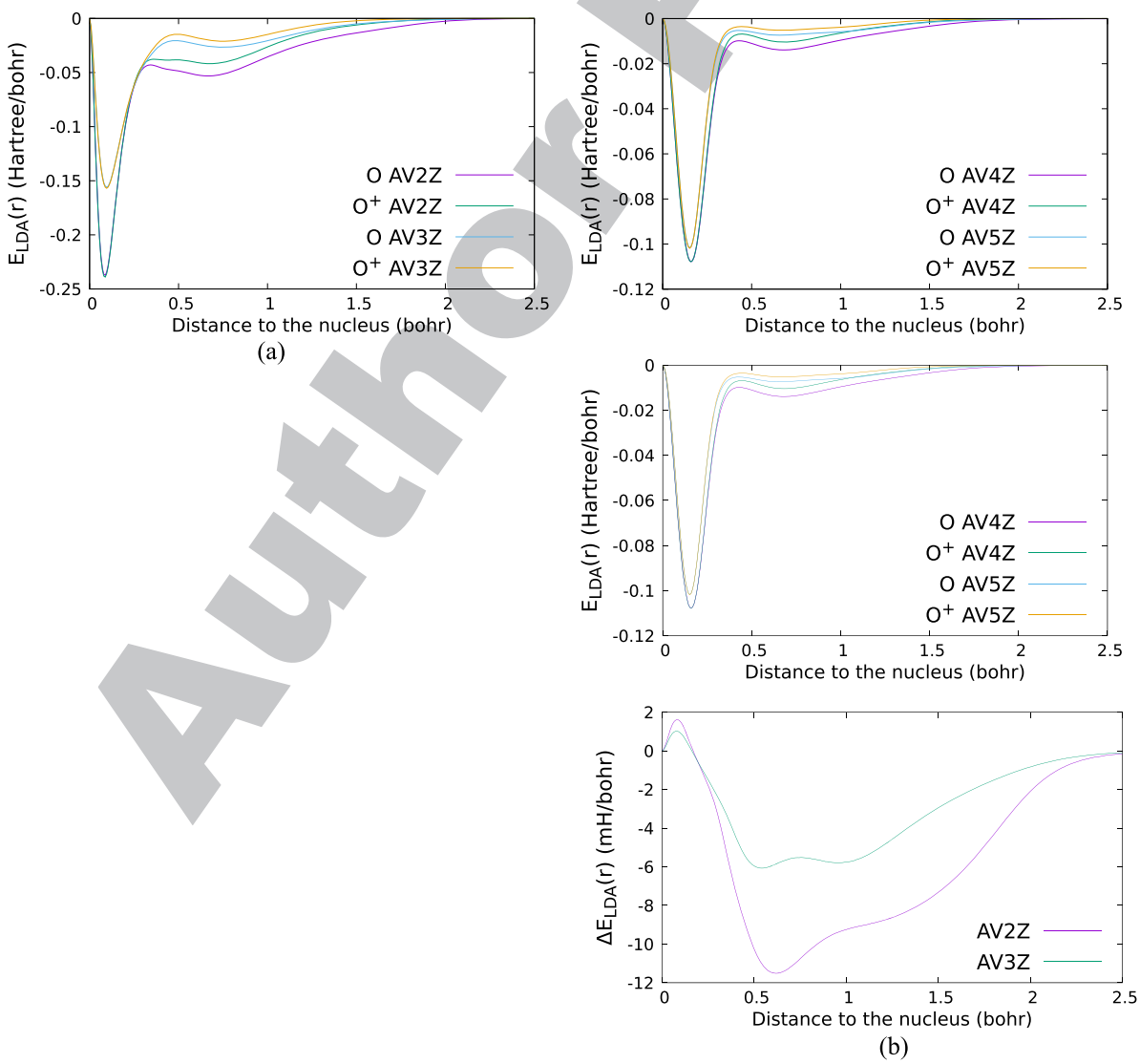


FIG. 7. Behavior of $\bar{E}^{\mathcal{B}}_{\text{LDA}}(r)$ for the AV2Z and AV3Z basis sets (left plot) and AV4Z and AV5Z basis sets (right plot) for the oxygen atom and its first cation.

