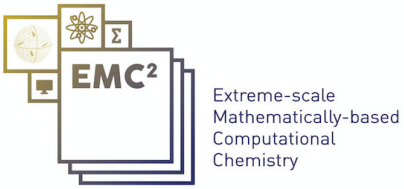

Sorbonne Université

École doctorale 388 : Chimie Physique et Chimie Analytique de Paris-Centre

Laboratoire de Chimie Théorique

The extension of the density-based basis-set correction method for calculations of molecular properties

par Diata Traore

Thèse de doctorat de Chimie Théorique

Dirigée par Julien Toulouse et Emmanuel Giner

Présentée et soutenue publiquement le 07 juillet 2023

Devant le jury composé de :

Delphine Cabaret,	Professeure, Sorbonne Université	Présidente du jury
Emmanuel Fromager,	Professeur, Université de Strasbourg	Rapporteur
Andreas Köhn,	Professeur, University of Stuttgart	Rapporteur
Pierre-François Loos,	Directeur de recherche, CNRS/Université Paul Sabatier	Examinateur
Pina Romaniello,	Chargée de recherche, CNRS/Université Paul Sabatier	Examinatrice
Julien Toulouse,	Maître de conférences, Sorbonne Université	Directeur de thèse
Emmanuel Giner,	Chargé de recherche, CNRS/Sorbonne Université	Co-encadrant

Résumé

La chimie quantique est la branche des sciences de la matière qui vise à améliorer la compréhension et la description des systèmes électroniques. Cet objectif encourage le développement de méthodes pour des études qualitatives et quantitatives adaptées à tel ou tel système (de l'atome à la protéine en passant par les cristaux) ou problème (des calculs d'énergies fondamentales aux propriétés spectroscopiques en passant par des calculs de dynamique). Ces méthodes donnent accès à une meilleure compréhension mais aussi des prédictions des résultats expérimentaux et des propriétés moléculaires.

Plus spécifiquement, cette thèse résume la poursuite du développement d'une méthode de calcul d'énergies et de propriétés moléculaires : la méthode de correction de base par la densité électronique[1] (voir les références [1–7] pour les travaux précédant cette thèse). Ce développement est motivé par le problème des calculs de structures électroniques et plus spécifiquement par celui des coûts de calcul importants qui limitent l'application des méthodes existantes à des systèmes de tailles limitées.

L'équation de Schrödinger

Pour quantifier les problèmes de structure électronique, la stratégie privilégiée est la résolution de l'équation de Schrödinger. Cette équation a valu au scientifique éponyme un prix Nobel de physique en 1933. Ce dernier appliqua pour la première fois son équation au problème des niveaux d'énergie de l'atome d'hydrogène en 1926 [8].

L'équation de Schrödinger fait intervenir une fonction d'onde qui dépend des coordonnées spatiales et de spin des électrons constituant le système et qui permet la description de structures électroniques à travers une liste de paramètres que le quanticien cherche à optimiser [9–11]. Dans ce contexte, la paramétrisation de la fonction d'onde peut passer par de nombreuses stratégies et l'approche privilégiée dans ce travail de thèse est la théorie des orbitales moléculaires où les électrons sont délocalisés sur la molécule par l'utilisation d'orbitales moléculaires. Dans l'approximation la plus simple[2], les paramètres du système sont les poids attribués à chaque orbitale atomique pour chacun des états monoélectroniques du système. Ces poids peuvent alors être lus comme des probabilités d'occupation des orbitales atomiques.

Dans l'approximation de Born-Oppenheimer[3], l'équation de Schrödinger indépendante du temps est définie comme

$$\hat{H}\Psi(\mathbf{x}) = E\Psi(\mathbf{x}), \tag{1}$$

où $\hat{H}$ est l'opérateur hamiltonien électronique tel que $\hat{H} = \hat{T} + \hat{W}_{\text{ee}} + \hat{V}_{\text{ne}} + \hat{V}_{\text{nn}}$ comprenant l'opérateur d'énergie cinétique des électrons et les interactions coulombiennes électron-électron, noyaux-électrons et noyaux-noyaux, $\Psi(\mathbf{x})$ est la fonction d'onde dépendant des coordonnées spatiales et de spin des N électrons du système, $\mathbf{x} = (x_1, x_2, ..., x_N)$, et E est l'énergie associée à

[1] Dans ce résumé, nous nous y référerons sous l'appellation «méthode de correction de base ».

[2] Hartree-Fock, en très simplifié.

[3] ou l'approximation des noyaux fixes.

la fonction d'onde $\Psi(\mathbf{x})$. Les opérateurs constituant l'hamiltonien et la définition de la fonction d'onde sont détaillées dans le corps de la thèse.

Les approximations de l'équation de Schrödinger : définition de l'énergie de corrélation

Une fois décidé l'objet de base pour la description d'un unique électron, l'orbitale moléculaire, le quanticien doit décrire la structure des N électrons du système à travers une unique fonction d'onde. Cette fonction d'onde doit encoder les propriétés des électrons dans un milieu électronique et ses interactions avec ce dernier. Plus concrètement, il doit prendre en compte le caractère fermionique des électrons, le principe d'exclusion de Pauli et toute autre forme de corrélation entre les électrons du système[4].

Une première approche est de considérer seulement l'aspect fermionique des électrons, c'est à dire le principe d'exclusion de Pauli. C'est l'approximation de Hartree-Fock (HF) qui permet cette description par l'utilisation d'un unique déterminant de Slater pour exprimer la fonction d'onde à N électrons où chaque orbitale moléculaire $\chi_i(\mathbf{x}_n)$ (i est le numéro de l'orbitale moléculaire et n est le $n$ème électron) est, au plus, occupée par un seul électron. Finalement, la fonction d'onde Hartree-Fock s'exprime comme

$$\Phi_{\mathrm{HF}}(\mathbf{x}) = \frac{1}{\sqrt{N}!}\begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & . & . & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & . & . & \chi_N(\mathbf{x}_2) \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & . & . & \chi_N(\mathbf{x}_N) \end{vmatrix}. \tag{2}$$

Dans cette approximation, l'équation de Schrödinger (1) peut être décomposée en N équations couplées où l'interaction entre électrons est traitée comme un champ moyen. Avec cette approximation du champ moyen, il est possible d'obtenir une bonne description qualitative de l'état fondamental où, *in fine*, les électrons occupent les orbitales moléculaires de plus basses énergies comme illustré sur la Fig. 1.

Cependant, pour une description quantitative exploitable, notamment pour des comparaisons avec les résultats expérimentaux, il est nécessaire d'aller au delà de l'approximation du champ moyen, en utilisant les méthodes post-Hartree-Fock où la fonction d'onde est généralement exprimée comme une combinaison linéaire de déterminants de Slater. Par exemple, en utilisant une approche d'interaction de configuration (CI), la fonction d'onde a pour expression

$$|\Psi_{\mathrm{CI}}\rangle = c_0|\Phi_{\mathrm{HF}}\rangle + \sum_{i}^{\text{déterminants}} c_i|\Phi_i\rangle \tag{3}$$

où $|\Phi_{\mathrm{HF}}\rangle$ est la fonction d'onde Hartree-Fock et $|\Phi_i\rangle$ sont des déterminants construits à partir du déterminant Hartree-Fock. Ces déterminants sont schématisés sur la Fig. 1. Les paramètres du système à optimiser deviennent alors les coefficients du développement CI.

[4]Pour les habitués au formalisme, nous regroupons dans l'expression «toute autre forme de corrélation», la corrélation qui vient du terme de Coulomb. Ce terme diverge lorsque les deux électrons considérés sont proches ce qui complique les calculs à courte portée. C'est pourquoi l'approximation de Hartree-Fock, où cette interaction est traitée par un champ moyen, est un point de départ naturel. Pour les non habitués au formalisme, cette notion est explicitée dans la thèse mais n'est pas nécessaire pour la compréhension de ce résumé.

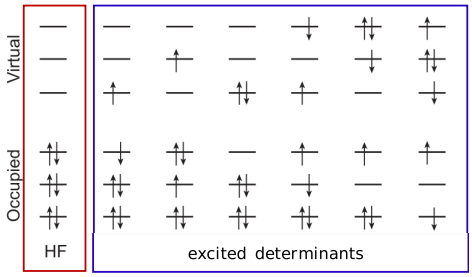

Figure 1: Schéma des états du système décrits par des déterminants de Slater. Chaque ligne représente une orbitale moléculaire. HF: déterminant solution de l'équation de Hartree-Fock; Excited determinants: déterminants construits à partir du déterminant Hartree-Fock où les électrons sont excités dans les orbitales virtuelles (inoccupées dans le déterminant Hartree-Fock) (Figure inspirée de la Réf. [12]).

En terme d'énergie, la différence observée avec l'énergie Hartree-Fock est appelée énergie de corrélation E_c, tel que

$$E_c = E_{PHF} - E_{HF}, \tag{4}$$

où E_{PHF} est l'énergie obtenue à partir d'une approximation post-Hartree-Fock pour la fonction d'onde et E_{HF} est l'énergie obtenue dans l'approximation Hartree-Fock. Parmi les méthodes post-Hartree-Fock, nous décrivons dans ce manuscrit les méthodes d'interaction de configuration et la méthode coupled-cluster.

La théorie de la fonctionnelle de la densité

La théorie de la fonctionnelle de la densité est un autre moyen d'aborder le problème de la structure électronique [13–15]. Son intérêt est qu'on y optimise la densité d'électrons, formalisée par une fonction à 3 coordonnées spatiales, au lieu de la fonction d'onde à $4N$ coordonnées d'espace et de spin. Cette réduction importante du nombre de degrés de liberté permet d'étendre le domaine d'application de la chimie quantique à des systèmes plus grands.

Il existe également des méthodes hybrides combinant les méthodes centrées sur la forme de la fonction d'onde et les approximations de la fonctionnelle de la densité. Deux d'entre elles sont au cœur du présent travail de thèse : la théorie de la fonctionnelle de la densité à séparation de portée [16] et la méthode de correction de base. Nous explicitons leurs formalismes mathématiques dans la suite de ce résumé.

La théorie de la fonctionnelle de la densité à séparation de portée

Qualitativement, la théorie de la fonctionnelle de la densité à séparation de portée divise l'espace des $3N$ coordonnées spatiales en deux domaines :

- Le domaine dit de «courte portée» correspond aux régions de l'espace où deux électrons sont proches. Cette région de l'espace nécessite une base de fonctions particulièrement grande (pour des coûts de calculs d'autant plus conséquents) pour donner suffisamment de flexibilité à la description des termes à deux électrons, notamment l'énergie de corrélation électronique ou la densité de paire, objets que nous définissons dans le corps de la présente thèse. Dans cette méthode, cette région de l'espace est décrite à l'aide d'une fonctionnelle de la densité, bénéficiant de sa représentation compacte de la corrélation électronique.

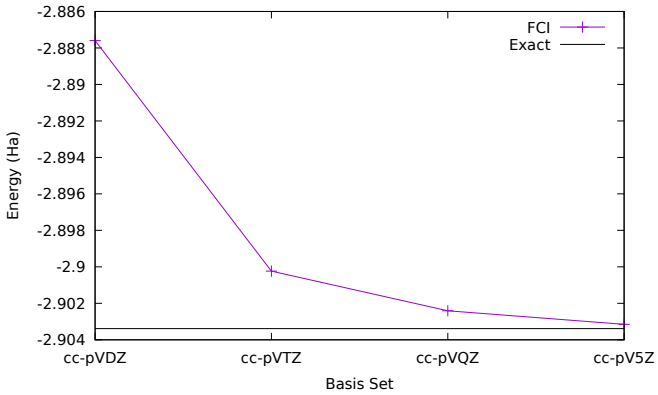

Figure 2: Convergence en base de l'énergie (en Hartree) de l'état fondamental de l'atome d'hélium. FCI est la méthode CI où l'on considère tous les déterminants accessibles à partir du déterminant Hartree Fock. cc-pVXZ sont les noms des bases de fonction classées par ordre croissant de nombre de fonctions de base les constituant.

- Le domaine dit «de longue portée» correspond au reste de l'espace. Les méthodes de fonctions d'onde permettent, en général, une meilleure description de l'interaction entre électrons dans ce domaine et sont donc utilisées pour cette partie de l'espace.

En pratique, des approximations pour la fonctionnelle de la densité ont été spécifiquement proposées pour le terme de courte portée qui ont permis d'appliquer la théorie de la fonctionnelle de la densité à séparation de portée à un large éventail de systèmes.

La méthode de correction de base par la densité électronique

L'objectif de la méthode de correction de base est d'améliorer la lente convergence en base des méthodes de fonction d'onde post-Hartree-Fock. En effet, pour ces méthodes, les énergies et propriétés moléculaires convergent lentement, bien que systématiquement, avec la taille de la base de fonction utilisée. Le caractère systématique de cette convergence (illustré sur la Fig. 2) permet d'estimer une extrapolation des quantités dans la limite de la base complète mais cette extrapolation est très rapidement limitée par le nombre de fonctions de base nécessaires qui augmente considérablement avec la taille du système.

Avec la méthode de correction de base, nous proposons une procédure permettant d'estimer l'erreur induite par le choix de base et cette erreur est estimée à partir d'une fonctionnelle de la densité.

Jusqu'à présent, le choix de la fonctionnelle de la densité repose sur la considération que la quantité limitante dans la convergence en base est l'énergie de corrélation à courte portée. Ce qui rappelle la philosophie de la théorie de la fonctionnelle de la densité à séparation de portée. C'est pourquoi nous approximons la fonctionnelle de la densité de la méthode de correction de base par des variantes de celle développée pour la théorie de la fonctionnelle de la densité à séparation de portée. Elles sont décrites qualitativement dans le tableau 1 et leurs expressions sont données dans le corps de la thèse.

C'est dans ce contexte que s'inscrit le début du présent travail de doctorat dans lequel nous implémentons et testons différentes approximations de cette fonctionnelle de la densité pour les calculs d'énergies d'états fondamentaux et d'états excités mais aussi de moments dipolaires.

Table 1: Description des fonctionnelles de corrélation de courte portée utilisées pour la fonctionnelle complémentaire de correction de base.

Fonctionnelle	Caractéristiques
$\bar{E}^{\mathrm{sr},\mu}_{\mathrm{c,LDA}}[n]$	Basée sur l'approximation de la densité locale (LDA), cette fonctionnelle de la densité pour la corrélation (c) de courte portée (sr) ne dépend que de la densité en chaque point de l'espace.
$\bar{E}^{\mathrm{sr},\mu}_{\mathrm{c,PBE\text{-}OT}}[n, n_2]$	Basée sur la variante PBE (du nom de ses créateurs Perdew, Burke et Ernzerhof) de l'approximation du gradient généralisé (GGA), cette fonctionnelle dépend de la densité et de son gradient en chaque point de l'espace mais aussi de la densité de paire à coalescence (OT) $n_2(\mathbf{r},\mathbf{r})$ (voir définition dans la thèse). Ajouter la dépendance en la densité de paire à coalescence permet d'ajouter de l'information sur l'interaction entre les électrons en chaque point de l'espace.
$\bar{E}^{\mathrm{sr},\mu}_{\mathrm{c,PBE\text{-}UEG}}[n]$	Cette variante de la fonctionnelle PBE-OT approxime la densité de paire à coalescence en chaque point de l'espace par celle du gaz d'électron uniforme de densité $n(\mathbf{r})$.

Correction de base par la densité par une équation auto-cohérente

Ce chapitre présente le travail publié dans la référence [E. Giner, D. Traore, B. Pradines, J. Toulouse, J. Chem. Phys. **155**, 204104 1-10 (2021)]. On y explicite l'algorithme de correction de base auto-cohérent.

Dans les travaux précédant cet article, la correction de base de l'énergie était calculée à posteriori à partir de la densité de la fonction d'onde:

$$E_0^{\mathcal{B}} = \langle \Psi^{\mathcal{B}} | \hat{H} | \Psi^{\mathcal{B}} \rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}], \tag{5}$$

où $\bar{E}^{\mathcal{B}}[n]$ est la fonctionnelle de la densité pour la correction de base (appelée fonctionnelle de correction de base dans la suite de ce résumé) et $\Psi^{\mathcal{B}}$ est la fonction d'onde dans la base $\mathcal{B}$.

Considérons une approche de fonctions d'onde fondée sur une minimisation traduite en terme d'équation de Lagrange, par exemple une approche d'interaction de configuration. L'algorithme de correction de base auto-cohérente traite la minimisation qui définie la méthode de correction de base en incluant la dérivée de la fonctionnelle de la densité dans l'équation d'Euler-Lagrange. La minimisation de l'énergie dans la base,

$$E_0^{\mathcal{B}} = \min_{\Psi^{\mathcal{B}}} \left\{ \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\mathrm{ee}} + \hat{V}_{\mathrm{ne}} | \Psi^{\mathcal{B}} \rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}] \right\}, \tag{6}$$

conduit alors à

$$\hat{H}^{\mathcal{B}}_{\mathrm{eff}}[n_{\Psi^{\mathcal{B}}_0}] | \Psi^{\mathcal{B}}_0 \rangle = \mathcal{E}^{\mathcal{B}}_0 | \Psi^{\mathcal{B}}_0 \rangle, \tag{7}$$

où $\mathcal{E}^{\mathcal{B}}_0$ est le multiplicateur de Lagrange associé à la contrainte de normalisation de la fonction d'onde $\Psi^{\mathcal{B}}_0$ et l'hamiltonien effectif est défini comme

$$\hat{H}^{\mathcal{B}}_{\mathrm{eff}}[n] = \hat{T}^{\mathcal{B}} + \hat{W}^{\mathcal{B}}_{\mathrm{ee}} + \hat{V}^{\mathcal{B}}_{\mathrm{ne}} + \hat{\bar{V}}^{\mathcal{B}}[n], \tag{8}$$

avec $\hat{\bar{V}}^{\mathcal{B}}[n]$ le potentiel effectif à 1 électron tel que

$$\hat{\bar{V}}^{\mathcal{B}}[n] = \int \mathrm{d}\mathbf{r}\ \bar{v}^{\mathcal{B}}[n](\mathbf{r})\ \hat{n}^{\mathcal{B}}(\mathbf{r}), \tag{9}$$

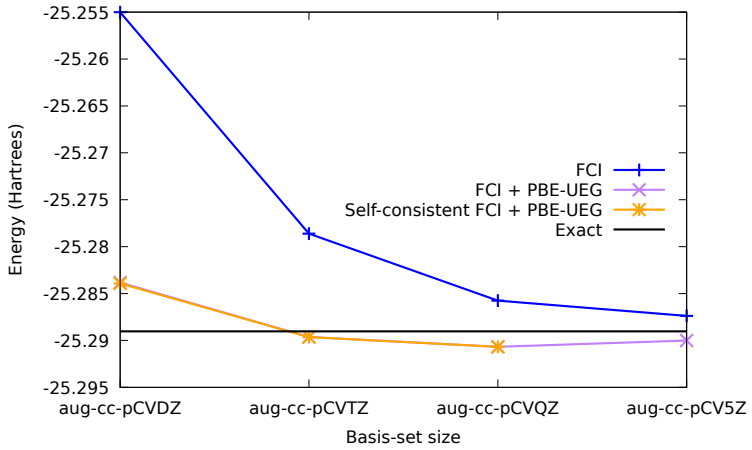

Figure 3: Convergence en base de l'énergie de la molécule BH. FCI est la méthode CI où l'on considère tous les déterminants accessibles à partir du déterminant Hartree Fock. FCI + PBE-UEG correspond à l'énergie obtenue avec une correction a posteriori. Self-consistent FCI + PBE-UEG correspond à l'énergie obtenue à partir de l'algorithme auto-cohérent. aug-cc-pVXZ sont les noms des bases de fonction classées par ordre croissant du nombre de fonctions de base constituant la base.

où $\bar{v}^{\mathcal{B}}[n](\mathbf{r}) = \delta\bar{E}^{\mathcal{B}}[n]/\delta n(\mathbf{r})$ et $\hat{n}^{\mathcal{B}}(\mathbf{r})$ est l'opérateur densité projeté sur la base $\mathcal{B}$.

En pratique, $\bar{E}^{\mathcal{B}}[n]$ et sa dérivée $\bar{v}^{\mathcal{B}}[n]$ sont approximées grâce à une fonctionnelle de la densité issue de la théorie de la fonctionnelle de la densité à séparation de portée.

De cette manière, le potentiel de correction de base intervient dans les itérations du calcul de fonction d'onde et des tests systématiques permettent d'éclaircir les résultats précédemment publiés sur l'approximation non-auto-cohérente (Eq. (5)). En effet, dans cet article nous montrons que les algorithmes auto-cohérent et non auto-cohérent donnent des résultats similaires pour les calculs d'énergie de l'état fondamental. Nous donnons un exemple dans la Fig. 3 pour la molécule BH où l'on voit que les courbes de convergence en base des énergies calculées avec les deux approches se superposent. Ainsi, nous illustrons la validité de l'approximation non auto-cohérente pour les calculs de cette quantité.

Dans cet article, nous présentons également nos travaux sur les calculs de moments dipolaires. Le moment dipolaire est une quantité plus sensible au niveau de précision de la densité que l'énergie de l'état fondamental et constitue un premier cadre simple pour comprendre l'effet des fonctionnelles de correction de base et leurs potentiels sur la convergence en base de la densité. Les résultats montrent des convergences en base nettement améliorées en présence du potentiel de correction de base (voir Fig. 4), encourageant ainsi l'extension de la méthode de correction de base aux propriétés moléculaires.

Correction de la base pour l'estimation de moments dipolaires à partir de la méthode *coupled-cluster*

Alors que dans l'article précédent nous avons appliqué la correction de base auto-cohérente aux calculs de moments dipolaires avec des approches variationnelles de type CI, nous propopons ici d'étendre le calcul de propriétés moléculaires aux approches non-variationnelles telles que le coupled cluster. Dans ce cadre, la formule du moment dipolaire incluant la correction de base est la suivante

$$d_0^{\mathcal{B}} = d^{\mathcal{B}} + \bar{d}^{\mathcal{B}}, \tag{10}$$

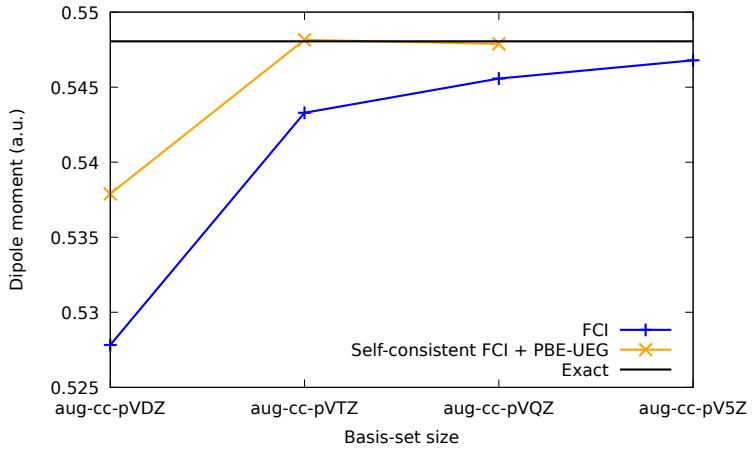

Figure 4: Convergence en base du moment dipolaire de la molécule BH. FCI est la méthode CI où l'on considère tous les déterminants accessibles à partir du déterminant Hartree Fock. Self-consistent FCI+PBE-UEG correspond au moment dipolaire obtenu à partir de l'algorithme auto-cohérent. aug-cc-pVXZ sont les noms des bases de fonction classés par ordre croissant du nombre de fonctions de base constituant la base.

où $d^{\mathcal{B}}$ est le moment dipolaire calculé à partir d'une méthode de fonctions d'onde avec une base $\mathcal{B}$ de fonctions et $\bar{d}^{\mathcal{B}}$ est le terme de correction. Dans l'article, nous expliquons comment ce dernier terme peut être calculé, indépendemment du premier, à partir de la densité Hartree-Fock alors que le premier terme peut être calculé à partir de n'importe quelle méthode corrélée, et plus particulièrement la méthode coupled-cluster. L'équation 10 peut être obtenue en considérant le moment dipolaire comme une dérivée de l'énergie définie comme 5 par rapport à un champ électrique extérieur. On peut donc approximer chaque terme de l'équation 10 par un calcul de différence d'énergies.

Les tests sur 14 molécules sont présentés dans la partie résultats de l'article (et sur la Fig. 5 de ce résumé) et montrent des résultats aussi encourageants que ceux présentés dans l'article précédent. La correction de base peut donc être calculée indépendamment du calcul de fonction d'onde dans la base. De plus, l'utilisation d'une fonction d'onde Hartree-Fock pour calculer la correction de base est suffisante pour améliorer significativement les résultats. Dans ce schéma, le calcul de la correction de base reste très marginal par rapport à celui du calcul de fonction d'onde, ce qui est très avantageux.

Suite à cette étude, nous nous sommes naturellement demandé comment étendre la méthode de correction de base à des propriétés plus générales que le moment dipolaire. En commençant par les propriétés découlant de l'équation de la réponse linéaire, nous devons faire intervenir la dérivée seconde de l'énergie pour laquelle la qualité de la fonctionnelle de correlation de courte portée n'est pas encore connue.

A ce stade, nous réfléchissions à chercher une forme de fonctionnelle de correction de base plus générale que la fonctionnelle de corrélation de courte portée et complètement (ou en tout cas plus) spécifique au problème de convergence en base (voir le chapitre 4). Cette problématique motiva le développement d'un modèle à une dimension proposant un cadre simple et contrôlé pour le développement de cette fonctionnelle mais aussi pour la compréhension générale de la méthode de correction de base.

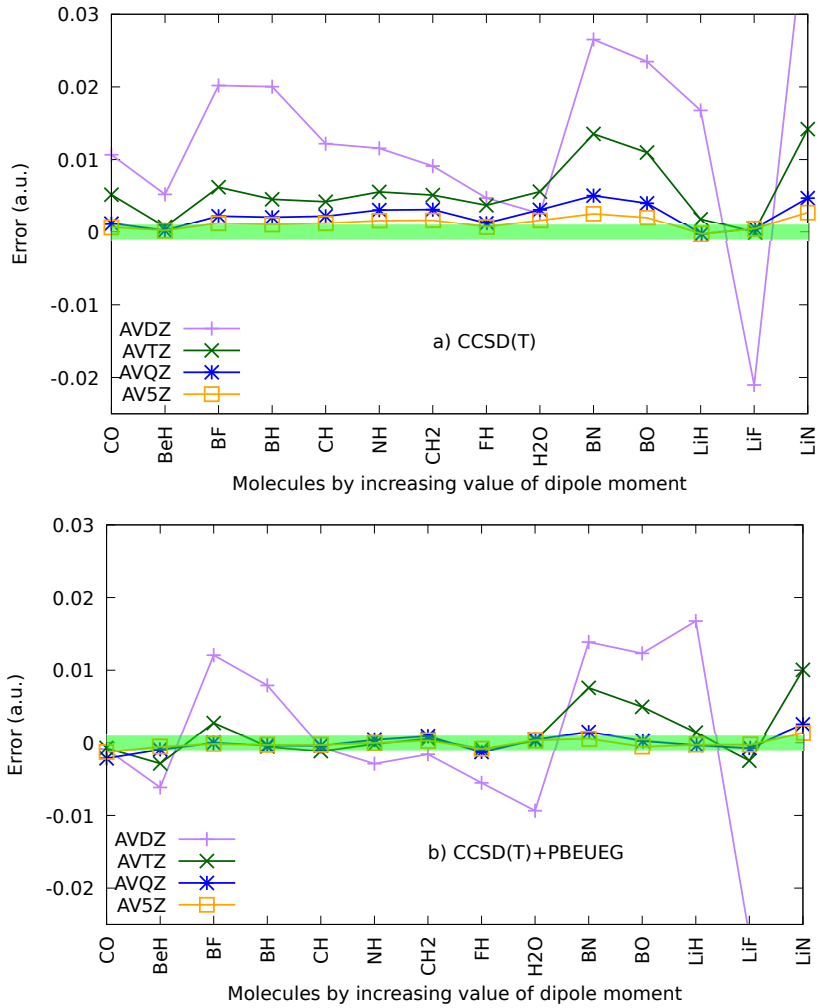

Figure 5: (a) Erreurs absolues des moment dipolaires calculés à partir d'énergies a) CCSD(T) et (b) CCSD(T)+PBEUEG pour 14 molécules. Les moments dipolaires sont comparés aux extrapolations en base dans la limite de la base complète des valeurs CCSD(T). La zone en vert indique une erreur de ±0.001 unités atomiques.

Un modèle à une dimension : un cadre rigoureux pour le développement de la méthode de correction de base

Ce chapitre présente le travail publié dans la référence [D. Traore, E. Giner, J. Toulouse, J. Chem. Phys **156**, 044113 1-13 (2022)]. On y explicite le développement d'un modèle à 2 électrons dans un espace à 1 dimension pour l'étude de la méthode de correction de base.

Comme spécifié dans le paragraphe précédent, l'extension de notre méthode aux propriétés électroniques nous a conduit au besoin de comprendre plus précisément l'effet de la correction de base et des approximations choisies pour son application aux systèmes réels. Dans le domaine du développement de méthodes, le quanticien est souvent amené à travailler avec des modèles simplifiés où il se concentrera spécifiquement sur le problème qui l'intéresse. Notre problème étant la convergence lente en base, nous étudions ici un système modèle qui reproduit cette problématique. Nous résumons les caractéristiques de ce modèle dans la liste qui suit.

L'hamiltonien

Après une étude bibliographique, nous avons décidé d'utiliser un potentiel δ-Dirac entre les électrons. Celui-ci permet de reproduire la convergence lente en base rencontrée en 3D. Dans cet article, l'hamiltonien a alors pour forme:

$$\hat{H} = \hat{T} + \hat{W}_{\text{ee}} + \hat{V}_{\text{ne}}, \tag{11}$$

avec

$$\hat{T} = -\frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2}{\partial x_i^2},\ \hat{W}_{\text{ee}} = \delta(x_1 - x_2),\ \hat{V}_{\text{ne}} = -Z\sum_{i=1}^{N}\delta(x_i)\ \ , \tag{12}$$

l'opérateur d'énergie cinétique, l'interaction entre électrons et l'interaction entre les électrons et le noyau, respectivement.

La base de fonctions d'onde

Nous avons testé différents types de fonctions de base (bases de Slater, de Gaussienne, ondes planes, ...) et avons décidé d'utiliser une base de fonctions de Hermite-Gauss. Elles ont l'avantage d'avoir des propriétés mathématiques intéressantes qui nous permettent de limiter nos temps de calculs et de ressembler aux genres de fonctions de bases que l'on utilise pour les systèmes à 3D en plus de reproduire une convergence systématique de l'énergie avec le nombre de fonctions de Hermite-Gauss utilisées.

Les méthodes de fonction d'onde disponibles

Le code réalisé permet actuellement de lancer des calculs Hartree-Fock, de théorie des perturbations (Møller-Plesset d'ordre 2) et d'utiliser la fonctionnelle de correction de base construite à partir de gaz d'électrons uniformes à 1D et de l'approximation locale de la densité (LDA). Pour avoir plus de détails sur le formalisme mathématique de cette correction de base, nous invitons le lecteur à lire le chapitre 4. Pour ce résumé, il est surtout important de préciser que la fonctionnelle développée ne dépend plus de la théorie de la fonctionnelle de la densité à séparation de portée.

Ce travail constitue un premier pas vers l'objectif d'une nouvelle fonctionnelle de correction de base pour les calculs en 3D. L'extension du formalisme développé dans cet article pour les systèmes 3D nécessitera de comprendre les propriétés des gaz d'électrons en 3D associés à une base fixée et de construire un algorithme adaptable aux bases et systèmes d'intérêt.

Extension de la méthode de correction de base à la théorie de la réponse linéaire

Dans l'optique d'étendre la méthode de correction de base aux propriétés moléculaires, nous souhaitons développer une méthode rigoureuse pour les calculs de ces quantités. En chimie quantique, l'un des formalismes à notre disposition est l'équation de réponse linéaire qui relie les paramètres variationnels du système à sa réponse à une perturbation $\hat{V}(t)$ dépendante du temps. Les fréquences $\omega_n^{\mathcal{B}}$ sont les quantités physiques les plus importantes obtenues par la résolution des équations de la réponse, car celles-ci peuvent être identifiée aux énergies d'excitation.

L'équation de réponse linéaire a pour expression,

$$\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}^* & \mathbf{A}^*\end{pmatrix}\begin{pmatrix}\mathbf{X}_n\\ \mathbf{Y}_n\end{pmatrix} = \omega_n^{\mathcal{B}}\begin{pmatrix}\mathbf{S} & \mathbf{0}\\ \mathbf{0} & -\mathbf{S}\end{pmatrix}\begin{pmatrix}\mathbf{X}_n\\ \mathbf{Y}_n\end{pmatrix}, \tag{13}$$

où $\mathbf{X}_n$ et $\mathbf{Y}_n$ sont les vecteurs propres, $\mathbf{S}$ est la matrice de recouvrement, et $\mathbf{A}$ et $\mathbf{B}$ sont des matrices dépendant de l'hamiltonien effectif mais aussi de la dérivée seconde de la fonctionnelle de correction de base. Cette dernière quantité n'a pas encore été implémentée et testée dans les précédentes études et nécessite une étude rigoureuse dans le cadre de l'utilisation de la fonctionnelle de corrélation de courte portée.

Dans l'article correspondant au chapitre 5, l'équation de réponse linéaire en présence de l'hamiltonien effectif a été testée sur un modèle à 1D avec potentiel harmonique et interaction δ entre électrons. Nous y montrons que les énergies d'excitation sont légèrement dégradées en comparaison avec l'utilisation d'un hamiltonien classique mais que les énergies totales des états excités, $E_n^{\mathcal{B}} = E_0^{\mathcal{B}} + \omega_n^{\mathcal{B}}$, sont nettement améliorées. Des tests préliminaires sur le béryllium nous ont montré les mêmes conclusions.

Si nous parvenons à comprendre l'effet de la correction de base sur les énergies d'excitation, l'étape suivante serait de calculer des propriétés moléculaires dépendant grandement de l'énergie de corrélation.

Conclusion générale

Cette thèse présente l'état du développement de la méthode de correction de base après trois années de doctorat. Après l'implémentation dans Quantum Package [17] et le test de l'équation auto-cohérente pour la correction de base, nous nous sommes intéressés à l'extension de la méthode aux propriétés moléculaires.

Nous nous sommes finalement focalisés sur la construction d'un modèle simplifié pour la compréhension des propriétés mathématiques de l'équation de correction de base et de la fonctionnelle complémentaire de correction de base.

Ce modèle devrait permettre d'entamer le développement d'une fonctionnelle complémentaire spécifique au problème de correction de base et d'éviter (ou confirmer) l'utilisation de la fonctionnelle de corrélation de courte portée développée pour la théorie de la fonctionnelle de la densité à séparation de portée. De plus, dans l'optique d'étendre notre domaine d'étude aux propriétés moléculaires, nous utilisons ce modèle pour l'extension de notre nouvelle méthode à la théorie de la réponse linéaire avant son application aux systèmes en 3D.

Bibliography

[1] E. Giner, B. Pradines, A. Ferté, R. Assaraf, A. Savin and J. Toulouse, J. Chem. Phys. **149**, 194301 (2018).

[2] A. Ferté, E. Giner and J. Toulouse, J. Chem. Phys. **150**, 084103 (2019).

[3] E. Giner, A. Scemama, J. Toulouse and P.-F. Loos, J. Chem. Phys. **151**, 144118 (2019).

[4] P.-F. Loos, B. Pradines, A. Scemama, J. Toulouse and E. Giner, J. Phys. Chem. Lett. **10**, 2931 (2019).

[5] E. Giner, A. Scemama, P.-F. Loos and J. Toulouse, J. Chem. Phys. **152**, 174104 (2020).

[6] P.-F. Loos, B. Pradines, A. Scemama, E. Giner and J. Toulouse, J. Chem. Theory Comput. **16**, 1018 (2020).

[7] Y. Yao, E. Giner, T. A. Anderson, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **155**, 204104 (2021).

[8] E. Schrödinger, Phys. Rev. **28**, 1049 (1926).

[9] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, Inc., 1996).

[10] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic Structure Theory* (John Wiley & Sons, LTD, 2000).

[11] F. Jensen, *Introduction to computational chemistry* (John Wiley & Sons, LTD, 2017).

[12] V. Gupta, *Principles and Applications of Quantum Chemistry* (Elsevier, Inc., 2015).

[13] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[14] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[15] J. Toulouse, Review of approximations for the exchange-correlation energy in density-functional theory, `https://arxiv.org/abs/2103.02645` (2021).

[16] J. Toulouse, F. Colonna and A. Savin, Phys. Rev. A **70**, 062505 (2004).

[17] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Véril, M. Caffarel, P.-F. Loos, E. Giner and A. Scemama, J. Chem. Theory. Comput. **15**, 3591 (2019).

Abstract

This thesis summarizes the development of a method for calculating energies and molecular properties: the density-based basis-set correction method. This development is motivated by the problem of electronic-structure calculations and more specifically by that of the large calculation costs which limit the application of standard methods to systems of limited size. With the basis-set correction method, we propose to estimate the finite basis-set error using functional of the density. The work presented in this manuscript begins with the extension of the basis-set correction method to a self-consistent formalism, allowing us to add the effect of the basis-set correction on the optimization of the wave function. We then apply the basis-set correction method to dipole moment calculations by showing how to apply the basis-set correction to coupled-cluster calculations. We also propose a one-dimensional model system for the development and understanding of the basis-set correction method and the approximations chosen for the basis-set correction functional. Finally, we use this model system for extending the basis-set correction method to linear-response calculations.

Contents

Introduction

Quantum chemistry aims to develop methods to describe the electronic structure of molecular systems. These methods permit predictions and deeper understanding of experiments and of molecular properties. At the heart of the quantum chemistry problem lies the Schrödinger equation for electronic systems [1, 2] where the Coulomb repulsion between electrons forbids finding analytical expressions for the solutions. The more problematic region of the space of electron coordinates to be described is when two electrons are close to each other because of the strong Coulomb repulsion: these are the so-called short-range correlation effects. Indeed, the wave-function parametrization needs to be flexible enough to describe the depletion of the wave function in these regions, and more specifically the so-called electron-electron cusp [3].

In practice, the wave function is expanded on a finite basis set and one can distinguish between two types of approximations: i) the error due to the use of a finite basis set, and ii) the error due to the wave-function method. There exists a hierarchy of both basis sets and wave-function methods with increasing accuracy and computational costs, which is pictorially summarized in the famous Pople diagram [4] (see Fig. 6).

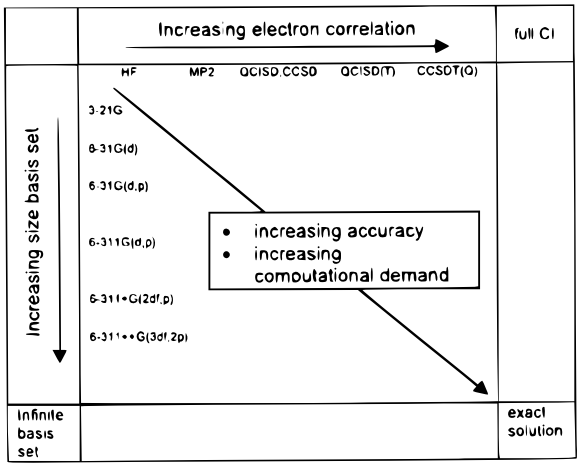

Figure 6: Pople diagram from Ref. [5]. The horizontal axis classifies some wave-function methods by accuracy levels. The vertical axis corresponds to basis sets with increasing sizes.

The simplest wave-function ansatz is the so-called Hartree-Fock approximation which substitutes the Coulomb repulsion term by a mean-field potential. Within the Hartree-Fock framework, one solves N coupled equations, one for each electron which permits a qualitative description of the electronic structure with a reasonable computational cost. However, to reach chemical accuracy (an error less than 1 kcal/mol for chemically relevant energy differences), one needs to go beyond the mean-field approximation and to consider the exact Coulomb repulsion, which leads to the post-Hartree-Fock methods. Among them, the full-configuration-interaction method is the exact solution within a given basis set and thus constitutes the target accuracy of the wave-function methods. There exists a wide range of approximations between the Hartree-Fock

approximation and the full-configuration-interaction method. Among the various approximated wave-function ansätze, we can cite truncated configuration interaction, perturbation theory, or the coupled cluster methods.

Regarding now the basis sets used in quantum chemistry, they are built such that energies and molecular properties are systematically improvable until reaching the complete-basis-set limit. Nevertheless, the complete-basis-set limit is numerically unreachable and the convergence to this limit is slow. There are essentially two ways to estimate the complete-basis-set limit: i) extrapolation schemes using basis sets of manageable sizes [6, 7], and ii) explicitly correlated methods which introduce the electron-electron cusp and the related short-range correlation effects [8, 9]. Nevertheless, these two approaches suffer from computational limitations: extrapolation schemes need increasingly large basis-set calculations which become rapidly unfeasible and explicitly correlated methods introduce auxiliary basis sets and a rather complicated mathematical framework.

An alternative to the manipulation of expensive wave-function methods is density-functional theory [10, 11]. In this theory, the exact energy can be described using only density functionals, among which the so-called exchange-correlation functional plays a crucial role. Nevertheless in practice the exchange-correlation density functional is not known and therefore approximations to the latter quantity have been derived. These functionals are computationally cheap and therefore allow for applications on very large systems. However, these approximated functionals are not systematically improvable and suffer from their uncontrolled description of long-range correlation effects such as strong correlation and dispersion forces.

Finally, the combination of density-functional and wave-function theories has lead to hybrid approaches among which we can cite the so-called range-separated density-functional theory [12]. In the latter, one uses a non-diverging long-range electron-electron interaction with a correlated wave-function method and the remaining short-range interaction is handled with a density functional.

In the present thesis, we expose the developments of a recently introduced hybrid method which tackles the problem of the slow basis-set convergence of wave-function methods: it is the density-based basis-set correction method [13]. In the latter, one considers a complementary functional of the density to estimate the correlation effects missing in a given finite basis set. As most of the missing correlation effect are short-range correlation effect, in practice the complementary functional is approximated by a specific type of range-separated density functional. Previous works [13–19] have focused mainly on corrections to the energy and the present work focuses on the extension to more generic molecular properties.

The manuscript is divided as follows.

Chapter 1 introduces some of the quantum chemistry methods involved in the present work.
After that, we discuss the slow basis-set convergence of the correlation energy which appears in wave-function based calculations. From that follows the introduction of methods to estimate the complete-basis-set limit. We end this chapter with a summary of the previous works on the density-based basis-set correction method.

Chapter 2 corresponds to the implementation and test of the self-consistent algorithm for the density-based basis-set correction method. With this procedure we go beyond the non-self-consistent approximation for energies. We apply the self-consistent basis-set correction method

to the dipole moment which is a property known to strongly depend on the density.

Chapter 3 corresponds to the extension of the correction of dipole moments to non-variational methods such as the widely used coupled-cluster method. The aim of this work is to propose an a-posteriori correction to the dipole moment and hence to avoid the self-consistent procedure.

Chapter 4 corresponds to the development of a one-dimensional model system to help formalizing the mathematical framework of the density-based basis-set correction method. By avoiding the approximation of the basis-set complementary functional by a density functional from another theory (namely range-separated density-functional theory) we allow for a clearer understanding of the method.

Motivated by the possibility of computing properties related to excited states, we propose in Chapter 5, an extension of the density-based basis-set correction method to the linear-response formalism. We explain how the basis-set correction potential in the Hamiltonian leads to a correlation kernel which depends on the second-order derivative of the complementary functional. This chapter includes also tests performed on a one-dimensional model.

Bibliography

[1] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, Inc., 1996).

[2] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic Structure Theory* (John Wiley & Sons, LTD, 2000).

[3] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).

[4] J. A. Pople, J. Chem. Phys. **43**, S229 (1965).

[5] S. Varis, *Molecular modelling of some explosives and propellants*, Tech. rep. (2013).

[6] T. Helgaker, W. Klopper, H. Koch and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[7] A. Halkier, W. Klopper, T. Helgaker and P. Jørgensen, J. Chem. Phys. **111**, 4424 (1999).

[8] L. Kong, F. A. Bischoff and E. F. Valeev, Chem. Rev. **112**, 75 (2012).

[9] C. Hättig, W. Klopper, A. Köhn and D. P. Tew, Chem. Rev. **112**, 4 (2012).

[10] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[11] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[12] J. Toulouse, F. Colonna and A. Savin, Phys. Rev. A **70**, 062505 (2004).

[13] E. Giner, B. Pradines, A. Ferté, R. Assaraf, A. Savin and J. Toulouse, J. Chem. Phys. **149**, 194301 (2018).

[14] E. Giner, A. Scemama, P.-F. Loos and J. Toulouse, J. Chem. Phys. **152**, 174104 (2020).

[15] P.-F. Loos, B. Pradines, A. Scemama, J. Toulouse and E. Giner, J. Phys. Chem. Lett. **10**, 2931 (2019).

[16] Y. Yao, E. Giner, J. Li, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **153**, 124117 (2020).

[17] Y. Yao, E. Giner, T. A. Anderson, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **155**, 204104 (2021).

[18] E. Giner, A. Scemama, J. Toulouse and P.-F. Loos, J. Chem. Phys. **151**, 144118 (2019).

[19] P.-F. Loos, B. Pradines, A. Scemama, E. Giner and J. Toulouse, J. Chem. Theory Comput. **16**, 1018 (2020).

Abbreviations

Abbreviation	Meaning
CBS	Complete basis set
c	correlation (energy)
CC	Coupled-cluster
CCSD	Coupled-cluster single and double excitations
(RO)CCSD(T)	(Restricted open-shell) Coupled-cluster single and double with perturbative triple excitations
CI	Configuration interaction
CIPSI	Configuration interaction using a perturbative selection made iteratively
CIS	Configuration interaction with single excitations
CISD	Configuration interaction with single and double excitations
DFA	Density-functional approximation
DFT	Density-functional theory
FCI	Full-configuration interaction
GGA	Generalized-gradient approximation (functional)
GTO	Gaussian-type orbital
(RO)HF	(Restricted open-shell) Hartree-Fock
Hxc	Hartree-exchange-correlation (energy)
KS-DFT	Kohn-Sham density-functional theory
LDA	Local-density approximation
lr	long-range
md	multi-determinantal
MP2	second-order Möller-Plesset perturbation theory
PBE	Perdew-Burke-Ernzerhof (functional)
PBE-OT	PBE with on-top pair density (functional)
PBE-UEG	PBE with uniform electron gaz on-top pair density (functional)
PT2	Second order perturbation theory
QP2	Quantum Package 2.0
RS-DFT	Range-separated density-function theory
sr	short-range
UEG	Uniform electron gaz
WFT	Wave-function theory
xc	Exchange-correlation (energy)

1 Methods

1.1 A brief review of wave-function methods

1.1.1 The quantum electronic problem

We consider the Schrödinger equation in the Born-Oppenheimer[1] and non-relativistic approximations for an N-electron system, in atomic units:

$$\hat{H}\Psi(\mathbf{x}) = E\Psi(\mathbf{x}), \tag{1.1}$$

with the electronic Hamiltonian operator $\hat{H} = \hat{T} + \hat{W}_{\text{ee}} + \hat{V}_{\text{ne}} + \hat{V}_{\text{nn}}$ where $\hat{T} = -\sum_{i=1}^{N} \Delta_{\mathbf{r}_i}/2$ is the kinetic-energy operator, $\hat{W}_{\text{ee}} = \sum_{i>j} 1/\|\mathbf{r}_i - \mathbf{r}_j\|$ is the Coulomb interaction between the electrons, $\hat{V}_{\text{ne}} = -\sum_i v_{\text{ne}}(\mathbf{r}_i)$ is the interaction between electrons and nuclei, $v_{\text{ne}}(\mathbf{r}_i) = Z_I/\|\mathbf{r}_i - \mathbf{R}_I\|$, $\hat{V}_{\text{nn}} = \sum_{I>J} Z_I Z_J/\|\mathbf{R}_I - \mathbf{R}_J\|$ is the nuclei-nuclei interaction, and Z_I is the Ith atomic number. In the latter definitions, lowercase letters designate electron coordinates and uppercase letters designate nuclear coordinates.

Here, $\Psi(\mathbf{x})$ is the N-electron wave function and $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, ..., \mathbf{x}_N)$ are the coordinates of the N electrons. For the ith electron, $\mathbf{x}_i = (\mathbf{r}_i, s_i)$ where $\mathbf{r}_i = (x_i, y_i, z_i)$ are its spatial coordinates and s_i its spin coordinates. Finally, E is the energy of the electronic state described by $\Psi(\mathbf{x})$.

Most of the present work focuses on the lowest-energy state commonly called the ground state. The energy and wave function associated with this state are denoted by E_0 and Ψ_0.

The variational principle states that any approximation $\tilde{\Psi}$ for Ψ_0 should lead to an energy $\tilde{E} \simeq \langle\tilde{\Psi}|\hat{H}|\tilde{\Psi}\rangle > E_0$. Therefore, given a family of approximate wave functions $\tilde{\Psi}$ with a restricted set of variational parameters, the best energy one can access is

$$\tilde{E}_0 = \min_{\tilde{\Psi}} \langle\tilde{\Psi}|\hat{H}|\tilde{\Psi}\rangle, \tag{1.2}$$

and $\tilde{E}_0 = E_0$ in the limit where the minimizing wave function $\tilde{\Psi}_0$ in Eq. (1.2) is the exact ground-state wave function, i.e $\tilde{\Psi}_0 = \Psi_0$.

1.1.2 Hartree-Fock approximation

In the Hartree-Fock approximation [1, 2], the wave function is approximated by a single Slater determinant. The properties of the determinant permit to recover an antisymmetric N-electron

[1] i.e. «when one can consider the electrons in a molecule to be moving in the field of fixed nuclei.» [1]

wave function and thus respects the fermionic character of the electronic system.

Therefore, the Hartree-Fock (HF) wave function is

$$\Phi_{\mathrm{HF}}(\mathbf{x}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & . & . & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & . & . & \chi_N(\mathbf{x}_2) \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & . & . & \chi_N(\mathbf{x}_N) \end{vmatrix}, \tag{1.3}$$

where $\chi_p(\mathbf{x}_i) \equiv \chi_p(\mathbf{r}_i, s_i)$ is the pth one-electron molecular spin-orbital for the ith electron at the spatial position $\mathbf{r}_i$ with spin s_i with associated energy ϵ_p. From Pauli's exclusion principle, one should note that each spin-orbital is occupied at most by one electron.

Using this framework, instead of solving the N-electron Schrödinger equation, one solves N coupled one-electron equations

$$\hat{f}\chi_p(\mathbf{x}) = \epsilon_p \chi_p(\mathbf{x}), \tag{1.4}$$

where the Fock operator is

$$\hat{f} = -\frac{1}{2}\Delta_{\mathbf{r}} - \sum_I \frac{Z_I}{\|\mathbf{r} - \mathbf{R}_I\|} + \hat{v}^{\mathrm{HF}}(\mathbf{x}). \tag{1.5}$$

The last term in Eq. (1.5) is the Hartree-Fock potential which is the mean-field potential seen by the electron due to the presence of the other electrons and its expression is

$$\hat{v}^{\mathrm{HF}}(\mathbf{x}) = \sum_k^{\mathrm{occ.}} \left(\hat{J}_k(\mathbf{x}) - \hat{K}_k(\mathbf{x}) \right), \tag{1.6}$$

where the sum is over occupied spin orbitals in the Hartree-Fock determinant.

From the latter equation, we define the Coulomb operator, $\hat{J}_k(\mathbf{x})$, such that

$$\hat{J}_k(\mathbf{x}_1)\chi_p(\mathbf{x}_1) = \left[\int \mathrm{d}\mathbf{x}_2 \, |\chi_k(\mathbf{x}_2)|^2 \|\mathbf{r}_1 - \mathbf{r}_2\|^{-1} \right] \chi_p(\mathbf{x}_1), \tag{1.7}$$

and the exchange operator, $\hat{K}_k(\mathbf{x})$, such that

$$\hat{K}_k(\mathbf{x}_1)\chi_p(\mathbf{x}_1) = \left[\int \mathrm{d}\mathbf{x}_2 \, \chi_k^*(\mathbf{x}_2) \|\mathbf{r}_1 - \mathbf{r}_2\|^{-1} \chi_p(\mathbf{x}_2) \right] \chi_k(\mathbf{x}_1). \tag{1.8}$$

As $\hat{v}^{\mathrm{HF}}(\mathbf{x})$ depends on all occupied molecular orbitals, equations (1.4) must be solved using a self-consistent field approach.

Using this approximation, one gets a good qualitative description of the molecular system (almost 99% of the total energy for a wide range of systems[2]). Nevertheless, for a quantitative description one needs to go beyond the Hartree-Fock approximation, i.e. using the so-called post-Hartree-Fock methods. These methods aim to get rid of the mean-field approximation, by solving the N-electron Schrödinger equation using the exact Coulombic interaction between the particles. These post-Hartree-Fock methods bring the effect of electronic correlation.

[2]that do not include strong or static correlation effects, which require a multi-determinantal description.

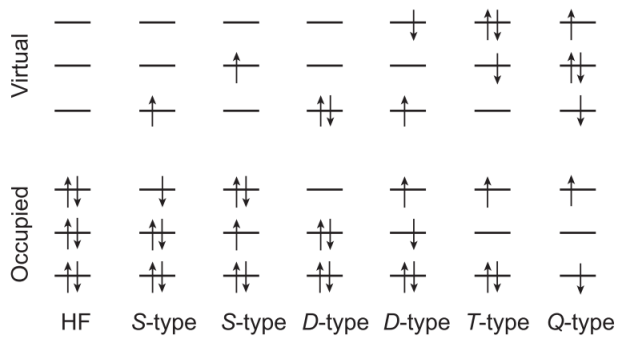

Figure 1.1: Representations of Slater determinants for a system with six electrons. HF: Hartree-Fock determinant ; S-type: determinants with a single excitation ; D-type: determinants with double excitations ; T-type: determinants with triple excitations, and so on. The occupied and virtual orbitals are those occupied and unoccupied in the Hartree-Fock determinant, respectively. (Figure from Ref. [3])

1.1.3 Post-Hartree-Fock methods

Truncated- and full- configuration interactions

The configuration interaction (CI) [2] wave function is a linear combination of Slater determinants generated by distributing one or more electrons in unoccupied molecular orbitals. The generation of Slater determinants is most easily written in the second quantization formalism.

One considers operators that withdraw one electron from a given molecular spin-orbital or add an electron to one. This operators are the annihilation and creation operators, $\hat{a}_i$ and $\hat{a}_i^\dagger$, respectively, where i is the spin-orbital index.

Thus, one writes, for a determinant $|\Phi_i^a\rangle$ describing a single excitation (S-type in Fig. 1.1) from the occupied spin-orbital i to the virtual spin-orbital a,

$$|\Phi_i^a\rangle = \hat{a}_a^\dagger \hat{a}_i |\Phi_{\mathrm{HF}}\rangle, \tag{1.9}$$

for a determinant $|\Phi_{ij}^{ab}\rangle$ describing a double excitation (D-type in Fig. 1.1) from the occupied spin-orbitals i and j to the virtual spin-orbitals a and b,

$$|\Phi_{ij}^{ab}\rangle = \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i |\Phi_{\mathrm{HF}}\rangle, \tag{1.10}$$

for a determinant $|\Phi_{ijk}^{abc}\rangle$ describing a triple excitation (T-type in Fig. 1.1),

$$|\Phi_{ijk}^{abc}\rangle = \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_c^\dagger \hat{a}_k \hat{a}_j \hat{a}_i |\Phi_{\mathrm{HF}}\rangle, \tag{1.11}$$

and so on.

The full configuration interaction (FCI) wave function $|\Psi_{\mathrm{FCI}}\rangle$ is the wave function obtained as the linear combination of all possible determinants we can build from the set of molecular orbitals:

$$|\Psi_{\mathrm{FCI}}\rangle = c_0|\Phi_{\mathrm{HF}}\rangle + \sum_{i}^{\text{occ.}} \sum_{a}^{\text{virt.}} c_i^a |\Phi_i^a\rangle + \sum_{i>j}^{\text{occ.}} \sum_{a>b}^{\text{virt.}} c_{ij}^{ab} |\Phi_{ij}^{ab}\rangle + \sum_{i>j>k}^{\text{occ.}} \sum_{a>b>c}^{\text{virt.}} c_{ijk}^{abc} |\Phi_{ijk}^{abc}\rangle + ... \tag{1.12}$$

where $\{c_i^a, c_{ij}^{ab}, c_{ijk}^{abc}, ...\}$ are the FCI coefficients, and occ. and virt. refer to occupied and virtual orbitals in the Hartree-Fock determinant, respectively. The FCI coefficients are obtained by solving the following minimization scheme:

$$E_{\mathrm{FCI}} = \min_{\Psi_{\mathrm{FCI}}} \langle \Psi_{\mathrm{FCI}} | \hat{H} | \Psi_{\mathrm{FCI}} \rangle, \tag{1.13}$$

where E_{FCI} is the FCI energy associated with the state $|\Psi_{\mathrm{FCI}}\rangle$. Due to the expectation value formulation, the FCI energy is variational.

In practice, one should notice the rapid increase of the dimension of the FCI space[3] with the number of electrons. For instance, in Fig. 1.5 we report the number of determinants in the FCI wave function for some simple systems. We read that the number of determinant scale exponentially with the basis set size. Therefore, one usually truncates the wave-function expansion Eq. (1.12) in terms of excitation levels. In this manner, we define several configuration interaction approximations. The most common and well known is the Configuration Interaction with Single and Double excitations (CISD) which leads to the wave function

$$|\Psi_{\mathrm{CISD}}\rangle = c_0|\Phi_{\mathrm{HF}}\rangle + \sum_{i}^{\text{occ.}}\sum_{a}^{\text{virt.}} c_i^a|\Phi_i^a\rangle + \sum_{i>j}^{\text{occ.}}\sum_{a>b}^{\text{virt.}} c_{ij}^{ab}|\Phi_{ij}^{ab}\rangle. \tag{1.14}$$

The CISD method has been mostly abandoned due to its lack of size extensivity, i.e., the energy does not scale linearly with the number of electrons.

Selected Configuration Interaction

The selected configuration interaction approximation that we will consider here is the CIPSI, Configuration Interaction using a Perturbative Selection made Iteratively [4, 5].

The CIPSI wave function is the linear combination of determinants $\{|I\rangle\}$ in a set $\mathcal{R}$ that are the more suited to build a wave function for a good quantitative description of the specific electronic system:

$$|\Psi_{\mathrm{CIPSI}}\rangle = \sum_{I\in\mathcal{R}} c_I|I\rangle. \tag{1.15}$$

In order to obtain Ψ_{CIPSI}, one starts with a substet $\mathcal{R}^{(0)}$ of $\mathcal{R}$. The set $\mathcal{R}^{(0)}$ gives the wave function

$$|\Psi^{(0)}\rangle = \sum_{I\in\mathcal{R}^{(0)}} c_I^{(0)}|I\rangle, \tag{1.16}$$

where the superscript (0) indicates that we are at the initial step of the procedure, where the determinants $|I\rangle$ are in the set $\mathcal{R}^{(0)}$. From here, we consider a kth step of the procedure such that $\mathcal{R}^{(0)}$ is a subset of $\mathcal{R}^{(k)}$ which is itself a subset of $\mathcal{R}$, and

$$|\Psi^{(k)}\rangle = \sum_{I\in\mathcal{R}^{(k)}} c_I^{(k)}|I\rangle. \tag{1.17}$$

At each step, we define the subset $\bar{\mathcal{R}}^{(k)}$ that contains all the determinants $|P\rangle$ that are not in $\mathcal{R}^{(k)}$.

We compute the contributions of each determinants $|P\rangle$ to the second-order perturbation energy,

$$\epsilon_{P,\mathrm{PT2}} = -\frac{|\langle\Psi^{(k)}|\hat{H}|P\rangle|^2}{\langle P|\hat{H}|P\rangle - \langle\Psi^{(k)}|\hat{H}|\Psi^{(k)}\rangle}, \tag{1.18}$$

and select a certain amount of determinants having the most important contributions in absolute value. The selected determinants are added to the set $\mathcal{R}^{(k)}$ to build the set $\mathcal{R}^{(k+1)}$ and therefore the wave function

$$|\Psi^{(k+1)}\rangle = \sum_{I\in\mathcal{R}^{(k+1)}} c_I^{(k+1)}|I\rangle. \tag{1.19}$$

[3] i.e., the number of Slater determinants. Increasing the size of the system leads to more excitations possible and thus to an higher number of FCI coefficients to optimize.

We also define the second-order perturbation theory at the step (k) as

$$E_{\mathrm{PT2}} = \sum_{P \in \bar{\mathcal{R}}^{(k)}} \epsilon_{P,\mathrm{PT2}}. \tag{1.20}$$

If E_{PT2} is greater than a a given threshold, one starts over at Eq. (1.17). At the end of the procedure, the CIPSI energy is evaluated as

$$E_{\mathrm{CIPSI}} = E_{\mathrm{var}} + E_{\mathrm{PT2}}, \tag{1.21}$$

where E_{var} is the variational energy,

$$E_{\mathrm{var}} = \langle \Psi_{\mathrm{CIPSI}} | \hat{H} | \Psi_{\mathrm{CIPSI}} \rangle. \tag{1.22}$$

The CIPSI algorithm is implemented in the Quantum Package 2.0 software [6] and we invite the reader seeking for more information to consult Ref. [7] and the Quantum Package website quantum-package.readthedocs.io.

Coupled-cluster theory

In subsequent chapters, we will use coupled-cluster approximations and mainly the CCSD(T) method, i.e, Coupled-Cluster Singles and Doubles with perturbative Triples [2,8], the gold standard of quantum chemistry. The name gold standard is due to its good trade-off between computational cost and accuracy for many molecular systems. In addition, the coupled-cluster method cures the size-extensivity problem encountered in truncated CI methods. For our purpose, we will not go into much considerations but we use this paragraph to qualitatively describe its properties.

The coupled-cluster (CC) wave function is expressed through an exponential ansatz, such that

$$|\Psi_{\mathrm{CC}}\rangle = e^{\hat{T}} |\Phi_{\mathrm{HF}}\rangle, \tag{1.23}$$

where $\hat{T}$ is the excitation operator. For example, for Coupled-Cluster Singles and Doubles (CCSD), $\hat{T} = \hat{T}_1 + \hat{T}_2$ where $\hat{T}_1$ and $\hat{T}_2$ satisfy the following relations:

$$\begin{aligned} \hat{T}_1 |\Phi_{\mathrm{HF}}\rangle &= \sum_{i}^{\text{occ.}} \sum_{a}^{\text{virt.}} t_i^a |\Phi_i^a\rangle, \\ \hat{T}_2 |\Phi_{\mathrm{HF}}\rangle &= \sum_{i>j}^{\text{occ.}} \sum_{a>b}^{\text{virt.}} t_{ij}^{ab} |\Phi_{ij}^{ab}\rangle. \end{aligned} \tag{1.24}$$

The amplitudes t_i^a, t_{ij}^{ab} are the coupled-cluster amplitudes defined through the equations

$$\begin{aligned} \langle \Phi_i^a | e^{-(\hat{T}_1+\hat{T}_2)} \hat{H} e^{(\hat{T}_1+\hat{T}_2)} | \Phi_{\mathrm{HF}} \rangle = 0, \\ \langle \Phi_{ij}^{ab} | e^{-(\hat{T}_1+\hat{T}_2)} \hat{H} e^{(\hat{T}_1+\hat{T}_2)} | \Phi_{\mathrm{HF}} \rangle = 0. \end{aligned} \tag{1.25}$$

However, in practice one computes the coupled-cluster energy using a projection on the Hartree-Fock determinant:

$$E_{\mathrm{CCSD}} = \langle \Phi_{\mathrm{HF}} | \hat{H} e^{(\hat{T}_1+\hat{T}_2)} | \Phi_{\mathrm{HF}} \rangle, \tag{1.26}$$

where the coupled-cluster amplitudes are found from Eqs. (1.25).

However, to reach chemical accuracy, it has been shown that one must consider up to triple excitations. In practice, one uses CCSD(T) where they are taken into account through a perturbation contribution computed on top of the CCSD calculation. Because it is not written as an expectation value, the CCSD(T) method is not variational.

1.2 A brief review of density-functional methods

1.2.1 Density-functional theory

Density-functional theory (DFT) is an alternative to wave-function approaches where one reformulates the ground-state energy as a minimization over the one-electron density $n(\mathbf{r})$. In fact, according to the Hohenberg-Kohn theorem [9], the energy of the ground state is a functional of its electronic density. The variational problem over the density can be derived from the variational problem over the wave function in Eq. (1.2):

$$\begin{aligned} E_0 &= \min_{\Psi} \langle \Psi | \hat{T} + \hat{W}_{\mathrm{ee}} + \hat{V}_{\mathrm{ne}} | \Psi \rangle \\ &= \min_{n} \left\{ \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi \rangle + \int \mathrm{d}\mathbf{r} v_{\mathrm{ne}}(\mathbf{r}) n(\mathbf{r}) \right\} \\ &= \min_{n} \left\{ F[n] + \int \mathrm{d}\mathbf{r} v_{\mathrm{ne}}(\mathbf{r}) n(\mathbf{r}) \right\} \end{aligned} \tag{1.27}$$

where $F[n]$ is the Levy-Lieb universal density functional,

$$F[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi \rangle, \tag{1.28}$$

and $\Psi \to n$ means that we minimize over wave functions which have the density n.

DFT gives access to energies and molecular properties often comparable with wave-function correlated methods in terms of accuracy, at a computational cost comparable with the Hartree-Fock approximation.

Kohn-Sham formulation of the Levy-Lieb universal density functional

The Kohn-Sham formulation [10] is the most common formulation where one considers a unique Slater determinant Φ and decomposes the Levy-Lieb universal density functional as

$$F[n] = \min_{\Phi \to n} \langle \Phi | \hat{T} | \Phi \rangle + E_{\mathrm{Hxc}}[n], \tag{1.29}$$

where $E_{\mathrm{Hxc}}[n]$ is the universal Hartree-exchange-correlation functional of the density,

$$E_{\mathrm{Hxc}}[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi \rangle - \min_{\Phi \to n} \langle \Phi | \hat{T} | \Phi \rangle, \tag{1.30}$$

for which one usually considers the following decomposition:

$$E_{\mathrm{Hxc}}[n] = E_{\mathrm{H}}[n] + E_{\mathrm{xc}}[n]. \tag{1.31}$$

In this decomposition, $E_{\mathrm{H}}[n]$ is the Hartree density functional defined by

$$E_{\mathrm{H}}[n] = \frac{1}{2} \iint \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \, \frac{n(\mathbf{r}_1) n(\mathbf{r}_2)}{\|\mathbf{r}_1 - \mathbf{r}_2\|}, \tag{1.32}$$

and describing the Coulombic interaction between electronic densities. The exchange and correlation density functional, $E_{\mathrm{xc}}[n]$, is also decomposed into $E_{\mathrm{xc}}[n] = E_{\mathrm{x}}[n] + E_{\mathrm{c}}[n]$. In the latter expression, the exchange density functional is defined by

$$E_{\mathrm{x}}[n] = \langle \Phi[n] | \hat{W}_{\mathrm{ee}} | \Phi[n] \rangle - E_{\mathrm{H}}[n], \tag{1.33}$$

where $\Phi[n]$ is the minimizing Slater determinant with density n in Eq. (1.29). Finally, $E_{\mathrm{c}}[n]$ is the correlation density functional and its expression is

$$E_{\mathrm{c}}[n] = \langle \Psi[n] | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi[n] \rangle - \langle \Phi[n] | \hat{T} + \hat{W}_{\mathrm{ee}} | \Phi[n] \rangle, \tag{1.34}$$

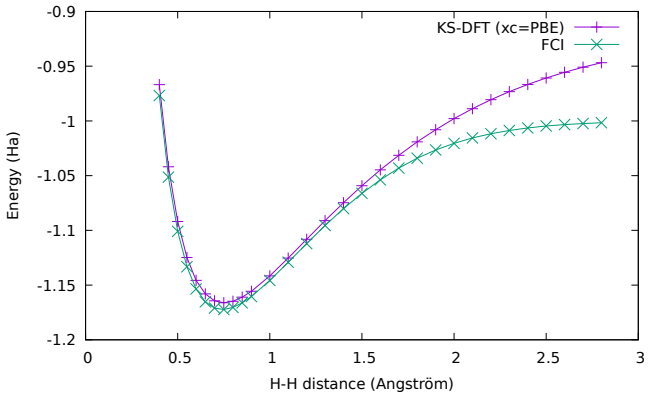

Figure 1.2: H_2 dissociation curve using cc-pVTZ basis-set. Kohn-Sham DFT (KS-DFT) with PBE exchange-correlation functional is compared with FCI.

where $\Psi[n]$ is the minimizing multi-determinantal wave function with density n in Eq. (1.28). In this context, the Kohn-Sham expression for the exact ground-state energy is

$$E_0 = \min_{\Phi} \left(\langle \Phi | \hat{T} + \hat{V}_{\text{ne}} | \Phi \rangle + E_{\text{Hxc}}[n_\Phi] \right), \tag{1.35}$$

where n_Φ is the density calculated from Φ.

Density-functional approximations

In practice, one does not have an explicit expression for $E_{\text{xc}}[n]$. Thus, one uses density-functional approximations (DFA).

There are two most common families of approximations: the local-density approximations (LDA) and the generalized-gradient approximations (GGA). The LDA functional is expressed as

$$E_{\text{xc}}^{\text{LDA}}[n] = \int \mathrm{d}\mathbf{r}\, e_{\text{xc}}^{\text{UEG}}(n(\mathbf{r})), \tag{1.36}$$

where $e_{\text{xc}}^{\text{UEG}}(n)$ is the exchange-correlation energy density of the infinite uniform-electron gas[4] (UEG). The GGA functionals are expressed as

$$E_{\text{xc}}^{\text{GGA}}[n] = \int \mathrm{d}\mathbf{r}\, e_{\text{xc}}^{\text{GGA}}(n(\mathbf{r}), \nabla n(\mathbf{r})), \tag{1.37}$$

where $e_{\text{xc}}^{\text{GGA}}(n, \nabla n)$ is an exchange-correlation energy density which also depends on the density gradient. The GGA functional that we employ in the present work are the Perdew-Burke-Ernzerhof (PBE) functional [11].

The major benefit in the definition of DFA is that it takes into account the short-range correlation in a compact way which enables fast basis convergences. However, nonlocal correlation effects such as strong correlation or dispersion forces are difficult to describe with LDA or GGA functionals. As an example, in Fig. 1.2, we show the dissociation curve of the dihydrogen molecule. Around the equilibrium distance, the density-functional method reproduces the FCI energies. But as we stretch the molecule, the approximate functional tends to overestimate ground-state energies.

For a more thorough introduction to the subject, we suggest the reader to consult Ref. [12] and references therein.

[4]The uniform electron gas are model systems where the density is uniform in space.

1.2.2 Range-separated density-functional theory

We focus in this paragraph on range-separated DFT (RS-DFT) [13] where the Coulomb repulsion is splitted in two parts i) the long-range repulsion is computed with a wave-function method, and ii) the short-range repulsion is computed with a functional of the density. In this method, one exploits the systematic and good convergence of the long-range interaction contribution to the energy using wave-function-based approaches, and the fact that density-functional approximations permit to use a simple formulation of the short-range repulsion. As we dedicate this section to RS-DFT, we keep the problem of the short-range repulsion treatment to Sec. 1.3.2 where we illustrate how the short-range repulsion is the bottleneck in wave-function calculations.

The RS-DFT method is based on the following decomposition of the Levy-Lieb universal density functional:

$$F[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{W}_{\rm ee}^{{\rm lr},\mu} | \Psi \rangle + \bar{E}_{\rm Hxc}^{{\rm sr},\mu}[n], \tag{1.38}$$

where $\hat{W}_{\rm ee}^{{\rm lr},\mu}$ is the long-range (lr) interaction,

$$\hat{W}_{\rm ee}^{{\rm lr},\mu} = \sum_{i<j} w_{\rm ee}^{{\rm lr},\mu}(\|\mathbf{r}_i - \mathbf{r}_j\|), \tag{1.39}$$

with $w_{\rm ee}^{{\rm lr},\mu}(r) = {\rm erf}(\mu r)/r$, and the error function erf removing the short-range divergence of the Coulomb interaction. On the left side of Fig. 1.3, we show the long-range interaction for different values of μ. We see that as we increase the value of μ, we tend toward the exact interaction $1/\|\mathbf{r}_1 - \mathbf{r}_2\|$. In practice, one often use a value around 0.5 for μ [14, 15].

The last term in Eq. (1.38), $\bar{E}_{\rm Hxc}^{{\rm sr},\mu}[n]$, is the complementary short-range (sr) Hartree-exchange-correlation (Hxc) functional of the density, and μ is the range-separation parameter.

Finally, the exact ground-state energy in the RS-DFT formulation is expressed as

$$E_0 = \min_{\Psi} \left\{ \langle \Psi | \hat{T} + \hat{V}_{\rm ne} + \hat{W}_{\rm ee}^{{\rm lr},\mu} | \Psi \rangle + \bar{E}_{\rm Hxc}^{{\rm sr},\mu}[n_\Psi] \right\}, \tag{1.40}$$

where n_Ψ is the density calculated from Ψ.

Usually, the sr-Hxc functional of the density is decomposed in $\bar{E}_{\rm Hxc}^{{\rm sr},\mu}[n] = E_{\rm H}^{{\rm sr},\mu}[n] + \bar{E}_{\rm xc}^{{\rm sr},\mu}[n]$. In the latter expression, $E_{\rm H}^{{\rm sr},\mu}[n]$ is the short-range Hartree density functional:

$$E_{\rm H}^{{\rm sr},\mu}[n] = \frac{1}{2} \iint {\rm d}\mathbf{r}_1 {\rm d}\mathbf{r}_2 \, n(\mathbf{r}_1) n(\mathbf{r}_2) w_{\rm ee}^{{\rm sr},\mu}(\|\mathbf{r}_1 - \mathbf{r}_2\|) \tag{1.41}$$

where $w_{\rm ee}^{{\rm sr},\mu}(r) = 1/r - w_{\rm ee}^{{\rm lr},\mu}(r)$. On the right side of Fig. 1.3, we illustrate how the short-range interaction reproduces the problematic short-range divergence of the exact $1/\|\mathbf{r}_1 - \mathbf{r}_2\|$ interaction. The second term in the RS-DFT energy (Eq. (1.40)), $\bar{E}_{\rm xc}^{{\rm sr},\mu}[n]$, is the short-range exchange and correlation density functional. This functional is decomposed to

$$\bar{E}_{\rm xc}^{{\rm sr},\mu}[n] = E_{\rm x}^{{\rm sr},\mu}[n] + \bar{E}_{\rm c}^{{\rm sr},\mu}[n]. \tag{1.42}$$

Here, we introduce the short-range exchange density functional

$$E_{\rm x}^{{\rm sr},\mu}[n] = \langle \Phi^{\rm KS}[n] | \hat{W}_{\rm ee}^{{\rm sr},\mu} | \Phi^{\rm KS}[n] \rangle - E_{\rm H}^{{\rm sr},\mu}[n], \tag{1.43}$$

where $\Phi^{\rm KS}[n]$ is the Kohn-Sham (KS) single determinant with density n and such that $\Phi^{\rm KS}[n] = \Psi^{\mu=0}[n]$, and the short-range interaction is defined by

$$\hat{W}_{\rm ee}^{{\rm sr},\mu} = \sum_{i<j} w_{\rm ee}^{{\rm sr},\mu}(\|\mathbf{r}_i - \mathbf{r}_j\|). \tag{1.44}$$

The short-range correlation density functional reads

$$\bar{E}_{\mathrm{c}}^{\mathrm{sr},\mu}[n] = \bar{E}_{\mathrm{Hxc}}^{\mathrm{sr},\mu}[n] - \langle \Phi^{\mathrm{KS}}[n] | \hat{W}_{\mathrm{ee}}^{\mathrm{sr},\mu} | \Phi^{\mathrm{KS}}[n] \rangle. \tag{1.45}$$

In Ref. [16], the authors proposed a multi-determinantal decomposition for $\bar{E}_{\mathrm{xc}}^{\mathrm{sr},\mu}[n]$:

$$\bar{E}_{\mathrm{xc}}^{\mathrm{sr},\mu}[n] = E_{\mathrm{x,md}}^{\mathrm{sr},\mu}[n] + \bar{E}_{\mathrm{c,md}}^{\mathrm{sr},\mu}[n]. \tag{1.46}$$

where we consider the multi-determinantal short-range exchange density functional

$$E_{\mathrm{x,md}}^{\mathrm{sr},\mu}[n] = \langle \Psi^{\mathrm{lr},\mu}[n] | \hat{W}_{\mathrm{ee}}^{\mathrm{sr},\mu} | \Psi^{\mathrm{lr},\mu}[n] \rangle - E_{\mathrm{H}}^{\mathrm{sr},\mu}[n], \tag{1.47}$$

where $\Psi^{\mathrm{lr},\mu}[n]$ is the multi-determinantal wave function with density n which minimizes the first term in Eq. (1.38), $\langle \Psi | \hat{T} + \hat{W}_{\mathrm{ee}}^{\mathrm{lr},\mu} | \Psi \rangle$, and the short-range interaction is defined by

$$\hat{W}_{\mathrm{ee}}^{\mathrm{sr},\mu} = \sum_{i<j} w_{\mathrm{ee}}^{\mathrm{sr},\mu}(\|\mathbf{r}_i - \mathbf{r}_j\|). \tag{1.48}$$

Finally, the multi-determinantal complementary short-range correlation density functional is

$$\bar{E}_{\mathrm{c,md}}^{\mathrm{sr},\mu}[n] = \langle \Psi[n] | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi[n] \rangle - \langle \Psi^{\mathrm{lr},\mu}[n] | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi^{\mathrm{lr},\mu}[n] \rangle. \tag{1.49}$$

Therefore, the RS-DFT with the multi-determinantal decomposition expression for the exact ground-state energy is

$$E_0 = \langle \Psi^{\mathrm{lr},\mu} | \hat{T} + \hat{V}_{\mathrm{ne}} + \hat{W}_{\mathrm{ee}} | \Psi^{\mathrm{lr},\mu} \rangle + \bar{E}_{\mathrm{c,md}}^{\mathrm{sr},\mu}[n_{\Psi^{\mathrm{lr},\mu}}], \tag{1.50}$$

where $\Psi^{\mathrm{lr},\mu}$ is the wave function which minimizes Eq. (1.40). In opposition with the decomposition in Eq. (1.42), the multi-determinantal decomposition treats the exact exchange functional and it does not depend on an intermediate Slater determinant.

In this variant of RS-DFT, a function of the density is used as an approximation for $\bar{E}_{\mathrm{c,md}}^{\mathrm{sr},\mu}[n]$. In the seminal work of Toulouse et al. [16], a local density approximation (LDA) is proposed:

$$\bar{E}_{\mathrm{c,md}}^{\mathrm{sr},\mu}[n] \simeq \bar{E}_{\mathrm{c,md,LDA}}^{\mathrm{sr},\mu}[n] = \int \mathrm{d}\mathbf{r}\, \bar{e}_{\mathrm{c,md}}^{\mathrm{sr},\mu,\mathrm{UEG}}(n(\mathbf{r})), \tag{1.51}$$

where $\bar{e}_{\mathrm{c,md}}^{\mathrm{sr},\mu,\mathrm{UEG}}(n)$ is the corresponding energy density of the UEG.

However, the LDA functional suffers from its locality assumption. Moreover, the authors in Ref. [16] showed that the asymptotic behavior (i.e. $\mu \to \infty$) of the multi-determinantal correlation functional only depend on the on-top pair density which is the probability density of having two electrons at the same spatial position. From this consideration comes the idea of GGA-based correlation functionals with explicit dependence on the on-top (OT) pair density: the PBE-OT and the PBE-UEG multi-determinant complementary short-range correlation functionals.

First, the PBE-OT multi-determinantal complementary short-range correlation density functional [17] is

$$\bar{E}_{\mathrm{c,md}}^{\mathrm{sr},\mu}[n] \simeq \bar{E}_{\mathrm{c,md,PBE\text{-}OT}}^{\mathrm{sr},\mu}[n] = \int \mathrm{d}\mathbf{r}\, \bar{e}_{\mathrm{c,md}}^{\mathrm{sr},\mu,\mathrm{PBE}}(n(\mathbf{r}), \nabla n(\mathbf{r}), n_2(\mathbf{r},\mathbf{r})), \tag{1.52}$$

where $n_2(\mathbf{r},\mathbf{r})$ is the on-top pair density from $\Psi^{\mathrm{lr},\mu}$ and $\bar{e}_{\mathrm{c,md}}^{\mathrm{sr},\mu,\mathrm{PBE}}$ is a function based on the PBE correlation functional.

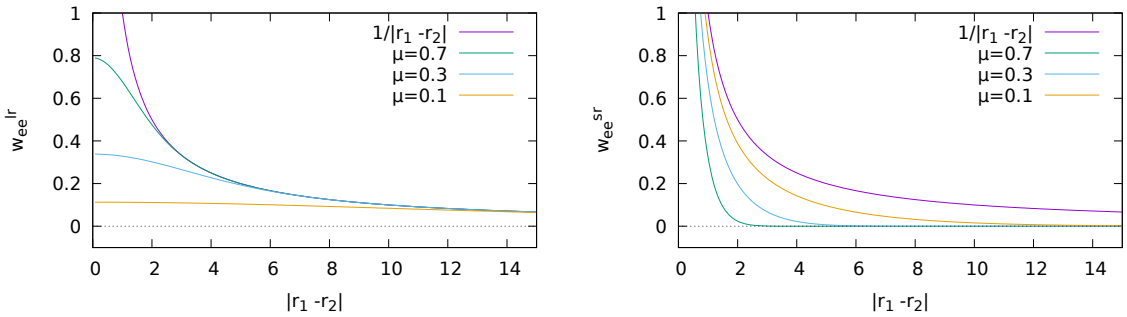

Figure 1.3: Left : $w_{\rm ee}^{{\rm lr},\mu}(|r_1 - r_2|) = {\rm erf}(\mu|r_1 - r_2|)/|r_1 - r_2|$, Right :$w_{\rm ee}^{{\rm sr},\mu}(|r_1 - r_2|) = (1 - {\rm erf}(\mu|r_1 - r_2|))/|r_1 - r_2|$.

Finally, the PBE-UEG multi-determinantal complementary short-range correlation density functional [18] uses $n_2^{\rm UEG}(\mathbf{r},\mathbf{r})$, the on-top pair density of the uniform electron gas having the density $n(\mathbf{r})$:

$$\bar{E}_{\rm c,md}^{{\rm sr},\mu}[n] \simeq \bar{E}_{\rm c,md,PBE\text{-}UEG}^{{\rm sr},\mu}[n] = \int {\rm d}\mathbf{r}\, \bar{e}_{\rm c,md}^{{\rm sr},\mu,{\rm PBE}}(n(\mathbf{r}), \nabla n(\mathbf{r}), n_2^{\rm UEG}(\mathbf{r},\mathbf{r})). \tag{1.53}$$

The on-top pair density of the uniform electron gas has been previously computed and therefore can be imported in the physical system of interest thanks to the usual LDA mapping. Therefore, the PBE-UEG functional is a good tradeoff between the LDA and the PBE-OT functionals.

For explicit expressions of the correlation energy per particle and related quantities, we invite the reader to consult the supplementary materials of Ref. [19] and Sec. 2.2.2 for the PBE-OT short-range correlation density functional.

1.3 The basis-set convergence of correlation effects in WFT

In this section, we address the problem of basis-set convergence of the correlation energy and molecular properties in wave-function theory methods. To explain it briefly, a quantum chemistry calculation is performed through an optimization of the wave function which is expanded on a finite set of basis functions (Sec. 1.3.1). However, to recover the exact energy and properties, one should make this expansion on a complete basis set (CBS), i.e. an infinite number of basis functions, which is computationally unfeasible.

In practice, the finite basis set leads to errors that mostly impact the electronic correlation description (Sec. 1.3.2). These errors add to the wave-function ansatz that has been chosen (truncated CI, CCSD, CCSD(T), ...) but can be quantified and removed through extrapolation techniques, or the use of a correlation factor (Sec. 1.3.3).

Depending on the system of interest, one chooses the type of basis functions better suited to its description while taking into account the computational cost. For molecular systems, the basis functions are centered on each nuclei. In this work, we use one family of basis sets: the Dunning (Gaussian based) basis sets.

In practice, there is no need to know the list of available basis sets and their technical details and one usually limits oneself to the family of basis sets suitable for the systems of interest. However, in the case of method development, it can be useful to have a general picture of the available basis sets to keep in mind their strengths and weaknesses. For that reason, I invite the reader seeking for an introduction to basis-set families to look for the dedicated book chapters in Ref. [2] or Refs. as [20,21]. Those who are only looking for the available basis sets for a given

atom are invited to have a look at the Basis-Set Exchange library[5].

1.3.1 Gaussian basis sets

One expands the one-electron molecular orbitals on atomic-centered basis functions. Considering the pth molecular orbital $\chi_p(\mathbf{x})$, its expansion in the basis set $\mathcal{B} = \{\phi_\mu\}$ is such that

$$\chi_p(\mathbf{x}) = f(s) \sum_\mu c_{\mu,p} \phi_\mu(\mathbf{r}). \tag{1.54}$$

The coefficients $c_{\mu,p}$ are found by solving the Fock equations (1.4), $f(s)$ is a spin function, and $\phi_\mu(\mathbf{r})$ are the so-called basis functions.

In the Dunning (or correlation-consistent) basis sets, the basis functions are linear combinations (or contractions) of primitive Gaussian-type orbitals (GTO). The number of primitive functions, $N_{\text{prim.}}$, depends on the atom considered and the basis set. Fortunately, the basis functions are all tabulated and one should not build them by hand. The basis functions are usually build as

$$\phi_\mu(\mathbf{r}) = (x - X_A)^a (y - Y_A)^b (z - Z_A)^c \sum_i d_{\mu,i} e^{-\alpha_{\mu,i} \|\mathbf{r} - \mathbf{R}_A\|^2} \tag{1.55}$$

where $\mathbf{r} = (x, y, z)$, and $\mathbf{R}_A = (X_A, Y_A, Z_A)$ are the coordinates of the nucleus A, and a, b, c are integers related to the angular momentum. The coefficients $d_{\mu,i}$ and exponents $\alpha_{\mu,i}$ are the one defined in libraries.

One finds the coefficents and exponents for Dunning basis sets in libraries such as Basis-Set Exchange and we will notice their astounding variety. We provide the most common basis sets in the following subsections.

Polarized-valence basis sets [22–24]

The nomenclature can be understood starting from the cc-pVXZ basis sets where "cc" stands for correlation consistent, "pV" for polarized valence, and "XZ" for X-zeta where $\mathrm{X} \in \{D, T, Q, 5, 6, 7, ...\}$ is called the cardinal number of the basis set. The basis-set size increases going from X=D to X=T, from X=T to X=Q, and so on.

Augmented basis sets [23–25]

For the computation of other molecular properties such as the dipole moments, we consider basis sets with diffuse orbitals which are Gaussians with small exponents $\alpha_{\mu,i}$. The Y-aug-cc-pVXZ basis sets are well suited in that case. There, "aug" indicates the use of GTO useful for the description of these diffuse orbitals and "Y" designates the level of this description[6].

Core-valence basis sets [27, 28]

In subsequent chapters, we also mention the use of the frozen-core approximation to decrease the computational cost. In this approximation, the core orbitals remain doubly occupied even at post-Hartree-Fock levels. For some systems where core electrons participate to the property of interest, one performs full-electron calculations. In this case, using Dunning basis sets, one

[5] www.basissetexchange.org

[6] Usually, one has access to Y=null, Y=D, and Y=T. For beyond, one can find a procedure in Ref. [26]. We used this Y-aug-cc-pVXZ basis sets when studying static polarizabilities where the convergence with respect to Y is also significant. But for real systems and applications, one should not go to higher values for "Y" as the optimization of the added GTO functions is less systematic than for the cc-pVXZ basis sets.

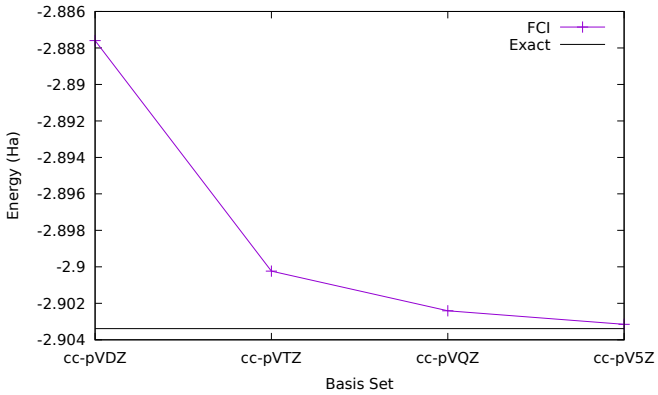

Figure 1.4: Convergence of the helium ground-state energy with the basis set calculated by FCI.

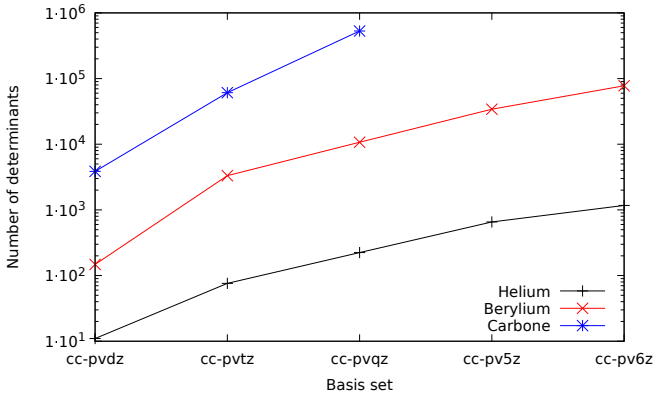

Figure 1.5: Number of Slater determinants in the FCI expansion for the helium, berylium and carbon atoms. We use a logarithmic scale to emphasize the exponential scaling of the dimension of the FCI space with the basis set size.

should consider the cc-p**C**VXZ basis sets where the **C** indicates the presence of basis functions adapted for the description of core correlation.

These are the three families of Dunning basis set mentioned in this work [29]. The main idea to keep in mind while reading the subsequent chapters is that by increasing the basis-set size, in a given family, one consistently increases the quality of description of the molecular system, especially the correlation effects which we discuss in the next paragraph.

1.3.2 Slow basis-set convergence of correlation

In Fig. 1.4, we compute the basis-set convergence of the helium ground-state energy. We notice that the convergence is systematic toward the exact value.

The systematic convergence appearing in Hartree-Fock, and post-Hartree-Fock leads to extrapolation schemes (Sec. 1.3.3) in order to estimate the CBS.

This behavior is generalizable molecular properties. However, as we increase the size of the system, computations with large basis sets become unfeasible and thus the CBS value can be soon unreachable using FCI.

The slow basis convergence illustrated in Fig. 1.4 was explained in the early years of quantum chemistry by Hylleraas [30]. As mentioned in Sec. 1.1, post Hartree-Fock methods use the exact divergent Coulomb interaction, $1/\|\mathbf{r}_i - \mathbf{r}_j\|$. The latter implies the occurrence of the electron-

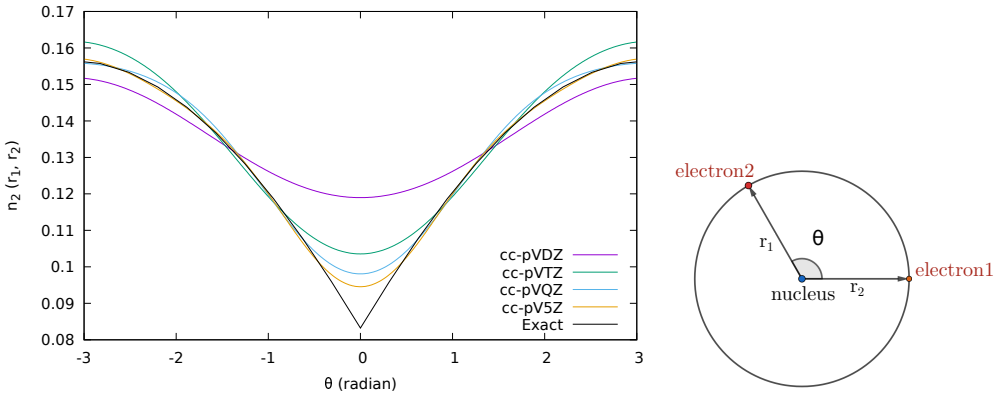

Figure 1.6: Electron pair density of the helium atom, $n_2(\mathbf{r}_1, \mathbf{r}_2) = 2|\Psi(\mathbf{r}_1, \mathbf{r}_2)|^2$ using Dunning basis sets compared to the exact behavior (data from Quantum Package).

electron cusp[7] in the exact wave function [31]. This translates on the spherically averaged wave function, $\tilde{\Psi}$, having a linear behavior in $\mathbf{r}_{12} = \|\mathbf{r}_1 - \mathbf{r}_2\|$ around $\mathbf{r}_1 = \mathbf{r}_2$ [32]:

$$\frac{\partial \tilde{\Psi}}{\partial r_{12}}\Big|_{r_{12}=0} = \frac{1}{2}\Psi(r_{12} = 0). \tag{1.56}$$

For instance, in Fig. 1.6 we compute the pair density $n_2(\mathbf{r}_1, \mathbf{r}_2)$ at the FCI level of the helium atom for different values of θ, the angle between the $\mathbf{r}_1$ and $\mathbf{r}_2$, and different basis sets. We emphasise that because we use the FCI, the only approximation is the basis set. For $\theta \simeq 0$, when the two electrons are close, the pair density converges slowly toward the exact one: this is an illustration of the slow convergence of the short-range correlation effect in WFT. One can also notice that all the wave functions expanded in a finite basis set do not have the cusp. For instance, even with the cc-pV5Z basis set, the FCI wave-function is far from the exact one. By contrast, the long-range behavior is well converged using the cc-pV5Z basis set. This example illustrates how post-Hartree-Fock methods (even FCI) struggle with the description of short-range correlation but are well suited for the convergence of long-range correlation.