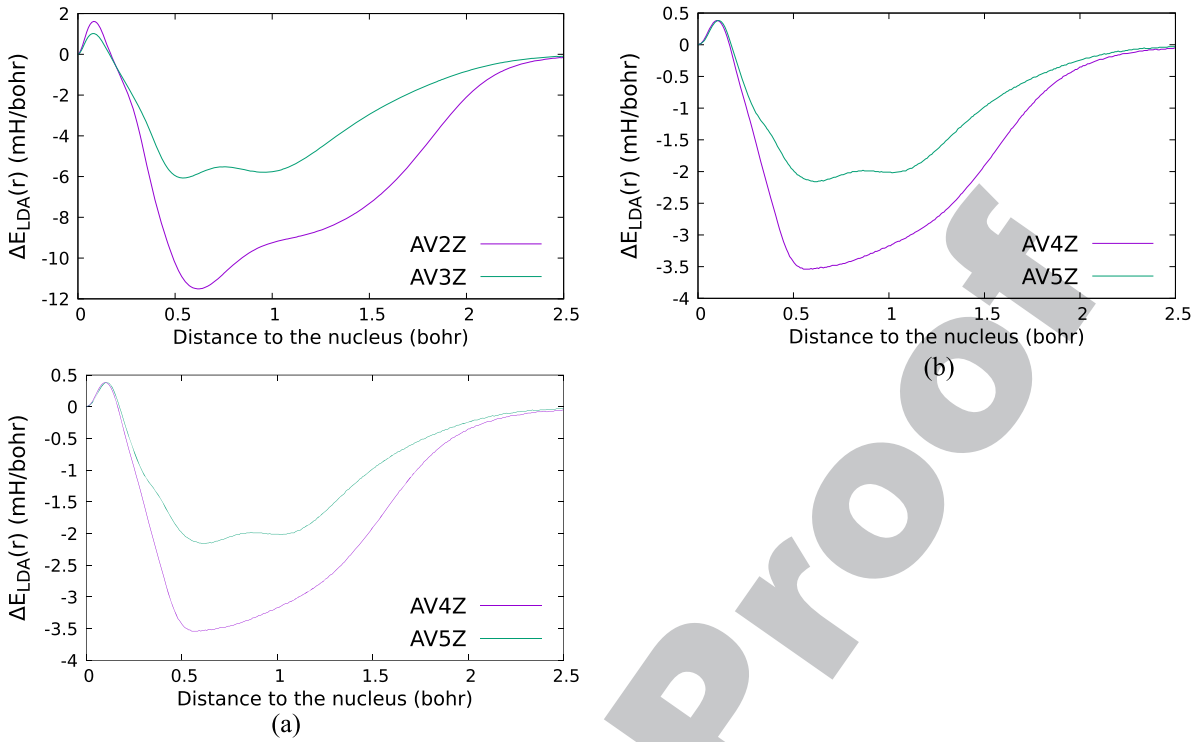


FIG. 8. Behavior of $\Delta\bar{E}^{\mathcal{B}}_{\mathrm{LDA}}(r)$ for the AV2Z and AV3Z basis sets (left plot) and AV4Z and AV5Z basis sets (right plot) for the oxygen atom.

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 Sec. II A) and a series of approximations (see Secs. 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 Sec. 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 Sec. 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 derive some of the important exact properties of the effective electron-electron interaction (see Appendix B and Sec. II D 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 Sec. II C), 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 Sec. II D 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 IR^3. The computation of $\mu(\mathbf{r})$ nonetheless requires the computation of two-electron 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 Sec. II E 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 μ 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 IR^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 Sec. III). Using FCI wave functions (see Sec. 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 Sec. III B) which are known to be challenging for WFT methods, even at the 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 Sec. 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 ON A BASIS SET

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

$$\hat{W}_{\rm ee} = \frac{1}{2} \iiiint \mathrm{d}\mathbf{X}_1\, \mathrm{d}\mathbf{X}_2\, \mathrm{d}\mathbf{X}_3\, \mathrm{d}\mathbf{X}_4 \delta(\mathbf{X}_1 - \mathbf{X}_4)\ \delta(\mathbf{X}_2 - \mathbf{X}_3) \times \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \hat{\Psi}^\dagger(\mathbf{X}_4)\hat{\Psi}^\dagger(\mathbf{X}_3)\hat{\Psi}(\mathbf{X}_2)\hat{\Psi}(\mathbf{X}_1), \tag{A1}$$

where $\hat{\Psi}(\mathbf{X})$ and $\hat{\Psi}^\dagger(\mathbf{X})$ are the 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 as

$$\hat{W}^{\mathcal{B}}_{\rm ee} = \frac{1}{2} \sum_{ijkl\,\in\,\mathcal{B}} V^{kl}_{ij}\ \hat{a}^\dagger_k \hat{a}^\dagger_l \hat{a}_j \hat{a}_i, \tag{A2}$$