1.3.3 Standard methods to estimate the complete basis-set limit

The interest of performing quantum chemistry calculations using wave-function approaches lies in the ability to obtain systematic accurate results. In this framework, one must keep in mind that the more accurate result within a given method (MP2, CC, ...) is provided by the CBS limit. To approximate this limit, one might consider extrapolation schemes or the use of a correlation factor.

Two-point extrapolation schemes

We define E^X as the ground-state energy computed using a Dunning basis set with cardinal number X and refer to this basis set through the labelling XZ instead of cc-pVXZ or cc-pCVXZ or aug-cc-pVXZ. We also consider E^{CBS} as the ground-state energy at the CBS limit. Both quantities are computed using the same correlated method.

The Hartree-Fock energy computed using the basis set XZ is E^X_{HF}, and the Hartree-Fock energy computed using a complete basis set is $E^{\mathrm{CBS}}_{\mathrm{HF}}$.

From these quantities, we define the correlation energy using the basis set XZ,

$$E^X_{\mathrm{c}} = E^X - E^X_{\mathrm{HF}}, \tag{1.57}$$

[7] We do not consider the nuclei-electron cusp in this section for which the principle is the same.

and the correlation energy at the CBS limit,

$$E_{\rm c}^{\rm CBS} = E^{\rm CBS} - E_{\rm HF}^{\rm CBS}. \tag{1.58}$$

In Ref. [33] the following relation between Eqs. (1.57) and (1.58) is found,

$$E_{\rm c}^{X} = E_{\rm c}^{\rm CBS} + AX^{-3}, \tag{1.59}$$

where A is a parameter to be fitted from which we deduce a two-point extrapolation scheme:

$$E_{\rm c}^{\rm CBS} = \frac{E_{\rm c}^{X} X^3 - E_{\rm c}^{X-1}(X-1)^3}{X^3 - (X-1)^3}, \tag{1.60}$$

where $E_{\rm c}^{X-1}$ is the correlation energy computed using a $(X-1)Z$ basis set.

In Ref. [34], this extrapolation scheme is extended to the dipole moment, d:

$$d_{\rm c}^{\rm CBS} = \frac{d_{\rm c}^{X} X^3 - d_{\rm c}^{X-1}(X-1)^3}{X^3 - (X-1)^3}, \tag{1.61}$$

where

$$d_{\rm c}^{X} = d^{X} - d_{\rm HF}^{X}, \tag{1.62}$$

and d^X and $d_{\rm HF}^X$ are the correlated dipole moment and Hartree-Fock dipole moment obtained in a XZ basis set, respectively.

Explicitly correlated wave-function approaches

A family of wave-function methods aiming to attenuate the finite basis-set error with the use of an explicit correlation function have been introduced: the so-called F12 methods [35, 36]. There, one adds geminal functions in addition to the conventional orbitals defined in a given basis set. These geminal functions aim to capture the correlation effects outside the basis set and are defined as

$$w_{ij}(\mathbf{x}_1, \mathbf{x}_2) = \hat{Q}_{12} f(r_{12}) \hat{S} \frac{1}{\sqrt{2}} (\phi_i(\mathbf{x}_1)\phi_j(\mathbf{x}_2) - \phi_i(\mathbf{x}_2)\phi_j(\mathbf{x}_1)) \tag{1.63}$$

where $f(r_{12})$ is the correlation factor and is commonly expressed as

$$f(r_{12}) = \frac{1}{\gamma}(1 - \exp(-\gamma r_{12})). \tag{1.64}$$

In the last equation, γ is a parameter which permits to control the spatial range to be considered in the correlation factor. In Fig. 1.7, we compute this correlation factor as a function of r_{12} for different values of γ. We notice that small values of γ lead to a broader correlation factor.

In the geminal function (Eq. (1.63)), $\hat{Q}_{12}$ is a strong-orthogonality projector. It ensures orthogonality between the geminals and the orbitals used for the expansion of the wave function in the basis set. It is commonly expressed as

$$\hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2)(1 - \hat{V}_1\hat{V}_2), \tag{1.65}$$

where $\hat{O}_n = \sum_i^{\rm occ.} \chi_i(\mathbf{x}_n)\chi_i^*(\mathbf{x}_n)$ and $\hat{V}_n = \sum_a^{\rm virt.} \chi_a(\mathbf{x}_n)\chi_a^*(\mathbf{x}_n)$ are projectors into the occupied and standard virtual spaces, respectively, and $n \in \{1, 2\}$.

Finally, $\hat{S}$ is a rational generator which ensures to recover s- and p-wave coalescence conditions.

Thus, the F12 approach is a rigorous basis-set correction method based on wave-function theory but suffers from the need of an auxiliary basis set that complicates significantly the computations.

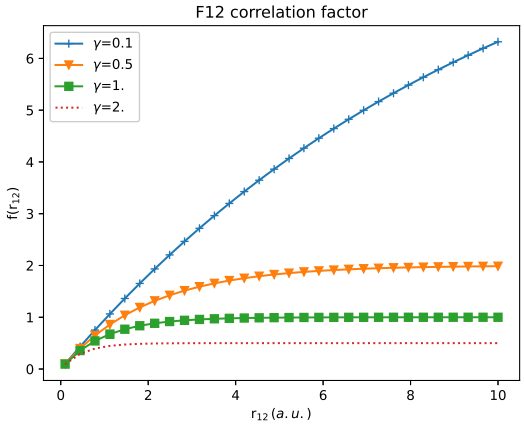

Figure 1.7: Correlation factor as a function of r_{12} for different values of γ. Small values of γ leads to wide range for $f(r_{12})$.

1.4 Density-based basis-set correction

In Secs. 1.1 and 1.2, we introduced wave-function and density-functional methods and emphasis was placed on their strengths and weaknesses. On the one side, wave-function methods offer applicability to a large variety of systems with sufficient accuracy but suffers from the computational cost that grows rapidly with the system size. On the other side, density-functional approximations are less systematic and accurate but allow applicability to larger systems.

In the so-called density-based basis-set correction approach [37], one focuses on the slow convergence of properties mostly due to the unreachable electron-electron cusp behavior. In Ref. [37], the authors propose a decomposition of the Levy-Lieb universal density functional into a basis-set interaction component and a complementary functional of the density that should recover the out-of-the-basis interaction features (see Sec. 1.4.1). Therefore, this approach must be seen as an alternative approach for accelerating basis convergence of wave-function methods. In the same reference, the authors proposed to use the RS-DFT functionals as approximations for this complementary functional based on the similarities between the physics described in both approaches (see Sec. 1.4.2). The density-based basis-set correction has mainly been applied to ground-state energies (see Sec. 1.4.4). In the subsequent chapters of this thesis, we extend this approach to molecular properties.

1.4.1 Decomposition of the Levy-Lieb universal density functional

Starting from the DFT ground-state energy minimization in Eq. (1.27), the density-based basis-set correction approach considers the following decomposition for the Levy-Lieb universal density functional, for density $n^{\mathcal{B}}$ representable in the basis set $\mathcal{B}$,

$$F[n^{\mathcal{B}}] = \min_{\Psi^{\mathcal{B}} \to n^{\mathcal{B}}} \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\text{ee}} | \Psi^{\mathcal{B}} \rangle + \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}], \tag{1.66}$$

where $\Psi^{\mathcal{B}}$ is a multi-determinantal wave function representable in the basis set $\mathcal{B}$, and $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}]$ is the complementary density functional which reads

$$\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}] = \min_{\Psi \to n^{\mathcal{B}}} \langle \Psi | \hat{T} + \hat{W}_{\text{ee}} | \Psi \rangle - \min_{\Psi^{\mathcal{B}} \to n^{\mathcal{B}}} \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\text{ee}} | \Psi^{\mathcal{B}} \rangle. \tag{1.67}$$

Thus, $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}]$ adds the contributions of the wave functions Ψ that are not representable in the basis set $\mathcal{B}$. In this context, the expression for an approximation of the ground-state energy

evaluated in the basis set $\mathcal{B}$ is

$$E_0^{\mathcal{B}} = \min_{n^{\mathcal{B}}} \left\{ F[n^{\mathcal{B}}] + \int \mathrm{d}\mathbf{r}\, v_{\mathrm{ne}}(\mathbf{r}) n^{\mathcal{B}}(\mathbf{r}) \right\}. \tag{1.68}$$

Since in the minimization the densities are restricted to be representable in the basis set, $E_0^{\mathcal{B}}$ is an upper bound to the exact ground-state energy: $E_0^{\mathcal{B}} \geq E_0$. Moreover, in the CBS limit, we have

$$\lim_{\mathcal{B}\to\mathrm{CBS}} E_0^{\mathcal{B}} = E_0. \tag{1.69}$$

Furthermore, the density converges quickly with respect to the basis set. Thus, $E_0^{\mathcal{B}}$ has a faster basis convergence to the CBS limit than the FCI ground-state energy.

From Eqs. (1.66) and $\langle \Psi^{\mathcal{B}} | \hat{V}_{\mathrm{ne}} | \Psi^{\mathcal{B}} \rangle = \int \mathrm{d}\mathbf{r}\, v_{\mathrm{ne}}(\mathbf{r}) n^{\mathcal{B}}(\mathbf{r})$, one can rewrite Eq. (1.68) as the following minimization over wave functions developped in the basis set $\mathcal{B}$,

$$E_0^{\mathcal{B}} = \min_{\Psi^{\mathcal{B}}} \left\{ \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\mathrm{ee}} + \hat{V}_{\mathrm{ne}} | \Psi^{\mathcal{B}} \rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}] \right\}. \tag{1.70}$$

1.4.2 Mapping with RS-DFT

In contrast with RS-DFT, $\bar{E}^{\mathcal{B}}[n]$ is not universal as it depends on the basis set and thus on the system. Another difference with RS-DFT is the use of the full Coulomb interaction $\hat{W}_{\mathrm{ee}}$ instead of the long-range interaction $\hat{W}_{\mathrm{ee}}^{\mathrm{lr},\mu}$. We remind that the wave function $\Psi^{\mathcal{B}}$ is cuspless and therefore could be associated with an effective non-divergent electron-electron interaction. In Ref. [37], the authors introduced an effective local two-electron interaction projected in the basis $\mathcal{B}$, $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)$, with expression[8]

$$W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2) = \frac{f_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)}{n_{2,\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)}, \tag{1.71}$$

where

$$f_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{ijklmn\in\mathcal{B}} V_{ij}^{kl} \Gamma_{kl}^{mn}[\Psi^{\mathcal{B}}] \phi_n(\mathbf{r}_2) \phi_m(\mathbf{r}_1) \phi_i(\mathbf{r}_1) \phi_j(\mathbf{r}_2), \tag{1.72}$$

$V_{ij}^{kl} = \langle ij|kl\rangle$ are the two-electron integrals, $\Gamma_{kl}^{mn}[\Psi^{\mathcal{B}}] = 2\langle \Psi^{\mathcal{B}} | \hat{a}_{m\downarrow}^\dagger \hat{a}_{n\uparrow}^\dagger \hat{a}_{l\uparrow} \hat{a}_{k\downarrow} | \Psi^{\mathcal{B}} \rangle$ is the opposite-spin two-body density matrix of $\Psi^{\mathcal{B}}$, and $n_{2,\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{ijkl\in\mathcal{B}} \phi_i(\mathbf{r}_1) \phi_j(\mathbf{r}_2) \Gamma_{ij}^{kl}[\Psi^{\mathcal{B}}] \phi_k(\mathbf{r}_1) \phi_l(\mathbf{r}_2)$ is the pair density associated with $\Psi^{\mathcal{B}}$. The effective interaction $W_{\Psi^{\mathcal{B}}}$ is non-divergent and tends to the exact W_{ee} interaction in the CBS limit.

This has naturally lead to the mapping of the non-divergent long-range interaction of RS-DFT (Eq. (1.39)) with $W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_2)$. Therefore, in Ref. [37], the authors proposed to use a RS-DFT correlation functional as an approximation for $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}]$. More specifically, the authors chose the multi-determinantal correlation functional as it is build from a multiconfiguration wave function. In this mapping, they link a local and basis-specific definition of the RS-DFT range-separation parameter μ to the non-divergent electron-electron interaction through the following definition:

$$\mu^{\mathcal{B}}(\mathbf{r}) = \frac{\sqrt{\pi}}{2} W_{\Psi^{\mathcal{B}}}(\mathbf{r}, \mathbf{r}). \tag{1.73}$$

The definition for the range-separation parameter comes from imposing the limit

$$\lim_{r_{12}\to 0} \frac{\mathrm{erf}(\mu(\mathbf{r}_1) r_{12})}{r_{12}} = W_{\Psi^{\mathcal{B}}}(\mathbf{r}_1, \mathbf{r}_1), \tag{1.74}$$

where $r_{12} = \|\mathbf{r}_1 - \mathbf{r}_2\|$.

[8] for electrons with opposite spins.

In this framework, the short-range correlation functionals of RS-DFT were used as approximations for $\bar{E}^{\mathcal{B}}[n]$ in Refs. [18,37]. Therefore, the aim of this approximation is to add correlation effects on top of the correlated wave function $\Psi^{\mathcal{B}}$. The first approximation is the LDA-like complementary functional,

$$\bar{E}^{\mathcal{B}}_{\mathrm{LDA}}[n] = \int \mathrm{d}\mathbf{r}\, e^{\mathrm{sr},\mu^{\mathcal{B}}(\mathbf{r}),\mathrm{UEG}}_{\mathrm{c,md}}(n(\mathbf{r}))\,, \tag{1.75}$$

where $e^{\mathrm{sr},\mu,\mathrm{UEG}}_{\mathrm{c,md}}(n)$ is the multi-determinant short-range correlation energy density of the UEG parametrized in Ref. [38].

Then, the PBE-UEG complementary functional was introduced as

$$\bar{E}^{\mathcal{B}}_{\text{PBE-UEG}}[n] = \int \mathrm{d}\mathbf{r}\bar{e}^{\mathrm{sr},\mu^{\mathcal{B}}(\mathbf{r}),\mathrm{PBE}}_{\mathrm{c,md}}(n(\mathbf{r}), \nabla n(\mathbf{r}), n_2^{\mathrm{UEG}}(\mathbf{r})), \tag{1.76}$$

where $\bar{e}^{\mathrm{sr},\mu,\mathrm{PBE}}_{\mathrm{c,md}}(n, \nabla n, n_2)$ is the multi-determinant short-range correlation energy density based on the PBE functional, $n_2^{\mathrm{UEG}}(\mathbf{r})$ is the UEG on-top pair density, with density $n(\mathbf{r})$,

$$n_2^{\mathrm{UEG}}(\mathbf{r}) = n(\mathbf{r})^2\left[1 - \zeta(\mathbf{r})\right] g_0(n(\mathbf{r})), \tag{1.77}$$

$\zeta(\mathbf{r})$ is the spin-polarization, and $g_0(n)$ is the pair-distribution function of the UEG taken from Ref. [39].

The final approximation is the PBE-OT density functional,

$$\bar{E}^{\mathcal{B}}_{\text{PBE-OT}}[n, n_2] = \int \mathrm{d}\mathbf{r}\bar{e}^{\mathrm{sr},\mu^{\mathcal{B}}(\mathbf{r}),\mathrm{PBE}}_{\mathrm{c,md}}(n(\mathbf{r}), \zeta(\mathbf{r}), \nabla n(\mathbf{r}), \mathring{n}_2(\mathbf{r})), \tag{1.78}$$

where $\mathring{n}_2(\mathbf{r})$ is the on-top pair density extrapolated to the limit $\mu \to \infty$ such that

$$\mathring{n}_2(\mathbf{r}) = n_2(\mathbf{r})\left(1 + \frac{2}{\sqrt{\pi}\mu(\mathbf{r})}\right)^{-1}, \tag{1.79}$$

where $n_2(\mathbf{r})$ is the on-top pair density associated with the multi-determinantal wave-function $\Psi^{\mathcal{B}}$.

The so-called spin-unpolarized PBE-OT (SU-PBE-OT) density functional was introduced by imposing $\zeta(\mathbf{r}) = 0$, so that the functional becomes S_z-invariant.

1.4.3 Self-consistent VS non-self-consistent

To solve the minimization over the wave function in Eq. (1.70), one can choose between an a-posteriori correction,

$$E_0^{\mathcal{B}} \simeq \langle \Psi^{\mathcal{B}}_{\mathrm{FCI}} | \hat{T} + \hat{V}_{\mathrm{ne}} + \hat{W}_{\mathrm{ee}} | \Psi^{\mathcal{B}}_{\mathrm{FCI}} \rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}_{\mathrm{FCI}}}], \tag{1.80}$$

where $\Psi^{\mathcal{B}}_{\mathrm{FCI}}$ is the FCI ground-state wave-function. This a-posteriori correction is what we also call the non-self consistent method. This idea comes from the fact that the basis-set correction is small, thus the complementary functional in Eq. (1.67) can be considered as a small perturbation. Therefore, the a-posteriori correction has to be understood as a first-order perturbed ground-state energy.

We can more generally consider the approximation to

$$E_0^{\mathcal{B}} \simeq E^{\mathcal{B}}_{\mathrm{WFT1}} + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}_{\mathrm{WFT2}}}], \tag{1.81}$$

where $E^{\mathcal{B}}_{\mathrm{WFT1}}$ is the ground-state energy evaluated from a wave-function approximation referred as WFT1 using a basis set $\mathcal{B}$, and the density $n_{\Psi^{\mathcal{B}}_{\mathrm{WFT2}}}$ used to compute the correction energy is

evaluated from another wave-function approximation referred as WFT2. Therefore, any post-Hartree-Fock method is suitable for WFT1 and studies, as the one published in Ref. [40], show that a cheaper density is sufficient for the evaluation of $\bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}_{\text{WFT2}}}]$. This translates on the possibility to use a cheap ansatz as Hartree-Fock for WFT2.

Another strategy is to write the Euler-Lagrange equation associated with Eq. (1.70) with the constraint $\langle \Psi^{\mathcal{B}} | \Psi^{\mathcal{B}} \rangle = 1$ and Lagrange multiplier $\mathcal{E}^{\mathcal{B}}$. This leads to the following eigenvalue equation:

$$\left(\hat{T}^{\mathcal{B}} + \hat{V}^{\mathcal{B}}_{\text{ne}} + \hat{W}^{\mathcal{B}}_{\text{ee}} + \hat{\bar{V}}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}] \right) |\Psi^{\mathcal{B}}\rangle = \mathcal{E}^{\mathcal{B}} |\Psi^{\mathcal{B}}\rangle, \tag{1.82}$$

where $\hat{\bar{V}}^{\mathcal{B}}[n] = \int \mathrm{d}\mathbf{r} \left(\delta \bar{E}^{\mathcal{B}}[n] / \delta n^{\mathcal{B}}(\mathbf{r}) \right) \hat{n}(\mathbf{r})$ is the basis-set correction potential operator. The superscript $\mathcal{B}$ in $\hat{T}^{\mathcal{B}}$, $\hat{V}^{\mathcal{B}}_{\text{ne}}$, $\hat{W}^{\mathcal{B}}_{\text{ee}}$, and $\hat{n}^{\mathcal{B}}(\mathbf{r})$ means that the operators are projected in the basis set $\mathcal{B}$. Because it depends on the density $n_{\Psi^{\mathcal{B}}}$, Eq. (1.82) must be solved using a self-consistent scheme. Finally, $E^{\mathcal{B}}_0$ can be obtained as

$$E^{\mathcal{B}}_0 = \mathcal{E}^{\mathcal{B}} - \langle \Psi^{\mathcal{B}} | \hat{\bar{V}}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}] | \Psi^{\mathcal{B}} \rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}]. \tag{1.83}$$

To apply the basis-set correction to excited states, we use the density $n^{\mathcal{B}}_k$ of the k^{th} state to compute the complementary energy. Therefore, the energy of the k^{th} state is

$$E^{\mathcal{B}}_k = E^{\mathcal{B}}_{k,\text{FCI}} + \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}_k]. \tag{1.84}$$

This assumes that the basis-set correction functional $\bar{E}^{\mathcal{B}}[n]$ does not depend on the state, which is an approximation. An extension of the basis-set correction method to the linear-response theory is introduced in Chap. 5.

1.4.4 Applications

The non-self-consistent method in Eq. (1.81) was employed in several published works. In Ref. [37], it is used for the computation of ground-state energies and ionization potentials of atomic systems up to $Z = 10$, using FCI and CIPSI energies with the LDA-like complementary functional in Eq. (1.75). This first application of the basis-set correction method has shown that chemically accurate results are obtained using a TZ basis set. However, the use of a LDA-like framework leads to a systematic overestimation of the correlation energy.

For applications on strongly correlated molecular systems, the authors in Ref. [41] used FCI and CIPSI energies with LDA (Eq. (1.75)) and PBE-OT (Eq. (1.78)) basis-set corrections. In this work, the authors focused on dissociation curves where one often encounters strong correlation behavior.

In Refs. [18,42], the method has been applied on a wide range of weakly correlated molecular systems, namely the G2 set of molecules. They use CCSD(T) and SHCI[9] energies in combination with the PBE-UEG and PBE-OT functionals for the computation of atomization energies.

Still in the spirit of the strong-correlation problem, Ref. [43] provides application of the basis-set correction method to the challenging case of transition-metal systems using SHCI energies in combination with the PBE-UEG and the PBE-OT basis-set corrections. In the latter work, the use of TZ basis set leads to chemically accurate total, ionization and dissociation energies, where most of the results need 5Z basis sets without the basis-set correction contribution.

[9] the semistochastic heat-bath configuration interaction is a selected CI method.

Encouraged by the positive conclusions of the latter works, in Ref. [44], the same strategy is applied to various types of excited states (valence, Rydberg, double excitations), vertical and adiabatic using CIPSI energies, and LDA, PBE-UEG and, PBE-OT basis-set corrections.

An application to a many-body perturturbation theory method was also proposed in Ref. [19]. In that context, one should consider the basis-set correction potential. In Ref. [19], the authors compute the ionization energies in a GW[10] formulation. Using the GW approach, the quantity of interest is no longer the wave function but the Green's function G which is a time-dependent generalization of the density matrix. This approach also considers a non-interacting quantity, the non-interacting Green's function, G_0. We also define the inverse of the projection of the latter quantities on the basis set $\mathcal{B}$, $(G^{\mathcal{B}})^{-1}$ and $(G_0^{\mathcal{B}})^{-1}$, which are linked through the following Dyson equality, in the context of the basis-set correction method:

$$(G^{\mathcal{B}})^{-1} = (G_0^{\mathcal{B}})^{-1} - \Sigma_{\text{Hxc}}^{\mathcal{B}}[G^{\mathcal{B}}] - \bar{\Sigma}^{\mathcal{B}}[n_{G^{\mathcal{B}}}]. \tag{1.85}$$

In Eq. (1.85), $\Sigma_{\text{Hxc}}^{\mathcal{B}}[G^{\mathcal{B}}]$ is the Hartree-exchange-correlation self-energy in the basis and $\bar{\Sigma}^{\mathcal{B}}[n_{G^{\mathcal{B}}}]$ is the self energy coming from the functional derivative of the complementary basis-correction density functional:

$$\bar{\Sigma}^{\mathcal{B}}[n](\mathbf{r}, \mathbf{r}') = \bar{v}^{\mathcal{B}}[n](\mathbf{r})\delta(\mathbf{r} - \mathbf{r}'), \tag{1.86}$$

where $\bar{v}^{\mathcal{B}}[n](\mathbf{r}) = \delta \bar{E}^{\mathcal{B}}[n]/\delta n(\mathbf{r})$. The author of Ref. [19] applied this method to ionization potentials and they show that their basis convergences are significantly speeded up for a wide range of small and larger molecules.

[10]The GW method is a Green function based method mainly used for electronic band structure calculations in solids (see Ref. [45]). For the understanding of this paragraph, one should only keep in mind the fact that the GW approximation is a many-body perturbation theory method.

Bibliography

[1] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, Inc., 1996).

[2] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic Structure Theory* (John Wiley & Sons, LTD, 2000).

[3] V. Gupta, *Principles and Applications of Quantum Chemistry* (Elsevier, Inc., 2015).

[4] B. Huron, J. P. Malrieu and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).

[5] S. Evangelisti, J.-P. Daudey and J.-P. Malrieu, Chem. Phys. **75**, 91 (1983).

[6] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Véril, M. Caffarel, P.-F. Loos, E. Giner and A. Scemama, J. Chem. Theory. Comput. **15**, 3591 (2019).

[7] E. Giner, *Méthodes d'interaction de configurations et Monte Carlo quantique : marier le meilleur des deux mondes*, Tech. rep. (oct 2014).

[8] F. Jensen, *Introduction to computational chemistry* (John Wiley & Sons, LTD, 2017).

[9] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[10] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[11] J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[12] J. Toulouse, Review of approximations for the exchange-correlation energy in density-functional theory, `https://arxiv.org/abs/2103.02645` (2021).

[13] J. Toulouse, F. Colonna and A. Savin, Phys. Rev. A **70**, 062505 (2004).

[14] I. C. Gerber and J. G. Angyán, Chem. Phys. Lett. **415**, 100 (2005).

[15] E. Fromager, J. Toulouse and H. J. A. Jensen, J. Chem. Phys. **126**, 074111 (2007).

[16] J. Toulouse, P. Gori-Giorgi and A. Savin, Theor. Chem. Acc **114**, 305 (2005).

[17] A. Ferté, E. Giner and J. Toulouse, J. Chem. Phys. **150**, 084103 (2019).

[18] P.-F. Loos, B. Pradines, A. Scemama, J. Toulouse and E. Giner, J. Phys. Chem. Lett. **10**, 2931 (2019).

[19] P.-F. Loos, B. Pradines, A. Scemama, E. Giner and J. Toulouse, J. Chem. Theory Comput. **16**, 1018 (2020).

[20] E. R. Davidson and D. Feller, Chem. Rev. **86**, 681 (1986).

[21] F. Jensen, Wiley Interdisciplinary Reviews: Computational Molecular Science **3**, 273 (2013).

[22] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[23] D. E. Woon and T. H. Dunning, J. Chem. Phys. **98**, 1358 (1993).

[24] B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning and A. K. Wilson, Theor. Chem. Acc **128**, 69 (2011).

[25] R. A. Kendall, T. H. Dunning and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[26] D. E. Woon and T. H. Dunning, J. Chem. Phys. **100**, 2975 (1994).

[27] D. E. Woon and T. H. Dunning, J. Chem. Phys. **103**, 4572 (1995).

[28] K. A. Peterson and T. H. Dunning, J. Chem. Phys. **117**, 10548 (2002).

[29] G. Hill, Correlation consistent basis set repository, `http://www.grant-hill.group.shef.ac.uk/ccrepo/index.html` (2016).

[30] E. A. Hylleraas, in *Advances in Quantum Chemistry* (Elsevier, 1964), vol. 1, pp. 1–33.

[31] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).

[32] R. T. Pack and W. B. Brown, J. Chem. Phys. **45**, 556 (1966).

[33] T. Helgaker, W. Klopper, H. Koch and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[34] A. Halkier, W. Klopper, T. Helgaker and P. Jørgensen, J. Chem. Phys. **111**, 4424 (1999).

[35] L. Kong, F. A. Bischoff and E. F. Valeev, Chem. Rev. **112**, 75 (2012).

[36] C. Hättig, W. Klopper, A. Köhn and D. P. Tew, Chem. Rev. **112**, 4 (2012).

[37] E. Giner, B. Pradines, A. Ferté, R. Assaraf, A. Savin and J. Toulouse, J. Chem. Phys. **149**, 194301 (2018).

[38] S. Paziani, S. Moroni, P. Gori-Giorgi and G. B. Bachelet, Phys. Rev. B **73**, 155111 (2006).

[39] P. Gori-Giorgi and A. Savin, Phys. Rev. A **73**, 032506 (2006).

[40] D. Traore, J. Toulouse and E. Giner, J. Chem. Phys. **156**, 174101 (2022).

[41] E. Giner, A. Scemama, P.-F. Loos and J. Toulouse, J. Chem. Phys. **152**, 174104 (2020).

[42] Y. Yao, E. Giner, J. Li, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **153**, 124117 (2020).

[43] Y. Yao, E. Giner, T. A. Anderson, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **155**, 204104 (2021).

[44] E. Giner, A. Scemama, J. Toulouse and P.-F. Loos, J. Chem. Phys. **151**, 144118 (2019).

[45] M. J. van Setten, F. Weigend and F. Evers, J. Chem. Theory. Comput. **9**, 232 (2013).

2

Self-consistent density-based basis-set correction: How much do we lower total energies and improve dipole moments?

This chapter corresponds to the article [E. Giner, D. Traore, B. Bradines, J. Toulouse, J. Chem. Phys. **155**, 044109 (2021)].

This work provides a self-consistent extension of the recently proposed density-based basis-set correction method for wave-function electronic-structure calculations [E. Giner, B. Pradines, A. Ferté, R. Assaraf, A. Savin and J. Toulouse, J. Chem. Phys. **149**, 194301 (2018)]. In contrast to the previously used approximation where the basis-set correction density functional was a posteriori added to the energy from a wave-function calculation, here the energy minimization is performed including the basis-set correction. Compared to the non-self-consistent approximation, this allows one to lower the total energy and change the wave function under the effect of the basis-set correction. This work addresses two main questions: i) What is the change in total energy compared to the non-self-consistent approximation, and ii) can we obtain better properties, namely dipole moments, with the basis-set corrected wave functions? We implement the present formalism with two different basis-set correction functionals and test it on different molecular systems. The main results of the study are that i) the total energy lowering obtained by the self-consistent approach is extremely small, which justifies the use of the non-self-consistent approximation, and ii) the dipole moments obtained from the basis-set corrected wave functions are improved, being already close to their complete-basis-set values with triple-zeta basis sets. Thus, the present study further confirms the soundness of the density-based basis-set correction scheme.

2.1 Introduction

One of the main limitations of electronic-structure calculations based on wave-function theory (WFT) is the slow convergence of the results with respect to the size of the one-electron basis set. This aspect is of fundamental importance in quantum chemistry as WFT methods have otherwise many interesting features. In particular, in a given basis set $\mathcal{B}$, WFT methods can

usually be systematically improved toward the exact solution provided by full configuration interaction (FCI). At the root of the slow convergence of WFT lies the singularity of the Coulomb interaction: because it becomes infinite at electron-electron coalescence points, it creates a derivative discontinuity of the wave function at these points, the so-called electron-electron cusp [1,2], which is not representable by the usual incomplete basis sets employed. To cure the slow convergence problem of WFT, two main approaches have emerged: basis-set extrapolation techniques [3,4] and explicitly correlated F12 methods [5–10]. Basis-set extrapolation techniques consist in exploiting the known asymptotic behavior of WFT properties as a function of the size of the basis set in order to estimate the complete-basis-set (CBS) limit based on several calculations with basis sets of increasing sizes. Explicitly correlated F12 methods consist in supplementing to the usual basis sets a correlation factor which restores the electron-electron cusp and accelerates the convergence toward the CBS limit. Although these F12 methods are increasingly popular in quantum chemistry [11], they have the drawback of needing rather complex three- and four-electron integrals [12] and, more generally, of involving a relatively complex mathematical formalism which makes the adaptation of a WFT method to its F12 version a non trivial task.

An alternative path to speed up the convergence of WFT calculations with respect to the size of the basis set has been recently proposed by some of the present authors [13] by exploiting the ability of range-separated density-functional theory (RSDFT) to recover the short-range correlation effects missing from an incomplete basis set. The central idea developed in Ref. [13] is to define a mapping between the electron-electron Coulomb interaction projected into an incomplete basis set $\mathcal{B}$ and the non-diverging long-range electron-electron interaction $\mathrm{erf}(\mu r_{12})/r_{12}$ used in RSDFT. The connection is done through the definition of a range-separation parameter μ which varies in space and automatically adapts to the basis set $\mathcal{B}$. Once this adaptive range-separation parameter is defined, one can use a special flavor of short-range correlation density functionals used in RSDFT for the estimation of the correlation energy missing in the considered basis set $\mathcal{B}$. An important property of this RSDFT-based approach is that the basis-set energy correction properly vanishes in the CBS limit. This strategy was successfully validated for the calculations of ionization potentials [13, 14], molecular atomization energies [15–17], and excitation energies [18].

All the previous applications of this method rely on a non-self-consistent approximation in which the basis-set energy correction is just added a posteriori to a good estimate of the FCI energy in a given basis set $\mathcal{B}$. In the present work, we go beyond this approximation and develop a self-consistent formalism in order to answer two distinct questions: i) How crude is the non-self-consistent approximation for total energies; and ii) can the self-consistent formalism yield effective wave functions with better properties.

The paper is organized as follows. In Section 2.2.1 we present the exact theory of the self-consistent basis-set correction scheme using either a functional of the density only or a functional of both the density and the on-top pair density and we recall the non-self-consistent approximation previously employed. In Section 2.2.2 we introduce our approximations of the unknown exact basis-set correction functional by short-range functionals. In Section 2.2.3 we explain how we solve the self-consistent basis-set correction equations by a selected configuration interaction algorithm. In Section 2.3, we report and discuss results on the total energies of the Be atom and the BH molecule, and on the dipole moments of the BH, FH, H_2O, and CH_2 molecules. Section 4.4.3 contains our conclusions. Unless otherwise specified, Hartree (Ha) atomic units (a.u.) are used throughout the paper.

2.2 Theory

2.2.1 Self-consistent basis-set correction

Basis-set correction as a functional of the density

We start by reviewing the scheme where the basis-set correction is written as a functional of the density, which was first developed in Ref. [13]. Given a basis set $\mathcal{B}$, the exact ground-state energy E_0 of an electronic system can be approximated by the energy $E_0^{\mathcal{B}}$ obtained by the following minimization over $\mathcal{B}$-representable one-electron densities $n^{\mathcal{B}}$, i.e. densities that can be obtained from a wave function $\Psi^{\mathcal{B}}$ belonging to the N-electron Hilbert space generated by the basis set $\mathcal{B}$,

$$E_0^{\mathcal{B}} = \min_{n^{\mathcal{B}}} \left\{ F[n^{\mathcal{B}}] + \int \mathrm{d}\mathbf{r}\; v_{\mathrm{ne}}(\mathbf{r})\, n^{\mathcal{B}}(\mathbf{r}) \right\}, \tag{2.1}$$

where $v_{\mathrm{ne}}(\mathbf{r})$ is the nuclei-electron interaction potential. In this expression, $F[n^{\mathcal{B}}]$ is the standard Levy-Lieb constrained-search universal density functional [19, 20] evaluated at $n^{\mathcal{B}}$

$$F[n^{\mathcal{B}}] = \min_{\Psi \to n^{\mathcal{B}}} \langle \Psi | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi \rangle, \tag{2.2}$$

where $\hat{T}$ and $\hat{W}_{\mathrm{ee}}$ are the kinetic-energy operator and the Coulomb electron-electron interaction operator, respectively, and the notation $\Psi \to n^{\mathcal{B}}$ means a N-electron wave function yielding the density $n^{\mathcal{B}}$. It is important to notice that the wave functions Ψ used in the definition of $F[n^{\mathcal{B}}]$ in Eq. (2.2) are not restricted to be expandable in the basis set $\mathcal{B}$ but should instead be thought of as expanded on a complete basis set. The minimizing density $n_0^{\mathcal{B}}$ in Eq. (2.1) can be considered as the best variational approximation to the exact ground-state density n_0. Importantly, when the basis set $\mathcal{B}$ is increased up to the CBS limit, the density $n_0^{\mathcal{B}}$ and energy $E_0^{\mathcal{B}}$ converge to the exact ground-state density n_0 and energy E_0, respectively,

$$\lim_{\mathcal{B}\to\mathrm{CBS}} n_0^{\mathcal{B}} = n_0 \quad \text{and} \quad \lim_{\mathcal{B}\to\mathrm{CBS}} E_0^{\mathcal{B}} = E_0. \tag{2.3}$$

Since the $\mathcal{B}$-representability restriction is only applied to densities and not to wave functions, the basis-set convergence of $E_0^{\mathcal{B}}$ to E_0 is much faster than in a usual WFT calculation.

We then decompose the universal Levy-Lieb density functional as

$$F[n^{\mathcal{B}}] = \min_{\Psi^{\mathcal{B}} \to n^{\mathcal{B}}} \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\mathrm{ee}} | \Psi^{\mathcal{B}} \rangle + \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}], \tag{2.4}$$

where $\Psi^{\mathcal{B}}$ designates wave functions restricted to the N-electron Hilbert space generated by the basis set $\mathcal{B}$, and $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}]$ is the complementary basis-set correction density functional required to make Eq. (2.4) exact. This basis-set correction functional $\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}]$ recovers the part of the energy that is missing in the first term of the right-hand side of Eq. (2.4) due to the basis-set restriction of the wave functions $\Psi^{\mathcal{B}}$. Inserting Eq. (2.4) into Eq. (2.1) and recombining the two minimizations, we can obtain $E_0^{\mathcal{B}}$ by the following minimization over $\mathcal{B}$-restricted wave functions $\Psi^{\mathcal{B}}$

$$E_0^{\mathcal{B}} = \min_{\Psi^{\mathcal{B}}} \left\{ \langle \Psi^{\mathcal{B}} | \hat{T} + \hat{W}_{\mathrm{ee}} + \hat{V}_{\mathrm{ne}} | \Psi^{\mathcal{B}} \rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}] \right\}, \tag{2.5}$$

where $\hat{V}_{\mathrm{ne}}$ is the nuclei-electron interaction operator and $n_{\Psi^{\mathcal{B}}}(\mathbf{r}) = \langle \Psi^{\mathcal{B}} | \hat{n}(\mathbf{r}) | \Psi^{\mathcal{B}} \rangle$ is the density of the wave function $\Psi^{\mathcal{B}}$, where we have introduced the density operator $\hat{n}(\mathbf{r}) = \sum_{\sigma\in\{\uparrow,\downarrow\}} \hat{\psi}_\sigma^\dagger(\mathbf{r}) \hat{\psi}_\sigma(\mathbf{r})$ written in real-space second quantization. The minimizing wave function $\Psi_0^{\mathcal{B}}$ in Eq. (2.5) satisfies the following self-consistent Schrödinger-like equation:

$$\hat{H}_{\mathrm{eff}}^{\mathcal{B}}[n_{\Psi_0^{\mathcal{B}}}] | \Psi_0^{\mathcal{B}} \rangle = \mathcal{E}_0^{\mathcal{B}} | \Psi_0^{\mathcal{B}} \rangle, \tag{2.6}$$

where $\mathcal{E}_0^{\mathcal{B}}$ is the Lagrange multiplier associated with the normalization constraint of the wave function $\Psi_0^{\mathcal{B}}$ and the effective Hamiltonian is defined for a given density $n^{\mathcal{B}}$ as

$$\hat{H}_{\text{eff}}^{\mathcal{B}}[n^{\mathcal{B}}] = \hat{T}^{\mathcal{B}} + \hat{W}_{\text{ee}}^{\mathcal{B}} + \hat{V}_{\text{ne}}^{\mathcal{B}} + \hat{\bar{V}}^{\mathcal{B}}[n^{\mathcal{B}}]. \tag{2.7}$$

In this expression, $\hat{T}^{\mathcal{B}}$, $\hat{W}_{\text{ee}}^{\mathcal{B}}$, and $\hat{V}_{\text{ne}}^{\mathcal{B}}$ are the kinetic, electron-electron, and electron-nuclei operators projected in the N-electron Hilbert space generated by the basis set $\mathcal{B}$, and $\hat{\bar{V}}^{\mathcal{B}}[n^{\mathcal{B}}]$ is the one-electron effective potential operator,

$$\hat{\bar{V}}^{\mathcal{B}}[n^{\mathcal{B}}] = \int \mathrm{d}\mathbf{r}\, \bar{v}^{\mathcal{B}}(\mathbf{r})\, \hat{n}^{\mathcal{B}}(\mathbf{r}), \tag{2.8}$$

where $\bar{v}^{\mathcal{B}}(\mathbf{r}) = \delta \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}]/\delta n^{\mathcal{B}}(\mathbf{r})$ and $\hat{n}^{\mathcal{B}}(\mathbf{r})$ is the density operator projected in the basis set $\mathcal{B}$.

Using real-valued spatial orthonormal orbitals $\{\phi_p\}$ spanning the same space as the basis set $\mathcal{B}$, the expression of the effective Hamiltonian in second quantization is

$$\hat{H}_{\text{eff}}^{\mathcal{B}}[n^{\mathcal{B}}] = \sum_{pq}^{\mathcal{B}} (h_{pq} + \bar{v}_{pq}^{\mathcal{B}}) \hat{E}_{pq} + \frac{1}{2} \sum_{pqrs}^{\mathcal{B}} w_{pqrs} \hat{e}_{pqrs}, \tag{2.9}$$

where $\hat{E}_{pq} = \hat{a}_{p\uparrow}^{\dagger} \hat{a}_{q\uparrow} + \hat{a}_{p\downarrow}^{\dagger} \hat{a}_{q\downarrow}$ and $\hat{e}_{pqrs} = \hat{E}_{pr}\hat{E}_{qs} - \delta_{qr}\hat{E}_{ps}$ are the spin-singlet one- and two-particle elementary operators, h_{pq} are the usual one-electron integrals, $w_{pqrs} = \langle pq|rs\rangle$ are the usual two-electron integrals, and $\bar{v}_{pq}^{\mathcal{B}}$ are the one-electron integrals associated with the effective potential $\bar{v}^{\mathcal{B}}(\mathbf{r})$,

$$\bar{v}_{pq}^{\mathcal{B}} = \int \mathrm{d}\mathbf{r} \phi_p(\mathbf{r}) \bar{v}^{\mathcal{B}}(\mathbf{r}) \phi_q(\mathbf{r}). \tag{2.10}$$

Note that in Eq. (2.9), we put $\mathcal{B}$ on top of the sum symbols to indicate that the sums run over all orbitals generated by the basis set $\mathcal{B}$.

Finally, note that we have considered the total density $n^{\mathcal{B}}$ for simplicity but the theory can be trivially extended to spin densities $n_{\uparrow}^{\mathcal{B}}$ and $n_{\downarrow}^{\mathcal{B}}$.

Basis-set correction as a functional of density and on-top pair density

We now extend the theory to allow for a basis-set correction functional depending on both the density $n_{\Psi^{\mathcal{B}}}(\mathbf{r})$ and the on-top pair density $n_{2,\Psi^{\mathcal{B}}}(\mathbf{r}) = \langle \Psi^{\mathcal{B}}|\hat{n}_2(\mathbf{r})|\Psi^{\mathcal{B}}\rangle$ of a wave function $\Psi^{\mathcal{B}}$, where we have introduced the on-top pair density operator $\hat{n}_2(\mathbf{r}) = \sum_{\sigma\in\{\uparrow,\downarrow\}} \sum_{\sigma'\in\{\uparrow,\downarrow\}} \hat{\psi}_{\sigma}^{\dagger}(\mathbf{r})\hat{\psi}_{\sigma'}^{\dagger}(\mathbf{r})\hat{\psi}_{\sigma'}(\mathbf{r})\hat{\psi}_{\sigma}(\mathbf{r})$. In the spirit of the generalized Kohn-Sham scheme [21] (see also Ref. [22]), we write the universal Levy-Lieb density functional as

$$F[n^{\mathcal{B}}] = \min_{\Psi^{\mathcal{B}} \to n^{\mathcal{B}}} \left\{ \langle \Psi^{\mathcal{B}}|\hat{T} + \hat{W}_{\text{ee}}|\Psi^{\mathcal{B}}\rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}, n_{2,\Psi^{\mathcal{B}}}] \right\}, \tag{2.11}$$

where $\bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}, n_{2,\Psi^{\mathcal{B}}}]$ can be any functional of $n_{\Psi^{\mathcal{B}}}$ and $n_{2,\Psi^{\mathcal{B}}}$ such that the minimization in Eq. (2.11) exactly gives $F[n^{\mathcal{B}}]$. Insertion into Eq. (2.1) leads to

$$E_0^{\mathcal{B}} = \min_{\Psi^{\mathcal{B}}} \left\{ \langle \Psi^{\mathcal{B}}|\hat{T} + \hat{W}_{\text{ee}} + \hat{V}_{\text{ne}}|\Psi^{\mathcal{B}}\rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}, n_{2,\Psi^{\mathcal{B}}}] \right\}, \tag{2.12}$$

and the minimizing wave function $\Psi_0^{\mathcal{B}}$ satisfies the following self-consistent Schrödinger-like equation:

$$\hat{H}_{\text{eff}}^{\mathcal{B}}[n_{\Psi_0^{\mathcal{B}}}, n_{2,\Psi_0^{\mathcal{B}}}]|\Psi_0^{\mathcal{B}}\rangle = \mathcal{E}_0^{\mathcal{B}}|\Psi_0^{\mathcal{B}}\rangle, \tag{2.13}$$

where the effective Hamiltonian is defined for a given $\mathcal{B}$-representable density $n^{\mathcal{B}}$ and on-top pair density $n_2^{\mathcal{B}}$ as

$$\hat{H}_{\text{eff}}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}] = \hat{T}^{\mathcal{B}} + \hat{W}_{\text{ee}}^{\mathcal{B}} + \hat{V}_{\text{ne}}^{\mathcal{B}} + \hat{\bar{V}}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}] + \hat{\bar{W}}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}]. \tag{2.14}$$

In this expression, $\hat{\bar{V}}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}]$ is the one-electron effective potential operator,

$$\hat{\bar{V}}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}] = \int \mathrm{d}\mathbf{r}\; \bar{v}^{\mathcal{B}}(\mathbf{r})\; \hat{n}^{\mathcal{B}}(\mathbf{r}), \tag{2.15}$$

where $\bar{v}^{\mathcal{B}}(\mathbf{r}) = \delta \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}]/\delta n^{\mathcal{B}}(\mathbf{r})$, and $\hat{\bar{W}}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}]$ is the two-electron effective interaction operator,

$$\hat{\bar{W}}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}] = \frac{1}{2} \int \mathrm{d}\mathbf{r}\; \bar{w}^{\mathcal{B}}(\mathbf{r})\; \hat{n}_2^{\mathcal{B}}(\mathbf{r}), \tag{2.16}$$

where $\bar{w}^{\mathcal{B}}(\mathbf{r}) = 2\delta \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}]/\delta n_2^{\mathcal{B}}(\mathbf{r})$ and $\hat{n}_2^{\mathcal{B}}(\mathbf{r})$ is the on-top pair density operator projected in the basis set $\mathcal{B}$. The second-quantized expression of the effective Hamiltonian is

$$\hat{H}_{\mathrm{eff}}^{\mathcal{B}}[n^{\mathcal{B}}, n_2^{\mathcal{B}}] = \sum_{pq}^{\mathcal{B}} (h_{pq} + \bar{v}_{pq}^{\mathcal{B}}) \hat{E}_{pq} + \frac{1}{2} \sum_{pqrs}^{\mathcal{B}} \left(w_{pqrs} + \bar{w}_{pqrs}^{\mathcal{B}} \right) \hat{e}_{pqrs}, \tag{2.17}$$

where, as before, $\bar{v}_{pq}^{\mathcal{B}}$ are the one-electron integrals associated with the effective potential $\bar{v}^{\mathcal{B}}(\mathbf{r})$, and $\bar{w}_{pqrs}^{\mathcal{B}}$ are the two-electron integrals associated with the effective interaction $\bar{w}^{\mathcal{B}}(\mathbf{r})$,

$$\bar{w}_{pqrs}^{\mathcal{B}} = \int \mathrm{d}\mathbf{r} \phi_p(\mathbf{r}) \phi_r(\mathbf{r}) \bar{w}^{\mathcal{B}}(\mathbf{r}) \phi_q(\mathbf{r}) \phi_s(\mathbf{r}). \tag{2.18}$$

Of course, since the effective Hamiltonians in Eqs. (2.7) and (2.14) are different, their respective ground-state wave function $\Psi_0^{\mathcal{B}}$ are also different, even though we used the same notation.

Non-self-consistent approximation

In previous works [13,15,17,18], the minimization in Eq. (2.5) or (2.12) was not performed but the minimizing wave function $\Psi_0^{\mathcal{B}}$ was simply approximated by the standard FCI wave function $\Psi_{\mathrm{FCI}}^{\mathcal{B}}$ (or an estimate of it) in the basis set $\mathcal{B}$

$$\Psi_0^{\mathcal{B}} \approx \Psi_{\mathrm{FCI}}^{\mathcal{B}}, \tag{2.19}$$

leading to the following approximation for $E_0^{\mathcal{B}}$, for the basic theory of Section 2.2.1:

$$E_0^{\mathcal{B}} \approx E_{\mathrm{FCI}}^{\mathcal{B}} + \bar{E}^{\mathcal{B}}[n_{\Psi_{\mathrm{FCI}}^{\mathcal{B}}}], \tag{2.20}$$

where $E_{\mathrm{FCI}}^{\mathcal{B}}$ is the standard FCI energy (or an estimate of it) in the basis set $\mathcal{B}$, and for the extended theory of Section 2.2.1,

$$E_0^{\mathcal{B}} \approx E_{\mathrm{FCI}}^{\mathcal{B}} + \bar{E}^{\mathcal{B}}[n_{\Psi_{\mathrm{FCI}}^{\mathcal{B}}}, n_{2,\Psi_{\mathrm{FCI}}^{\mathcal{B}}}]. \tag{2.21}$$

The approximation in Eq. (2.19) is in fact equivalent to approximating the minimizing density $n_0^{\mathcal{B}}$ in Eq. (2.1) by the standard FCI ground-state density $n_{\Psi_{\mathrm{FCI}}^{\mathcal{B}}}$,

$$n_0^{\mathcal{B}}(\mathbf{r}) \approx n_{\Psi_{\mathrm{FCI}}^{\mathcal{B}}}(\mathbf{r}), \tag{2.22}$$

which seems intuitively a reasonable approximation as one expects $n_{\Psi_{\mathrm{FCI}}^{\mathcal{B}}}$ and $n_0^{\mathcal{B}}$ to be both close to the exact density n_0, and the encouraging numerical results obtained for energies with this non-self-consistent approximation tend to confirm the validity of Eq. (2.22). Nevertheless, in the present study, we will investigate the quantitative effect on energies and dipole moments of performing the minimization in Eq. (2.5) or (2.12).

2.2.2 Approximations for the basis-set correction functional $\bar{E}^{\mathcal{B}}$

Local range-separation parameter

As originally proposed in Ref. [13], the basis-set correction functional $\bar{E}^{\mathcal{B}}$ can be mapped to the so-called short-range correlation functional with multideterminant reference introduced in Ref. [23] in the context of RSDFT. This mapping relies on the definition of a local range-separation parameter $\mu^{\mathcal{B}}(\mathbf{r})$ [13],

$$\mu^{\mathcal{B}}(\mathbf{r}) = \frac{\sqrt{\pi}}{2} W^{\mathcal{B}}(\mathbf{r}) \tag{2.23}$$

which provides a local measure of the incompleteness of the basis set $\mathcal{B}$. It is defined such that the long-range electron-electron interaction of RSDFT, $w^{\mathrm{lr}}(r_{12}) = \mathrm{erf}(\mu r_{12})/r_{12}$, coincides at coalescence (i.e., at $r_{12} = 0$) with an effective interaction representing the Coulomb electron-electron interaction projected in the basis set $\mathcal{B}$. The expression of this effective interaction at coalescence is [13]

$$W^{\mathcal{B}}(\mathbf{r}) = \begin{cases} \frac{f^{\mathcal{B}}_{\Psi^{\mathcal{B}}_{\mathrm{loc}}}(\mathbf{r})}{n_{2,\Psi^{\mathcal{B}}_{\mathrm{loc}}}(\mathbf{r})}, & \text{if } n_{2,\Psi^{\mathcal{B}}_{\mathrm{loc}}}(\mathbf{r}) \neq 0, \\ \infty, & \text{otherwise,} \end{cases} \tag{2.24}$$

with

$$f^{\mathcal{B}}_{\Psi^{\mathcal{B}}_{\mathrm{loc}}}(\mathbf{r}) = \sum_{pqrstu}^{\mathcal{B}} w_{pqrs} \Gamma_{rstu} \phi_p(\mathbf{r}) \phi_q(\mathbf{r}) \phi_t(\mathbf{r}) \phi_u(\mathbf{r}), \tag{2.25}$$

where $\Gamma_{pqrs} = 2\langle \Psi^{\mathcal{B}}_{\mathrm{loc}} | \hat{a}^{\dagger}_{r\downarrow} \hat{a}^{\dagger}_{s\uparrow} \hat{a}_{q\uparrow} \hat{a}_{p\downarrow} | \Psi^{\mathcal{B}}_{\mathrm{loc}} \rangle$ is the opposite-spin two-electron density matrix of some "localizing" wave function $\Psi^{\mathcal{B}}_{\mathrm{loc}}$, and $n_{2,\Psi^{\mathcal{B}}_{\mathrm{loc}}}(\mathbf{r})$ is its associated on-top pair density,

$$n_{2,\Psi^{\mathcal{B}}_{\mathrm{loc}}}(\mathbf{r}) = \sum_{pqrs}^{\mathcal{B}} \Gamma_{pqrs} \phi_p(\mathbf{r}) \phi_q(\mathbf{r}) \phi_r(\mathbf{r}) \phi_s(\mathbf{r}). \tag{2.26}$$

The wave function $\Psi^{\mathcal{B}}_{\mathrm{loc}}$ is only used to localize the effective interaction projected in the basis set $\mathcal{B}$. The local range-separation parameter is very weakly dependent on this wave function $\Psi^{\mathcal{B}}_{\mathrm{loc}}$. It should be thought of as essentially dependent on the basis set $\mathcal{B}$. Importantly, in the CBS limit the effective interaction goes to the Coulomb interaction which diverges at coalescence, and consequently, the local range-separation parameter goes to infinity,

$$\lim_{\mathcal{B} \to \mathrm{CBS}} \mu^{\mathcal{B}}(\mathbf{r}) = \infty, \tag{2.27}$$

independently of $\Psi^{\mathcal{B}}_{\mathrm{loc}}$, which is fundamental to guarantee the correct behavior of the theory in the CBS limit.

Approximate basis-set correction functionals from short-range functionals

Approximations for the basis-set correction functional $\bar{E}^{\mathcal{B}}$ are obtained by using the previously defined local range-separation parameter in short-range correlation functionals. Specifically, for the basis-set correction functional in Eq. (2.5) we use the so-called PBE-UEG basis-set correction functional (UEG stands for "uniform electron gas") [15],

$$\bar{E}^{\mathcal{B}}_{\text{PBE-UEG}}[n] = \int d\mathbf{r}\, \bar{e}^{\mathrm{sr,PBE}}_{\mathrm{c,md}}(n(\mathbf{r}), \nabla n(\mathbf{r}), n_2^{\mathrm{UEG}}(\mathbf{r}), \mu(\mathbf{r})), \tag{2.28}$$

and, for the basis-set correction functional in Eq. (2.12), we use the so-called PBE-OT basis-set correction functional (OT stands for "on-top") [16],

$$\bar{E}^{\mathcal{B}}_{\text{PBE-OT}}[n, n_2] = \int d\mathbf{r}\, \bar{e}^{\mathrm{sr,PBE}}_{\mathrm{c,md}}(n(\mathbf{r}), \nabla n(\mathbf{r}), \mathring{n}_2(\mathbf{r}), \mu(\mathbf{r})), \tag{2.29}$$

where we have dropped the superscript $\mathcal{B}$ in the density, in the on-top pair density, and in the local range-separation parameter for simplicity. In these expressions, the short-range (sr) correlation energy density with multideterminant (md) reference $\bar{e}^{\mathrm{sr,PBE}}_{\mathrm{c,md}}(n, \nabla n, n_2, \mu)$ has the following generic form in terms of the density n, the density gradient ∇n, the on-top pair density n_2, and the range-separation parameter μ [24],

$$\bar{e}^{\mathrm{sr,PBE}}_{\mathrm{c,md}}(n, \nabla n, n_2, \mu) = \frac{e^{\mathrm{PBE}}_{\mathrm{c}}(n, \nabla n)}{1 + \beta(n, \nabla n, n_2)\, \mu^3}, \tag{2.30}$$

$$\beta(n, \nabla n, n_2) = \frac{e^{\mathrm{PBE}}_{\mathrm{c}}(n, \nabla n)}{c\, n_2}, \tag{2.31}$$

where $e^{\mathrm{PBE}}_{\mathrm{c}}(n, \nabla n)$ is the usual Perdew-Burke-Ernzerhof (PBE) correlation energy density [25] and $c = (2\sqrt{\pi}(1 - \sqrt{2}))/3$.

The key difference between the PBE-UEG and PBE-OT functional is the on-top pair density used. The PBE-UEG functional uses $n_2^{\mathrm{UEG}}(\mathbf{r})$ which is an estimate of the exact on-top pair density using a parametrization of the on-top pair density of the uniform electron gas (UEG) at density $n(\mathbf{r})$,

$$n_2^{\mathrm{UEG}}(\mathbf{r}) = n(\mathbf{r})^2 g_0(n(\mathbf{r})), \tag{2.32}$$

where the on-top pair-distribution function $g_0(n)$ is taken from Eq. (46) of Ref. [26]. By contrast, the PBE-OT functional uses $\mathring{n}_2(\mathbf{r})$ which is an estimate of the exact on-top pair density obtained from extrapolating the input on-top pair density $n_2(\mathbf{r})$ of the wave function $\Psi^{\mathcal{B}}$ to the limit $\mu \to \infty$ (see Ref. [26]),

$$\mathring{n}_2(\mathbf{r}) = \left(1 + \frac{2}{\sqrt{\pi}\mu(\mathbf{r})}\right)^{-1} n_2(\mathbf{r}). \tag{2.33}$$

As shown in Ref. [16], the difference between the two flavors of on-top pair densities comes from the treatment of strong correlation. While $n_2^{\mathrm{UEG}}(\mathbf{r})$ is a good approximation of the exact on-top pair density for weakly correlated situations, when strong-correlation effects are present, it fails to represent the large depletion of the exact on-top pair density and in this case $\mathring{n}_2(\mathbf{r})$ provides a much better approximation of the exact on-top pair density.

The explicit expression of the one-electron effective potential associated with the PBE-UEG functional,

$$\bar{v}^{\mathcal{B}}_{\mathrm{PBE\text{-}UEG}}(\mathbf{r}) = \frac{\delta \bar{E}^{\mathcal{B}}_{\mathrm{PBE\text{-}UEG}}[n]}{\delta n(\mathbf{r})}, \tag{2.34}$$

was already given in a previous work [14]. The corresponding potential for the PBE-OT functional,

$$\bar{v}^{\mathcal{B}}_{\mathrm{PBE\text{-}OT}}(\mathbf{r}) = \frac{\delta \bar{E}^{\mathcal{B}}_{\mathrm{PBE\text{-}OT}}[n, n_2]}{\delta n(\mathbf{r})}, \tag{2.35}$$

has a very similar expression, with the simplification that the on-top pair density n_2 in Eq. (2.29) is taken as independent of the density whereas n_2^{UEG} in Eq. (2.28) is a function of the density. For the PBE-OT functional, we have in addition the two-electron effective interaction

$$\bar{w}^{\mathcal{B}}_{\mathrm{PBE\text{-}OT}}(\mathbf{r}) = \frac{\delta \bar{E}^{\mathcal{B}}_{\mathrm{PBE\text{-}OT}}[n, n_2]}{\delta n_2(\mathbf{r})}. \tag{2.36}$$

Its explicit expression is

$$\bar{w}^{\mathcal{B}}_{\mathrm{PBE\text{-}OT}}(\mathbf{r}) = \frac{\partial \bar{e}^{\mathrm{sr,PBE}}_{\mathrm{c,md}}}{\partial n_2}(n(\mathbf{r}), \nabla n(\mathbf{r}), \mathring{n}_2(\mathbf{r}), \mu(\mathbf{r})) \frac{\partial \mathring{n}_2(\mathbf{r})}{\partial n_2(\mathbf{r})} \tag{2.37}$$

where

$$\frac{\partial \mathring{n}_2(\mathbf{r})}{\partial n_2(\mathbf{r})} = \left(1 + \frac{2}{\sqrt{\pi}\mu(\mathbf{r})}\right)^{-1}, \tag{2.38}$$

and

$$\frac{\partial \bar{e}^{\mathrm{sr,PBE}}_{\mathrm{c,md}}(n, \nabla n, n_2, \mu)}{\partial n_2} = \frac{\bar{e}^{\mathrm{sr,PBE}}_{\mathrm{c,md}}(n, \nabla n, n_2, \mu)^2\, \mu^3}{c\, (n_2)^2}. \tag{2.39}$$

Note that the localizing wave function $\Psi^{\mathcal{B}}_{\mathrm{loc}}$ used in the definition of the local range-separation parameter (see Section 2.2.2) is taken as fixed, i.e. we do not consider variations of $\Psi^{\mathcal{B}}_{\mathrm{loc}}$. Importantly, since the local range-separation parameter diverges in the CBS limit [Eq. (2.27)] and since the correlation energy per particle $\bar{e}^{\mathrm{sr,PBE}}_{\mathrm{c,md}}(n, \nabla n, n_2, \mu)$ in Eq. (2.30) vanishes for $\mu \to \infty$, both basis-set correction functionals correctly vanish in the CBS limit,

$$\lim_{\mathcal{B}\to\mathrm{CBS}} \bar{E}^{\mathcal{B}}_{\mathrm{PBE\text{-}UEG}}[n] = 0, \tag{2.40}$$

$$\lim_{\mathcal{B}\to\mathrm{CBS}} \bar{E}^{\mathcal{B}}_{\mathrm{PBE\text{-}OT}}[n, n_2] = 0, \tag{2.41}$$

i.e., the CBS limit is unaltered by the correction.

Frozen-core version of the basis-set correction functionals

When the wave function $\Psi^{\mathcal{B}}$ is calculated in the frozen-core approximation, we use the frozen-core version of the basis-set correction functionals introduced in Ref. [15]. The basis-set correction functionals become

$$\bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}] \to \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}_{\mathrm{val}}}], \tag{2.42}$$

and

$$\bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}, n_{2,\Psi^{\mathcal{B}}}] \to \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}_{\mathrm{val}}}, n_{2,\Psi^{\mathcal{B}}_{\mathrm{val}}}], \tag{2.43}$$

where $\Psi^{\mathcal{B}}_{\mathrm{val}}$ is the wave function $\Psi^{\mathcal{B}}$ truncated to the "valence" orbital space (i.e., with all core orbitals removed). Accordingly, the local range-separation parameter becomes

$$\mu^{\mathcal{B}}(\mathbf{r}) \to \mu^{\mathcal{B}}_{\mathrm{val}}(\mathbf{r}) \tag{2.44}$$

where

$$\mu^{\mathcal{B}}_{\mathrm{val}}(\mathbf{r}) = \frac{\sqrt{\pi}}{2} W^{\mathcal{B}}_{\mathrm{val}}(\mathbf{r}). \tag{2.45}$$

The valence-only effective interaction at coalescence is

$$W^{\mathcal{B}}_{\mathrm{val}}(\mathbf{r}) = \begin{cases} \dfrac{f^{\mathcal{B}}_{\Psi^{\mathcal{B}}_{\mathrm{loc,val}}}(\mathbf{r})}{n_{2,\Psi^{\mathcal{B}}_{\mathrm{loc,val}}}(\mathbf{r})}, & \text{if } n_{2,\Psi^{\mathcal{B}}_{\mathrm{loc,val}}}(\mathbf{r}) \neq 0, \\ \infty, & \text{otherwise,} \end{cases} \tag{2.46}$$

where

$$f^{\mathcal{B}}_{\Psi^{\mathcal{B}}_{\mathrm{loc,val}}}(\mathbf{r}) = \sum_{pq}^{\mathcal{B}} \sum_{rstu}^{\mathcal{B}_{\mathrm{val}}} w_{pqrs} \Gamma_{rstu} \phi_p(\mathbf{r}) \phi_q(\mathbf{r}) \phi_t(\mathbf{r}) \phi_u(\mathbf{r}), \tag{2.47}$$

where $\Gamma_{pqrs} = 2\langle \Psi^{\mathcal{B}}_{\mathrm{loc,val}} | \hat{a}^\dagger_{r\downarrow} \hat{a}^\dagger_{s\uparrow} \hat{a}_{q\uparrow} \hat{a}_{p\downarrow} | \Psi^{\mathcal{B}}_{\mathrm{loc,val}} \rangle$ is the opposite-spin two-electron density matrix of the localizing wave function truncated to the valence orbital space $\Psi^{\mathcal{B}}_{\mathrm{loc,val}}$, and $n_{2,\Psi^{\mathcal{B}}_{\mathrm{loc,val}}}(\mathbf{r})$ is its associated on-top pair density

$$n_{2,\Psi^{\mathcal{B}}_{\mathrm{loc,val}}}(\mathbf{r}) = \sum_{pqrs}^{\mathcal{B}_{\mathrm{val}}} \Gamma_{pqrs} \phi_p(\mathbf{r}) \phi_q(\mathbf{r}) \phi_r(\mathbf{r}) \phi_s(\mathbf{r}). \tag{2.48}$$

In Eqs. (2.47) and (2.48), the indication "$\mathcal{B}_{\text{val}}$" on top of the sum symbols means that the sums are only over valence (i.e., non-core) orbitals. It is noteworthy that $\mu^{\mathcal{B}}_{\text{val}}(\mathbf{r})$ still fulfills Eq. (2.27) and thus the frozen-core versions of the basis-set correction functionals still correctly vanish in the CBS limit.

Correspondingly, the frozen-core versions of the effective Hamiltonians in Eqs. (2.7) and (2.14) are simply obtained by setting to zero all the one-electron effective integrals $\bar{v}^{\mathcal{B}}_{pq}$ and the two-electron effective integrals $\bar{w}^{\mathcal{B}}_{pqrs}$ if at least one orbital involved in the integral is a core orbital.

2.2.3 Selected configuration interaction to solve the self-consistent eigenvalue equations

To solve the self-consistent basis-set correction eigenvalue equations [Eqs. (2.6) and (2.13)], we use an adaptation of the configuration-interaction perturbatively selected iteratively (CIPSI) algorithm [27–29], similar to the computational strategy already used in the context of RSDFT [24].

To solve Eq. (2.6) for a given basis set $\mathcal{B}$, we start at the first iteration, denoted as $k-1$, from a guess wave function $\Psi_0^{\mathcal{B},(k-1)}$ (usually a CIPSI ground-state wave function for the standard Hamiltonian),

$$|\Psi_0^{\mathcal{B},(k-1)}\rangle = \sum_{\mathrm{I}\in\mathcal{R}^{(k-1)}} c_{\mathrm{I}}^{(k-1)}\ |\mathrm{I}\rangle, \tag{2.49}$$

where $\mathcal{R}^{(k-1)}$ denotes a set of Slater determinants. We then use the density of this wave function to form the following effective Hamiltonian at the next iteration k:

$$\hat{H}^{\mathcal{B},(k)}_{\text{eff}} = \hat{H}^{\mathcal{B}}_{\text{eff}}[n_{\Psi_0^{\mathcal{B},(k-1)}}], \tag{2.50}$$

and we want to find the associated ground-state wave function $\Psi^{\mathcal{B},(k)}$,

$$\hat{H}^{\mathcal{B},(k)}_{\text{eff}}|\Psi^{\mathcal{B},(k)}\rangle = \mathcal{E}^{\mathcal{B},(k)}|\Psi^{\mathcal{B},(k)}\rangle. \tag{2.51}$$

This wave function is obtained by the CIPSI algorithm as

$$|\Psi^{\mathcal{B},(k)}\rangle = \sum_{\mathrm{I}\in\mathcal{R}^{(k)}} c_{\mathrm{I}}\ |\mathrm{I}\rangle, \tag{2.52}$$

where $\mathcal{R}^{(k)}$ is the new set of Slater determinants at iteration k. According the CIPSI algorithm, the set $\mathcal{R}^{(k)}$ is obtained by repeatedly adding to a reference wave function $\Psi^{\mathcal{B},(k)}_{\text{ref}}$ the determinants K having the largest second-order perturbation theory (PT2) contributions $|\mathcal{E}^{(k)}_{\mathrm{K,PT2}}|$ with

$$\mathcal{E}^{(k)}_{\mathrm{K,PT2}} = -\frac{|\langle\Psi^{\mathcal{B},(k)}_{\text{ref}}|\hat{H}^{\mathcal{B},(k)}_{\text{eff}}|\mathrm{K}\rangle|^2}{\langle\mathrm{K}|\hat{H}^{\mathcal{B},(k)}_{\text{eff}}|\mathrm{K}\rangle - \langle\Psi^{\mathcal{B},(k)}_{\text{ref}}|\hat{H}^{\mathcal{B},(k)}_{\text{eff}}|\Psi^{\mathcal{B},(k)}_{\text{ref}}\rangle}, \tag{2.53}$$

iteratively doubling the number of determinants in $\Psi^{\mathcal{B},(k)}_{\text{ref}}$ until the absolute value of the total PT2 energy correction due to the missing determinants,

$$|\mathcal{E}^{(k)}_{\mathrm{PT2}}| = \left|\sum_{\mathrm{K}\notin\mathcal{R}^{(k)}} \mathcal{E}^{(k)}_{\mathrm{K,PT2}}\right|, \tag{2.54}$$

is smaller than a given threshold. To reduce the cost of the evaluation of the PT2 contribution, the semi-stochastic multi-reference approach of Garniron *et al.* [30] is used, adopting the technical specifications recommended in that work.

This determines the set of determinants $\mathcal{R}^{(k)}$ which is then fixed for the rest of the iteration k. The energy $E_0^{\mathcal{B},(k)}$ for this iteration is then determined according to the minimization in Eq. (2.5)

$$E_0^{\mathcal{B},(k)} = \min_{\Psi^{\mathcal{B},(k)}} \left\{ \langle \Psi^{\mathcal{B},(k)} | \hat{T} + \hat{W}_{\text{ee}} + \hat{V}_{\text{ne}} | \Psi^{\mathcal{B},(k)} \rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B},(k)}}] \right\}, \tag{2.55}$$

which amounts to solving the iterative equation,

$$\hat{H}_{\text{eff}}^{\mathcal{B}}[n_{\Psi_0^{\mathcal{B},(k)}}] | \Psi_0^{\mathcal{B},(k)} \rangle = \mathcal{E}_0^{\mathcal{B},(k)} | \Psi_0^{\mathcal{B},(k)} \rangle, \tag{2.56}$$

for the optimal coefficients of the determinants $\{c_{\text{I}}^{(k)}\}$ leading to the minimizing wave function at iteration k,

$$| \Psi_0^{\mathcal{B},(k)} \rangle = \sum_{\text{I} \in \mathcal{R}^{(k)}} c_{\text{I}}^{(k)} \, | \text{I} \rangle. \tag{2.57}$$

The iterations over k are repeated until the variation of $E_0^{\mathcal{B},(k)}$ is smaller than a given threshold. The evaluation of the dipole moment is obtained as the expectation value of the dipole operator over the converged wave function $\Psi_0^{\mathcal{B}}$.

The same approach is used for solving Eq. (2.13) which involves the on-top pair density in addition to the density.

2.3 Computation of total energies and dipole moments

2.3.1 Computational details

We study the total ground-state energies of the Be atom and BH molecule together with the dipole moments of the BH, FH, H_2O molecules in their ground states and of the CH_2 molecule in its lowest spin-singlet state. We report standard CIPSI (i.e., near FCI) results without the basis-set correction (referred to as "CIPSI"), as well as CIPSI results including the basis-set correction using the PBE-UEG and PBE-OT functionals with or without self-consistency. The non-self-consistent calculations are referred to as "CIPSI+PBE-UEG and "CIPSI+PBE-OT", whereas the self-consistent calculations are referred to as "SC CIPSI+PBE-UEG" and "SC CIPSI+PBE-OT" where SC stands for self-consistent. The orbitals used for all converged CIPSI calculations are the natural orbitals obtained from a standard CIPSI calculation stopped at a total PT2 energy correction smaller in absolute value than 0.001 Ha. For the localizing wave function $\Psi_{\text{loc}}^{\mathcal{B}}$ involved in the definition of the local range-separation parameter $\mu^{\mathcal{B}}(\mathbf{r})$ (see Section 2.2.2), we choose either a single Slater determinant (SD) built from the natural orbitals of the largest CIPSI wave function (which we refer as μ_{SD}) or the largest CIPSI wave function (which we refer as μ_{CIPSI}). We use the Dunning correlation-consistent basis-set family [31–35]. We perform both non-frozen-core calculations using the core-valence aug-cc-pCV*X*Z basis sets, and frozen-core calculations (with the 1s orbitals of non-hydrogen atoms frozen) using the valence aug-cc-pV*X*Z basis sets and the corresponding frozen-core version of the basis-set correction (see Section 2.2.2). All the CIPSI calculations have been performed with Quantum Package [36].

We also report the dipole moment at the coupled cluster singles doubles perturbative triples [CCSD(T)] level which were taken from Ref. [37] for the BH and FH molecules, and obtained using linear-response calculations from the Dalton software [38, 39] for the CH_2 and H_2O molecules. The molecular geometries are taken from Ref. [37] for BH and FH, and from Ref. [40] for H_2O and CH_2.

2.3.2 Total energies of the Be atom and the BH molecule

Tables 2.1 and 2.2 report the total energies of the Be atom and the BH molecule, respectively, calculated using the aug-cc-pCV*X*Z basis sets with CIPSI without any basis-set correction and

Table 2.1: Total ground-state energies (in Ha) of the Be atom calculated using the aug-cc-pCVXZ (ACVXZ) basis sets (with X = D, T, Q) with CIPSI without the frozen-core approximation and including different basis-set corrections with or without self-consistency. The energy lowering $\Delta E_{\rm SC}$ (in μHa) from the non-self-consistent to the self-consistent version of the basis-set correction is reported in square brackets.

	ACVDZ	ACVTZ	ACVQZ
CIPSI	-14.6519225	-14.6623971	-14.6655767
CIPSI+PBE-UEG$_{\mu_{\rm SD}}$	-14.6683617	-14.6686314	-14.6681020
SC CIPSI+PBE-UEG$_{\mu_{\rm SD}}$ [$\Delta E_{\rm SC}$]	-14.6683878 [-26.1]	-14.6686354 [-4.0]	-14.6681026 [-0.6]
CIPSI+PBE-UEG$_{\mu_{\rm CIPSI}}$	-14.6677035	-14.6683762	-14.6679643
SC CIPSI+PBE-UEG$_{\mu_{\rm CIPSI}}$ [$\Delta E_{\rm SC}$]	-14.6677395 [-36.0]	-14.6683874 [-11.2]	-14.6679681 [-3.8]
CIPSI+PBE-OT$_{\mu_{\rm SD}}$	-14.6663376	-14.6678846	-14.6677982
SC CIPSI+PBE-OT$_{\mu_{\rm SD}}$ [$\Delta E_{\rm SC}$]	-14.6663741 [-36.5]	-14.6678956 [-11.0]	-14.6678020 [-3.8]
CIPSI+PBE-OT$_{\mu_{\rm CIPSI}}$	-14.6659463	-14.6677128	-14.6677140
SC CIPSI+PBE-OT$_{\mu_{\rm CIPSI}}$ [$\Delta E_{\rm SC}$]	-14.6659748 [-28.5]	-14.6677223 [-9.5]	-14.6677171 [-3.1]

Exact non-relativistic total energy[a]

-14.6673565

[a]From Ref. [41].

Table 2.2: Total ground-state energies (in Ha) of the BH molecule calculated using the aug-cc-pCVXZ (ACVXZ) basis sets (with X = D, T, Q, 5) with CIPSI without the frozen-core approximation and including different basis-set corrections with or without self-consistency. The energy lowering $\Delta E_{\rm SC}$ (in μHa) from the non-self-consistent to the self-consistent version of the basis-set correction is reported in square brackets.

	ACVDZ	ACVTZ	ACVQZ	ACV5Z
CIPSI	-25.2550150	-25.2786179	-25.2857583	-25.2873779
CIPSI+PBE-UEG$_{\mu_{\rm SD}}$	-25.2838179	-25.2896303	-25.2906772	-25.2900079
SC CIPSI+PBE-UEG$_{\mu_{\rm SD}}$ [$\Delta E_{\rm SC}$]	-25.2839270 [-109.1]	-25.2896471 [-16.8]	-25.2906804 [-3.2]	–
CIPSI+PBE-UEG$_{\mu_{\rm CIPSI}}$	-25.2825079	-25.2890245	-25.2903975	–
SC CIPSI+PBE-UEG$_{\mu_{\rm CIPSI}}$ [$\Delta E_{\rm SC}$]	-25.2826090 [-101.1]	-25.2890400 [-15.5]	–	–
CIPSI+PBE-OT$_{\mu_{\rm SD}}$	-25.2798297	-25.2880486	-25.2899774	–
SC CIPSI+PBE-OT$_{\mu_{\rm SD}}$ [$\Delta E_{\rm SC}$]	-25.2800391 [-209.4]	-25.2881008 [-52.2]	-25.2899937 [-16.3]	–
CIPSI+PBE-OT$_{\mu_{\rm CIPSI}}$	-25.2789738	-25.2876363	-25.2897883	–
SC CIPSI+PBE-OT$_{\mu_{\rm CIPSI}}$ [$\Delta E_{\rm SC}$]	-25.2791600 [-186.2]	-25.2876809 [-44.6]	–	–

CIPSI total energy extrapolated to the CBS limit

-25.289032

Table 2.3: Dipole moment (in a.u.) of the ground state of the BH molecule calculated using the aug-cc-pCVXZ (ACVXZ) basis sets (with X = D, T, Q, 5) by Hartree-Fock (HF), CCSD(T), and CIPSI without the frozen-core approximation and including different self-consistent basis-set corrections. For the CIPSI calculations, the PT2 energy correction $|\mathcal{E}_{\rm PT2}|$ (in Ha) is reported in square brackets. Extrapolations to the CBS limit are given in the last column.

	ACVDZ	ACVTZ	ACVQZ	ACV5Z	CBS
HF	0.68796	0.68570	0.68489	0.68530	
CCSD(T)[a]	0.52968	0.54649	0.54984	0.55125	0.55271
CIPSI [$\|\mathcal{E}_{\rm PT2}\|$]	0.52758 [2×10^{-6}]	0.54388 [4×10^{-6}]	0.54789 [4×10^{-6}]	0.54975 [2×10^{-5}]	0.55142
CIPSI extrapolated to $\mathcal{E}_{\rm PT2}\to 0$	0.52757	0.54386	0.54790	0.54967	0.55126
SC CIPSI+PBE-UEG$_{\mu_{\rm SD}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.53658 [2×10^{-6}]	0.54835 [4×10^{-6}]	0.55040 [4×10^{-5}]		
SC CIPSI+PBE-UEG$_{\mu_{\rm CIPSI}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.53758 [2×10^{-6}]	0.54812 [4×10^{-6}]			
SC CIPSI+PBE-OT$_{\mu_{\rm SD}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.54340 [2×10^{-6}]	0.55092 [4×10^{-6}]	0.55093 [1×10^{-5}]		
SC CIPSI+PBE-OT$_{\mu_{\rm CIPSI}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.54333 [2×10^{-6}]	0.54973 [8×10^{-6}]			

[a]From Ref. [37].

Table 2.4: Dipole moment (in a.u.) of the ground state of the BH molecule calculated using the aug-cc-pVXZ (AVXZ) basis sets (with X = D, T, Q, 5) by Hartree-Fock (HF), CCSD(T), and CIPSI with the frozen-core approximation and including different self-consistent basis-set corrections. For the CIPSI calculations, the PT2 energy correction $|\mathcal{E}_{\rm PT2}|$ (in Ha) is reported in square brackets. Extrapolations to the CBS limit are given in the last column.

	AVDZ	AVTZ	AVQZ	AV5Z	CBS
HF	0.68796	0.68650	0.68494	0.68496	
CCSD(T)	0.52939	0.54500	0.54724	0.54843	0.54966
CIPSI [$\|\mathcal{E}_{\rm PT2}\|$]	0.52782 [3×10^{-8}]	0.54334 [1×10^{-7}]	0.54563 [4×10^{-7}]	0.54691 [1×10^{-6}]	0.54823
SC CIPSI+PBE-UEG$_{\mu_{\rm SD}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.53791 [4×10^{-7}]	0.54815 [7×10^{-7}]	0.54790 [3×10^{-6}]	0.54815 [3×10^{-6}]	
SC CIPSI+PBE-OT$_{\mu_{\rm SD}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.54512 [4×10^{-7}]	0.55029 [5×10^{-7}]	0.54880 [2×10^{-8}]		

with different basis-set corrections. It can be observed that the total energies obtained with the basis-set corrections converge much faster toward the estimated exact total energies than the total energies obtained without any basis-set correction. For the Be atom, all the basis-set corrected total energies from the aug-cc-pCVDZ to the aug-cc-pCVQZ basis sets have errors below 1.6 mHa $\approx$ 1 kcal/mol compared to the estimated exact total energy, whereas without basis-set correction such an accuracy is not even reached with the aug-cc-pCVQZ basis set. Similar trends are observed for the BH molecule: all the basis-set corrected total energies have errors below 1 kcal/mol already from the aug-cc-pCVTZ basis set, whereas without basis-set correction such an accuracy is reached only with the aug-cc-pCV5Z basis set.

Focusing now on the differences between the various basis-set corrections, we can notice that i) using the local range-separation parameter $\mu_{\rm SD}$ gives larger basis-set corrections than using $\mu_{\rm CIPSI}$, ii) the PBE-UEG functional gives larger basis-set corrections than the PBE-OT functional. As regards the effect of the self-consistency, it is remarkable to notice that self-consistency lowers the total energy by a very small fraction of the total basis-set correction (typically less than 1%), whatever the choice of functional or local range-separation parameter. These results thus validate the previously introduced non-self-consistent approximation to the basis-set correction (see Section 2.2.1) for energy calculations.

2.3.3 Dipole moments of the BH, FH, H_2O, and CH_2 molecules

As seen from Section 2.3.2, the self-consistency of the basis-set correction does not lead to significant changes of the total energies. Nevertheless, one can wonder if the effective wave

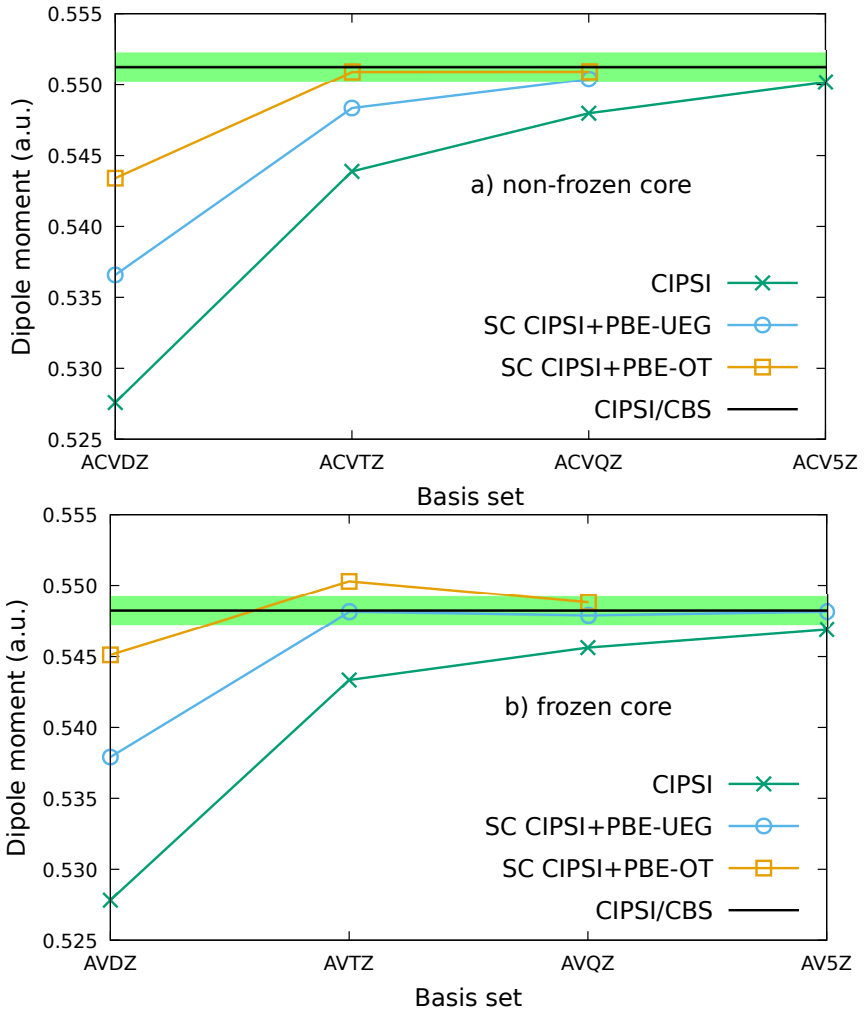

Figure 2.1: Basis-set convergence of the dipole moment of the ground state of the BH molecule calculated using the aug-cc-pCVXZ (ACVXZ) and aug-cc-pVXZ (AVXZ) basis sets (with X = D, T, Q, 5) by CIPSI including different self-consistent basis-set corrections without the frozen-core approximation (a) and with the frozen-core approximation (b). The self-consistent basis-set corrections are for the local range-separation parameter μ_{SD}. The green area indicates an error of ±0.001 a.u. around the CIPSI CBS value.

functions obtained with the self-consistent basis-set correction provide better properties.

We choose to investigate this aspect through the computation of the dipole moments of the BH, FH, H_2O, and CH_2 molecules for several reasons: i) the basis-set convergence of dipole moments with correlated wave-function methods is known to be slow (see, e.g., Refs. [37,40,42]), ii) these molecules are sufficiently small to have near-CBS reference values, iii) the BH and CH_2 molecules exhibit a non-trivial mixture of both strong and weak correlation due to the 2s-2p near degeneracy in the boron and carbon atoms, whereas the FH and H_2O molecules are dominated by weak correlation effects.

Dipole moment of the BH molecule

We start by a detailed analysis on the BH molecule. The slow convergence of the dipole moment of the BH molecule with respect to the size of the basis set has been illustrated at various correlation levels including CCSD(T) by Halkier *et al.* [37], and we report in Table 2.3 their CCSD(T) results at the non-frozen core level. In order to have a reasonable estimate of the dipole moment in the CBS limit, we use the two-point X^{-3} extrapolation proposed in Ref. [37] using the aug-cc-pCVQZ and aug-cc-pCV5Z basis sets. At the CCSD(T) level, an error of about 0.001 a.u. with respect to the CBS limit is not even reached with the aug-cc-pCV5Z basis set, while the error is about 0.006 a.u. and 0.003 a.u. with the aug-cc-pCVTZ and aug-cc-pCVQZ basis set, respectively, showing indeed the quite slow convergence of the dipole moment of BH.

We also report in Table 2.3 the dipole moment at the CIPSI level together with the value of the PT2 energy correction $|\mathcal{E}_{\rm PT2}|$ associated with the variational wave function for which the dipole moment have been calculated. The values of $|\mathcal{E}_{\rm PT2}|$ are all below 2×10^{-5} Ha, which was found to be mandatory to converge the CIPSI dipole moments to a precision of about 0.0001 a.u. for the BH molecule. These represent therefore rather large calculations involving about 10^8 Slater determinants for the largest aug-cc-pCV5Z basis set. In order to obtain an estimation of the error of the dipole moment at the CIPSI level in a given basis set with respect to FCI, we extrapolate the CIPSI dipole moment to $|\mathcal{E}_{\rm PT2}| \to 0$ using a linear extrapolation as a function of $|\mathcal{E}_{\rm PT2}|$ (similar to the proposal of Holmes *et al.* [43] for the total energy) using different values of $|\mathcal{E}_{\rm PT2}|$. As one can see from Table 2.3, for any basis set, the difference between the dipole moment calculated by CIPSI with the smallest $|\mathcal{E}_{\rm PT2}|$ available and the dipole moment extrapolated with respect to $|\mathcal{E}_{\rm PT2}|$ are negligible. Thus, with the values of $|\mathcal{E}_{\rm PT2}|$ used, the CIPSI dipole moments provide reliable estimates of the FCI dipole moments. As regards basis-set errors, similar to the CCSD(T) results, an error of about 0.001 a.u. on the CIPSI dipole moment with respect to the CBS limit is not even reached using the aug-cc-pCV5Z basis set, illustrating once more the slow convergence of the dipole moment with respect to the size of the basis set.

From Table 2.3 one can also notice that there is a small deviation between the CCSD(T) and the CIPSI dipole moments which is of about 0.002 a.u. for the aug-cc-pCVXZ basis sets with X= D, T, Q and of about 0.001 a.u. for the aug-cc-pCV5Z basis set. In addition, the deviation between the CCSD(T) and CIPSI dipole moments extrapolated to the CBS limit is about 0.001 a.u., showing that CCSD(T) provides a very accurate value for the dipole moment of the BH molecule.

Coming now to the self-consistent basis-set correction calculations, we report in Table 2.3 the dipole moments obtained using the PBE-UEG and PBE-OT functionals using the local range-separation parameters $\mu_{\rm SD}$ and $\mu_{\rm CIPSI}$, and a graphical representation of the data for $\mu_{\rm SD}$ is given in Figure 2.1 (panel a). Clearly, the basis-set correction strongly accelerates the convergence to the CBS limit. More specifically, it can be observed that i) in a given basis set, all basis-set corrections significantly improve the dipole moment with respect to the CIPSI value, ii) the PBE-OT functional gives more accurate results than the PBE-UEG functional, iii) an error of about 0.001 a.u. on the dipole moment is obtained already with the aug-cc-pCVTZ basis set

Table 2.5: Dipole moment (in a.u.) of the ground state of the FH molecule calculated using the aug-cc-pVXZ (AVXZ) basis sets (with X = D, T, Q, 5) by Hartree-Fock (HF), CCSD(T), and CIPSI with the frozen-core approximation and including different self-consistent basis-set corrections. For the CIPSI calculations, the PT2 energy correction $|\mathcal{E}_{\rm PT2}|$ (in Ha) is reported in square brackets. Extrapolations to the CBS limit are given in the last column.

	AVDZ	AVTZ	AVQZ	AV5Z	CBS
HF	0.75976	0.75750	0.75634	0.75617	
CCSD(T)	0.70342	0.70465	0.70707	0.70794	0.70903
CIPSI $_{[|\mathcal{E}_{\rm PT2}|]}$	0.70249 $_{[9\times10^{-6}]}$	0.70406 $_{[1\times10^{-4}]}$	0.70662 $_{[1\times10^{-4}]}$		
CIPSI extrapolated wrt $\mathcal{E}_{\rm PT2}$	0.70248	0.70391	0.70646		
SC CIPSI+PBE-UEG$_{\mu_{\rm SD}}$ $_{[|\mathcal{E}_{\rm PT2}|]}$	0.71326 $_{[3\times10^{-5}]}$	0.70873 $_{[1\times10^{-4}]}$			
SC CIPSI+PBE-OT$_{\mu_{\rm SD}}$ $_{[|\mathcal{E}_{\rm PT2}|]}$	0.71362 $_{[2\times10^{-5}]}$	0.70915 $_{[1\times10^{-4}]}$			

when using the PBE-OT functional. The result i) shows that, although the self-consistency does not lead to significant improvement on the total energy (see Table 2.2), it is crucial to yield effective wave functions providing better properties, illustrating the impact and the accuracy of the effective basis-set correction potentials employed. The result ii) shows that the use of the on-top pair density of the wave function rather than that estimated from the UEG gives a better approximation of the exact on-top pair density, which can be understood as a signature of strong-correlation effects.