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

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

$$\hat{a}^\dagger_i = \int \mathrm{d}\mathbf{X}\ \phi_i(\mathbf{X})\ \hat{\Psi}^\dagger(\mathbf{X}). \tag{A4}$$

Therefore, by defining

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

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

$$\hat{W}^{\mathcal{B}}_{\rm ee} = \frac{1}{2} \iiiint \mathrm{d}\mathbf{X}_1\, \mathrm{d}\mathbf{X}_2\, \mathrm{d}\mathbf{X}_3\, \mathrm{d}\mathbf{X}_4\, w^{\mathcal{B}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) \times \hat{\Psi}^\dagger(\mathbf{X}_4)\hat{\Psi}^\dagger(\mathbf{X}_3)\hat{\Psi}(\mathbf{X}_2)\hat{\Psi}(\mathbf{X}_1). \tag{A6}$$

In the limit of a complete basis set (written as "$\mathcal{B} \to \infty$"), $\hat{W}^{\mathcal{B}}_{\rm ee}$ coincides with $\hat{W}_{\rm ee}$

$$\lim_{\mathcal{B}\to\infty} \hat{W}^{\mathcal{B}}_{\rm ee} = \hat{W}_{\rm ee}, \tag{A7}$$

which implies that

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

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

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

which, by introducing the two-body density matrix,

$$n^{(2)}_\Psi(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) = \langle\Psi|\hat{\Psi}^\dagger(\mathbf{X}_4)\hat{\Psi}^\dagger(\mathbf{X}_3)\hat{\Psi}(\mathbf{X}_2)\hat{\Psi}(\mathbf{X}_1)|\Psi\rangle, \tag{A10}$$

turns into

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

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

$$\langle\Psi^{\mathcal{B}}|\hat{W}^{\mathcal{B}}_{\rm ee}|\Psi^{\mathcal{B}}\rangle = \frac{1}{2} \iiiint \mathrm{d}\mathbf{X}_1\, \mathrm{d}\mathbf{X}_2\, \mathrm{d}\mathbf{X}_3\, \mathrm{d}\mathbf{X}_4\, w^{\mathcal{B}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) \times n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4), \tag{A12}$$

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

$$n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4) = \sum_{mnpq\,\in\,\mathcal{B}} \phi_p(\mathbf{X}_4)\phi_q(\mathbf{X}_3)\phi_n(\mathbf{X}_2)\phi_m(\mathbf{X}_1)\ \Gamma^{pq}_{mn}[\Psi^{\mathcal{B}}], \tag{A13}$$

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

$$\Gamma^{pq}_{mn}[\Psi^{\mathcal{B}}] = \langle\Psi^{\mathcal{B}}|\hat{a}^\dagger_p \hat{a}^\dagger_q \hat{a}_n \hat{a}_m|\Psi^{\mathcal{B}}\rangle. \tag{A14}$$

By integrating over $\mathbf{X}_3$ and $\mathbf{X}_4$ in Eq. (A12), we obtain

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

where we introduced the function

$$f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \sum_{ijklmn\,\in\,\mathcal{B}} V^{kl}_{ij}\ \Gamma^{mn}_{kl}[\Psi^{\mathcal{B}}]\phi_n(\mathbf{X}_2)\phi_m(\mathbf{X}_1)\phi_i(\mathbf{X}_1)\phi_j(\mathbf{X}_2). \tag{A16}$$

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

$$\langle\Psi^{\mathcal{B}}|\hat{W}^{\mathcal{B}}_{\rm ee}|\Psi^{\mathcal{B}}\rangle = \langle\Psi^{\mathcal{B}}|\hat{W}_{\rm ee}|\Psi^{\mathcal{B}}\rangle, \tag{A17}$$

which translates into

$$\frac{1}{2} \iint \mathrm{d}\mathbf{X}_1\, \mathrm{d}\mathbf{X}_2\, f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) = \frac{1}{2} \iint \mathrm{d}\mathbf{X}_1\, \mathrm{d}\mathbf{X}_2\, \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\, n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1, \mathbf{X}_2) \tag{A18}$$

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

$$W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = \frac{f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2)}{n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2)}, \tag{A19}$$

one can rewrite Eq. (A18) as

$$\iint \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2\; W_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2)\; n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = \iint \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2\; \frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\; n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2). \tag{A20}$$