Finally, we also report in Table 2.4 the dipole moments obtained with the frozen-core approximation using the aug-cc-pVXZ basis sets, and the corresponding graphical representation is given in Figure 2.1 (panel b). Again, the basis-set correction with either the PBE-UEG or PBE-OT functional yields a faster basis-set convergence of the dipole moment that in the standard CIPSI calculations. However, in this case, the convergence toward the CBS limit is slightly better when using the PBE-UEG functional. The PBE-OT functional slightly overestimates the dipole moment by 0.002 a.u. and 0.0006 a.u. with the aug-cc-pVTZ and aug-cc-pVQZ basis sets, respectively, whereas the PBE-UEG functional yields a deviation below 0.0005 a.u. from the aug-cc-pVTZ to the aug-cc-pV5Z basis set.

2.3.4 Dipole moments of the FH, H_2O, and CH_2 molecules

We now pursue our analysis on the FH, H_2O, and CH_2 molecules using only the frozen-core approximation. The basis-set convergence of the dipole moments of these molecules was studied in Refs. [37,42] at the CCSD(T) level. In Tables 2.5, 2.6, and 2.7, we report CCSD(T) and CIPSI results with the aug-cc-pVXZ basis sets. The CBS estimates are obtained from a two-point X^{-3} extrapolation using the aug-cc-pVQZ and aug-cc-pV5Z basis sets at the CCSD(T) level, and also at the CIPSI level in the case of CH_2. One can notice that, at the CCSD(T) or CIPSI level, an error of about 0.001 a.u. with respect to the estimated CBS limit is not even reached with the aug-cc-pV5Z basis set. In addition, in the case of CH_2, there is a significant discrepancy between the extrapolated CCSD(T) and CIPSI dipole moments, which might be due to some strong-correlation effects that are mistreated at the CCSD(T) level.

We also report in Tables 2.5, 2.6, and 2.7, the results using the self-consistent basis-set correction. In contrast with the BH molecule, the effective wave functions obtained with the PBE-UEG and PBE-OT functionals always yield very similar dipole moments for the FH and H_2O molecules. This can be explained by the fact that these molecules at their equilibrium geometries are weakly correlated systems for which the on-top pair density based on the UEG is a good approximation. For the FH, H_2O, and CH_2 molecules, the dipole moment is overestimated with the aug-cc-pVDZ basis set using both functionals, but the results with the aug-cc-pVTZ

Table 2.6: Dipole moment (in a.u.) of the ground state of the H_2O molecule calculated using the aug-cc-pVXZ (AVXZ) basis sets (with X = D, T, Q, 5) by Hartree-Fock (HF), CCSD(T), and CIPSI with the frozen-core approximation and including different self-consistent basis-set corrections. For the CIPSI calculations, the PT2 energy correction $|\mathcal{E}_{\rm PT2}|$ (in Ha) is reported in square brackets. Extrapolations to the CBS limit are given in the last column.

	AVDZ	AVTZ	AVQZ	AV5Z	CBS
HF	0.78670	0.78038	0.77955	0.77956	
CCSD(T)	0.72703	0.72364	0.72695	0.72815	0.72941
CIPSI [$\|\mathcal{E}_{\rm PT2}\|$]	0.72610 [3×10^{-5}]	0.72294 [2×10^{-4}]			
SC CIPSI+PBE-UEG$_{\mu_{\rm SD}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.73809 [2×10^{-5}]	0.72818 [2×10^{-5}]			
SC CIPSI+PBE-OT$_{\mu_{\rm SD}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.73840 [2×10^{-4}]	0.72855 [1×10^{-4}]			

Table 2.7: Dipole moment (in a.u.) of the lowest spin-singlet state of the CH_2 molecule calculated using the aug-cc-pVXZ (AVXZ) basis sets (with X = D, T, Q, 5) by Hartree-Fock (HF), CCSD(T), and CIPSI with the frozen-core approximation and including different self-consistent basis-set corrections. For the CIPSI calculations, the PT2 energy correction $|\mathcal{E}_{\rm PT2}|$ (in Ha) is reported in square brackets. Extrapolations to the CBS limit are given in the last column.

	AVDZ	AVTZ	AVQZ	AV5Z	CBS
HF	0.74878	0.74478	0.74355	0.74353	
CCSD(T)	0.65614	0.66009	0.66211	0.66310	0.66416
CIPSI [$\|\mathcal{E}_{\rm PT2}\|$]	0.65120 [2×10^{-5}]	0.65446 [3×10^{-5}]	0.65643 [4×10^{-5}]	0.65780 [1×10^{-4}]	0.65926
SC CIPSI+PBE-UEG$_{\mu_{\rm SD}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.66249 [2×10^{-5}]	0.65958 [3×10^{-5}]	0.65890 [3×10^{-5}]		
SC CIPSI+PBE-OT$_{\mu_{\rm SD}}$ [$\|\mathcal{E}_{\rm PT2}\|$]	0.66527 [2×10^{-5}]	0.66055 [4×10^{-5}]	0.65932 [5×10^{-4}]		

basis set are already very close to the estimated CBS limit. From a quantitative point of view, for the FH molecule with the aug-cc-pVDZ basis set the error with respect to the CBS dipole moment is reduced from about 0.007 a.u. at the CIPSI level to about 0.004 a.u. with the basis-set correction, whereas with the aug-cc-pVTZ basis set the error is reduced from 0.005 a.u. to below 0.0005 a.u.. For the CH_2 molecule, with the aug-cc-pVDZ basis set the error with respect to the CBS extrapolated CIPSI value is about 0.008 a.u. whereas it is about 0.003 a.u. and 0.006 a.u. using the PBE-UEG and PBE-OT basis-set corrections, respectively. With the aug-cc-pVTZ basis set, the error at the CIPSI level is still of about 0.005 a.u., whereas it is below 0.001 a.u. for both the PBE-UEG and PBE-OT functionals. Finally, in the case of the H_2O molecule, the results using the basis-set correction are actually worst than the CIPSI ones when using the aug-cc-pVDZ basis set, the error increasing from about 0.003 a.u. for CIPSI to about 0.009 a.u. with the basis-set correction. One should nevertheless keep in mind that the convergence of the dipole moment of H_2O is non monotonic at the CCSD(T) level, the dipole moment obtained with the aug-cc-pVDZ basis set being closer to the CBS limit than the ones obtained using the aug-cc-pVTZ or aug-cc-pVQZ basis sets. Therefore, the seemingly good values obtained at the CCSD(T) and CIPSI levels using the aug-cc-pVDZ basis set are likely to be due to a compensation of errors. With the aug-cc-pVTZ basis set, the expected trend is recovered, with CIPSI giving an error of about 0.006 a.u. and the basis-set correction reducing this error to about 0.001 a.u..

2.4 Conclusion

In the present work, we have established the fully self-consistent density-based basis-set correction scheme [13]. Differently from previous works where a non-self-consistent approximation was

used [13–18], here the energy is minimized in the presence of the basis-set correction functional which i) guarantees to get a lower total energy with respect to the non-self-consistent approximation, and ii) allows the wave function to change under the presence of the basis-set correction. We have tested this scheme on a few atomic and molecular systems (Be, BH, FH, H_2O, and CH_2) with CIPSI wave functions and two different basis-set correction functionals PBE-UEG and PBE-OT. While PBE-UEG is a functional of the density, PBE-OT uses in addition the on-top pair density of the wave function as an independent variable.

The main results are that i) the lowering in total energy is extremely small compared to the non-self-consistent approximations (typically less than 1%), which thus justifies this approximation for energy-only calculations, and ii) the wave functions obtained from the self-consistent basis-set correction scheme yield dipole moments which converge much faster with respect to the size of the basis set than standard wave-function calculations, being already close to the CBS value with a triple-zeta basis set. This study thus tend to demonstrate that the density-based basis-set correction scheme is not only useful for energy calculations but also for calculations of response properties.

Acknowledgement

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme Grant agreement No. 810367 (EMC2).

Data availability

The data that support the findings of this study are available within the article and from the corresponding author upon reasonable request.

Bibliography

[1] E. A. Hylleraas, Z. Phys. **54**, 347 (1929).

[2] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).

[3] T. Helgaker, W. Klopper, H. Koch and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[4] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).

[5] S. Ten-no, Theor. Chem. Acc. **131**, 1070 (2012).

[6] S. Ten-no and J. Noga, WIREs Comput. Mol. Sci. **2**, 114 (2012).

[7] C. Hättig, W. Klopper, A. Köhn and D. P. Tew, Chem. Rev. **112**, 4 (2012).

[8] L. Kong, F. A. Bischoff and E. F. Valeev, Chem. Rev. **112**, 75 (2012).

[9] A. Grüneis, S. Hirata, Y.-Y. Ohnishi and S. Ten-no, J. Chem. Phys. **146**, 080901 (2017).

[10] Q. Ma and H.-J. Werner, WIREs Comput. Mol. Sci. **8**, e1371 (2018).

[11] D. P. Tew, W. Klopper, C. Neiss and C. Hattig, Phys. Chem. Chem. Phys. **9**, 1921 (2007).

[12] G. M. J. Barca and P.-F. Loos, J. Chem. Phys. **147**, 024103 (2017).

[13] E. Giner, B. Pradines, A. Ferté, R. Assaraf, A. Savin and J. Toulouse, J. Chem. Phys. **149**, 194301 (2018).

[14] P.-F. Loos, B. Pradines, A. Scemama, E. Giner and J. Toulouse, J. Chem. Theory Comput. **16**, 1018 (2020).

[15] P.-F. Loos, B. Pradines, A. Scemama, J. Toulouse and E. Giner, J. Phys. Chem. Lett. **10**, 2931 (2019).

[16] E. Giner, A. Scemama, P.-F. Loos and J. Toulouse, J. Chem. Phys. **152**, 174104 (2020).

[17] Y. Yao, E. Giner, J. Li, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **153**, 124117 (2020).

[18] E. Giner, A. Scemama, J. Toulouse and P.-F. Loos, J. Chem. Phys. **151**, 144118 (2019).

[19] M. Levy, Proc. Natl. Acad. Sci. U. S. A. **76**, 6062 (1979).

[20] E. H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983).

[21] A. Seidl, A. Görling, P. Vogl, J. A. Majewski and M. Levy, Phys. Rev. B **53**, 3764 (1996).

[22] J. Toulouse, Review of approximations for the exchange-correlation energy in density-functional theory, `https://arxiv.org/abs/2103.02645` (2021).

[23] J. Toulouse, P. Gori-Giorgi and A. Savin, Theor. Chem. Acc **114**, 305 (2005).

[24] A. Ferté, E. Giner and J. Toulouse, J. Chem. Phys. **150**, 084103 (2019).

[25] J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[26] P. Gori-Giorgi and A. Savin, Phys. Rev. A **73**, 032506 (2006).

[27] B. Huron, J. P. Malrieu and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).

[28] E. Giner, A. Scemama and M. Caffarel, Can. J. Chem. **91**, 879 (2013).

[29] E. Giner, A. Scemama and M. Caffarel, J. Chem. Phys. **142**, 044115 (2015).

[30] Y. Garniron, A. Scemama, P.-F. Loos and M. Caffarel, J. Chem. Phys. **147**, 034101 (2017).

[31] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[32] D. E. Woon and T. H. Dunning, J. Chem. Phys. **98**, 1358 (1993).

[33] D. E. Woon and T. H. Dunning, J. Chem. Phys. **103**, 4572 (1995).

[34] K. A. Peterson and T. H. Dunning, J. Chem. Phys. **117**, 10548 (2002).

[35] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson and T. L. Windus, J. Chem. Inf. Model. **59**, 4814 (2019).

[36] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Véril, M. Caffarel, P.-F. Loos, E. Giner and A. Scemama, J. Chem. Theory. Comput. **15**, 3591 (2019).

[37] A. Halkier, W. Klopper, T. Helgaker and P. Jørgensen, J. Chem. Phys. **111**, 4424 (1999).

[38] K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, et al., WIREs Comput. Mol. Sci. **4**, 269 (2014).

[39] Dalton, a molecular electronic structure program, Release v2020.0 (2020), see http://daltonprogram.org.

[40] D. Hait and M. Head-Gordon, J. Chem. Theory. Comput. **14**, 1969 (2018).

[41] I. Hornyák, L. Adamowicz and S. Bubin, Phys. Rev. A **100**, 032504 (2019).

[42] K. L. Bak, J. Gauss, T. Helgaker, P. Jørgensen and J. Olsen, Chem. Phys. Lett. **319**, 563 (2000).

[43] A. A. Holmes, C. J. Umrigar and S. Sharma, J. Chem. Phys. **147**, 164111 (2017).

3

Basis-set correction for coupled-cluster estimation of dipole moments

This chapter corresponds to the article [D. Traore, J. Toulouse, E. Giner, J. Chem. Phys. **156**, 174101 (2022)]

The present work proposes an approach to obtain a basis-set correction based on density-functional theory (DFT) for the computation of molecular properties in wave-function theory (WFT). This approach allows one to accelerate the basis-set convergence of any energy derivative of a non-variational WFT method, generalizing previous works on the DFT-based basis-set correction where either only ground-state energies could be computed with non-variational wave functions [J. Phys. Chem. Lett. **10**, 2931 (2019)] or properties could be computed as expectation values over variational wave functions [J. Chem. Phys. **155**, 044109 (2021)]. This work focuses on the basis-set correction of dipole moments in coupled-cluster with single, double, and perturbative triple excitations (CCSD(T)), which is numerically tested on a set of fourteen molecules with dipole moments covering two orders of magnitude. As the basis-set correction relies only on Hartree-Fock densities, its computational cost is marginal with respect to the one of the CCSD(T) calculations. Statistical analysis of the numerical results shows a clear improvement of the basis convergence of the dipole moment with respect to the usual CCSD(T) calculations.

3.1 Introduction

Quantum chemistry aims to provide theoretical methods to predict molecular properties starting from the many-body quantum mechanical problem. To solve this problem a wide range of methods were developed in the last few decades mainly based on wave-function theory (WFT) and density-functional theory (DFT). The purpose of both approaches is to accurately treat correlation effects, or in other terms, the quantum effects which go beyond a mean-field description such as Hartree-Fock (HF). In the context of WFT, there exists a wide range of methods of increasing computational cost – ranging from Møller-Plesset perturbation theory [1] to coupled-cluster methods [2] – which in principle systematically converge to the full configuration interaction (FCI) limit which is the exact solution within a given basis set. Nevertheless, the accuracy of the results of a WFT method – even at the FCI level – strongly depends on the quality of the

basis set, mainly because of the slow convergence of the wave function near the electron-electron coalescence point [3, 4]. The combination of the slow basis-set convergence and the strong increase of the computational cost with both the size of the basis set and the number of electrons makes it very difficult to obtain well converged WFT calculations on large systems.

There are mainly two approaches to tackle the basis-set convergence problem of WFT: basis-set extrapolation techniques [5, 6] and explicitly correlated F12 methods [7–12]. The basis-set extrapolation techniques rely on a known asymptotic behavior of the correlation energy with the size of the basis set but requires WFT calculations with basis sets of increasing sizes, which makes their application limited to small or medium system size. The F12 methods accelerate the basis-set convergence of the results thanks to the inclusion of a correlation factor explicitly depending on electron-electron distances and restoring Kato's electron-electron cusp condition [4]. Although F12 methods improve indeed the results (typically, energy differences obtained with a F12 method using a triple-zeta basis set are as accurate as the ones obtained with the corresponding uncorrected WFT method using a quintuple-zeta basis set [13]), the F12 methods necessarily induce computational overheads due to the large auxiliary basis sets required to resolve three- and four-electron integrals [14].

An alternative path has been recently introduced by some of the present authors in Ref. [15] where a rigorous framework was proposed to correct for the basis-set incompleteness of WFT using DFT. A central idea of this work is the fact that the Coulomb electron-electron interaction projected in an incomplete basis set is non-divergent and quite similar to the long-range interaction used in range-separated DFT (RSDFT). A basis-set correction density functional can then be built from RSDFT short-range correlation functionals using a local range-separation parameter which automatically adapts to the basis set used. This results in a relatively cheap way of correcting the basis-set incompleteness of WFT, which has the desirable property of leading to an unaltered complete-basis-set (CBS) limit. Two versions of this theory were proposed: (i) a non self-consistent version where the basis-set correction functional is evaluated with any accurate approximation of the FCI density and then simply added to an approximation of the FCI energy in a given basis set [15]; and (ii) a recently introduced self-consistent variant were the energy is minimized in the presence of the basis-set correction functional and therefore allows for the wave function to be changed by the DFT correction [16]. The efficiency of the non-self-consistent approach for computing total energies and chemically relevant energy differences of relatively large magnitudes (such as ionization potentials [15, 17], molecular atomization energies [18–21], or excitation energies [22]) has been well established in previous works on a quite wide range of atomic and molecular systems including light to transition-metal elements, and it was numerically shown that the self-consistent framework does not give any significant improvement of total energies [16].

The main advantage of the self-consistent formulation is nevertheless to allow for the computation of first-order properties as expectation values over the minimized wave function thanks to the variational property of the theory. In Ref. [16] the present authors have focused on the dipole moments which are known to exhibit a slow convergence with respect to the size of the basis set [23–25]. It was shown that the dipole moments computed at near FCI level with the self-consistent basis-set correction method are close to the CBS limit in triple-zeta basis sets, which contrasts with the slow basis-set convergence of the usual WFT approaches. The drawback of the self-consistent framework is nonetheless to require a self-consistent variational WFT calculation, which therefore excludes its application to non-variational approaches such as coupled-cluster with singles, doubles, and perturbative triple excitations (CCSD(T)).

In the present work, we propose to overcome this limitation and target the computation of first-order molecular properties as energy derivatives of the non-self-consistent basis-set correction approach. We apply this strategy to the computation of dipole moments at the CCSD(T) level and propose a cheap computational strategy for the basis-set correction which uses only

densities at the HF level, similarly to what have been done in the context of atomization energies in Ref. [18].

This paper is organized as follows. In Sec. 3.2, we introduce the theory of the basis-set correction extended to the computation of dipole moments. In Sec. 3.3, we provide computational details of our study on a set of fourteen molecules with dipole moments covering two orders of magnitude. The numerical results are discussed in Sec. 3.4, and compared for some molecules with the fully self-consistent formalism of Ref. [16]. Detailed results, as well as the molecular geometries used, are available in the Supplementary material (Appendix C).

3.2 Theory

3.2.1 Dipole moment from the self-consistent basis-set correction

In this section, we generalize the framework of the basis-set correction to the presence of a static external electric field. Consider the Hamiltonian of a N-electron system under an external electric field $\boldsymbol{\epsilon} = \epsilon \boldsymbol{u}$ of strength ϵ along a direction $\boldsymbol{u}$,

$$\hat{H}(\epsilon) = \hat{H}_0 - \epsilon \hat{d}, \tag{3.1}$$

where $\hat{H}_0$ is the Hamiltonian of the system without the electric field,

$$\hat{H}_0 = \hat{T} + \hat{V}_{\text{ne}} + \hat{W}_{\text{ee}}, \tag{3.2}$$

with $\hat{T}$ being the kinetic-energy operator, $\hat{V}_{\text{ne}}$ being the electron-nuclei interaction operator, and $\hat{W}_{\text{ee}}$ the electron-electron interaction operator, and $\hat{d} = \hat{\boldsymbol{d}} \cdot \boldsymbol{u}$ where $\hat{\boldsymbol{d}}$ is the total (electron+nuclear) dipole-moment operator,

$$\hat{\boldsymbol{d}} = -\sum_{i=1}^{N} \boldsymbol{r}_i + \sum_{A=1}^{N_{\text{nuclei}}} Z_A \boldsymbol{R}_A, \tag{3.3}$$

where $\boldsymbol{r}_i$ are the electron coordinates, and Z_A and $\boldsymbol{R}_A$ are the nuclei charges and coordinates.

In the basis-set correction formalism [15,16,19], the ground-state energy $E_0(\epsilon)$ of the Hamiltonian in Eq. (3.1) is approximated by

$$E_0^{\mathcal{B}}(\epsilon) = \min_{\Psi^{\mathcal{B}}} \left\{ \langle \Psi^{\mathcal{B}} | \hat{H}(\epsilon) | \Psi^{\mathcal{B}} \rangle + \bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}] \right\}, \tag{3.4}$$

where the minimization is performed over the set of N-electron wave functions $\Psi^{\mathcal{B}}$ expanded on the N-electron Hilbert space generated by the one-electron basis set $\mathcal{B}$ and $\bar{E}^{\mathcal{B}}[n_{\Psi^{\mathcal{B}}}]$ is the basis-set correction functional evaluated at the density $n_{\Psi^{\mathcal{B}}}$ of $\Psi^{\mathcal{B}}$. The energy functional $\bar{E}^{\mathcal{B}}[n]$ (introduced in Ref. [15]) compensates for the restriction on the wave functions $\Psi^{\mathcal{B}}$ due to the incompleteness of the basis set $\mathcal{B}$. The restriction coming from the basis set $\mathcal{B}$ in Eq. (3.4) then applies only to densities $n_{\Psi^{\mathcal{B}}}$. Roughly speaking, since the density converges much faster than the wave function with respect to the basis set, $E_0^{\mathcal{B}}(\epsilon)$ is a much better approximation to the exact energy $E_0(\epsilon)$ than the corresponding FCI ground-state energy $E_{\text{FCI}}^{\mathcal{B}}(\epsilon)$ calculated with the same basis set $\mathcal{B}$. Moreover, in the CBS limit, $\bar{E}^{\mathcal{B}}[n]$ vanishes and thus $E_0^{\mathcal{B}}(\epsilon)$ correctly converges to the exact energy $E_0(\epsilon)$.

From the basis-set corrected energy $E_0^{\mathcal{B}}(\epsilon)$ in Eq. (3.4), one can define the corresponding basis-set corrected dipole moment $d^{\mathcal{B}}$ as the first-order derivative with respect to the electric field

$$d^{\mathcal{B}} = -\left. \frac{\mathrm{d}E_0^{\mathcal{B}}(\epsilon)}{\mathrm{d}\epsilon} \right|_{\epsilon=0}. \tag{3.5}$$

It is important to stress here that $d^{\mathcal{B}}$ is different from the FCI dipole moment $d^{\mathcal{B}}_{\text{FCI}}$ with the same basis set $\mathcal{B}$, as the former is taken as the derivative of $E^{\mathcal{B}}_0(\epsilon)$ which contains the basis-set correction functional $\bar{E}^{\mathcal{B}}[n]$. Similarly to the case of the energy, we expect $d^{\mathcal{B}}$ to have a faster basis-set convergence than $d^{\mathcal{B}}_{\text{FCI}}$.

Since $E^{\mathcal{B}}_0(\epsilon)$ is stationary with respect to $\Psi^{\mathcal{B}}$, the Hellmann-Feynman theorem applies and gives $d^{\mathcal{B}}$ as a simple expectation value

$$d^{\mathcal{B}} = \langle \Psi^{\mathcal{B}}_0(\epsilon = 0) | \hat{d} | \Psi^{\mathcal{B}}_0(\epsilon = 0) \rangle, \tag{3.6}$$

where $\Psi^{\mathcal{B}}_0(\epsilon = 0)$ is the minimizing wave function of the self-consistent equation in Eq. (3.4) at $\epsilon = 0$. This was the approach used in Ref. [16].

Note that the Hellmann-Feynman theorem applies because the basis set $\mathcal{B}$ is independent from the perturbing electric field. By contrast, if we wanted to calculate the gradient of the energy with respect to nuclear coordinates, we would have to take into account the dependence of the atomic basis set $\mathcal{B}$ on the nuclei coordinates. In that case, the Hellmann-Feynman theorem would not apply and we would have to consider additional Pulay terms coming from the dependence of both the wave function $\Psi^{\mathcal{B}}_0$ and the basis-set correction energy $\bar{E}^{\mathcal{B}}$ on the nuclei coordinates.

3.2.2 Dipole moment from the non-self-consistent basis-set correction

As initially proposed in Ref. [15] for the case without the electric field, one can avoid the minimization in Eq. (3.4) and approximate the energy $E^{\mathcal{B}}_0(\epsilon)$ by approximating the minimizing wave function $\Psi^{\mathcal{B}}_0(\epsilon)$ in Eq. (3.4) by the FCI wave function $\Psi^{\mathcal{B}}_{\text{FCI}}(\epsilon)$ in a given basis set $\mathcal{B}$. This leads to the following estimation of the ground-state energy

$$E^{\mathcal{B}}_0(\epsilon) \approx E^{\mathcal{B}}_{\text{FCI}}(\epsilon) + \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}_{\text{FCI}}(\epsilon)], \tag{3.7}$$

where $n^{\mathcal{B}}_{\text{FCI}}(\epsilon)$ is the ground state FCI density obtained in the presence of the electric field of strength ϵ. The corresponding non-self-consistent basis-set corrected dipole moment is thus

$$d^{\mathcal{B}} \approx d^{\mathcal{B}}_{\text{FCI}} + \bar{d}^{\mathcal{B}}, \tag{3.8}$$

where

$$\bar{d}^{\mathcal{B}} = -\left. \frac{\mathrm{d}\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}_{\text{FCI}}(\epsilon)]}{\mathrm{d}\epsilon} \right|_{\epsilon=0} \tag{3.9}$$

is the non-self-consistent basis-set correction to the dipole moment.

As obtaining both the dipole moment and the density at FCI level is often computationally prohibitive, we follow Ref. [18] and approximate the FCI energy by the CCSD(T) energy and the FCI density by the HF density (in the presence of the electric field)

$$E^{\mathcal{B}}_0(\epsilon) \approx E^{\mathcal{B}}_{\text{CCSD(T)}}(\epsilon) + \bar{E}^{\mathcal{B}}[n^{\mathcal{B}}_{\text{HF}}(\epsilon)]. \tag{3.10}$$

Within these approximations, the basis-set corrected dipole moment in Eq. (3.8) becomes

$$d^{\mathcal{B}} \approx d^{\mathcal{B}}_{\text{CCSD(T)}} + \bar{d}^{\mathcal{B}}, \tag{3.11}$$

where $d^{\mathcal{B}}_{\text{CCSD(T)}}$ is the dipole moment at the CCSD(T) level and the basis-set correction $\bar{d}^{\mathcal{B}}$ is

$$\bar{d}^{\mathcal{B}} = -\left. \frac{\mathrm{d}\bar{E}^{\mathcal{B}}[n^{\mathcal{B}}_{\text{HF}}(\epsilon)]}{\mathrm{d}\epsilon} \right|_{\epsilon=0}. \tag{3.12}$$

We approximate the basis-set correction functional $\bar{E}^{\mathcal{B}}[n]$ with the so-called (spin-dependent) PBEUEG energy functional introduced in Ref. [18] where the local range-separation parameter $\mu^{\mathcal{B}}(\boldsymbol{r})$ is obtained using the HF wave function in the basis set $\mathcal{B}$ as proposed in Refs. [15] and [18]. The results obtained with Eq. (3.11) with the PBEUEG approximation for $\bar{E}^{\mathcal{B}}[n]$ evaluated at the HF density will be referred to as CCSD(T)+PBEUEG.

In practice, we calculate the CCSD(T) dipole moment and the basis-set correction to the dipole moment using a finite-difference approximation for the energy derivatives with respect to the electric field

$$d^{\mathcal{B}}_{\mathrm{CCSD(T)}} \simeq -\frac{E^{\mathcal{B}}_{\mathrm{CCSD(T)}}(\epsilon) - E^{\mathcal{B}}_{\mathrm{CCSD(T)}}(-\epsilon)}{2\epsilon}, \tag{3.13}$$

and

$$\bar{d}^{\mathcal{B}} \simeq -\frac{\bar{E}^{\mathcal{B}}\left[n^{\mathcal{B}}_{\mathrm{HF}}(\epsilon)\right] - \bar{E}^{\mathcal{B}}\left[n^{\mathcal{B}}_{\mathrm{HF}}(-\epsilon)\right]}{2\epsilon}, \tag{3.14}$$

using a finite field strength of $\epsilon = 10^{-4}$ a.u., as suggested in Ref. [23].

3.3 Computational details

The computation of the basis-set correction to the dipole moment $\bar{d}^{\mathcal{B}}$ were performed using the Quantum Package program [26] and the CCSD(T) dipole moment were obtained with the Gaussian program [27]. We used the augmented Dunning basis sets (Refs. [28, 29]) aug-cc-pVXZ (abbreviated as AVXZ in the tables and figures of the paper) where X is the cardinal number of the basis set $\mathrm{X} \in \{\mathrm{D}, \mathrm{T}, \mathrm{Q}, 5\}$. As no core-valence functions are used, the frozen-core approximation is used throughout this paper where the 1s orbital is kept frozen for the elements from Li to F.

The tests are done on a set of $n = 14$ molecules among which six open-shell molecules, for which we use restricted open-shell CCSD(T) (ROCCSD(T)) energies and restricted open-shell HF (ROHF) densities, and eight closed-shell molecules. Experimental geometries used for the computations are taken from Ref. [25] for the entire set except in the case of BH and FH for which the geometries are taken from Ref. [23]. We also report the results obtained in Ref. [16] for the BH, FH, CH_2, and H_2O molecules using the self-consistent formalism [Eq. (3.6)] at near-FCI level in order to compare with the present non-self-consistent formalism.

The accuracy of the dipole moments obtained with a given basis set and a given level of approximation is evaluated with respect to the CBS limit of the CCSD(T) dipole moments, $d^{\mathrm{CBS}}_{\mathrm{CCSD(T)}}$, which are evaluated as in Ref. [23]. In particular, the CBS results are computed as follows

$$d^{\mathrm{CBS}}_{\mathrm{CCSD(T)}} = d^{\mathrm{CBS}}_{\mathrm{HF}} + d^{\mathrm{CBS}}_{\mathrm{c}}, \tag{3.15}$$

where $d^{\mathrm{CBS}}_{\mathrm{c}}$ is the CBS limit of the correlation contribution to the CCSD(T) dipole moment which is computed using the following two-point X^{-3} extrapolation formula

$$d^{\mathrm{CBS}}_{\mathrm{c}} = \frac{d^{\mathrm{X}}_{\mathrm{c}} X^3 - d^{(\mathrm{X-1})}_{\mathrm{c}} (X-1)^3}{X^3 - (X-1)^3}, \tag{3.16}$$

with

$$d^{\mathrm{X}}_{\mathrm{c}} = d^{\mathrm{X}}_{\mathrm{CCSD(T)}} - d^{\mathrm{X}}_{\mathrm{HF}}, \tag{3.17}$$

where $d^{\mathrm{X}}_{\mathrm{c}}$ and $d^{\mathrm{X}}_{\mathrm{HF}}$ are the correlation and HF contributions, respectively, to the CCSD(T) dipole moment using the aug-cc-pVXZ basis set. In the present calculations, we evaluate Eq. (3.16) at $X = 5$ and we estimate the CBS limit of HF dipole moment $d^{\mathrm{CBS}}_{\mathrm{HF}}$ in Eq. (3.15) simply by using the HF dipole moment in the aug-cc-pV5Z basis set. For all the systems studied here, the

HF dipole moments are converged with an accuracy better than 0.001 a.u. (as measured by the difference between the aug-cc-pVQZ and aug-cc-pV5Z dipole moments).

At a given level of calculation in a basis set $\mathcal{B}$ we report the error on the dipole moment with respect to the CBS reference $\Delta^{\mathcal{B}} = d^{\mathcal{B}} - d^{\text{CBS}}_{\text{CCSD(T)}}$ and the relative error $\Delta^{\mathcal{B}}_{\text{rel}} = \Delta^{\mathcal{B}}/d^{\text{CBS}}_{\text{CCSD(T)}}$. To statistically analyze the results, we also calculate the normal distribution function of the errors for a given basis set $\mathcal{B}$,

$$\rho(\Delta^{\mathcal{B}}) = \frac{1}{\Delta^{\mathcal{B}}_{\text{std}}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\Delta^{\mathcal{B}} - \bar{\Delta}^{\mathcal{B}}}{\Delta^{\mathcal{B}}_{\text{std}}}\right)^2\right], \tag{3.18}$$

where $\bar{\Delta}^{\mathcal{B}} = (\sum_{i=1}^{n} \Delta^{\mathcal{B}}_i)/n$ is the mean error (ME) and $\Delta^{\mathcal{B}}_{\text{std}} = \sqrt{\sum_{i=1}^{n}(\Delta^{\mathcal{B}}_i - \bar{\Delta}^{\mathcal{B}})^2/(n-1)}$ is the root-mean-square deviation (RMSD).

3.4 Results and discussion

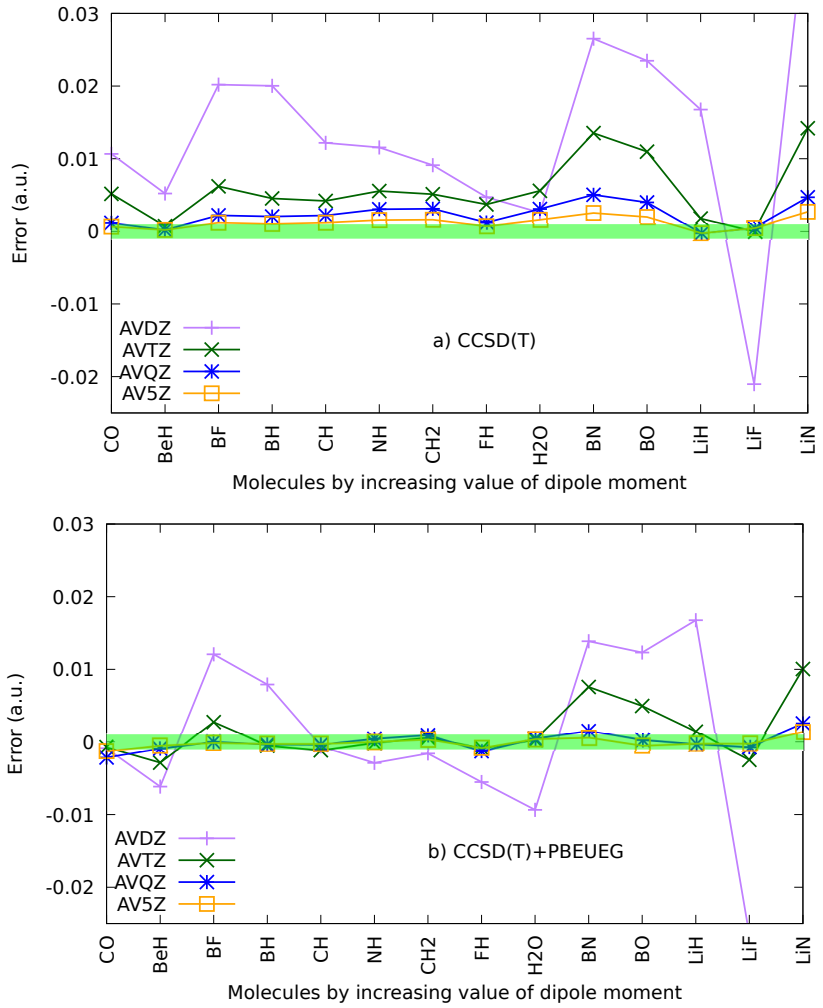

Figure 3.1: (a) CCSD(T) and (b) CCSD(T)+PBEUEG errors on the dipole moments of 14 molecules compared to CCSD(T)/CBS reference values. The green area indicates an error of ±0.001 a.u..

In Table 3.1, we report the dipole moments at various levels of approximations (HF, CCSD(T), CCSD(T) + PBEUEG) with different basis sets, as well as the CCSD(T)/CBS reference values,

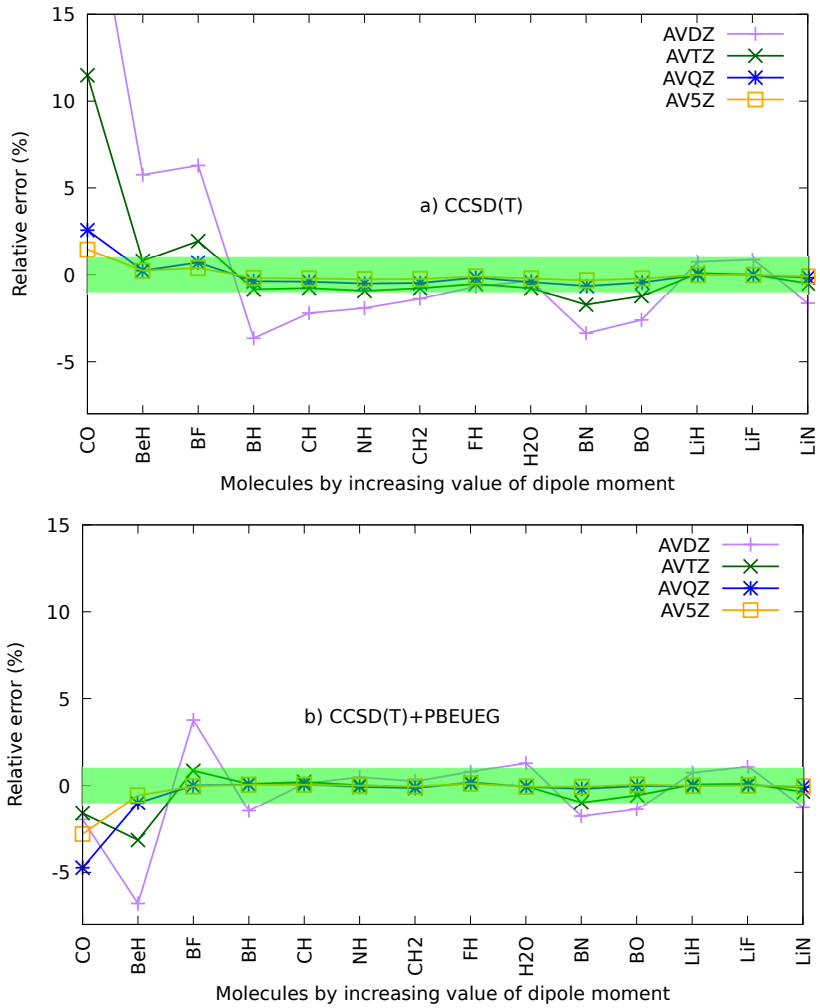

Figure 3.2: (a) CCSD(T) and (b) CCSD(T)+PBEUEG relative errors on the dipole moments of 14 molecules compared to CCSD(T)/CBS reference values. The green area indicates an error of ±1%.

for the set of 14 molecules. Note the wide range of magnitudes of the dipole moments (from 0.04485 a.u. for CO to 2.78718 a.u. for LiN). The mean error (ME), mean absolute error (MAE), mean absolute relative error (MARE), maximal absolute error (MAX), and root-mean-square deviation (RMSD) obtained with CCSD(T) and CCSD(T)+PBEUEG are reported in Table 3.2. The graphical representations of this data are provided in Figs. 3.1 and 3.2 for the errors and relative errors, and in Fig. 3.3 for the normal distributions of errors.

Analyzing first the results at the CCSD(T) level in Table 3.2, one can notice that, as expected, the ME and MAE systematically decrease with the size of the basis set. Moreover, as noted in previous studies [18], not only the average values of the errors but also the RMSD tends to decrease with the basis-set size. Nevertheless, the improvement of the results is rather slow as a MAE below 0.001 a.u. is not reached even with the aug-cc-pV5Z basis set, illustrating the slow convergence of properties with respect to the basis set at the CCSD(T) level. Regarding the relative errors in Fig. 3.2, not surprisingly, the largest errors with respect to the CBS reference come from the molecules with smallest dipole moments (*i.e.* CO and BeH). More quantitatively, an aug-cc-pVQZ basis set is needed to obtain a MARE smaller than 1%.

Going from CCSD(T) to CCSD(T)+PBEUEG, one observes a systematic decrease of the

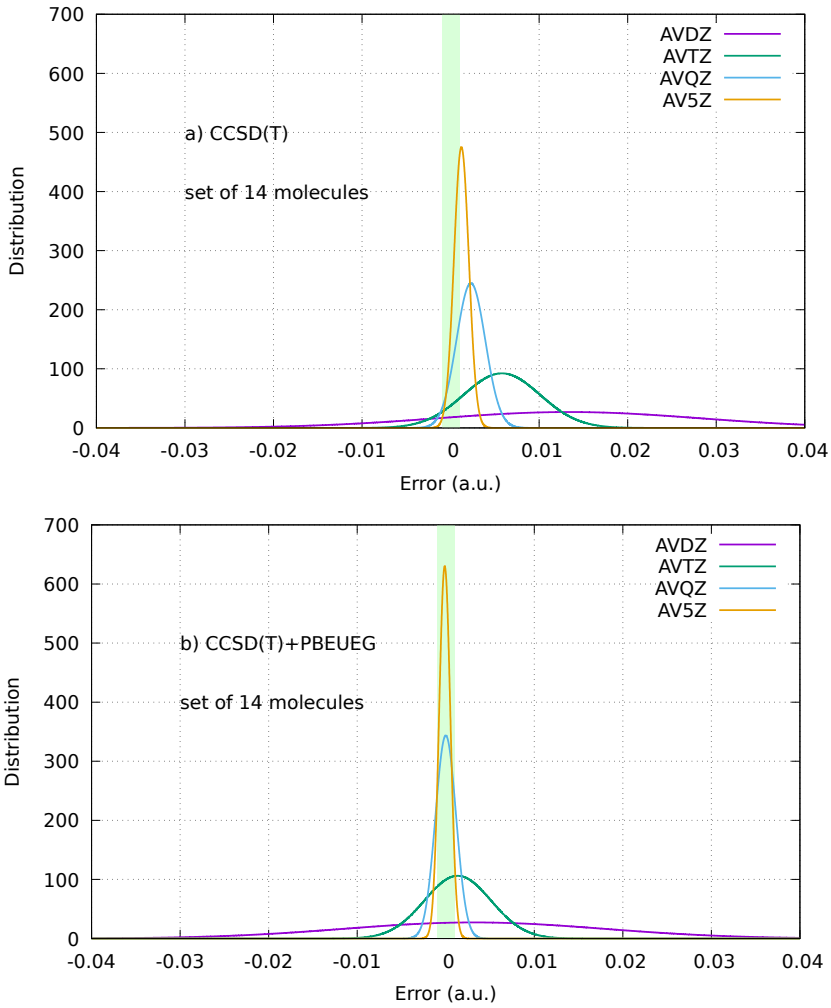

Figure 3.3: (a) CCSD(T) and (b) CCSD(T)+PBEUEG normal distribution of errors on the dipole moments of 14 molecules compared to CCSD(T)/CBS reference values. The green area indicates an error of ± 0.001 a.u..

MAE, ME, MARE, and RMSD. Focusing on the MAE, an error below 0.001 a.u. is reached with the aug-cc-pVQZ basis set, whereas such an accuracy is not even reached at the CCSD(T) level with the aug-cc-pV5Z basis set. Qualitatively, for the aug-cc-pVTZ basis set and larger basis sets, the MAEs obtained with CCSD(T)+PBEUEG with a basis set of cardinal number X are comparable to the MAEs obtained with CCSD(T) with a basis set of cardinal number X+1. Regarding the MARE, an error below 1% is reached with CCSD(T)+PBEUEG already with the aug-cc-pVTZ basis set. One nevertheless observes that the effects of the basis-set correction on the RMSD is very weak. From the plots of Fig. 3.1 one notes that even if the basis-set correction systematically improves the results for the aug-cc-pVTZ basis set, its effect is less impressive when there is both a large error and a large dipole moment (*i.e.* for BN, BO, and LiN).

In order to further demonstrate the validity of the different approximations leading to the CCSD(T) + PBEUEG method, we conclude this study by a comparison with the self-consistent basis-set correction formalism of Ref. [16], as well as different flavors of non-self-consistent approximations. In Ref. [16], the self-consistent method referred to as SC CIPSI+PBEUEG was introduced, which can be considered as the nearly exact theory within our framework thanks to the use of near-FCI (CIPSI, see [26] and references therein) wave functions. In all the calculations reported below, the absolute value of the second-order perturbative contribution to

the energy in the CIPSI calculations is below 10^{-4} a.u., which implies that the CIPSI energy and density can indeed be considered as near-FCI quantities. We also consider two different levels of non-self-consistent approximations: (i) CIPSI+PBEUEG@CIPSI where the CIPSI energy is corrected with the PBEUEG functional evaluated at the CIPSI density, and (ii) CCSD(T)+PBEUEG@CIPSI where the CIPSI energy is approximated by the CCSD(T) energy but the PBEUEG functional is still evaluated at the CIPSI density. Therefore, we have a hierarchy of approximations for the basis-set correction method using the PBEUEG functional: SC CIPSI+PBEUEG as the exact self-consistent theory, then CIPSI+PBEUEG@CIPSI as the exact non-self-consistent theory, then CCSD(T)+PBEUEG@CIPSI where only the WFT energy part is approximated with respect to CIPSI+PBEUEG@CIPSI, and finally CCSD(T)+PBEUEG where both the WFT energy and the density are approximated with respect to CIPSI+PBEUEG@CIPSI.

We report in Table 3.3 the results obtained with these different levels of theory for the dipole moments of the BH, CH_2, FH, and H_2O molecules. The results obtained with the self-consistent method SC CIPSI+PBEUEG are in close agreement with that obtained with the different non-self-consistent approximations, the largest discrepancy being less than 0.006 a.u. for BH in the aug-cc-pVDZ basis set with CCSD(T)+PBEUEG@CIPSI. Comparing the two methods at the extremities of our hierarchy of approximations, one can notice that the absolute deviation between CCSD(T)+PBEUEG and SC CIPSI+PBEUEG in a given basis set is never larger than 0.001 a.u. for FH and H_2O, and the discrepancy slightly increases up to 0.006 and 0.003 a.u. in the case of CH_2 and BH, respectively. Nevertheless, as originally reported in Ref. [16] and apparent from Table 3.3, discrepancies of the same order of magnitude also appear between the uncorrected CIPSI and CCSD(T) results in the case of the CH_2 and BH molecules. This suggests that the main source of differences between the CCSD(T)+PBEUEG and SC CIPSI+PBEUEG methods actually comes from the parent WFT theory. Focussing now specifically on the effect of the density in the non-self-consistent basis-set correction, one can notice that the use of either a HF or CIPSI density does not significantly change the results, as the largest deviation between CCSD(T)+PBEUEG@CIPSI and CCSD(T)+PBEUEG are about 0.002 a.u. in the case of CH_2 in the aug-cc-pVDZ basis set. These results illustrate the validity of the different approximations leading to the CCSD(T)+PBEUEG approach and are encouraging considering that the latter has a much lower computational cost with respect to the self-consistent basis-set formalism. Indeed, CCSD(T)+PBEUEG relies only on a standard CCSD(T) calculation and HF calculations for the basis-set correction which is of negligible computational cost with respect to CCSD(T).

Table 3.1: HF, CCSD(T), and CCSD(T)+PBEUEG dipole moments in atomic units. For the open-shell systems, we use the spin-restricted open-shell (RO) version of these methods.

	AVDZ	AVTZ	AVQZ	AV5Z	CBS
CO					
HF	-0.10199	-0.10499	-0.10433	-0.10421	
CCSD(T)	0.05550	0.05000	0.04600	0.04550	0.04485
CCSD(T) + PBEUEG	0.04398	0.04414	0.04273	0.04360	
BeH					
ROHF	0.11017	0.11076	0.11199	0.11218	
ROCCSD(T)	0.09550	0.09100	0.09050	0.09050	0.09030
ROCCSD(T) + PBEUEG	0.08416	0.08746	0.08941	0.08980	
BF					
HF	0.34436	0.33390	0.33314	0.33328	
CCSD(T)	0.34100	0.32700	0.32300	0.32200	0.32081
CCSD(T) + PBEUEG	0.33287	0.32351	0.32082	0.32068	
BH					
HF	0.68796	0.68649	0.68494	0.68496	
CCSD(T)	0.52950	0.54500	0.54750	0.54850	0.54953
CCSD(T)+PBEUEG	0.54162	0.55002	0.54986	0.54980	
CH					
ROHF	0.62348	0.62000	0.61871	0.61858	
ROCCSD(T)	0.54150	0.54950	0.55150	0.55250	0.55368
ROCCSD(T) + PBEUEG	0.55427	0.55481	0.55405	0.55396	
NH					
ROHF	0.63850	0.63505	0.63381	0.63384	
ROCCSD(T)	0.59350	0.59950	0.60200	0.60350	0.60504
ROCCSD(T) + PBEUEG	0.60792	0.60519	0.60464	0.60506	
CH_2 (singlet)					
HF	0.74877	0.74477	0.74355	0.74353	
CCSD(T)	0.65600	0.66000	0.66200	0.66350	0.66510
CCSD(T) + PBEUEG	0.66666	0.66455	0.66420	0.66478	
FH					
HF	0.75976	0.75751	0.75634	0.75617	
CCSD(T)	0.70350	0.70450	0.70700	0.70750	0.70820
CCSD(T) + PBEUEG	0.71371	0.70903	0.70946	0.70900	
H_2O					
HF	0.78671	0.78039	0.77956	0.77956	
CCSD(T)	0.72700	0.72400	0.72650	0.72800	0.72957
CCSD(T) + PBEUEG	0.73891	0.72930	0.72912	0.72920	
BN					
ROHF	1.13451	1.13862	1.13831	1.13840	
ROCCSD(T)	0.76250	0.77550	0.78400	0.78650	0.78902
ROCCSD(T) + PBEUEG	0.77517	0.78145	0.78756	0.78846	
BO					
ROHF	1.17803	1.18533	1.18527	1.18539	
ROCCSD(T)	0.88300	0.89550	0.90250	0.90450	0.90647
ROCCSD(T) + PBEUEG	0.89417	0.90153	0.90622	0.90698	
LiH					
HF	2.37055	2.36235	2.36153	2.36129	
CCSD(T)	2.32500	2.31000	2.30800	2.30800	2.30825
CCSD(T)+PBEUEG	2.32501	2.30965	2.30795	2.30802	
LiF					
HF	2.56111	2.54103	2.53949	2.53905	
CCSD(T)	2.50400	2.48300	2.48250	2.48250	2.48297
CCSD(T) + PBEUEG	2.50942	2.48542	2.48367	2.48321	
LiN					
ROHF	2.90309	2.90379	2.90372	2.90317	
ROCCSD(T)	2.74200	2.77300	2.78250	2.78450	2.78718
ROCCSD(T) + PBEUEG	2.75215	2.77714	2.78464	2.78583	

Table 3.2: Mean error (ME), mean absolute error (MAE), mean absolute relative error (MARE), maximal absolute error (MAX), and root-mean-square deviation (RMSD) (in atomic units) for the CCSD(T) and CCSD(T)+PBEUEG dipole moments of 14 molecules. See Fig. 3.3 for the corresponding plots of the normal distributions of errors.

	AVDZ	AVTZ	AVQZ	AV5Z
ME				
CCSD(T)	0.01336	0.00579	0.00229	0.00122
CCSD(T)+PBEUEG	0.00319	0.00135	0.000004	-0.00012
MAE				
CCSD(T)	0.01637	0.00579	0.00233	0.00125
CCSD(T)+PBEUEG	0.01080	0.00258	0.00086	0.00049
MARE (in %)				
CCSD(T)	3.9	1.5	0.5	0.3
CCSD(T)+PBEUEG	1.6	0.6	0.4	0.3
MAX				
CCSD(T)	0.04518 (LiN)	0.01418 (LiN)	0.00502 (BN)	0.00268 (LiN)
CCSD(T)+PBEUEG	0.03504 (LiN)	0.01004 (LiN)	0.00254 (LiN)	0.00136 (LiN)
RMSD				
CCSD(T)	0.01484	0.00432	0.00163	0.00084
CCSD(T)+PBEUEG	0.01464	0.00376	0.00116	0.00063

Table 3.3: Dipole moments obtained with near-FCI (CIPSI) and CCSD(T) calculations, and with the self-consistent basis-set correction method (SC CIPSI+PBEUEG) of Ref. [16] and different non-self-consistent basis-set correction methods (CIPSI+PBEUEG@CIPSI, CCSD(T)+PBEUEG@CIPSI, and CCSD(T)+PBEUEG). Estimated CBS values using Eq. (3.16) with $X = 5$ are reported when computations could be done with the aug-cc-pV5Z basis set. [a] From Ref. [16]. [b] Results non available due to the computational requirement.

	AVDZ	AVTZ	AVQZ	AV5Z	CBS
BH					
CIPSI[a]	0.52782	0.54334	0.54563	0.54691	0.54823
SC CIPSI+PBEUEG[a]	0.53791	0.54815	0.54790	0.54815	
CIPSI+PBEUEG@CIPSI	0.54139	0.54852	0.54795	0.54825	
CCSD(T)	0.52950	0.54500	0.54750	0.54850	0.54953
CCSD(T)+PBEUEG@CIPSI	0.54307	0.55023	0.54988	0.54996	
CCSD(T)+PBEUEG	0.54162	0.55002	0.54986	0.54980	
CH_2 (singlet)					
CIPSI[a]	0.65120	0.65446	0.65643	0.65780	0.65926
SC CIPSI+PBEUEG[a]	0.66249	0.65958	0.65890	—[b]	
CIPSI+PBEUEG@CIPSI	0.66382	0.66029	0.65952	—[b]	
CCSD(T)	0.65600	0.66000	0.66200	0.66350	0.66510
CCSD(T)+PBEUEG@CIPSI	0.66874	0.66556	0.66448	—[b]	
CCSD(T)+PBEUEG	0.66666	0.66455	0.66420	0.66478	
FH					
CIPSI[a]	0.70249	0.70406	0.70662	—[b]	—[b]
SC CIPSI+PBEUEG[a]	0.71326	0.70873	—[b]	—[b]	
CIPSI+PBEUEG@CIPSI	0.71329	0.71188	—[b]	—[b]	
CCSD(T)	0.70350	0.70450	0.70700	0.70750	0.70820
CCSD(T)+PBEUEG@CIPSI	0.71425	0.71209	—[b]	—[b]	
CCSD(T)+PBEUEG	0.71371	0.70903	0.70946	0.70900	
H_2O					
CIPSI[a]	0.72610	0.72294	—[b]	—[b]	—[b]
SC CIPSI+PBEUEG[a]	0.73809	0.72818	—[b]	—[b]	
CIPSI+PBEUEG@CIPSI	0.73656	0.72762	—[b]	—[b]	
CCSD(T)	0.72700	0.72400	0.72650	0.72800	0.72957
CCSD(T)+PBEUEG@CIPSI	0.73734	0.72819	0.72872	—[b]	
CCSD(T)+PBEUEG	0.73891	0.72930	0.72912	0.72920	

3.5 Conclusion

In the present study, we have proposed an extension of the recently introduced non-self-consistent basis-set correction of CCSD(T) ground-state energies [18] to the computation of properties as energy derivatives, focussing here on the dipole moment. The theory relies on the originally proposed DFT-based basis-set correction approach [15] which accelerates the basis-set convergence to the unaltered CBS limit. Numerical tests on a set of 14 molecules (including both closed and open-shell) with dipole moments spanning two orders of magnitude have been carried in order

to obtain a representative study of the performance of the present approach.

Although this study aims at correcting the basis-set convergence of the CCSD(T) dipole moments, it can be formally generalized to any wave-function method and any energy derivative with respect to a static perturbation. In its present form, the basis-set correction relies only on HF calculations, which makes the basis-set correction essentially computationally free compared to the correlated wave-function calculation. This approach is an alternative to the recently proposed self-consistent basis-set correction [16] which allows for the computation of first-order properties through expectation values over an energy-minimized wave function. In contrast with the self-consistent formalism, the present approach does not require a variational wave function, which considerably extends the domain of application of the basis-set correction.

Regarding now the numerical results, we have shown that the present approach significantly accelerates the basis-set convergence of CCSD(T) dipole moments. Typically, the error obtained in a basis set of cardinal X with the basis-set correction is comparable to the error of the uncorrected CCSD(T) calculation with cardinal number $X+1$. We also compared the present non-self-consistent basis-set correction with the self-consistent formalism of Ref. [16] and showed that the two theories agree within a few milli-atomic units, illustrating the soundness of the approximations leading to the non-self-consistent approach.

Considering the generality, the global performance, and the small computational cost of the present approach, it could be an alternative to explicitly correlated approaches for calculation of molecular properties. In the near future we will extend the method to higher-order static properties, such as static polarizabilities, and also to more general dynamic properties, leading in particular to the possibility of accelerating the basis-set convergence of excitation energies.

Supplementary information

The preset work comes with supplementary information containing: i) all the geometries of the molecules studied here, ii) a graphical representation of the convergence of the dipole moment at CCSD(T) and CCSD(T)+PBEUEG levels for each system studied, iii) the absolute and relative errors with respect to the estimated CBS and the CCSD(T) and CCSD(T)+PBEUEG levels.

Acknowledgement

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme Grant agreement No. 810367 (EMC2).

Author Declarations

The authors have no conflicts to disclose.

Data Availability

The data presented in this study are available in the supplementary materials associated with the present paper.

Bibliography

[1] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (Oct 1934).

[2] R. J. Bartlett and M. Musiaλ, Rev. Mod. Phys. **79**, 291 (Feb 2007).

[3] E. A. Hylleraas, Z. Phys. **54**, 347 (1929).

[4] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).

[5] T. Helgaker, W. Klopper, H. Koch and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[6] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).

[7] S. Ten-no, Theor. Chem. Acc. **131**, 1070 (2012).

[8] S. Ten-no and J. Noga, WIREs Comput. Mol. Sci. **2**, 114 (2012).

[9] C. Hättig, W. Klopper, A. Köhn and D. P. Tew, Chem. Rev. **112**, 4 (2012).

[10] L. Kong, F. A. Bischoff and E. F. Valeev, Chem. Rev. **112**, 75 (2012).

[11] A. Grüneis, S. Hirata, Y.-Y. Ohnishi and S. Ten-no, J. Chem. Phys. **146**, 080901 (2017).

[12] Q. Ma and H.-J. Werner, WIREs Comput. Mol. Sci. **8**, e1371 (2018).

[13] D. P. Tew, W. Klopper, C. Neiss and C. Hattig, Phys. Chem. Chem. Phys. **9**, 1921 (2007).

[14] G. M. J. Barca and P.-F. Loos, J. Chem. Phys. **147**, 024103 (2017).

[15] E. Giner, B. Pradines, A. Ferté, R. Assaraf, A. Savin and J. Toulouse, J. Chem. Phys. **149**, 194301 (2018).

[16] E. Giner, D. Traore, B. Pradines and J. Toulouse, J. Chem. Phys. **155**, 044109 (2021).

[17] P.-F. Loos, B. Pradines, A. Scemama, E. Giner and J. Toulouse, J. Chem. Theory Comput. **16**, 1018 (2020).

[18] P.-F. Loos, B. Pradines, A. Scemama, J. Toulouse and E. Giner, J. Phys. Chem. Lett. **10**, 2931 (2019).

[19] E. Giner, A. Scemama, P.-F. Loos and J. Toulouse, J. Chem. Phys. **152**, 174104 (2020).

[20] Y. Yao, E. Giner, J. Li, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **153**, 124117 (2020).

[21] Y. Yao, E. Giner, T. A. Anderson, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **155**, 204104 (2021).

[22] E. Giner, A. Scemama, J. Toulouse and P.-F. Loos, J. Chem. Phys. **151**, 144118 (2019).

[23] A. Halkier, W. Klopper, T. Helgaker and P. Jørgensen, J. Chem. Phys. **111**, 4424 (1999).

[24] K. L. Bak, J. Gauss, T. Helgaker, P. Jørgensen and J. Olsen, Chem. Phys. Lett. **319**, 563 (2000).

[25] D. Hait and M. Head-Gordon, J. Chem. Theory. Comput. **14**, 1969 (2018).

[26] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Véril, M. Caffarel, P.-F. Loos, E. Giner and A. Scemama, J. Chem. Theory. Comput. **15**, 3591 (2019).

[27] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, Gaussian˜16 Revision C.01 (2016), Gaussian Inc. Wallingford CT.

[28] R. A. Kendall, T. H. Dunning and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[29] B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning and A. K. Wilson, Theor. Chem. Acc **128**, 69 (2011).

4

Basis-set correction based on density-functional theory: Rigorous framework for a one dimensional model

This chapter corresponds to the article [D. Traore, E. Giner, J. Toulouse, J. Chem. Phys **156**, 044113 (2022)].

We reexamine the recently introduced basis-set correction theory based on density-functional theory, which consists in correcting the basis-set incompleteness error of wave-function methods using a density functional. We use a one-dimensional model Hamiltonian with delta-potential interactions which has the advantage of making easier to perform a more systematic analysis than for three-dimensional Coulombic systems while keeping the essence of the slow basis convergence problem of wave-function methods. We provide some mathematical details about the theory and propose a new variant of basis-set correction which has the advantage of being suited to the development of an adapted local-density approximation. We show indeed how to develop a local-density approximation for the basis-set correction functional which is automatically adapted to the basis set employed, without resorting to range-separated density-functional theory as in previous works, but using instead a finite uniform electron gas whose electron-electron interaction is projected on the basis set. The work puts the basis-set correction theory on firmer grounds and provides an interesting strategy for the improvement of this approach.

4.1 Introduction

In electronic-structure theory of atoms, molecules or solids, one of the main limitations of standard correlated wave-function computational methods for solving the Schrödinger equation is the slow convergence of the energy and other properties with respect to the size of the one-electron basis set employed (see, e.g., Refs. [1–4]). This slow basis convergence originates from the singular behavior of the repulsive Coulomb electron-electron interaction at small interelectronic distances, which creates a depletion in the wave function with a characteristic derivative discontinuity at electron-electron coalescence — the infamous electron-electron cusp [5, 6].

The two main approaches for dealing with this problem are (i) extrapolation to the complete-basis-set limit by using increasingly large basis sets [1, 2], and (ii) explicitly correlated methods which incorporate in the wave function a correlation factor reproducing the electron-electron

cusp (see, e.g., Ref. [7]). Recently, some of the present authors introduced an alternative basis-set correction scheme based on density-functional theory (DFT) [8]. This latter scheme consists in correcting the energy calculated by a wave-function method with a finite basis set by a density functional incorporating the short-range electron correlation effects missing in the basis set. This basis-set correction scheme was further developed and tested in Refs. [9–15], demonstrating that it successfully accelerates the basis convergence of wave-function methods for various properties and systems.

The advantages of the basis-set correction scheme is its conceptual simplicity and computational efficiency. In practice, however, the limpidity of this approach is somewhat diminished by the fact that in all previously cited works the basis-set correction functional was approximated by short-range correlation functionals borrowed from range-separated DFT [16–20]), relying on an approximate mapping between the basis-set correction theory and range-separated DFT.

In the present work, we reexamine more closely the basis-set correction theory. For this, we use a one-dimensional (1D) model Hamiltonian with delta-potential interactions [21–24] which has the advantage of making easier to perform a more systematic analysis than for three-dimensional (3D) Coulombic systems, while keeping the essence of the slow basis convergence problem of wave-function methods. After introducing the 1D model and discussing its relevance in Section 5.3, we present the basis-set correction theory in some mathematical details in Section 4.3. In particular, we introduce a new variant of basis-set correction which has the advantage of being suited for development of an adapted local-density approximation (LDA). In Section 4.4, we show indeed how to develop a LDA for the basis-set correction functional which is automatically adapted to the basis set employed without resorting to range-separated DFT but using instead a finite uniform electron gas (UEG) whose electron-electron interaction is projected on the basis set. Section 4.4.3 contains our conclusion and outlook. Hartree atomic units (a.u.) are used throughout this work.

4.2 One-dimensional model system

4.2.1 Description of the model

We consider $N = 2$ electrons in a 1D He-like atom with delta-potential interactions described by the Hamiltonian [21–24]

$$H = T + W_{\mathrm{ee}} + V_{\mathrm{ne}}, \tag{4.1}$$

where

$$T = -\frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2}{\partial x_i^2},\ W_{\mathrm{ee}} = \delta(x_1 - x_2),\ V_{\mathrm{ne}} = -Z\sum_{i=1}^{N}\delta(x_i) \tag{4.2}$$

are the kinetic-energy operator, the Dirac-delta electron-electron interaction, and the Dirac-delta nucleus-electron potential with nuclear charge $Z = 2$, respectively. Since we will be only interested in the spin-singlet ground state, we can ignore spin and antisymmetry, and thus work on the one-electron Hilbert space $\mathcal{h} = L^2(\mathbb{R}, \mathbb{C})$ and the two-electron (non-antisymmetrized) tensor-product Hilbert space $\mathcal{H} = \mathcal{h} \otimes \mathcal{h}$. The ground-state energy can be expressed as

$$E_0 = \min_{\Psi\in\mathcal{W}} \langle \Psi, H\Psi\rangle, \tag{4.3}$$

where $\mathcal{W}$ is the set of all admissible wave functions

$$\mathcal{W} = \left\{\Psi \in \mathcal{H} \mid \Psi \in H^1(\mathbb{R}^2, \mathbb{C}),\ \langle\Psi,\Psi\rangle = 1\right\}, \tag{4.4}$$

where $H^1(\mathbb{R}^2,\mathbb{C}) = \{f \in L^2(\mathbb{R}^2,\mathbb{C}) \mid \partial_i f \in L^2(\mathbb{R}^2,\mathbb{C}), i = 1,2\}$ is the first-order Sobolev space and $\langle\,,\rangle$ designates the L^2 inner product.

The ground state of this 1D He-like atom with delta-potential interactions can be considered as a model for the ground state of the real 3D He atom. Indeed, it can be shown that the ground state of a generalization to arbitrary dimension D of the electronic Hamiltonian of the He atom with Coulomb-potential interactions exactly reduces for $D = 1$, after appropriate scaling of the energies and distances, to the ground state of the Hamiltonian in Eq. (4.1) [23, 25–28]. Our main interest in this model is that it gives an electron-electron cusp (or derivative discontinuity) condition identical to the familiar 3D one [5,6], i.e. for small interelectronic distances $x_{12} = x_1 - x_2$ the exact ground-state wave function behaves as

$$\Psi_0(x_1,x_2) = \Psi_0(x_1,x_1)\left(1 + \frac{1}{2}|x_{12}| + O(x_{12}^2)\right). \tag{4.5}$$

When using a finite one-electron basis set, we thus expect a slow convergence with the basis size very similar to the slow convergence observed in 3D quantum systems with the Coulomb electron-electron interaction. This is why we prefer this model to other 1D quantum systems (see, e.g., Refs. [29, 30]).

Another neat fact about the present model is that it can be solved analytically at the Hartree-Fock (HF) level [31]. The total HF ground-state energy is

$$E_{\rm HF} = -Z^2 + \frac{Z}{2} - \frac{1}{12} = -3.083333... \text{ a.u.}, \tag{4.6}$$

and the HF (doubly) occupied spatial orbital is

$$\forall x \in \mathbb{R},\ \phi_1(x) = 2\beta\sqrt{\gamma}\frac{e^{-\beta|x|}}{1-\gamma\, e^{-2\beta|x|}}, \tag{4.7}$$

with $\beta = Z - 1/2 = 3/2$ and $\gamma = 1/(4Z-1) = 1/7$. The exact ground-state energy cannot be calculated analytically but has been accurately estimated numerically [21, 24] to be $E_0 = -3.155390$ a.u..

4.2.2 Full-configuration interaction in a basis set

We now consider full-configuration-interaction (FCI) calculations in a finite one-electron basis set $\mathcal{B} \subset H^1(\mathbb{R},\mathbb{C})$. To have a systematically improvable basis set, we use Hermite (or Hermite-Gaussian) basis functions

$$\forall x \in \mathbb{R},\ f_n^\alpha(x) = N_n^\alpha\, H_n(\sqrt{2\alpha}x)\, e^{-\alpha x^2}, \tag{4.8}$$

where n is a natural number, H_n are the Hermite polynomials, $N_n^\alpha = (2^n n!)^{-1/2}(2\alpha/\pi)^{1/4}$ is the normalization factor, and $\alpha > 0$ is a real constant. It is well known that the set $\{f_n^\alpha\}_{n=0,\ldots,n_{\max}}$ converges to a complete orthonormal basis of $L^2(\mathbb{R},\mathbb{C})$ in the limit $n_{\max} \to \infty$ for any fixed exponent α. We deliberately use the same exponent in all basis functions, namely $\alpha = 11.5$, instead of multiple exponents in order to avoid optimizing them. Except for that, this basis set is quite similar to the Gaussian-type-orbital basis sets widely used in quantum chemistry. Since we are not interested in the convergence of the HF energy with this basis set (which is slow, see Appendix 4.5.1) but only in the convergence of the FCI correlation energy, we add the exact occupied HF orbital ϕ_1 given in Eq. (4.7) to the basis set. Our final basis set is thus

$$\mathcal{B} = \left\{\phi_1\right\} \cup \left\{f_n^\alpha\right\}_{n=0,\ldots,n_{\max}} \equiv \left\{\chi_\mu\right\}_{\mu=1,\ldots,M}, \tag{4.9}$$

and contains $M = n_{\max} + 2$ basis functions: $\chi_1 = \phi_1$, $\chi_2 = f_0^\alpha$, ..., $\chi_M = f_{n_{\max}}^\alpha$.

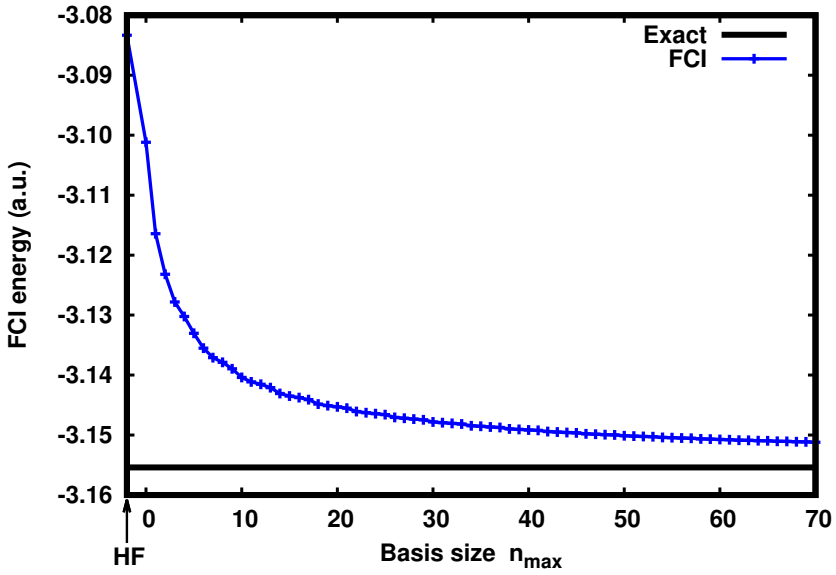

Figure 4.1: FCI ground-state energy $E^{\mathcal{B}}_{\rm FCI}$ [Eq. (4.10)] of the 1D He-like atom as a function of the basis size $n_{\rm max}$. The exact energy is taken from Ref. [24].

We introduce now $\mathcal{h}^{\mathcal{B}} = \text{span}(\mathcal{B})$ as the M-dimensional one-electron Hilbert space spanned by this basis set $\mathcal{B}$, and $\mathcal{H}^{\mathcal{B}} = \mathcal{h}^{\mathcal{B}} \otimes \mathcal{h}^{\mathcal{B}}$ the corresponding two-electron Hilbert space of dimension M^2. The FCI ground-state energy for this basis set $\mathcal{B}$ is

$$E^{\mathcal{B}}_{\rm FCI} = \min_{\Psi \in \mathcal{W}^{\mathcal{B}}} \langle \Psi, H\Psi \rangle, \tag{4.10}$$

where $\mathcal{W}^{\mathcal{B}}$ is the set of normalized wave functions restricted to $\mathcal{H}^{\mathcal{B}}$

$$\mathcal{W}^{\mathcal{B}} = \left\{ \Psi \in \mathcal{H}^{\mathcal{B}} \mid \langle \Psi, \Psi \rangle = 1 \right\}. \tag{4.11}$$

In practice, we proceed as follows. For each basis size $n_{\rm max}$, we first perform a HF calculation [3, 32]. The nucleus-electron integrals are just $\langle \chi_\mu, v_{\rm ne} \chi_\nu \rangle = -Z\chi_\mu(0)\chi_\nu(0)$, and the kinetic integrals $\langle \chi_\mu, t\chi_\nu \rangle = (1/2)\int_{\mathbb{R}} \chi'_\mu(x)\chi'_\nu(x)\mathrm{d}x$ and the two-electron integrals $\langle \chi_\mu\chi_\nu, W_{\rm ee}\chi_\lambda\chi_\sigma \rangle = \int_{\mathbb{R}} \chi_\mu(x)\chi_\nu(x)\chi_\lambda(x)\chi_\sigma(x)\mathrm{d}x$ are calculated by Romberg numerical integration [33]. We then use the obtained HF orbitals $\{\phi_i\}_{i=1,\ldots,M}$ to expand the FCI ground-state wave function as

$$\Psi^{\mathcal{B}}_{\rm FCI}(x_1, x_2) = \sum_{i=1}^{M} \sum_{i=1}^{M} c_{i,j} \phi_i(x_1) \phi_j(x_2). \tag{4.12}$$

The FCI coefficients $c_{i,j}$ and the associated FCI ground-state energy $E^{\mathcal{B}}_{\rm FCI}$ are found by diagonalization of the Hamiltonian. Parity inversion symmetry is exploited in all our calculations.

In Fig. 4.1, we report the FCI ground-state energy $E^{\mathcal{B}}_{\rm FCI}$ as a function of the basis size $n_{\rm max}$. We observe a quite slow convergence of $E^{\mathcal{B}}_{\rm FCI}$ with $n_{\rm max}$ toward the exact ground-state energy E_0. A numerical fit from $n_{\rm max} = 50$ and 70 gives the following power-law convergence

$$E^{\mathcal{B}}_{\rm FCI} \underset{n_{\rm max}\to\infty}{\sim} E_0 + \frac{A}{n^b_{\rm max}}, \tag{4.13}$$

gives $A \approx 0.077$ a.u. and $b \approx 0.68$. Equivalently, in terms of the number of basis functions M, this convergence law can be stated as $1/M^b$. Note that since our basis set includes the exact

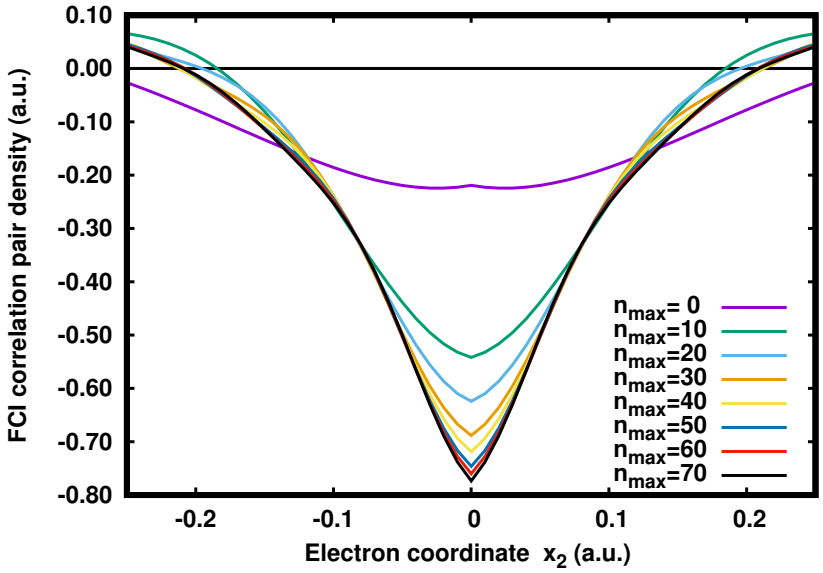

Figure 4.2: FCI correlation pair density $\rho^{\mathcal{B}}_{2,\mathrm{c}}(x_1, x_2)$ of the 1D He-like atom at $x_1 = 0$ as a function of x_2 for different basis size $n_{\max}$.

HF occupied orbital this slow convergence is entirely due to the correlation energy. According to the analysis given in Appendix 4.5.2, we theoretically expect $b = 0.5$. The difference most likely means that larger values of $n_{\max}$ are needed to reach the asymptotic regime. Note that for the 3D Coulomb case it is well known that the correlation energy exhibits a cubic-law convergence with respect to either the maximal angular momentum or the maximal principal quantum number of the basis set, or equivalently a $1/M$ convergence law in terms of the number of basis functions [3, 34]. In the present work, the use of a basis of Hermite functions with a single exponent thus leads to an even slower convergence rate.

To illustrate further the slow basis convergence, we also calculate the FCI correlation pair density

$$\rho^{\mathcal{B}}_{2,\mathrm{c}}(x_1, x_2) = 2\left(|\Psi^{\mathcal{B}}_{\mathrm{FCI}}(x_1, x_2)|^2 - |\Phi_{\mathrm{HF}}(x_1, x_2)|^2\right), \tag{4.14}$$

where $\Phi_{\mathrm{HF}}(x_1, x_2) = \phi_1(x_1)\phi_1(x_2)$ is the HF wave function. In Fig. 4.2, this quantity is plotted with respect to the second electron coordinate x_2 for a fixed value of the first electron coordinate $x_1 = 0$. The convergence of $\rho^{\mathcal{B}}_{2,\mathrm{c}}(x_1, x_2)$ with $n_{\max}$ is again slow and reminiscent of the well-known slow basis convergence of the correlation pair density for the 3D Coulomb electron-electron interaction (see, e.g., Refs. [3, 35, 36]). Note that the small derivative discontinuity seen on all the curves of Fig. 4.2 at $x_2 = 0$ is due to the fact that we include in our basis set the exact HF orbital in Eq. (4.7) which has itself an electron-nucleus cusp, namely $\phi_1(x) = \phi_1(0)(1-Z|x|+O(x^2))$. The electron-electron cusp condition in Eq. (4.5) is only recovered for large $n_{\max}$.

In conclusion, the present 1D model adequately captures the main characteristics of the slow basis convergence problem of standard quantum-chemistry wave-function methods, and it is thus appropriate for applying our basis-set correction approach.

4.3 Basis-set correction theory

We now develop the basis-set correction theory based on DFT for the present 1D model which aims at removing the basis-set incompleteness error in the FCI ground-state energy. This requires to develop an extension of standard DFT from the usual complete-basis-set setting to the case

of the incomplete finite one-electron basis set $\mathcal{B}$. In such a finite basis set, it is known that the original Hohenberg-Kohn theorem [37] does not hold anymore [38–40], in the sense that there is an infinite number of local potentials which give, after projection in a finite basis set, the same ground-state density [41]. However, we will show that we can still define density functionals associated with a finite basis set.

4.3.1 Density-functional theory for the one-dimensional model

We start by reviewing some useful definitions of standard DFT specialized to the 1D model. For mathematically oriented reviews of DFT, see for instance Refs. [42–49], and in particular Ref. [50] which encompasses the 1D case.

Working on the same Hilbert space $\mathcal{H}$ as before, we now consider the following 1D Hamiltonian for $N = 2$ electrons with a general external potential v

$$H[v] = T + W_{\rm ee} + V, \tag{4.15}$$

where again $T = -(1/2)\sum_{i=1}^{N} \partial^2/\partial x_i^2$ and $W_{\rm ee} = \delta(x_1 - x_2)$, and $V = \sum_{i=1}^{N} v(x_i)$ is now a general external potential operator. We will still take admissible wave functions to be in the space $\mathcal{W}$ given in Eq. (4.4). The convex set of densities representable by a wave function $\Psi \in \mathcal{W}$, the so-called N-representable densities, is then [42, 50]

$$\begin{aligned} \mathcal{R} &= \{\rho \mid \exists\, \Psi \in \mathcal{W}, \rho_\Psi = \rho\} \\ &= \left\{\rho \in L^1(\mathbb{R}) \mid \rho \geq 0,\ \int_{\mathbb{R}} \rho(x){\rm d}x = N,\ \ \sqrt{\rho} \in H^1(\mathbb{R})\right\}, \end{aligned} \tag{4.16}$$

where $\rho_\Psi(x_1) = N\int_{\mathbb{R}} |\Psi(x_1, x_2)|^2 {\rm d}x_2$ is the density of the wave function Ψ. We have $\mathcal{R} \subset \mathcal{X}$ where $\mathcal{X}$ is the Banach space $\mathcal{X} = C_0(\mathbb{R}) \cap L^1(\mathbb{R})$ with $C_0(\mathbb{R})$ the space of continuous functions vanishing at infinity. Therefore, the space of external potentials v that we can consider is the continuous dual space of $\mathcal{X}$, i.e. $\mathcal{V} = \mathcal{X}' = M(\mathbb{R}) + L^\infty(\mathbb{R})$ where $M(\mathbb{R})$ is the space of bounded Radon measures. Note that the set $\mathcal{V}$ includes the external potential considered in Section 5.3, i.e. $v_{\rm ne}(x) = -Z\delta(x)$. For $v \in \mathcal{V}$, we then define the ground-state energy as

$$E_0[v] = \inf_{\Psi \in \mathcal{W}} \langle \Psi, H[v]\Psi\rangle. \tag{4.17}$$

The Levy-Lieb density functional [42, 51] is defined as a constrained-search over wave functions yielding the density ρ

$$\forall \rho \in \mathcal{R},\ F[\rho] = \min_{\Psi \in \mathcal{W}_\rho} \langle \Psi, (T + W_{\rm ee})\Psi\rangle, \tag{4.18}$$

where $\mathcal{W}_\rho = \{\Psi \in \mathcal{W} \mid \rho_\Psi = \rho\}$. It gives the ground-state energy as

$$E_0[v] = \inf_{\rho \in \mathcal{R}} \left(F[\rho] + (v, \rho)\right), \tag{4.19}$$

where we have introduced the notation $(v, \rho) = \int_{\mathbb{R}} v(x)\rho(x){\rm d}x$. If a minimizing density ρ_0 exists in Eq. (4.19) then it is an exact ground-state density for the potential v.

One can also define the Lieb density functional [42], which is the Legendre–Fenchel convex-conjugate of $E_0[v]$

$$\forall \rho \in \mathcal{R},\ F_{\rm L}[\rho] = \sup_{v \in \mathcal{V}} \left(E_0[v] - (v, \rho)\right). \tag{4.20}$$

Just like the Levy-Lieb functional F, the Lieb functional F_{L} gives the exact ground-state energy as $E_0[v] = \inf_{\rho\in\mathcal{R}}(F_{\mathrm{L}}[\rho] + (v,\rho))$. In general, the functionals F and F_{L} are different, the Lieb functional being in fact the lower semi-continuous convex envelope (lscv) of the Levy-Lieb functional, i.e.

$$F_{\mathrm{L}} = \mathrm{lscv}(F) \leq F. \tag{4.21}$$

It turns out that the Lieb functional can also be expressed as a generalization of the Levy-Lieb functional in which the constrained search is extended from pure-state wave functions to ensemble density matrices [42, 52]. This implies that $F[\rho] = F_{\mathrm{L}}[\rho]$ for densities ρ which are densities of a *non-degenerate* ground state of the Hamiltonian $H[v]$ for some potential v. In the present case of two spin-singlet electrons, the ground state is always non-degenerate and thus the Levy-Lieb and Lieb functionals are identical, i.e. $F = F_{\mathrm{L}}$.

4.3.2 First variant of basis-set correction

The first variant of basis-set correction corresponds to the one introduced for the 3D Coulombic case in Ref. [8] and further developed in Refs. [9–12, 14]. We consider the Hamiltonian $H[v]$ in Eq. (4.15) on the two-electron Hilbert space $\mathcal{H}^{\mathcal{B}}$ associated with the basis set $\mathcal{B}$. For $v \in \mathcal{V}$, the FCI ground-state energy is

$$E^{\mathcal{B}}_{\mathrm{FCI}}[v] = \min_{\Psi\in\mathcal{W}^{\mathcal{B}}} \langle \Psi, H[v]\Psi\rangle, \tag{4.22}$$

where $\mathcal{W}^{\mathcal{B}}$, given in Eq. (4.11), is the set of normalized wave functions restricted to $\mathcal{H}^{\mathcal{B}}$. We define the corresponding Levy-Lieb density functional for the basis set $\mathcal{B}$ as

$$\forall \rho \in \mathcal{R}^{\mathcal{B}},\ F^{\mathcal{B}}[\rho] = \min_{\Psi\in\mathcal{W}^{\mathcal{B}}_{\rho}} \langle \Psi, (T + W_{\mathrm{ee}})\Psi\rangle, \tag{4.23}$$

where $\mathcal{W}^{\mathcal{B}}_{\rho} = \left\{\Psi \in \mathcal{W}^{\mathcal{B}} \mid \rho_{\Psi} = \rho\right\}$ and $\mathcal{R}^{\mathcal{B}}$ is the set of densities representable by a wave function $\Psi \in \mathcal{W}^{\mathcal{B}}$

$$\mathcal{R}^{\mathcal{B}} \quad = \quad \left\{\rho \mid \exists\, \Psi \in \mathcal{W}^{\mathcal{B}}, \rho_{\Psi} = \rho\right\}. \tag{4.24}$$

A priori, this set is not convex and not easily characterized. The FCI ground-state energy can be expressed as

$$E^{\mathcal{B}}_{\mathrm{FCI}}[v] = \min_{\rho\in\mathcal{R}^{\mathcal{B}}} \left(F^{\mathcal{B}}[\rho] + (v,\rho)\right). \tag{4.25}$$

We now decompose the exact Levy-Lieb density functional $F[\rho]$ in Eq. (4.18) as

$$\forall \rho \in \mathcal{R}^{\mathcal{B}},\ F[\rho] = F^{\mathcal{B}}[\rho] + \bar{E}^{\mathcal{B}}[\rho], \tag{4.26}$$

where $\bar{E}^{\mathcal{B}}[\rho]$ is the complementary basis-set correction density functional

$$\bar{E}^{\mathcal{B}}[\rho] = \langle \Psi[\rho], (T + W_{\mathrm{ee}})\Psi[\rho]\rangle - \langle \Psi^{\mathcal{B}}[\rho], (T + W_{\mathrm{ee}})\Psi^{\mathcal{B}}[\rho]\rangle, \tag{4.27}$$

and $\Psi[\rho]$ and $\Psi^{\mathcal{B}}[\rho]$ are minimizing wave functions in Eqs. (4.18) and (4.23), respectively. Clearly, since $\mathcal{W}^{\mathcal{B}}_{\rho} \subset \mathcal{W}_{\rho}$, we have $\forall \rho \in \mathcal{R}^{\mathcal{B}}, F^{\mathcal{B}}[\rho] \geq F[\rho]$, and thus $\bar{E}^{\mathcal{B}}[\rho] \leq 0$. Since the decomposition in Eq. (4.26) is defined only for $\rho \in \mathcal{R}^{\mathcal{B}}$, we cannot recover the exact ground-state energy $E_0[v]$.

Instead, we can obtain the following approximate energy obtained by restricting the minimization in Eq. (4.19) to densities ρ in $\mathcal{R}^{\mathcal{B}}$

$$\begin{aligned} E_0^{\mathcal{B}}[v] &= \min_{\rho\in\mathcal{R}^{\mathcal{B}}} \left(F[\rho] + (v,\rho)\right) \\ &= \min_{\rho\in\mathcal{R}^{\mathcal{B}}} \left(\min_{\Psi\in\mathcal{W}_\rho^{\mathcal{B}}} \langle\Psi, (T+W_{\rm ee})\Psi\rangle + \bar{E}^{\mathcal{B}}[\rho] + (v,\rho) \right) \\ &= \min_{\Psi\in\mathcal{W}^{\mathcal{B}}} \left(\langle\Psi, (T+W_{\rm ee}+V)\Psi\rangle + \bar{E}^{\mathcal{B}}[\rho_\Psi] \right), \end{aligned} \tag{4.28}$$

or, designating by $\Psi_0^{\mathcal{B}} \in \mathcal{W}^{\mathcal{B}}$ a minimizing wave function in Eq. (4.28),

$$E_0^{\mathcal{B}}[v] = \langle\Psi_0^{\mathcal{B}}, (T+W_{\rm ee}+V)\Psi_0^{\mathcal{B}}\rangle + \bar{E}^{\mathcal{B}}[\rho_{\Psi_0^{\mathcal{B}}}]. \tag{4.29}$$

It is easy to see that $E_0[v] \leq E_0^{\mathcal{B}}[v] \leq E_{\rm FCI}^{\mathcal{B}}[v]$. For a given basis set $\mathcal{B}$, the functional $\bar{E}^{\mathcal{B}}[\rho]$ provides a (self-consistent) basis-set correction to the FCI energy so that $E_0^{\mathcal{B}}[v]$ is a better approximation to $E_0[v]$ than $E_{\rm FCI}^{\mathcal{B}}[v]$ is. Moreover, as the basis set is increased toward completeness, $E_0^{\mathcal{B}}[v]$ should converge much faster to $E_0[v]$ than $E_{\rm FCI}^{\mathcal{B}}[v]$ does, since, roughly speaking, densities ρ typically converge faster than wave functions Ψ with respect to the basis set.

For simplicity, instead of performing the minimization in Eq. (4.28), one may use a non-self-consistent approximation consisting in using the FCI ground-state wave function $\Psi_{\rm FCI}^{\mathcal{B}}$ in place of $\Psi_0^{\mathcal{B}}$, giving what we will call a "FCI+DFT" energy

$$E_{\rm FCI+DFT}^{\mathcal{B}}[v] = \langle\Psi_{\rm FCI}^{\mathcal{B}}, (T+W_{\rm ee}+V)\Psi_{\rm FCI}^{\mathcal{B}}\rangle + \bar{E}^{\mathcal{B}}[\rho_{\Psi_{\rm FCI}^{\mathcal{B}}}], \tag{4.30}$$

which is an upper bound of $E_0^{\mathcal{B}}[v]$, i.e. $E_{\rm FCI+DFT}^{\mathcal{B}}[v] \geq E_0^{\mathcal{B}}[v]$. Again, as the basis set is increased toward completeness, $E_{\rm FCI+DFT}^{\mathcal{B}}[v]$ should converge much faster to $E_0[v]$ than $E_{\rm FCI}^{\mathcal{B}}[v]$ does.

One inconvenience of this basis-set correction scheme is that, for a given finite basis set $\mathcal{B}$, it does not give the exact ground-state energy, even in principle if we knew the exact complementary basis-set correction functional $\bar{E}^{\mathcal{B}}[\rho]$. This is due to the fact that $F^{\mathcal{B}}[\rho]$ is defined only on the restricted set of densities $\mathcal{R}^{\mathcal{B}}$. Another related inconvenience is that since $\bar{E}^{\mathcal{B}}[\rho]$ is defined only on this restricted set of densities, it is not clear how to define the LDA for it. Defining the LDA would indeed require to consider uniform densities, but uniform densities are not in $\mathcal{R}^{\mathcal{B}}$. Even though uniform densities are not in $\mathcal{R}$ either, they can be approached with densities from $\mathcal{R}$ [53,54], so it would be preferable to have a complementary basis-set correction functional defined on the entire set $\mathcal{R}$. This would also permit to connect in principle the basis-set correction scheme to the exact ground-state energy. This is what is achieved by the second variant of basis-set correction.

4.3.3 Second variant of basis-set correction

For the second variant of basis-set correction, we work on the full Hilbert space $\mathcal{H}$ (not restricted to the basis set $\mathcal{B}$), and define the following Hamiltonian

$$H^{{\rm w}\mathcal{B}}[v] = T + W_{\rm ee}^{\mathcal{B}} + V, \tag{4.31}$$

where the kinetic-energy operator T and the external potential operator V are still defined as before, but the electron-electron interaction operator is now projected in the basis set $\mathcal{B}$

$$W_{\rm ee}^{\mathcal{B}} = P^{\mathcal{B}} W_{\rm ee} P^{\mathcal{B}}, \tag{4.32}$$

where $P^{\mathcal{B}}$ is the orthogonal projector onto the basis-set-restricted Hilbert space $\mathcal{H}^{\mathcal{B}}$. The notation "w$\mathcal{B}$" is to remind us that only $W_{\rm ee}$ is projected. In fact, $W_{\rm ee}^{\mathcal{B}}$ is a complicated non-local

two-electron operator. Using an orthonormal orbital basis $\{\phi_i\}_{i=1,\ldots,M}$ spanning the same space as $\mathcal{B}$, its integral kernel can be written as

$$W_{\rm ee}^{\mathcal{B}}(x_1,x_2;x_1',x_2') = \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M}\sum_{l=1}^{M} \phi_i(x_1)\phi_j(x_2)\,\langle\phi_i\phi_j, W_{\rm ee}\phi_k\phi_l\rangle\,\phi_k^*(x_1')\phi_l^*(x_2'), \tag{4.33}$$

where $\langle\phi_i\phi_j, W_{\rm ee}\phi_k\phi_l\rangle = \int_{\mathbb{R}} \phi_i^*(x)\phi_j^*(x)\phi_k(x)\phi_l(x)\mathrm{d}x$ are the two-electron integrals in the orbital basis $\{\phi_i\}$. For $v \in \mathcal{V}$, the associated ground-state energy is

$$E_0^{\mathrm{w}\mathcal{B}}[v] = \inf_{\Psi\in\mathcal{W}} \langle\Psi, H^{\mathrm{w}\mathcal{B}}[v]\Psi\rangle. \tag{4.34}$$

Clearly, if we were to restrict the minimization in Eq. (4.34) to the set $\mathcal{W}^{\mathcal{B}}$, $E_0^{\mathrm{w}\mathcal{B}}[v]$ would reduce to $E_{\rm FCI}^{\mathcal{B}}[v]$ [Eq. (4.22)]. Therefore, we have $E_0^{\mathrm{w}\mathcal{B}}[v] \leq E_{\rm FCI}^{\mathcal{B}}[v]$. Moreover, because $W_{\rm ee}$ is a positive operator, one would expect that projecting it in a finite basis should typically decrease the ground-state energy, i.e. $E_0^{\mathrm{w}\mathcal{B}}[v] \leq E_0[v]$, but this may not be generally true.