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 needs to study only $f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2)$. By explicitiing the two-electron integrals, $f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2)$ can be written as

$$f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = \sum_{ijklmn\,\in\,\mathcal{B}} \Gamma^{mn}_{kl}[\Psi^{\mathcal{B}}]\phi_n(\mathbf{X}_2)\phi_m(\mathbf{X}_1)\phi_i(\mathbf{X}_1)\phi_j(\mathbf{X}_2) \times \iint \mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}'\phi_k(\mathbf{X})\phi_l(\mathbf{X}')\phi_i(\mathbf{X})\phi_j(\mathbf{X}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}, \tag{B1}$$

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

$$f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = \sum_{mnkl\,\in\,\mathcal{B}} \Gamma^{mn}_{kl}[\Psi^{\mathcal{B}}]\phi_n(\mathbf{X}_2)\phi_m(\mathbf{X}_1) \times \int \mathrm{d}\mathbf{X}\left(\sum_{i\,\in\,\mathcal{B}}\phi_i(\mathbf{X}_1)\phi_i(\mathbf{X})\right)\phi_k(\mathbf{X}) \times \int \mathrm{d}\mathbf{X}'\left(\sum_{j\,\in\,\mathcal{B}}\phi_j(\mathbf{X}_2)\phi_j(\mathbf{X}')\right)\phi_l(\mathbf{X}')\,\frac{1}{|\mathbf{r}-\mathbf{r}'|}. \tag{B2}$$

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

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

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 $\mathbb{R}$, $g^{\mathcal{B}}$ is its component in $\mathcal{B}$, and $g^{\perp}$ is the orthogonal component

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

then

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

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

$$\lim_{\mathcal{B}\to\infty}\int \mathrm{d}\mathbf{X}'\;\delta^{\mathcal{B}}(\mathbf{X}_2-\mathbf{X}')\phi_l(\mathbf{X}')\;\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \phi_l(\mathbf{X}_2)\;\frac{1}{|\mathbf{r}-\mathbf{r}_2|} \tag{B6}$$

and

$$\lim_{\mathcal{B}\to\infty}\int \mathrm{d}\mathbf{X}\;\delta^{\mathcal{B}}(\mathbf{X}_1-\mathbf{X})\phi_k(\mathbf{X})\int \mathrm{d}\mathbf{X}'\;\frac{\delta^{\mathcal{B}}(\mathbf{X}_2-\mathbf{X}')\phi_l(\mathbf{X}')}{|\mathbf{r}-\mathbf{r}'|} = \phi_k(\mathbf{X}_1)\;\phi_l(\mathbf{X}_2)\;\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}. \tag{B7}$$

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

$$\lim_{\mathcal{B}\to\infty} f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = \sum_{klmn\,\in\,\mathcal{B}}\Gamma^{mn}_{kl}[\Psi^{\mathcal{B}}]\phi_m(\mathbf{X}_1)\;\phi_n(\mathbf{X}_2) \times \phi_l(\mathbf{X}_2)\;\phi_k(\mathbf{X}_1)\;\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}, \tag{B8}$$

which is nothing but

$$\lim_{\mathcal{B}\to\infty} f_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2) = n^{(2)}_{\Psi^{\mathcal{B}}}(\mathbf{X}_1,\mathbf{X}_2)\;\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}. \tag{B9}$$

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 for all points $(\mathbf{X}_1,\mathbf{X}_2)$