We then define the corresponding Levy-Lieb functional for all densities $\rho \in \mathcal{R}$ as

$$\forall \rho \in \mathcal{R},\ F^{\mathrm{w}\mathcal{B}}[\rho] = \min_{\Psi\in\mathcal{W}_\rho} \langle\Psi, (T + W_{\rm ee}^{\mathcal{B}})\Psi\rangle. \tag{4.35}$$

For the same reasons as before, comparison with Eq. (4.23) shows that, for $\rho \in \mathcal{R}^{\mathcal{B}}$, $F^{\mathrm{w}\mathcal{B}}[\rho] \leq F^{\mathcal{B}}[\rho]$. The ground-state energy $E_0^{\mathrm{w}\mathcal{B}}[v]$ can be written as

$$E_0^{\mathrm{w}\mathcal{B}}[v] = \inf_{\rho\in\mathcal{R}} \left(F^{\mathrm{w}\mathcal{B}}[\rho] + (v,\rho)\right). \tag{4.36}$$

We now decompose the exact Levy-Lieb density functional $F[\rho]$ as

$$\forall \rho \in \mathcal{R},\ F[\rho] = F^{\mathrm{w}\mathcal{B}}[\rho] + \bar{E}_{\rm Hxc}^{\mathrm{w}\mathcal{B}}[\rho], \tag{4.37}$$

which defines the complementary Hartree-exchange-correlation (Hxc) basis-set correction functional $\bar{E}_{\rm Hxc}^{\mathrm{w}\mathcal{B}}[\rho]$. Analogously to what is done in multideterminant range-separated DFT [18, 20, 55–58], the functional $\bar{E}_{\rm Hxc}^{\mathrm{w}\mathcal{B}}[\rho]$ can be decomposed as

$$\bar{E}_{\rm Hxc}^{\mathrm{w}\mathcal{B}}[\rho] = \bar{E}_{\rm Hx,md}^{\mathrm{w}\mathcal{B}}[\rho] + \bar{E}_{\rm c,md}^{\mathrm{w}\mathcal{B}}[\rho], \tag{4.38}$$

where $\bar{E}_{\rm Hx,md}^{\mathrm{w}\mathcal{B}}[\rho]$ is the Hartree-exchange (Hx) contribution defined as the expectation value of the complementary interaction $\bar{W}_{\rm ee}^{\mathcal{B}} = W_{\rm ee} - W_{\rm ee}^{\mathcal{B}}$ over the minimizing multideterminant (md) wave function $\Psi^{\mathrm{w}\mathcal{B}}[\rho]$ (that we will assume to be unique up to a global phase factor) in Eq. (4.35)

$$\bar{E}_{\rm Hx,md}^{\mathrm{w}\mathcal{B}}[\rho] = \langle\Psi^{\mathrm{w}\mathcal{B}}[\rho], \bar{W}_{\rm ee}^{\mathcal{B}}\Psi^{\mathrm{w}\mathcal{B}}[\rho]\rangle, \tag{4.39}$$

and $\bar{E}_{\rm c,md}^{\mathrm{w}\mathcal{B}}[\rho]$ is the remaining correlation (c) contribution

$$\bar{E}_{\rm c,md}^{\mathrm{w}\mathcal{B}}[\rho] = \langle\Psi[\rho], (T + W_{\rm ee})\Psi[\rho]\rangle - \langle\Psi^{\mathrm{w}\mathcal{B}}[\rho], (T + W_{\rm ee})\Psi^{\mathrm{w}\mathcal{B}}[\rho]\rangle. \tag{4.40}$$

Clearly, since $\Psi[\rho]$ minimizes $\langle\Psi, (T + W_{\rm ee})\Psi\rangle$, we have $\bar{E}_{\rm c,md}^{\mathrm{w}\mathcal{B}}[\rho] \leq 0$. Since the decomposition in Eq. (4.37) is defined for all densities $\rho \in \mathcal{R}$, we can obtain the exact ground-state energy as

$$\begin{aligned} E_0[v] &= \inf_{\rho\in\mathcal{R}} \left(F^{\mathrm{w}\mathcal{B}}[\rho] + \bar{E}_{\rm Hxc}^{\mathrm{w}\mathcal{B}}[\rho] + (v,\rho)\right) \\ &= \inf_{\rho\in\mathcal{R}} \left(\min_{\Psi\in\mathcal{W}_\rho} \langle\Psi, (T + W_{\rm ee}^{\mathcal{B}})\Psi\rangle + \bar{E}_{\rm Hxc}^{\mathrm{w}\mathcal{B}}[\rho] + (v,\rho)\right) \\ &= \inf_{\Psi\in\mathcal{W}} \left(\langle\Psi, (T + W_{\rm ee}^{\mathcal{B}} + V)\Psi\rangle + \bar{E}_{\rm Hxc}^{\mathrm{w}\mathcal{B}}[\rho_\Psi]\right). \end{aligned} \tag{4.41}$$

For potentials v for which there exists a minimizing wave function $\Psi_0^{\mathrm{w}\mathcal{B}} \in \mathcal{W}$ in Eq. (4.41), this wave function yields a exact ground-state density ρ_0, i.e. $\rho_{\Psi_0^{\mathrm{w}\mathcal{B}}} = \rho_0$. It is in fact the minimizing wave function in Eq. (4.35) giving the density ρ_0, i.e. $\Psi_0^{\mathrm{w}\mathcal{B}} = \Psi^{\mathrm{w}\mathcal{B}}[\rho_0]$. Using this fact and combining Eqs. (4.41), (4.38), and (4.39), we can express the exact ground-state energy as

$$E_0[v] = \langle \Psi_0^{\mathrm{w}\mathcal{B}}, (T + W_{\mathrm{ee}} + V)\Psi_0^{\mathrm{w}\mathcal{B}} \rangle + \bar{E}_{\mathrm{c,md}}^{\mathrm{w}\mathcal{B}}[\rho_{\Psi_0^{\mathrm{w}\mathcal{B}}}]. \tag{4.42}$$

Thus, this second variant of basis-set correction leads to an energy expression similar to Eq. (4.29) obtained for the first variant of basis-set correction, with the functional $\bar{E}_{\mathrm{c,md}}^{\mathrm{w}\mathcal{B}}[\rho]$ replacing the functional $\bar{E}^{\mathcal{B}}[\rho]$ and the wave function $\Psi_0^{\mathrm{w}\mathcal{B}}$ replacing the wave function $\Psi_0^{\mathcal{B}}$. One advantage of this second variant of basis-set correction is that it gives the exact ground-state energy $E_0[v]$. The price to pay is that the minimization in Eq. (4.41) is more complicated than in Eq. (4.28) since it is over general wave functions $\Psi \in \mathcal{W}$ and not simply wave functions restricted to $\mathcal{W}^{\mathcal{B}}$. Moreover, the minimization involves not only the functional $\bar{E}_{\mathrm{c,md}}^{\mathrm{w}\mathcal{B}}[\rho]$ but also the functional $\bar{E}_{\mathrm{Hx,md}}^{\mathrm{w}\mathcal{B}}[\rho]$ in Eq. (4.39) which did not appear in the first variant of basis-set correction.

Similarly to the first variant of basis-set correction, we can define a non-self-consistent approximation consisting in using the FCI ground-state wave function $\Psi_{\mathrm{FCI}}^{\mathcal{B}}$ in place of $\Psi_0^{\mathrm{w}\mathcal{B}}$ in Eq. (4.42), giving an alternative "FCI+DFT" energy

$$E_{\mathrm{FCI+DFT}}^{\mathrm{w}\mathcal{B}}[v] = \langle \Psi_{\mathrm{FCI}}^{\mathcal{B}}, (T + W_{\mathrm{ee}} + V)\Psi_{\mathrm{FCI}}^{\mathcal{B}} \rangle + \bar{E}_{\mathrm{c,md}}^{\mathrm{w}\mathcal{B}}[\rho_{\Psi_{\mathrm{FCI}}^{\mathcal{B}}}], \tag{4.43}$$

which is quite similar but not equivalent to Eq. (4.30) since the complementary basis-set correction functional is different. Like for the first variant of basis-set correction, when the basis set is increased toward completeness, $E_{\mathrm{FCI+DFT}}^{\mathrm{w}\mathcal{B}}[v]$ should converge much faster to $E_0[v]$ than $E_{\mathrm{FCI}}^{\mathcal{B}}[v]$ does.

Finally, we define the Lieb density functional for this second variant of basis-set correction

$$\forall \rho \in \mathcal{R},\ F_{\mathrm{L}}^{\mathrm{w}\mathcal{B}}[\rho] = \sup_{v \in \mathcal{V}} \left(E_0^{\mathrm{w}\mathcal{B}}[v] - (v, \rho) \right). \tag{4.44}$$

Like in standard DFT, by the theory of Legendre–Fenchel transformations, this Lieb functional $F_{\mathrm{L}}^{\mathrm{w}\mathcal{B}}$ must be the lower semi-continuous convex envelope of the Levy-Lieb functional $F^{\mathrm{w}\mathcal{B}}$, i.e.

$$F_{\mathrm{L}}^{\mathrm{w}\mathcal{B}} = \mathrm{lscv}(F^{\mathrm{w}\mathcal{B}}) \leq F^{\mathrm{w}\mathcal{B}}. \tag{4.45}$$

One could also write down this Lieb functional as a constrained-search over ensemble density matrices, and again we should have $F^{\mathrm{w}\mathcal{B}}[\rho] = F_{\mathrm{L}}^{\mathrm{w}\mathcal{B}}[\rho]$ for densities ρ which are densities of a non-degenerate ground state of the Hamiltonian $H^{\mathrm{w}\mathcal{B}}[v]$ for some potential v. As already mentioned, in the present case of two spin-singlet electrons, the ground state is always non-degenerate and thus the Levy-Lieb and Lieb functionals are identical, i.e. $F^{\mathrm{w}\mathcal{B}} = F_{\mathrm{L}}^{\mathrm{w}\mathcal{B}}$. As we will see in Section 4.4.2, the definition in Eq. (4.44) is useful to calculate the functional in practice since it involves a unconstrained maximization over potentials v whereas the definition in Eq. (4.35) involves a potentially more complicated constrained minimization over wave functions yielding a fixed density ρ.

In summary, the advantage of the second variant of basis-set correction over the first variant is that it is connected with the exact ground-state energy [via Eq. (4.41) or (4.42)] and that it involves a complementary basis-set correction functional which is defined for all densities in $\mathcal{R}$. In the next section, we exploit this in order to construct a LDA for the functional $\bar{E}_{\mathrm{c,md}}^{\mathrm{w}\mathcal{B}}[\rho]$.

4.4 Local-density approximation from finite uniform-electron gas

In standard DFT, the LDA is based on the infinite UEG. Essentially, for the 1D case, calculating the energy per particle of the infinite UEG amounts to plugging an uniform density

$\rho_{\rm unif} : x \mapsto \rho_0 \in (0, +\infty)$ in the density functional $F[\rho]$ and taking the thermodynamic limit, i.e. $\lim_{N\to\infty} F[\rho_{\rm unif}]/N$. One difficulty is that a non-zero uniform density function $\rho_{\rm unif}$ defined on the entire real line $\mathbb{R}$ is obviously not N-representable. For the 3D case, the infinite UEG was rigorously mathematically defined in Refs. [53, 54] by first convoluting the uniform density with a function of compact support (so that the convoluted density is N-representable) and then taking the thermodynamic limit $N \to \infty$ (after removing the Hartree energy which is divergent for the Coulombic 3D case). Here, in the spirit of Ref. [59], we will instead consider a *finite* UEG, i.e., for a finite electron number N.

4.4.1 Finite uniform-electron gas for the complete-basis-set case

To define a 1D finite UEG, we generalize the standard DFT of Section 4.3.1 from the real line $\mathbb{R}$ to a finite interval $\Omega_a = (-a/2, a/2)$ of length a. Hence, the one-electron Hilbert space is $\hbar_a = L^2(\Omega_a, \mathbb{C})$ and the two-electron Hilbert space is $\mathcal{H}_a = \hbar_a \otimes \hbar_a$. For external local potentials $v \in \mathcal{V}_a$ (where the space $\mathcal{V}_a$ will be specified below) and $N = 2$ electrons, we define the ground-state energy of the Hamiltonian $H[v] = T + W_{\rm ee} + V$ [Eq. (4.15)] restricted to the Hilbert space $\mathcal{H}_a$ as

$$E_{0,a}[v] = \inf_{\Psi\in\mathcal{W}_a} \langle \Psi, H[v]\Psi\rangle_a, \tag{4.46}$$

with the set of admissible wave functions

$$\mathcal{W}_a = \left\{\Psi \in \mathcal{H}_a \mid \Psi \in H^1_{\rm per}(\Omega_a^2, \mathbb{C}),\ \langle\Psi,\Psi\rangle_a = 1\right\}, \tag{4.47}$$

where $\langle\Psi_1,\Psi_2\rangle_a = \int_{\Omega_a^2} \Psi_1^*(x_1,x_2)\Psi_2(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2$ is the inner product on $\mathcal{H}_a$ and $H^1_{\rm per}$ designates the set of functions in H^1 with periodic boundary conditions on the domain (see, e.g., Ref. [60]), which can be defined as $H^1_{\rm per}(\Omega_a^2, \mathbb{C}) = \left\{\Psi_{|\Omega_a^2} \mid \Psi \in H^1_{\rm loc}(\mathbb{R}^2, \mathbb{C}),\ \Psi \text{ is } a\mathbb{Z}^2\text{-periodic}\right\}$ where $\Psi_{|\Omega_a^2}$ designates the restriction of Ψ to Ω_a^2 and $H^1_{\rm loc}(\mathbb{R}^2, \mathbb{C})$ is the local first-order Sobolev space.

The corresponding Levy-Lieb density functional is

$$\forall \rho \in \mathcal{R}_a,\ F_a[\rho] = \min_{\Psi\in\mathcal{W}_{a,\rho}} \langle\Psi, (T + W_{\rm ee})\Psi\rangle_a, \tag{4.48}$$

where $\mathcal{W}_{a,\rho} = \{\Psi \in \mathcal{W}_a,\ \rho_\Psi = \rho\}$ and $\mathcal{R}_a$ is the set of N-representable densities on Ω_a

$$\begin{aligned}\mathcal{R}_a &= \{\rho \mid \exists\, \Psi \in \mathcal{W}_a, \rho_\Psi = \rho\} \\ &= \left\{\rho \in L^1(\Omega_a) \mid \rho \geq 0,\ \int_{\Omega_a} \rho(x)\mathrm{d}x = N,\ \sqrt{\rho} \in H^1_{\rm per}(\Omega_a)\right\},\end{aligned} \tag{4.49}$$

and $H^1_{\rm per}(\Omega_a) = \{f \in H^1(\Omega_a) \mid \lim_{x\to -a/2} f(x) = \lim_{x\to a/2} f(x)\}$. We have $\mathcal{R}_a \subset \mathcal{X}_a$ where $\mathcal{X}_a$ is the Banach space $\mathcal{X}_a = C_{\rm per}(\Omega_a) \cap L^1(\Omega_a)$ with $C_{\rm per}(\Omega_a)$ the space of continuous functions on Ω_a with periodic boundary conditions. Therefore, the space of external potentials to consider is the continuous dual space of $\mathcal{X}_a$, i.e. $\mathcal{V}_a = \mathcal{X}_a' = M_{\rm per}(\Omega_a) + L^\infty(\Omega_a)$ where $M_{\rm per}(\Omega_a)$ is the space of bounded Radon measures on Ω_a with periodic boundary conditions. In the limit of an infinite interval ($a \to \infty$), we recover the standard theory on the real line $\mathbb{R}$, i.e. $\lim_{a\to\infty} F_a[\rho] = F[\rho]$ where $F[\rho]$ is the standard Levy-Lieb functional defined in Eq. (4.18). Similarly, we could generalize the Lieb density functionals in Eq. (4.20) to the finite interval Ω_a.

We now define a *finite* UEG (fUEG) by considering the uniform density $\rho_{\rm unif} : x \mapsto \rho_0 = N/a$ on the interval Ω_a for the fixed electron number $N = 2$. Note that $\rho_{\rm unif}$ is in fact the unique uniform density belonging to $\mathcal{R}_a$. The energy per particle of this finite UEG is

$$\varepsilon_{\rm fUEG}(\rho_0) = \frac{F_a[\rho_{\rm unif}]}{N}, \tag{4.50}$$

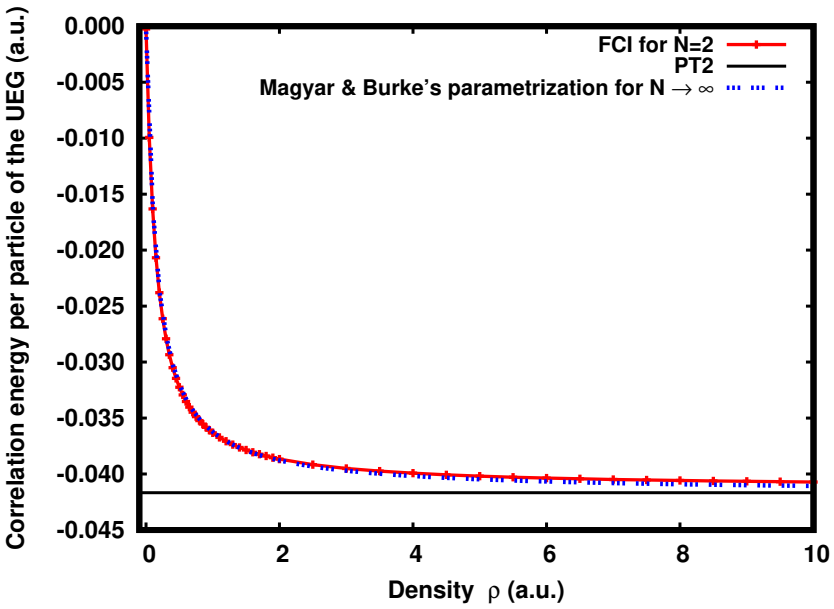

Figure 4.3: FCI correlation energy per particle $\varepsilon_{\text{c,fUEG}}(\rho)$ [Eqs. (4.50) and (4.52)] of the 1D finite UEG as a function of the density ρ for $N = 2$ electrons and a plane-wave cutoff $n_{\text{max}}^{\text{pw}} = 60$. The exact PT2 correlation energy per particle, which is independent of N and ρ, is indicated as a horizontal line. The correlation energy per particle of the infinite UEG ($N \to \infty$) as parametrized by Magyar and Burke [24] from essentially numerically exact Bethe-ansatz results is also plotted for comparison.

and is a function of the only variable ρ_0 since N is fixed and $a = N/\rho_0$. The value of $F_a[\rho_{\text{unif}}]$ corresponds to the ground-state energy of the two-electron Hamiltonian with zero external potential

$$H_{\text{fUEG}} = H[0] = T + W_{\text{ee}}, \tag{4.51}$$

with periodic boundary conditions on Ω_a, provided that the ground-state density is the uniform density ρ_{unif} (i.e., no translational symmetry breaking). We note that in Refs. [61–64] 1D finite UEGs mapped to a ring were introduced using the Coulomb electron-electron interaction. Here, instead we use the Dirac-delta electron-electron interaction and we do not work on a ring.

For a given density ρ_0 in the range $[0, 10]$ a.u. and for the fixed electron number $N = 2$ and interval length $a = N/\rho_0$, we calculate the ground-state energy by performing a FCI calculation using a one-electron plane-wave (pw) orthonormal basis $\{p_n\}_{n\in\mathbb{Z},\, |n|\leq n_{\text{max}}^{\text{pw}}}$ where $p_n(x) = (1/\sqrt{a})e^{ik_n x}$ and $k_n = 2\pi n/a$. The one-electron kinetic integrals $\langle p_{n_1}, t p_{n_2}\rangle_a = (2\pi^2 n_1^2/a^2)\delta_{n_1,n_2}$ and the two-electron integrals $\langle p_{n_1} p_{n_2}, W_{\text{ee}} p_{n_3} p_{n_4}\rangle_a = (1/a)\delta_{n_1+n_2,n_3+n_4}$ are trivial. We use a plane-wave cutoff $n_{\text{max}}^{\text{pw}} = 60$ which leads to FCI energies converged to at least 1 mhartree (and in fact generally better than that). As usual, the finite UEG energy per particle can be decomposed as

$$\varepsilon_{\text{fUEG}}(\rho) = t_{\text{s,fUEG}}(\rho) + \varepsilon_{\text{H,fUEG}}(\rho) + \varepsilon_{\text{x,fUEG}}(\rho) + \varepsilon_{\text{c,fUEG}}(\rho), \tag{4.52}$$

with the non-interacting kinetic energy per particle $t_{\text{s,fUEG}}(\rho) = 0$ (since $N = 2$ the only occupied orbital is the constant plane wave $p_0(x) = 1/\sqrt{a}$ which has a zero kinetic energy), the Hartree energy per particle $\varepsilon_{\text{H,fUEG}}(\rho) = \rho/2$, the exchange energy per particle $\varepsilon_{\text{x,fUEG}}(\rho) = -\rho/4$, and the correlation energy per particle $\varepsilon_{\text{c,fUEG}}(\rho)$.

The correlation energy per particle $\varepsilon_{\text{c,fUEG}}(\rho)$ is plotted in Fig. 4.3. In the high-density limit, $\varepsilon_{\text{c,fUEG}}(\rho)$ tends to the correlation energy per particle in second-order perturbation theory (PT2)

with respect to the electron-electron interaction W_{ee}

$$\lim_{\rho\to\infty} \varepsilon_{\text{c,fUEG}}(\rho) = \varepsilon_{\text{c,fUEG}}^{\text{PT2}} = -\frac{1}{24}, \tag{4.53}$$

which is a constant independent of ρ. It turns out that this constant is the same for $N = 2$ and $N \to \infty$ [24]. In the low-density limit, $\varepsilon_{\text{c,fUEG}}(\rho)$ goes to zero linearly with ρ (see Ref. [65])

$$\varepsilon_{\text{c,fUEG}}(\rho) \underset{\rho\to 0}{\sim} -\frac{\rho}{4}, \tag{4.54}$$

so as to exactly cancel out the Hartree and exchange energies per particle. This is due to the fact that, in this limit, the probability density of finding the electrons at the same point of space is zero and thus the Dirac-delta electron-electron interaction has not effect. This is the 1D version of the strong-interaction limit of DFT [66–69]. Equation (4.54) is also true for $N \to \infty$ [24, 70], and is in fact true independently of N [65]. In Fig. 4.3, we also show the correlation energy per particle of the infinite UEG ($N \to \infty$) as parametrized by Magyar and Burke [24] from essentially numerically exact Bethe-ansatz results. Not only the correlation energies per particle for $N = 2$ and $N \to \infty$ agree well for small and large densities, as they should since they have the same N-independent asymptotic behaviors [Eqs. (4.53) and (4.54)], but they also agree very well for intermediate densities (the maximal deviation between the two curves being about 0.4 mhartree), showing that the thermodynamic limit $N \to \infty$ is essentially already reached at $N = 2$ for this 1D UEG. Hence, there is no need considering 1D UEGs with $N > 2$ electrons. This must be due to the very short-range nature of the Dirac-delta electron-electron interaction. For the 1D UEG with the Coulomb interaction, the correlation energy per particle depends much more strongly on the electron number [62].

4.4.2 Finite uniform-electron gas for the incomplete-basis-set case

We now generalize the second variant of basis-set correction of Section 4.3.3 from the real line $\mathbb{R}$ to a finite interval $\Omega_a = (-a/2, a/2)$ of length a. For $v \in \mathcal{V}_a$ and $N = 2$ electrons, we define the ground-state energy of the restriction to the Hilbert space $\mathcal{H}_a$ of the Hamiltonian $H^{\text{w}\mathcal{B}}[v] = T + W_{\text{ee}}^{\mathcal{B}} + V$ [Eq. (4.31)], featuring the electron-electron interaction projected in the basis set $\mathcal{B}$ used for the 1D He-like atom, as

$$E_{0,a}^{\text{w}\mathcal{B}}[v] = \inf_{\Psi\in\mathcal{W}_a} \langle \Psi, H^{\text{w}\mathcal{B}}[v]\Psi\rangle_a. \tag{4.55}$$

The corresponding Levy-Lieb density functional is

$$\forall \rho \in \mathcal{R}_a,\ F_{\text{a}}^{\text{w}\mathcal{B}}[\rho] = \min_{\Psi\in\mathcal{W}_{a,\rho}} \langle \Psi, (T + W_{\text{ee}}^{\mathcal{B}})\Psi\rangle_a, \tag{4.56}$$

and the corresponding Lieb density functional is

$$\forall \rho \in \mathcal{R}_a,\ F_{\text{L,a}}^{\text{w}\mathcal{B}}[\rho] = \sup_{v\in\mathcal{V}_a} \left(E_{0,a}^{\text{w}\mathcal{B}}[v] - (v,\rho)_a\right), \tag{4.57}$$

where $(v,\rho)_a = \int_{\Omega_a} v(x)\rho(x)\mathrm{d}x$. Again, in the limit of an infinite interval ($a \to \infty$), we recover the theory of Section 4.3.3.

For a given basis set $\mathcal{B}$, we now define an associated finite UEG by considering the uniform density $\rho_{\text{unif}} : x \mapsto \rho_0 = N/a$ on the interval Ω_a for the fixed electron number $N = 2$. The kinetic + electron-electron energy per particle of this $\mathcal{B}$-dependent finite UEG is

$$f_{\text{fUEG}}^{\text{w}\mathcal{B}}(\rho_0) = \frac{F_a^{\text{w}\mathcal{B}}[\rho_{\text{unif}}]}{N}, \tag{4.58}$$

where

$$F_a^{\mathrm{w}\mathcal{B}}[\rho_{\mathrm{unif}}] = \langle \Psi^{\mathrm{w}\mathcal{B}}[\rho_{\mathrm{unif}}], (T + W_{\mathrm{ee}}^{\mathcal{B}})\Psi^{\mathrm{w}\mathcal{B}}[\rho_{\mathrm{unif}}]\rangle_a, \tag{4.59}$$

and $\Psi^{\mathrm{w}\mathcal{B}}[\rho_{\mathrm{unif}}]$ is the ground-state wave function (assumed to be unique up to a global phase factor) of the two-electron Hamiltonian

$$H_{\mathrm{fUEG}}^{\mathrm{w}\mathcal{B}} = T + W_{\mathrm{ee}}^{\mathcal{B}} + V^{\mathrm{w}\mathcal{B}}, \tag{4.60}$$

with periodic boundary conditions on Ω_a and with $V^{\mathrm{w}\mathcal{B}} = \sum_{i=1}^{N} v^{\mathrm{w}\mathcal{B}}(x_i)$ where $v^{\mathrm{w}\mathcal{B}}(x)$ is the local potential (that we assume to exist and which is defined up to an additive constant) which enforces the constraint that the ground-state wave function $\Psi^{\mathrm{w}\mathcal{B}}[\rho_{\mathrm{unif}}]$ yields indeed the uniform density ρ_{unif}. Since the projected electron-electron interaction $W_{\mathrm{ee}}^{\mathcal{B}}$ breaks translation invariance, the addition of the potential $v^{\mathrm{w}\mathcal{B}}$ is necessary to restore a uniform density. This is in contrast with the UEG for the complete-basis-set case for which no external potential was necessary to obtain a uniform density [Eq. (4.51)]. To conveniently obtain the potential $v^{\mathrm{w}\mathcal{B}}$, we use the fact that, since the two-electron finite UEG has a non-degenerate ground state, the Levy-Lieb functional $F_a^{\mathrm{w}\mathcal{B}}$ and the Lieb functional $F_{\mathrm{L,a}}^{\mathrm{w}\mathcal{B}}$ are identical. The potential $v^{\mathrm{w}\mathcal{B}}$ then just corresponds to the maximizing potential in Eq. (4.57) for $\rho = \rho_{\mathrm{unif}}$ (see Refs. [71–73])

$$v^{\mathrm{w}\mathcal{B}} = \underset{v\in\mathcal{V}_a}{\mathrm{argmax}} \left(E_{0,a}^{\mathrm{w}\mathcal{B}}[v] - (v, \rho_{\mathrm{unif}})_a \right). \tag{4.61}$$

For a given basis set $\mathcal{B}$, for a given density ρ_0 in the range $[0, 10]$ a.u., and for the fixed electron number $N = 2$ and interval length $a = N/\rho_0$, we calculate the energy $E_{0,a}^{\mathrm{w}\mathcal{B}}[v]$ by performing a FCI calculation using a plane-wave orthonormal basis $\{p_n\}_{n\in\mathbb{Z},\,|n|\leq n_{\mathrm{max}}^{\mathrm{pw}}}$. The one-electron kinetic integrals are still $\langle p_{n_1}, t p_{n_2}\rangle_a = (2\pi^2 n_1^2/a^2)\delta_{n_1,n_2}$. The integrals of the electron-electron interaction projected in the basis set $\mathcal{B}$ [see Eq. (4.33)] can be calculated as

$$\langle p_{n_1} p_{n_2}, W_{\mathrm{ee}}^{\mathcal{B}} p_{n_3} p_{n_4}\rangle_a = \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M}\sum_{l=1}^{M} S_{i,n_1}^{*} S_{j,n_2}^{*} \langle \phi_i\phi_j, W_{\mathrm{ee}}\phi_k\phi_l\rangle\, S_{k,n_3} S_{l,n_4}, \tag{4.62}$$

where $\langle \phi_i\phi_j, W_{\mathrm{ee}}\phi_k\phi_l\rangle$ are the two-electron integrals in terms of the HF orbitals $\{\phi_i\}_{i=1,\ldots,M}$ of the 1D He-like atom expanded in the basis set $\mathcal{B}$ (see Section 4.2.2) and $S_{i,n} = \langle \phi_i, p_n\rangle_a = \int_{\Omega_a} \phi_i^*(x) p_n(x)\mathrm{d}x$ are the overlap integrals between the HF orbitals and the plane-wave basis functions. The potential v to optimize in Eq. (4.61) is also expanded on the same plane-wave basis set

$$v(x) = \sum_{n\in\mathbb{Z},|n|\leq n_{\mathrm{max}}^{\mathrm{pw}}} c_n p_n(x), \tag{4.63}$$

with coefficients $c_n \in \mathbb{R}$ and we impose $c_{-n} = c_n$ in order to have a real-valued and parity-even potential. The one-electron potential integrals, needed to calculate $E_{0,a}^{\mathrm{w}\mathcal{B}}[v]$, are

$$\langle p_{n_1}, v p_{n_2}\rangle_a = \int_{\Omega_a} p_{n_1}^*(x) v(x) p_{n_2}(x)\mathrm{d}x = \frac{c_{n_1-n_2}}{\sqrt{a}}, \tag{4.64}$$

and the second term in Eq. (4.61) is simply $(v, \rho_{\mathrm{unif}})_a = \sqrt{a} c_0 \rho_0$. Finally, for the optimization of the potential, it is useful to have the derivative of $F[v] = E_{0,a}^{\mathrm{w}\mathcal{B}}[v] - (v, \rho_{\mathrm{unif}})_a$ with respect to the coefficient c_n. Using the Hellmann-Feynman theorem, we find

$$\begin{aligned} \frac{\partial F[v]}{\partial c_n} &= (p_n, \rho_{\Psi_v})_a - (p_n, \rho_{\mathrm{unif}})_a \\ &= \frac{1}{\sqrt{a}} \sum_{n_1\in\mathbb{Z},|n_1|\leq n_{\mathrm{max}}^{\mathrm{pw}}} \gamma_{n_1,n_1+n} - \sqrt{a}\rho_0\delta_{n,0}, \end{aligned} \tag{4.65}$$

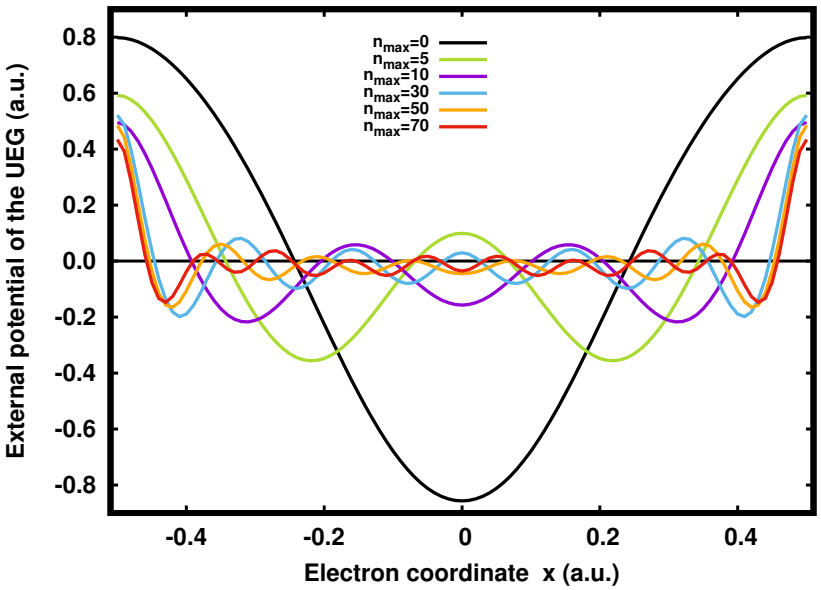

Figure 4.4: External optimized potential $v^{\mathrm{w}\mathcal{B}}(x)$ [Eq. (4.61)] keeping a uniform density $\rho_0 = 2$ a.u. for the 1D finite UEG for $N = 2$ electrons and for different sizes $n_{\max}$ of the basis set $\mathcal{B}$ of the 1D He-like atom.

where we have used $\rho_{\Psi_v}(x) = \sum_{n_1\in\mathbb{Z},|n_1|\leq n_{\max}^{\mathrm{pw}}} \sum_{n_2\in\mathbb{Z},|n_2|\leq n_{\max}^{\mathrm{pw}}} \gamma_{n_1,n_2} p_{n_1}(x) p_{n_2}^*(x)$ with the one-particle reduced density matrix γ of the ground-state wave function Ψ_v of $H^{\mathrm{w}\mathcal{B}}[v]$. In practice, we use a plane-wave cutoff of $n_{\max}^{\mathrm{pw}} = 30$. We can use a smaller cutoff than the cutoff used for the complete-basis-set UEG in Section 4.4.1 since the FCI energy $E_{0,a}^{\mathrm{w}\mathcal{B}}[v]$ has a fast convergence with $n_{\max}^{\mathrm{pw}}$ due to the presence of the projected electron-electron interaction $W_{\mathrm{ee}}^{\mathcal{B}}$. To optimize the coefficients $\{c_n\}$ of the potential, we use the conjugate gradient method [33]. Since the term $n = 0$ in Eq. (4.63) is just an arbitrary constant, we keep the coefficient c_0 fixed to 0. With a zero potential v, the FCI density $\rho_{\Psi_{v=0}}(x)$ can deviate from the target density ρ_0 by about as much as 0.2 a.u. for the basis set $\mathcal{B}$ of smallest size (i.e., $M = 1$). With our optimized potentials $v^{\mathrm{w}\mathcal{B}}$, the density $\rho_{\Psi_{v^{\mathrm{w}\mathcal{B}}}}(x)$ deviates from the target density ρ_0 to at most about 10^{-4} a.u..

As an aside, it might be worthwhile to stress here that in the Hamiltonian in Eq. (4.60), the kinetic-energy operator T is not projected in the basis set $\mathcal{B}$. We observed that if T is also projected in the basis set $\mathcal{B}$, then the high-lying states of T collapse to the lower part of the spectrum, which inevitably pollutes the nature of the ground state of the finite UEG Hamiltonian and renders numerically impossible to find a potential restoring a uniform density. This is why in the second-variant of basis-set correction in Section 4.3.3 we have decided to project only the electron-electron interaction W_{ee} in the basis set $\mathcal{B}$.

The optimized potentials $v^{\mathrm{w}\mathcal{B}}(x)$ obtained from Eq. (4.61) are plotted in Fig. 4.4 for the example of the target density $\rho_0 = 2$ a.u. and for different sizes $n_{\max}$ of the basis set $\mathcal{B}$ of the 1D He-like atom introduced in Section 4.2.2. To compensate for the breaking of translation invariance of the projected electron-electron interaction, all potentials show oscillations with maximum amplitude on the edges of the interval. As expected, when $n_{\max}$ increases, the amplitude of the potential decreases, as it must eventually go to zero in the complete-basis-set limit $n_{\max} \to \infty$.

Once the FCI ground-state wave function $\Psi^{\mathrm{w}\mathcal{B}}[\rho_{\mathrm{unif}}] = \Psi_{v^{\mathrm{w}\mathcal{B}}}$ for the optimal potential $v^{\mathrm{w}\mathcal{B}}$ has been obtained, we calculate the following energy per particle using this wave function

$$\varepsilon_{\mathrm{fUEG}}^{\mathrm{w}\mathcal{B}}(\rho_0) = \frac{\langle \Psi^{\mathrm{w}\mathcal{B}}[\rho_{\mathrm{unif}}], (T + W_{\mathrm{ee}}) \Psi^{\mathrm{w}\mathcal{B}}[\rho_{\mathrm{unif}}] \rangle_a}{N}, \tag{4.66}$$

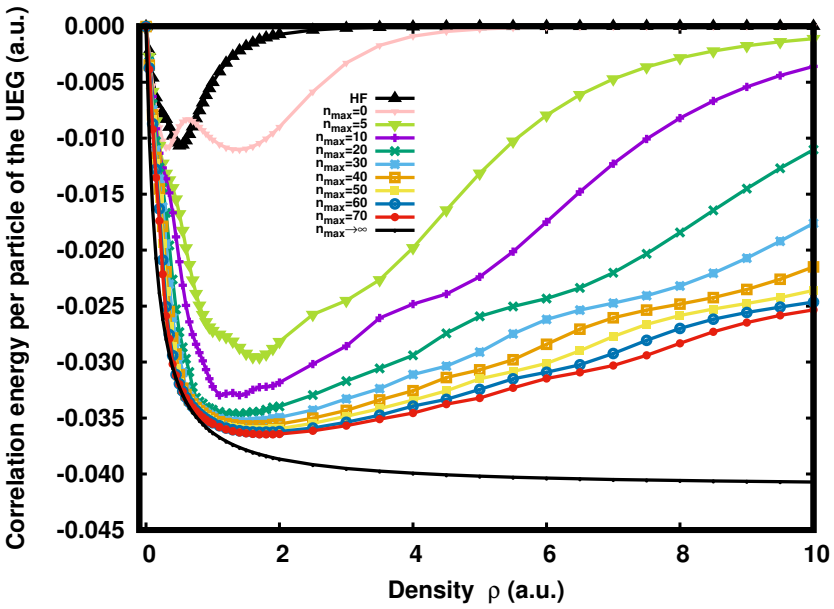

Figure 4.5: FCI correlation energy per particle $\varepsilon_{\text{c,fUEG}}^{\text{w}\mathcal{B}}(\rho)$ [Eqs. (4.66) and (4.67)] of the 1D finite UEG as a function of the density ρ for $N = 2$ electrons and for different sizes n_{max} of the basis set $\mathcal{B}$ of the 1D He-like atom. The curve labelled by "HF" corresponds to the limiting case where the basis set $\mathcal{B}$ contains only the exact HF occupied orbital.

which we can decompose in the same way as in Eq. (4.52)

$$\varepsilon_{\text{fUEG}}^{\text{w}\mathcal{B}}(\rho) = t_{\text{s,fUEG}}(\rho) + \varepsilon_{\text{H,fUEG}}(\rho) + \varepsilon_{\text{x,fUEG}}(\rho) + \varepsilon_{\text{c,fUEG}}^{\text{w}\mathcal{B}}(\rho), \tag{4.67}$$

with the same kinetic, Hartree, and exchange contributions as in Eq. (4.52), and a new correlation energy per particle $\varepsilon_{\text{c,fUEG}}^{\text{w}\mathcal{B}}(\rho)$. This latter quantity is plotted in Fig. 4.5 for different sizes n_{max} of the basis set $\mathcal{B}$ of the 1D He-like atom. As expected, when n_{max} increases, $\varepsilon_{\text{c,fUEG}}^{\text{w}\mathcal{B}}(\rho)$ becomes more negative and gets closer to the correlation energy per particle $\varepsilon_{\text{c,fUEG}}(\rho)$ of the complete-basis-set limit $n_{\text{max}} \to \infty$. For finite n_{max}, it can be observed that, in the high-density limit, the correlation energy per particle $\varepsilon_{\text{c,fUEG}}^{\text{w}\mathcal{B}}(\rho)$ goes to zero, unlike in Eq. (4.53). This is due to the fact that, as the density increases, the relevant electron-electron distances contributing to the correlation energy become smaller and the basis set $\mathcal{B}$ is unable to resolve the Dirac-delta electron-electron interaction at a fine enough distance scale.

Finally, we calculate the complementary multideterminant correlation energy per particle of the finite UEG for the basis set $\mathcal{B}$ [see Eq. (4.40)]

$$\bar{\varepsilon}_{\text{c,md,fUEG}}^{\text{w}\mathcal{B}}(\rho) = \varepsilon_{\text{fUEG}}(\rho) - \varepsilon_{\text{fUEG}}^{\text{w}\mathcal{B}}(\rho), \tag{4.68}$$

which is plotted in Fig. 4.6 for different basis sizes n_{max}. As n_{max} increases, the magnitude of $\bar{\varepsilon}_{\text{c,md,fUEG}}^{\text{w}\mathcal{B}}(\rho)$ decreases and must eventually go to zero in the complete-basis-set limit $n_{\text{max}} \to \infty$. The magnitude is largest for high densities since $\bar{\varepsilon}_{\text{c,md,fUEG}}^{\text{w}\mathcal{B}}(\rho)$ must compensate for the inability of the basis set $\mathcal{B}$ to represent the Dirac-delta electron-electron interaction at a small distance scale. Perhaps surprisingly, there is also a local maximum of the magnitude of $\bar{\varepsilon}_{\text{c,md,fUEG}}^{\text{w}\mathcal{B}}(\rho)$ at small densities. This is due to the fact that, at small densities, $\bar{\varepsilon}_{\text{c,md,fUEG}}^{\text{w}\mathcal{B}}(\rho)$ does not exactly cancel out the Hartree and exchange energies per particle, in contrast to the complete-basis-set case [Eq. (4.54)]. Again, this must come from the inability of the basis set $\mathcal{B}$ to represent the Dirac-delta electron-electron interaction sufficiently precisely to give a zero probability density

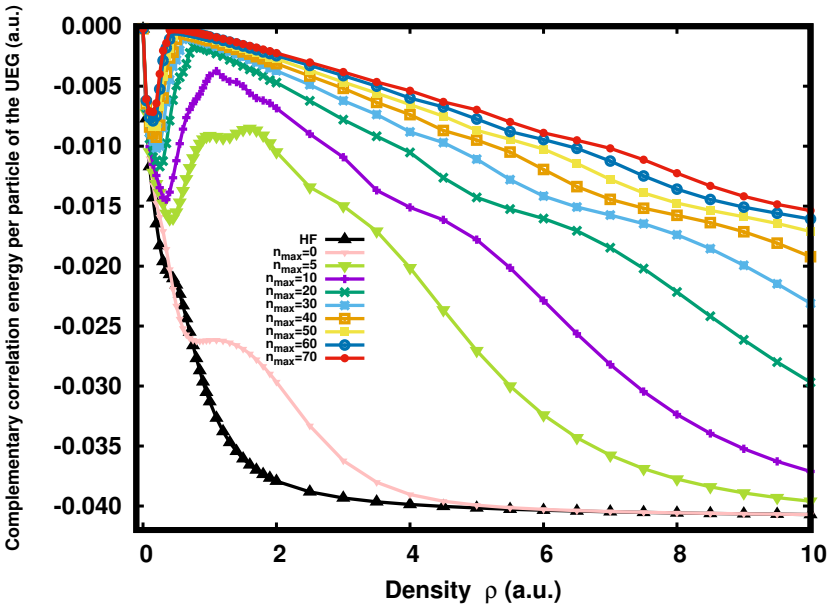

Figure 4.6: FCI complementary correlation energy per particle $\bar{\varepsilon}^{\mathrm{w}\mathcal{B}}_{\mathrm{c,md,fUEG}}(\rho)$ [Eq. (4.68)] of the 1D finite UEG as a function of the density ρ for $N = 2$ electrons and for different sizes $n_{\max}$ of the basis set $\mathcal{B}$ of the 1D He-like atom. The curve labelled by "HF" corresponds to the limiting case where the basis set $\mathcal{B}$ contains only the exact HF occupied orbital.

of finding the electrons as the same point of space in the low-density limit. Interestingly, in between the small and the large-density regimes, for $n_{\max} \geq 5$, the magnitude of $\bar{\varepsilon}^{\mathrm{w}\mathcal{B}}_{\mathrm{c,md,fUEG}}(\rho)$ passes through a minimum. In particular, for $n_{\max} = 70$, $\bar{\varepsilon}^{\mathrm{w}\mathcal{B}}_{\mathrm{c,md,fUEG}}(\rho)$ is almost zero at around $\rho \approx 0.5$ a.u., which means that the basis set $\mathcal{B}$ accurately captures the effect of the Dirac-delta electron-electron interaction at this density.

4.4.3 Finite local-density approximation

We can now define the *finite* LDA (fLDA) for the complementary multideterminant correlation density functional $\bar{E}^{\mathrm{w}\mathcal{B}}_{\mathrm{c,md}}[\rho]$ [Eq. (4.40)] involved in the second-variant of basis-set correction using the previously determined complementary correlation energy per particle $\bar{\varepsilon}^{\mathrm{w}\mathcal{B}}_{\mathrm{c,md,fUEG}}(\rho)$ of the 1D finite UEG

$$\bar{E}^{\mathrm{w}\mathcal{B}}_{\mathrm{c,md,fLDA}}[\rho] = \int_{\mathbb{R}} \rho(x) \bar{\varepsilon}^{\mathrm{w}\mathcal{B}}_{\mathrm{c,md,fUEG}}(\rho(x)) \mathrm{d}x. \tag{4.69}$$

We recall that in standard LDA the functional of an inhomogeneous system for a finite electron number N is approximated using the UEG for infinite electron number $N \to \infty$. Here, instead, in the present finite LDA the functional of the inhomogeneous system is approximated using the UEG of the same electron number N. The use of this finite LDA in lieu of the standard LDA should not be seen as a crucial point for the basis-set correction theory but more like a convenient alternative. For a sufficiently short-range complementary interaction $\bar{W}^{\mathcal{B}}_{\mathrm{ee}} = W_{\mathrm{ee}} - W^{\mathcal{B}}_{\mathrm{ee}}$, the LDA should not depend much on the electron number used in the definition of the underlying UEG.