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

[1] J. Goldstone, Proc. R. Soc. A **239**, 267 (1957).
[2] I. Lindgren, Phys. Scr. **32**, 291 (1985).
[3] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
[4] C. F. Bender and E. R. Davidson, Phys. Rev. **183**, 23 (1969).
[5] B. Huron, J. Malrieu, and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).
[6] R. J. Buenker and S. D. Peyerimholf, Theor. Chim. Acta **35**, 33 (1974).
[7] R. J. Buenker, S. D. Peyerimholf, and P. J. Bruna, *Computational Theoretical Organic Chemistry* (Reidel, Dordrecht, 1981), p. 55.
[8] S. Evangelisti, J.-P. Daudey, and J.-P. Malrieu, Chem. Phys. **75**, 91 (1983).
[9] R. J. Harrison, J. Chem. Phys. **94**, 5021 (1991).
[10] A. A. Holmes, N. M. Tubman, and C. J. Umrigar, J. Chem. Theory Comput. **12**, 3674 (2016).
[11] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).
[12] E. Hylleraas, Z. Phys. **54**, 347 (1929).
[13] C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, Chem. Rev. **112**, 4 (2012).
[14] L. Kong, F. A. Bischoff, and E. F. Valeev, Chem. Rev. **112**, 75 (2012).
[15] A. Grüneis, S. Hirata, Y.-Y. Ohnishi, and S. Ten-no, J. Chem. Phys. **146**, 080901 (2017).
[16] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[17] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[18] A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).
[19] L. Goerigk and S. Grimme, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 576 (2014).
[20] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
[21] J. Toulouse, F. Colonna, and A. Savin, Phys. Rev. A **70**, 062505 (2004).
[22] O. Franck, B. Mussard, E. Luppi, and J. Toulouse, J. Chem. Phys. **142**, 074107 (2015).
[23] J. G. Ángyán, I. C. Gerber, A. Savin, and J. Toulouse, Phys. Rev. A **72**, 012510 (2005).
[24] E. Goll, H.-J. Werner, and H. Stoll, Phys. Chem. Chem. Phys. **7**, 3917 (2005).
[25] J. Toulouse, I. C. Gerber, G. Jansen, A. Savin, and J. G. Ángyán, Phys. Rev. Lett. **102**, 096404 (2009).
[26] B. G. Janesko, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. **130**, 081105 (2009).
[27] T. Leininger, H. Stoll, H.-J. Werner, and A. Savin, Chem. Phys. Lett. **275**, 151 (1997).
[28] E. Fromager, J. Toulouse, and H. J. A. Jensen, J. Chem. Phys. **126**, 074111 (2007).
[29] E. Fromager, R. Cimiraglia, and H. J. A. Jensen, Phys. Rev. A **81**, 024502 (2010).
[30] E. D. Hedegård, S. Knecht, J. S. Kielberg, H. J. A. Jensen, and M. Reiher, J. Chem. Phys. **142**, 224108 (2015).
[31] L. Kronik, T. Stein, S. Refaely-Abramson, and R. Baer, J. Chem. Theory Comput. **8**, 1515 (2012).
[32] A. V. Krukau, G. E. Scuseria, J. P. Perdew, and A. Savin, J. Chem. Phys. **129**, 124103 (2008).

[33]T. M. Henderson, B. G. Janesko, G. E. Scuseria, and A. Savin, Int. J. Quantum Chem. **109**, 2023 (2009).
[34]J. Toulouse, P. Gori-Giorgi, and A. Savin, Theor. Chem. Acc. **114**, 305 (2005).
[35]J. C. Slater, Phys. Rev. **81**, 385 (1951).
[36]S. Paziani, S. Moroni, P. Gori-Giorgi, and G. B. Bachelet, Phys. Rev. B **73**, 155111 (2006).
[37]J. Rubio, J. Novoa, and F. Illas, Chem. Phys. Lett. **126**, 98 (1986).
[38]R. Cimiraglia and M. Persico, J. Comput. Chem. **8**, 39 (1987).
[39]C. Angeli and M. Persico, Theor. Chem. Acc. **98**, 117 (1997).
[40]C. Angeli, R. Cimiraglia, and J.-P. Malrieu, Chem. Phys. Lett. **317**, 472 (2000).
[41]E. Giner, A. Scemama, and M. Caffarel, Can. J. Chem. **91**, 879 (2013).
[42]A. Scemama, T. Applencourt, E. Giner, and M. Caffarel, J. Chem. Phys. **141**, 244110 (2014).
[43]E. Giner, A. Scemama, and M. Caffarel, J. Chem. Phys. **142**, 044115 (2015).
[44]E. Giner, R. Assaraf, and J. Toulouse, Mol. Phys. **114**, 910 (2016).
[45]P. S. Epstein, Phys. Rev. **28**, 695 (1926).
[46]R. K. Nesbet, Proc. R. Soc. A **230**, 312 (1955).
[47]Y. Garniron, A. Scemama, P.-F. Loos, and M. Caffarel, J. Chem. Phys. **147**, 034101 (2017).
[48]A. Scemama, T. Applencourt, Y. Garniron, E. Giner, G. David, and M. Caffarel, Quantum package V1.0, 2016.
[49]G. H. Booth and A. Alavi, J. Chem. Phys. **132**, 174104 (2010).
[50]S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia, and C. Froese Fischer, Phys. Rev. A **47**, 3649 (1993).