We then correct the FCI energy of the 1D He-like atom using this fLDA functional in the non-self-consistent approximation introduced in Eq. (4.43), obtaining what we will call the "FCI+fLDA" energy

$$E^{\mathrm{w}\mathcal{B}}_{\mathrm{FCI+fLDA}} = \langle \Psi^{\mathcal{B}}_{\mathrm{FCI}}, (T + W_{\mathrm{ee}} + V_{\mathrm{ne}}) \Psi^{\mathcal{B}}_{\mathrm{FCI}} \rangle + \bar{E}^{\mathrm{w}\mathcal{B}}_{\mathrm{c,md,fLDA}}[\rho_{\Psi^{\mathcal{B}}_{\mathrm{FCI}}}]. \tag{4.70}$$

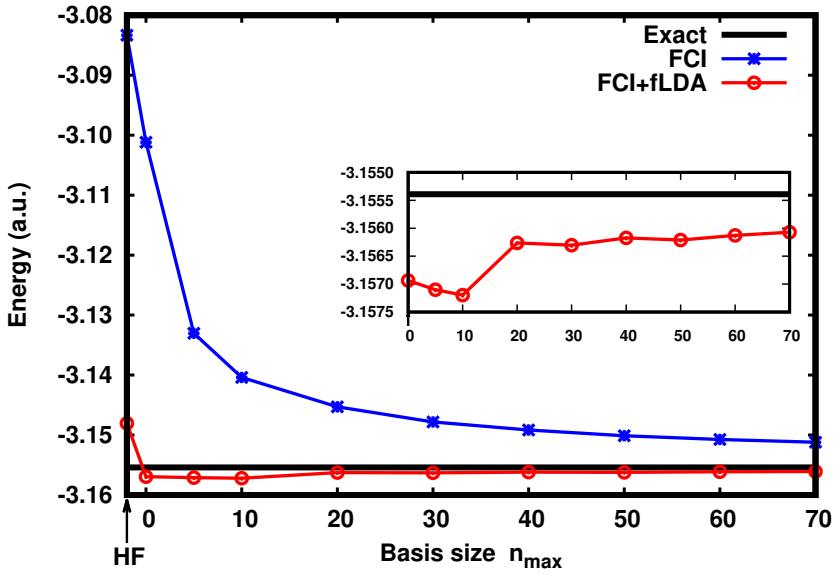

Figure 4.7: FCI ground-state energy $E_{\mathrm{FCI}}^{\mathcal{B}}$ [Eq. (4.10)] and FCI+fLDA ground-state energy $E_{\mathrm{FCI+fLDA}}^{\mathrm{w}\mathcal{B}}$ [Eq. (4.70)] of the 1D He-like atom as a function of the basis size $n_{\max}$. The first point labelled by "HF" corresponds to the limiting case where the basis set $\mathcal{B}$ contains only the exact HF occupied orbital (in this case, FCI simply reduces to HF). The exact energy is taken from Ref. [24].

In practice, we calculate $\bar{E}_{\mathrm{c,md,fLDA}}^{\mathrm{w}\mathcal{B}}[\rho_{\Psi_{\mathrm{FCI}}^{\mathcal{B}}}]$ by numerical integration using cubic interpolation between calculated values of $\bar{\varepsilon}_{\mathrm{c,md,fUEG}}^{\mathrm{w}\mathcal{B}}(\rho)$. The FCI densities of the 1D He-like atom take values from 0 to about 3.5 a.u..

In Fig. 4.7, the FCI+fLDA energy is plotted as a function of the basis size $n_{\max}$. It is clear that the basis-set correction provides a spectacular improvement of the FCI energy. For example, for $n_{\max} = 0$, the FCI energy is about 55 mhartree above the exact energy while the FCI+fLDA energy is only 1.5 mhartree below the exact energy. For $n_{\max} \geq 20$, the FCI+fLDA energy is within 1 mhartree of the exact energy. We note that the residual error must come from the fact that in Eq. (4.42) the functional $\bar{E}_{\mathrm{c,md}}^{\mathrm{w}\mathcal{B}}[\rho]$ is approximated with the fLDA functional and also that the wave function $\Psi_0^{\mathrm{w}\mathcal{B}}$ is approximated by the FCI wave function $\Psi_{\mathrm{FCI}}^{\mathcal{B}}$.

Conclusion

In this work, we have reexamined the recently introduced DFT-based basis-set correction theory on a 1D model with delta-potential interactions, which is a convenient setting to carefully study the slow basis convergence problem of quantum-chemistry wave-function methods. We provided mathematical details about the formulation of the theory, as well as a new variant of basis-set correction which has the advantage that the basis-set correction functional is defined for all N-representable densities. This allowed us to define a LDA for the basis-set correction functional, not based on range-separated DFT as in all previous works, but directly on a 1D finite UEG adapted to the basis set employed. We showed that this approach is very effective to correct for the basis-set incompleteness error in the FCI ground-state energy. We believe that the present work puts the basis-set correction theory on firmer grounds.

Future efforts will focus on the extension of the present work to 3D Coulombic systems. The extension of the theory is straightforward. What remains to be seen is whether the present work

adapts well to the standard Gaussian-type-orbital basis sets used in quantum chemistry and whether we can still construct an accurate LDA for the basis-set correction functional based on a 3D UEG with the Coulomb electron-electron interaction projected in the basis set.

Acknowledgements

We thank Éric Cancès, Louis Garrigue, Paola Gori-Giorgi, Muhammad Hassan, Antoine Levitt, Pierre-Francois Loos, and Mi-Song Dupuy for useful comments on the manuscript. We particularly thank Éric Cancès for pointing out to us the appropriate space $\mathcal{V}$ of potentials to consider and Antoine Levitt for giving us the correct definition of the space $H^1_{\mathrm{per}}(\Omega_a^2, \mathbb{C})$. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme Grant agreement No. 810367 (EMC2).

Author Declarations

The authors have no conflicts to disclose.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

4.5 Appendix: Convergence rate of the expectation value of a Dirac-delta potential in a basis of Hermite functions

4.5.1 One-electron Dirac-delta potential

Let us consider the 1D hydrogen-like Hamiltonian

$$h = -\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - Z\delta(x), \tag{4.71}$$

with nuclear charge $Z \in (0, +\infty)$. The ground-state wave function is (see, e.g., Refs. [74, 75])

$$\forall x \in \mathbb{R},\ \varphi(x) = \sqrt{Z}e^{-Z|x|}, \tag{4.72}$$

which exhibits a cusp identical to the 3D Coulombic case. We expand φ in the orthonormal basis of Hermite functions $\{f_n^\alpha\}_{n\in\mathbb{N}}$ [Eq. (4.8)]

$$\forall x \in \mathbb{R},\ \varphi(x) = \sum_{n=0}^{\infty} c_n f_n^\alpha(x), \tag{4.73}$$

with coefficients $c_n = \int_{\mathbb{R}} f_n^\alpha(x)\varphi(x)\mathrm{d}x$, which are non-zero only for even integers n. Using the following asymptotic equivalent of the unnormalized Hermite functions, for fixed x, [76]

$$H_n(\sqrt{2\alpha}x)e^{-\alpha x^2} \underset{n\to\infty}{\sim} \frac{2^n}{\sqrt{\pi}}\Gamma\left(\frac{n+1}{2}\right)\cos\left(x\sqrt{4\alpha n} - \frac{n\pi}{2}\right), \tag{4.74}$$

and the well-known asymptotic equivalent of the gamma function

$$\Gamma(z) \underset{z\to\infty}{\sim} \sqrt{2\pi}\, z^{z-1/2}e^{-z}, \tag{4.75}$$

we obtain the following asymptotic equivalent of the normalized Hermite functions

$$f_n^\alpha(x) \underset{n\to\infty}{\sim} \sqrt{\frac{2}{\pi}} \frac{\alpha^{1/4}}{n^{1/4}} \cos\left(x\sqrt{4\alpha n} - \frac{n\pi}{2}\right). \tag{4.76}$$

Writing $n = 2p$ with $p \in \mathbb{N}$, we then find the leading term of the asymptotic expansion of the coefficients c_{2p} by integrating over x

$$c_{2p} \underset{p\to\infty}{\sim} \frac{Z^{3/2}}{\sqrt{2\pi}\,\alpha^{3/4}} \frac{(-1)^p}{(2p)^{5/4}}. \tag{4.77}$$

This is perfectly consistent with the analysis of Refs. [77, 78] which shows that an exponentially decaying function φ having a square-integrable first weak derivative (i.e., $\varphi \in H^1(\mathbb{R})$) but a non-square-integrable second weak derivative (i.e., $\varphi \notin H^2(\mathbb{R})$) must have Hermite expansion coefficients c_n going to zero as n^{-k} with $k \in (1, 3/2]$. The leading term of the asymptotic expansion of $c_{2p} f_{2p}^\alpha(x)$ is thus

$$c_{2p} f_{2p}^\alpha(x) \underset{p\to\infty}{\sim} \frac{Z^{3/2}}{\pi\sqrt{\alpha}} \frac{1}{(2p)^{3/2}} \cos\left(x\sqrt{8\alpha p}\right), \tag{4.78}$$

and, in particular at $x = 0$,

$$c_{2p} f_{2p}^\alpha(0) \underset{p\to\infty}{\sim} \frac{Z^{3/2}}{\pi\sqrt{\alpha}} \frac{1}{(2p)^{3/2}}. \tag{4.79}$$

Calling $\tilde{\varphi}$ the best approximation (in the sense of the L^2 norm) to φ obtained with a maximal quantum number $n_{\max}$, i.e.

$$\forall x \in \mathbb{R},\ \tilde{\varphi}(x) = \sum_{n=0}^{n_{\max}} c_n f_n^\alpha(x), \tag{4.80}$$

we find that $\tilde{\varphi}(0)$ converges slowly to the exact value $\varphi(0) = \sqrt{Z}$ as

$$\tilde{\varphi}(0) \underset{n_{\max}\to\infty}{\sim} \varphi(0) - \frac{Z^{3/2}}{\pi\sqrt{\alpha}} \frac{1}{n_{\max}^{1/2}}, \tag{4.81}$$

and the expectation value of the Dirac-delta potential $v_{\text{ne}}(x) = -Z\delta(x)$ has a similar convergence behavior in $1/n_{\max}^{1/2}$

$$\begin{aligned} \langle \tilde{\varphi}, v_{\text{ne}} \tilde{\varphi} \rangle &= -Z\tilde{\varphi}(0)^2 \\ &\underset{n_{\max}\to\infty}{\sim} -Z\varphi(0)^2 + \frac{2Z^{5/2}\varphi(0)}{\pi\sqrt{\alpha}} \frac{1}{n_{\max}^{1/2}}. \end{aligned} \tag{4.82}$$

We also expect the total energy to converge as $1/n_{\max}^{1/2}$ in a basis of Hermite functions.

4.5.2 Two-electron Dirac-delta interaction

Let us consider the Hamiltonian of the 1D two-electron Hooke's atom [24]

$$H = -\frac{1}{2}\frac{\partial^2}{\partial x_1^2} - \frac{1}{2}\frac{\partial^2}{\partial x_2^2} + \frac{1}{2}\omega^2 x_1^2 + \frac{1}{2}\omega^2 x_2^2 + \delta(x_1 - x_2), \tag{4.83}$$

where ω is the angular frequency parameter of the external harmonic potential. In contrast to the 1D He-like atom [Eq. (4.1)], the 1D two-electron Hooke's atom has the advantage to be

solvable in terms of special functions. Indeed, changing the variables to the center-of-mass (cm) coordinate $X = (x_1 + x_2)/2$ and the relative (rel) coordinate $x_{12} = (x_1 - x_2)$ makes the Hamiltonian separable

$$H = h_{\rm cm} + h_{\rm rel}, \tag{4.84}$$

where

$$h_{\rm cm} = -\frac{1}{4}\frac{\partial^2}{\partial X^2} + \omega^2 X^2, \tag{4.85}$$

and

$$h_{\rm rel} = -\frac{\partial^2}{\partial x_{12}^2} + \frac{1}{4}\omega^2 x_{12}^2 + \delta(x_{12}). \tag{4.86}$$

The total ground-state energy is then

$$E_0 = \mathcal{E}_0 + \varepsilon_0, \tag{4.87}$$

where $\mathcal{E}_0 = \omega/2$ is the ground-state energy of $h_{\rm cm}$ and ε_0 is the ground-state energy of $h_{\rm rel}$ which is found from the equation [24, 79]

$$2\sqrt{2\omega}\,\frac{\Gamma\left(-\frac{\nu_0}{2}+\frac{1}{2}\right)}{\Gamma\left(-\frac{\nu_0}{2}\right)} = -1, \tag{4.88}$$

where $\nu_0 = \varepsilon_0/\omega - 1/2$. For example, for $\omega = 1$ a.u., we have $\varepsilon_0 = 0.806746$ a.u. [24]. The ground-state wave function is

$$\forall (X, x_{12}) \in \mathbb{R}^2,\ \Psi(X, x_{12}) = \Phi(X)\psi(x_{12}), \tag{4.89}$$

where the center-of-mass wave function is just given by the first Hermite function $\Phi(X) = f_0^{2\omega}(X)$ and the relative wave function is given by [79]

$$\forall x_{12} \in \mathbb{R},\ \psi(x_{12}) = cD_{\nu_0}(\sqrt{\omega}|x_{12}|), \tag{4.90}$$

where c is a real-valued normalization constant and $x \mapsto D_\nu(x)$ with $\nu \in \mathbb{R}$ is the parabolic cylinder function [76]. The relative wave function has the same cusp as in Eq. (4.5), i.e. $\psi(x_{12}) = \psi(0)[1 + (1/2)|x_{12}| + O(x_{12}^2)]$.

Let us consider now the expansion of the wave function Ψ in the tensor-product orthonormal basis of Hermite functions $\{(x_1, x_2) \mapsto f_{n_1}^{\alpha}(x_1) f_{n_2}^{\alpha}(x_2)\}_{(n_1,n_2)\in\mathbb{N}^2}$. Due to invariance of the harmonic-oscillator Hamiltonian to a rotation of coordinates, the same space is spanned by the rotated orthonormal basis $\{(X, x_{12}) \mapsto f_{n_1}^{2\alpha}(X) f_{n_2}^{\alpha/2}(x_{12})\}_{(n_1,n_2)\in\mathbb{N}^2}$ (see Refs. [77, 78]). This means that the relative wave function ψ is independently expanded as

$$\forall x_{12} \in \mathbb{R},\ \psi(x_{12}) = \sum_{n=0}^{\infty} d_n f_n^{\alpha/2}(x_{12}), \tag{4.91}$$

with coefficients $d_n = \int_{\mathbb{R}} f_n^{\alpha/2}(x_{12})\,\psi(x_{12})\mathrm{d}x_{12}$, which are non-zero only for even integers n. Using Eq. (4.76) and with the help of Mathematica [80], we find the leading term of the asymptotic expansion of the coefficients d_{2p}

$$d_{2p} \underset{p\to\infty}{\sim} -\frac{(-1)^p\, c}{\alpha^{3/4} 2^{\mu_0+1/4}\Gamma(\mu_0)}\frac{1}{(2p)^{5/4}}. \tag{4.92}$$

where $\mu_0 = -\nu_0/2 + 1/2$. Introducing the best approximation to ψ obtained with a maximal quantum number $n_{\max}$

$$\forall x_{12} \in \mathbb{R},\ \tilde{\psi}(x_{12}) \quad = \quad \sum_{n=0}^{n_{\max}} d_n f_n^{\alpha/2}(x_{12}), \tag{4.93}$$

we find that $\tilde{\psi}(0)$ converges slowly to the exact value $\psi(0) = c\sqrt{2\pi}/[2^{\mu_0}\Gamma(\mu_0)]$ as

$$\tilde{\psi}(0) \quad \underset{n_{\max}\to\infty}{\sim} \quad \psi(0) + \frac{c}{\sqrt{\pi\alpha}\, 2^{\mu_0}\Gamma(\mu_0)} \frac{1}{n_{\max}^{1/2}}, \tag{4.94}$$

and the expectation value of the Dirac-delta interaction $W_{\text{ee}} = \delta(x_{12})$ has a similar convergence behavior in $1/n_{\max}^{1/2}$

$$\begin{aligned} \langle \tilde{\psi}, W_{\text{ee}} \tilde{\psi} \rangle \quad &= \quad \tilde{\psi}(0)^2 \\ &\underset{n_{\max}\to\infty}{\sim} \quad \psi(0)^2 + \frac{2c\psi(0)}{\sqrt{\pi\alpha}\, 2^{\mu_0}\Gamma(\mu_0)} \frac{1}{n_{\max}^{1/2}}. \end{aligned} \tag{4.95}$$

We thus see that the two-electron energy converges as $1/n_{\max}^{1/2}$ in a basis of Hermite functions.

Bibliography

[1] T. Helgaker, W. Klopper, H. Koch and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[2] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).

[3] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic Structure Theory* (John Wiley & Sons, LTD, 2000).

[4] J. J. Shepherd, A. Grüneis, G. H. Booth, G. Kresse and A. Alavi, Phys. Rev. B **86**, 035111 (2012).

[5] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).

[6] R. T. Pack and W. B. Brown, J. Chem. Phys. **45**, 556 (1966).

[7] C. Hättig, W. Klopper, A. Köhn and D. P. Tew, Chem. Rev. **112**, 4 (2012).

[8] E. Giner, B. Pradines, A. Ferté, R. Assaraf, A. Savin and J. Toulouse, J. Chem. Phys. **149**, 194301 (2018).

[9] P.-F. Loos, B. Pradines, A. Scemama, J. Toulouse and E. Giner, J. Phys. Chem. Lett. **10**, 2931 (2019).

[10] E. Giner, A. Scemama, J. Toulouse and P.-F. Loos, J. Chem. Phys. **151**, 144118 (2019).

[11] P.-F. Loos, B. Pradines, A. Scemama, E. Giner and J. Toulouse, J. Chem. Theory Comput. **16**, 1018 (2020).

[12] E. Giner, A. Scemama, P.-F. Loos and J. Toulouse, J. Chem. Phys. **152**, 174104 (2020).

[13] Y. Yao, E. Giner, J. Li, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **153**, 124117 (2020).

[14] E. Giner, D. Traore, B. Pradines and J. Toulouse, J. Chem. Phys. **155**, 044109 (2021).

[15] Y. Yao, E. Giner, T. A. Anderson, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **155**, 204104 (2021).

[16] A. Savin, *Recent Developments of Modern Density Functional Theory* (Elsevier, Amsterdam, 1996), pp. 327–357.

[17] J. Toulouse, F. Colonna and A. Savin, Phys. Rev. A **70**, 062505 (2004).

[18] J. Toulouse, P. Gori-Giorgi and A. Savin, Theor. Chem. Acc **114**, 305 (2005).

[19] S. Paziani, S. Moroni, P. Gori-Giorgi and G. B. Bachelet, Phys. Rev. B **73**, 155111 (2006).

[20] A. Ferté, E. Giner and J. Toulouse, J. Chem. Phys. **150**, 084103 (2019).

[21] C. M. Rosenthal, J. Chem. Phys. **55**, 2474 (1971).

[22] I. R. Lapidus, Am. J. Phys. **43**, 790 (1975).

[23] D. R. Herrick and F. H. Stillinger, Phys. Rev. A **11**, 42 (1975).

[24] R. J. Magyar and K. Burke, Phys. Rev. A **70**, 032508 (Sep 2004).

[25] D. R. Herschbach, J. Chem. Phys. **84**, 838 (1986).

[26] D. J. Doren and D. R. Herschbach, Phys. Rev. A **34**, 2654 (1986).

[27] D. J. Doren and D. R. Herschbach, J. Chem. Phys. **87**, 433 (1987).

[28] D. R. Herschbach, J. Avery and O. Goscinski, eds., *Dimensional Scaling in Chemical Physics* (Springer Netherlands, 1993).

[29] P.-F. Loos, C. J. Ball and P. M. W. Gill, Phys. Chem. Chem. Phys. **17**, 3196 (2015).

[30] C. J. Ball, P.-F. Loos and P. M. W. Gill, Phys. Chem. Chem. Phys. **19**, 3987 (2017).

[31] Y. Nogami, M. Vallières and W. van Dijk, Am. J. Phys. **44**, 886 (1976).

[32] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, Inc., 1996).

[33] William H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical recipes 3rd edition* (Cambridge University Press, Cambridge, England, 2007), 3rd edn.

[34] W. Klopper, K. L. Bak, P. Jørgensen, J. Olsen and T. Helgaker, J. Phys. B: At Mol. Opt. Phys. **32**, R103 (1999).

[35] D. P. Tew, W. Klopper and T. Helgaker, J. Comput. Chem. **28**, 1307 (2007).

[36] O. Franck, B. Mussard, E. Luppi and J. Toulouse, J. Chem. Phys. **142**, 074107 (2015).

[37] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[38] J. E. Harriman, Phys. Rev. A **27**, 632 (Feb 1983).

[39] A. Görling and M. Ernzerhof, Phys. Rev. A **51**, 4501 (Jun 1995).

[40] R. Pino, O. Bokanowski, E. V. Ludeña and R. L. Boada, Theor. Chem. Acc. **123**, 189 (2009).

[41] V. N. Staroverov, G. E. Scuseria and E. R. Davidson, J. Chem. Phys. **124**, 141103 (2006).

[42] E. H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983).

[43] R. M. Dreizler and E. K. U. Gross, *Density functional theory* (Springer, Berlin, Germany, 1996), 1st edn.

[44] H. Eschrig, *The fundamentals of density functional theory* (Edition Am Gutenbergplatz Leipzig, 2003).

[45] A. Anantharaman and E. Cancès, Annales de l'Institut Henri Poincaré, Analyse non linéaire **26**, 2425 (2009).

[46] E. Engel and R. M. Dreizler, *Density functional theory*, Theoretical and mathematical physics (Springer, Berlin, Germany, 2011), 2011th edn.

[47] S. Kvaal, U. Ekström, A. M. Teale and T. Helgaker, J. Chem. Phys. **140**, 18A518 (2014).

[48] Helgaker, T, Principles of Density-Functional Theory, Lecture at the GDR CORREL Mini-School on Mathematics in Electronic Structure Theory (Université Pierre et Marie Curie, 2017), `https://trygvehelgaker.no/Presentations/Paris_2017.pdf`.

[49] Cancès, E, A mathematical introduction to density functional theory and Kohn-Sham models.

[50] M. Lewin, E. H. Lieb and R. Seiringer, Universal Functionals in Density Functional Theory (2019).

[51] M. Levy, Proc. Natl. Acad. Sci. U. S. A. **76**, 6062 (1979).

[52] S. M. Valone, J. Chem. Phys. **73**, 4653 (1980).

[53] M. Lewin, E. H. Lieb and R. Seiringer, Journal de l'École polytechnique — Mathématiques **5**, 79 (2018).

[54] M. Lewin, E. H. Lieb and R. Seiringer, Pure and Applied Analysis **2**, 35 (Jan. 2020).

[55] P. Gori-Giorgi and A. Savin, Int. J. Quantum Chem. **109**, 1950 (2009).

[56] A. Stoyanova, A. M. Teale, J. Toulouse, T. Helgaker and E. Fromager, J. Chem. Phys. **139**, 134113 (2013).

[57] E. Rebolini, J. Toulouse, A. M. Teale, T. Helgaker and A. Savin, Mol. Phys. **113**, 1740 (2015).

[58] E. Rebolini, A. M. Teale, T. Helgaker, A. Savin and J. Toulouse, Mol. Phys. **116**, 1443.

[59] P. M. W. Gill and P.-F. Loos, Theor. Chem. Acc **131** (2011).

[60] M. Lewin, Théorie spectrale & mécanique quantique, Lecture - Citer comme : Mathieu Lewin, Théorie spectrale et mécanique quantique, cours de l'École Polytechnique, 2018.

[61] P.-F. Loos and P. M. W. Gill, Phys. Rev. Lett. **108**, 083002 (2012).

[62] P.-F. Loos and P. M. W. Gill, J. Chem. Phys. **138**, 164124 (2013).

[63] P.-F. Loos, C. J. Ball and P. M. W. Gill, J. Chem. Phys. **140**, 18A524 (2014).

[64] P.-F. Loos, Phys. Rev. A **89**, 052523 (2014).

[65] E. Räsänen, M. Seidl and P. Gori-Giorgi, Phys. Rev. B **83**, 195111 (2011).

[66] M. Seidl, J. P. Perdew and M. Levy, Phys. Rev. A **59**, 51 (1999).

[67] M. Seidl, Phys. Rev. A **60**, 4387 (1999).

[68] M. Seidl, P. Gori-Giorgi and A. Savin, Phys. Rev. A **75**, 042511 (2007).

[69] P. Gori-Giorgi and M. Seidl, Phys. Chem. Chem. Phys. **12**, 14405 (2010).

[70] E. Räsänen, S. Pittalis, K. Capelle and C. R. Proetto, Phys. Rev. Lett. **102**, 206406 (2009).

[71] F. Colonna and A. Savin, J. Chem. Phys. **110**, 2828 (1999).

[72] Q. Wu and W. Yang, J. Chem. Phys. **118**, 2498 (2003).

[73] A. M. Teale, S. Coriani and T. Helgaker, J. Chem. Phys. **130**, 104111 (2009).

[74] D. J. Griffiths and D. F. Schroeter, *Introduction to quantum mechanics* (Cambridge university press, 2018).

[75] S. Geltman, Journal of Atomic, Molecular, and Optical Physics **2011**, 1 (2011).

[76] M. Abramowitz and I. A. Stegun, eds., *Handbook of mathematical functions*, Dover Books on Mathematics (Dover Publications, Mineola, NY, 1965).

[77] S. Kvaal, M. Hjorth-Jensen and H. M. Nilsen, Phys. Rev. B **76**, 085421 (2007).

[78] S. Kvaal, Phys. Rev. B **80**, 045321 (2009).

[79] T. Busch and G. Huyet, J. Phys. B: At Mol. Opt. Phys. **36**, 2553 (2003).

[80] Wolfram Research Inc., Mathematica, Version 12, Champaign, IL, 2020.

5

Density-based basis-set correction applied to linear response properties

This chapter corresponds to the article [D. Traore, E. Giner, J. Toulouse, J. Chem. Phys. **158**, 234107 (2023)]

The basis-set correction method based on density-functional theory consists in correcting the energy calculated by a wave-function method with a given basis set by a density functional. This basis-set correction density functional incorporates the short-range electron correlation effects missing in the basis set. This results in accelerated basis convergences of ground-state energies to the complete-basis-set limit. In this work, we extend the basis-set correction method to a linear-response formalism for calculating excited-state energies. We give the general linear-response equations, as well as the more specific equations for configuration-interaction wave functions. As a proof of concept, we apply this approach to the calculations of excited-state energies in a one-dimensional two-electron model system with harmonic potential and a Dirac-delta electron-electron interaction. The results obtained with full-configuration-interaction wave functions expanded in a basis of Hermite functions and a local-density-approximation basis-set correction functional show that the present approach does not help in accelerating the basis convergence of excitation energies. However, we show that it significantly accelerates basis convergences of excited-state total energies.

5.1 Introduction

One of the main limitations of standard electronic-structure wave-function computational methods is their slow convergence of ground- and excited-state energies and other properties with respect to the one-electron basis set (see, e.g., Refs. [1–4]). This slow convergence can be traced back to the non-smoothness of the exact eigenfunctions of the Schrödinger Hamiltonian with repulsive Coulomb electron-electron interaction [5], namely the electron-electron cusp condition [6, 7].

There are two main approaches for dealing with this slow basis convergence problem. The first approach consists in extrapolating the results to the complete-basis-set (CBS) limit by using increasingly large basis sets [1, 2]. This approach is very common for estimating the CBS limit

of the ground-state energy but has also been used for estimating the CBS limit of excited-state energies and properties (see, e.g., Refs. [8–11]). The second approach consists in using explicitly correlated R12 or F12 methods which incorporate in the wave function a correlation factor reproducing the electron-electron cusp (see, e.g., Refs. [12–15]). The vast majority of R12/F12 methods have been applied to ground-state energy calculations but linear-response extensions have also been proposed for excitation energies and dynamic response properties [16–20].

Recently, some of the present authors introduced an alternative basis-set correction method based on density-functional theory (DFT) [21]. It consists in correcting the energy calculated by a wave-function method (such as configuration interaction or coupled-cluster) with a given basis set by an adapted basis-set correction density functional incorporating the short-range electron correlation effects missing in the basis set, resulting in an accelerated convergence to the CBS limit. This basis-set correction method was further developed and validated on atomization energies [22–24] and dissociation energy curves [25]. The method was also extended to calculations of ionization potentials within the GW approach [26] and to calculations of dipole moments [27,28]. It was also proposed to extend the method to calculations of excitation energies using a straightforward state-specific approach in which the same basis-set correction functional is evaluated from the density of each state [29]. Even though the last approach was shown to be able to accelerate the basis convergence of electronic excitation energies of molecular systems, it is based on the a-priori questionable assumption that one can use the same basis-set correction functional for all states.

In the present work, we extend the basis-set correction method to a linear-response formalism, providing a more rigorous framework for calculating excitation energies. Moreover, it allows for calculations of response properties such as dynamic polarizabilities. As a first proof of concept, we apply this approach to calculations of excitation energies in a one-dimensional (1D) model system consisting of two electrons in a harmonic potential with a Dirac-delta two-electron interaction [30, 31]. We previously used a similar 1D model system in Ref. [32] to study with some mathematical rigor the basis-set correction method. The relevance of this 1D model for quantum chemistry lies in the fact that the Dirac-delta two-electron interaction induces a slow basis convergence quite similar to the one observed with the standard two-electron Coulomb interaction in three-dimensional (3D) systems.

The paper is organized as follows. In Sec. 5.2, we formulate the general linear-response theory for the DFT-based basis-set correction scheme, and we give explicit expressions for configuration-interaction wave functions. In Sec. 5.3, we apply the linear-response DFT-based basis-set correction theory to the 1D model system and we discuss the results. Finally, Sec. 5.4 contains our conclusions. Hartree atomic units are used throughout this work.

5.2 Linear-response DFT-based basis-set correction

In this section, we derive the general linear-response equations for the DFT-based basis-set correction approach. We consider a finite one-electron basis set $\mathcal{B} \subset H^1(\mathbb{R}^3 \times \{\uparrow, \downarrow\}, \mathbb{C})$ where H^1 is the first-order Sobolev space. The corresponding one-electron Hilbert space spanned by this basis set is denoted by $h^{\mathcal{B}} = \text{span}(\mathcal{B})$ and the corresponding N-electron Hilbert space is given by the N-fold antisymmetric tensor product of $h^{\mathcal{B}}$, i.e. $\mathcal{H}^{\mathcal{B}} = \bigwedge^N h^{\mathcal{B}}$.

5.2.1 General ground-state optimization

We consider a general parametrized wave function $|\Psi(\mathbf{p})\rangle \in \mathcal{H}^{\mathcal{B}}$ with M complex-valued parameters $\mathbf{p} = (p_1, p_2, ..., p_M) \in \mathbb{C}^M$. For example, these parameters could be configuration-interaction coefficients or orbital-rotation parameters. For convenience, we work with the intermediately

normalized wave function (see, e.g., Ref. [33])

$$|\bar{\Psi}(\mathbf{p})\rangle = \frac{|\Psi(\mathbf{p})\rangle}{\langle\Psi_0|\Psi(\mathbf{p})\rangle}, \tag{5.1}$$

where $|\Psi_0\rangle = |\Psi(\mathbf{p}^0)\rangle$ is the current wave function obtained for the current parameters $\mathbf{p} = \mathbf{p}^0$. The current wave function is taken as normalized to unity, i.e. $\langle\Psi_0|\Psi_0\rangle = 1$. The advantage of this intermediate normalization is that the first- and second-order derivatives of $|\bar{\Psi}(\mathbf{p})\rangle$ with respect to $\mathbf{p}$ at $\mathbf{p} = \mathbf{p}^0$,

$$|\bar{\Psi}_I\rangle = \left.\frac{\partial|\bar{\Psi}(\mathbf{p})\rangle}{\partial p_I}\right|_{\mathbf{p}=\mathbf{p}^0} \text{ and } |\bar{\Psi}_{I,J}\rangle = \left.\frac{\partial^2|\bar{\Psi}(\mathbf{p})\rangle}{\partial p_I \partial p_J}\right|_{\mathbf{p}=\mathbf{p}^0}, \tag{5.2}$$

are orthogonal to $|\Psi_0\rangle$, i.e. $\langle\bar{\Psi}_I|\Psi_0\rangle = 0$ and $\langle\bar{\Psi}_{I,J}|\Psi_0\rangle = 0$. This simplifies the derivation of the equations.

In the DFT-based basis-set correction approach [21], we introduce the following ground-state energy expression for a N-electron system with Hamiltonian $\hat{H}$

$$E^{\mathcal{B}}(\mathbf{p}) = \frac{\langle\bar{\Psi}(\mathbf{p})|\hat{H}|\bar{\Psi}(\mathbf{p})\rangle}{\langle\bar{\Psi}(\mathbf{p})|\bar{\Psi}(\mathbf{p})\rangle} + \bar{E}^{\mathcal{B}}[\rho_{\bar{\Psi}(\mathbf{p})}], \tag{5.3}$$

where $\bar{E}^{\mathcal{B}}[\rho]$ is the basis-set correction density functional evaluated at the density of $\bar{\Psi}(\mathbf{p})$

$$\rho_{\bar{\Psi}(\mathbf{p})}(\mathbf{r}) = \frac{\langle\bar{\Psi}(\mathbf{p})|\hat{\rho}(\mathbf{r})|\bar{\Psi}(\mathbf{p})\rangle}{\langle\bar{\Psi}(\mathbf{p})|\bar{\Psi}(\mathbf{p})\rangle}, \tag{5.4}$$

where $\hat{\rho}(\mathbf{r})$ is the density operator at point $\mathbf{r}$. The self-consistent basis-set corrected ground-state energy is then [27]

$$E_0^{\mathcal{B}} = \min_{\mathbf{p}\in\mathbb{C}^M} E^{\mathcal{B}}(\mathbf{p}). \tag{5.5}$$

The role of the density functional $\bar{E}^{\mathcal{B}}[\rho]$ is to accelerate the basis convergence without altering the CBS limit. The latter point is guaranteed by imposing that $\bar{E}^{\mathcal{B}}[\rho]$ vanishes in the CBS limit, i.e. $\lim_{\mathcal{B}\to\text{CBS}} \bar{E}^{\mathcal{B}}[\rho] = 0$.

In practice, the minimization can be done by iteratively solving an effective Schrödinger equation [27], or, more generally, using for example the Newton method in which the current parameters are iteratively updated using the parameters changes $\mathbf{\Delta p} = \mathbf{p} - \mathbf{p}^0$ found by solving the linear equations (see, e.g., Ref. [34])

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{\Delta p} \\ \mathbf{\Delta p}^* \end{pmatrix} = -\begin{pmatrix} \mathbf{g} \\ \mathbf{g}^* \end{pmatrix}, \tag{5.6}$$

where * designates the complex conjugate and $\mathbf{g}$ is the energy gradient vector

$$g_I = \left.\frac{\partial E^{\mathcal{B}}(\mathbf{p})}{\partial p_I^*}\right|_{\mathbf{p}=\mathbf{p}^0} = \langle\bar{\Psi}_I|\hat{H}^{\mathcal{B}}_{\text{eff}}|\Psi_0\rangle, \tag{5.7}$$

with the effective Hamiltonian

$$\hat{H}^{\mathcal{B}}_{\text{eff}} = \hat{H} + \hat{\bar{V}}^{\mathcal{B}}[\rho_{\Psi_0}], \tag{5.8}$$

involving the basis-set correction potential operator

$$\hat{\bar{V}}^{\mathcal{B}}[\rho] = \int_{\mathbb{R}^3} \bar{v}^{\mathcal{B}}[\rho](\mathbf{r})\hat{\rho}(\mathbf{r})\mathrm{d}\mathbf{r}, \tag{5.9}$$

with $\bar{v}^{\mathcal{B}}[\rho](\mathbf{r}) = \delta\bar{E}^{\mathcal{B}}[\rho]/\delta\rho(\mathbf{r})$. In Eq. (5.6), $\mathbf{A}$ and $\mathbf{B}$ are the energy Hessian matrices

$$\begin{aligned} A_{I,J} &= \left.\frac{\partial^2 E^{\mathcal{B}}(\mathbf{p})}{\partial p_I^* \partial p_J}\right|_{\mathbf{p}=\mathbf{p}^0} \\ &= \langle\bar{\Psi}_I|\hat{H}^{\mathcal{B}}_{\text{eff}} - \mathcal{E}^{\mathcal{B}}_0|\bar{\Psi}_J\rangle + K_{I,J}, \end{aligned} \tag{5.10}$$

where $\mathcal{E}^{\mathcal{B}}_0 = \langle\Psi_0|\hat{H}^{\mathcal{B}}_{\text{eff}}|\Psi_0\rangle$ is the energy of the effective Hamiltonian for the current wave function Ψ_0, and

$$\begin{aligned} B_{I,J} &= \left.\frac{\partial^2 E^{\mathcal{B}}(\mathbf{p})}{\partial p_I^* \partial p_J^*}\right|_{\mathbf{p}=\mathbf{p}^0} \\ &= \langle\bar{\Psi}_{I,J}|\hat{H}^{\mathcal{B}}_{\text{eff}}|\Psi_0\rangle + L_{I,J}, \end{aligned} \tag{5.11}$$

involving the basis-set correction kernel contributions

$$K_{I,J} = \int_{\mathbb{R}^3\times\mathbb{R}^3} \bar{f}^{\mathcal{B}}[\rho_{\Psi_0}](\mathbf{r},\mathbf{r}')\langle\bar{\Psi}_I|\hat{\rho}(\mathbf{r})|\Psi_0\rangle\langle\Psi_0|\hat{\rho}(\mathbf{r}')|\bar{\Psi}_J\rangle \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}', \tag{5.12}$$

and

$$L_{I,J} = \int_{\mathbb{R}^3\times\mathbb{R}^3} \bar{f}^{\mathcal{B}}[\rho_{\Psi_0}](\mathbf{r},\mathbf{r}')\langle\bar{\Psi}_I|\hat{\rho}(\mathbf{r})|\Psi_0\rangle\langle\bar{\Psi}_J|\hat{\rho}(\mathbf{r}')|\Psi_0\rangle \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}', \tag{5.13}$$

with $\bar{f}^{\mathcal{B}}[\rho](\mathbf{r},\mathbf{r}') = \delta^2\bar{E}^{\mathcal{B}}[\rho]/\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')$.

At the end of the optimization, provided that we have reached the global energy minimum, the current parameters $\mathbf{p}^0$ are the optimal ground-state parameters. To make clear the link with our previous work [27], we note that the present energy minimization is equivalent to solving the following effective Schrödinger equation projected in the basis of the first-order wave-function derivatives $\{|\bar{\Psi}_I\rangle\}_{I=1,\ldots,M}$

$$\langle\bar{\Psi}_I|\hat{H}^{\mathcal{B}}_{\text{eff}} - \mathcal{E}^{\mathcal{B}}_0|\Psi_0\rangle = 0. \tag{5.14}$$

Since $\langle\bar{\Psi}_I|\Psi_0\rangle = 0$, Eq. (5.14) is indeed equivalent to having a zero energy gradient, i.e. $g_I = \langle\bar{\Psi}_I|\hat{H}^{\mathcal{B}}_{\text{eff}}|\Psi_0\rangle = 0$.

5.2.2 General linear-response equations

Starting from the optimal ground state, we now add a time-dependent perturbation operator $\hat{V}(t)$ to the Hamiltonian,

$$\hat{H}(t) = \hat{H} + \hat{V}(t), \tag{5.15}$$

where $\hat{V}(t)$ is chosen as a periodic monochromatic electric-dipole interaction of frequency ω

$$\hat{V}(t) = -\hat{\mathbf{d}}\cdot\boldsymbol{\epsilon}^+ e^{-\mathrm{i}\omega t} - \hat{\mathbf{d}}\cdot\boldsymbol{\epsilon}^- e^{+\mathrm{i}\omega t}, \tag{5.16}$$

where $\hat{\mathbf{d}} = -\int_{\mathbb{R}^3}\mathbf{r}\,\hat{\rho}(\mathbf{r})\mathrm{d}\mathbf{r}$ is the dipole-moment operator, and $\boldsymbol{\epsilon}^+$ and $\boldsymbol{\epsilon}^-$ are the electric-field strengths for the positive and negative frequency terms (taken as different for intermediate derivations but ultimately we must have $\boldsymbol{\epsilon}^+ = \boldsymbol{\epsilon}^-$).

The wave function $|\bar{\Psi}(\mathbf{p}(t))\rangle$ will now depend on time through the parameters $\mathbf{p}(t) = \mathbf{p}^0 + \boldsymbol{\Delta}\mathbf{p}(t)$ where $\mathbf{p}^0$ are the optimal ground-state parameters and $\boldsymbol{\Delta}\mathbf{p}(t)$ are the time variations of the parameters which can be written as

$$\boldsymbol{\Delta}\mathbf{p}(t) = \mathbf{p}^+ e^{-\mathrm{i}\omega t} + \mathbf{p}^- e^{+\mathrm{i}\omega t}, \tag{5.17}$$

where $\mathbf{p}^+ \in \mathbb{C}^M$ and $\mathbf{p}^- \in \mathbb{C}^M$ are the Fourier components. The ground-state energy expression in Eq. (5.3) is replaced by the quasi-energy expression [35–40]

$$Q^{\mathcal{B}}(\boldsymbol{\epsilon}^+, \boldsymbol{\epsilon}^-; \mathbf{p}^+, \mathbf{p}^-) = \frac{1}{T} \int_0^T \left\{ \frac{\langle \bar{\Psi}(\mathbf{p}(t)) | \hat{H}(t) - i\frac{\partial}{\partial t} | \bar{\Psi}(\mathbf{p}(t)) \rangle}{\langle \bar{\Psi}(\mathbf{p}(t)) | \bar{\Psi}(\mathbf{p}(t)) \rangle} + \bar{E}^{\mathcal{B}}[\rho_{\bar{\Psi}(\mathbf{p}(t))}] \right\} \mathrm{d}t, \tag{5.18}$$

where $T = 2\pi/\omega$ is the period. Note that, in the definition of the quasi-energy in Eq. (5.18), the same basis-set correction density functional $\bar{E}^{\mathcal{B}}[\rho]$ used for the ground-state calculation is employed, which is known as the adiabatic approximation. This approximation is almost always used in time-dependent DFT calculations of excitation energies (see, e.g, Refs. [41, 42]). Due to this approximation, the basis-set correction contribution to the quasi-energy is a local functional of time. Overcoming this approximation would require the complicated task of developing a quasi-energy basis-set correction contribution having the form a non-local functional of time, i.e. depending on all the time history. We do not attempt to do that in the present work. The optimal quasi-energy $Q_0^{\mathcal{B}}(\boldsymbol{\epsilon}^+, \boldsymbol{\epsilon}^-)$ is a stationary value of $Q^{\mathcal{B}}(\boldsymbol{\epsilon}^+, \boldsymbol{\epsilon}^-; \mathbf{p}^+, \mathbf{p}^-)$ with respect to variations of the parameters $\mathbf{p}^+$ and $\mathbf{p}^-$, which we write as

$$Q_0^{\mathcal{B}}(\boldsymbol{\epsilon}^+, \boldsymbol{\epsilon}^-) \in \underset{(\mathbf{p}^+, \mathbf{p}^-) \in \mathbb{C}^{2M}}{\mathrm{stat}} Q^{\mathcal{B}}(\boldsymbol{\epsilon}^+, \boldsymbol{\epsilon}^-; \mathbf{p}^+, \mathbf{p}^-), \tag{5.19}$$

where "stat" refers to the set of stationary values. In the zero electric-field limit ($\boldsymbol{\epsilon}^+ = \boldsymbol{\epsilon}^- = \mathbf{0}$), vanishing parameters $\mathbf{p}^+ = \mathbf{p}^- = \mathbf{0}$ are optimal and the corresponding optimal quasi-energy reduces to the ground-state energy, i.e. $Q_0^{\mathcal{B}}(\mathbf{0}, \mathbf{0}) = E_0^{\mathcal{B}}$.

The optimal quasi-energy allows one to define the dynamic dipole polarizability tensor as

$$\alpha_{i,j}^{\mathcal{B}}(\omega) = - \left. \frac{\partial^2 Q_0^{\mathcal{B}}(\boldsymbol{\epsilon}^+, \boldsymbol{\epsilon}^-)}{\partial \epsilon_i^- \partial \epsilon_j^+} \right|_{\boldsymbol{\epsilon}^\pm = \mathbf{0}}, \tag{5.20}$$

where i and j refer to 3D Cartesian components. Calculating this second-order derivative using the chain rule via the optimal parameters $\mathbf{p}^+$ and $\mathbf{p}^-$ (which implicitly depend on $\boldsymbol{\epsilon}^+$ and $\boldsymbol{\epsilon}^-$) leads to (see, e.g., Refs. [39, 40])

$$\alpha_{i,j}^{\mathcal{B}}(\omega) = \begin{pmatrix} \mathbf{V}_i \\ \mathbf{V}_i^* \end{pmatrix}^\dagger \begin{pmatrix} \boldsymbol{\Lambda}(\omega) & \boldsymbol{\Xi} \\ \boldsymbol{\Xi}^* & \boldsymbol{\Lambda}(-\omega)^* \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{V}_j \\ \mathbf{V}_j^* \end{pmatrix}, \tag{5.21}$$

where $\boldsymbol{\Lambda}(\omega)$ and $\boldsymbol{\Xi}$ are the quasi-energy Hessian matrices

$$\Lambda_{I,J}(\omega) = \left. \frac{\partial^2 Q^{\mathcal{B}}(\boldsymbol{\epsilon}^+, \boldsymbol{\epsilon}^-; \mathbf{p}^+, \mathbf{p}^-)}{\partial p_I^{+*} \partial p_J^+} \right|_{\substack{\mathbf{p}^\pm = \mathbf{0} \\ \boldsymbol{\epsilon}^\pm = \mathbf{0}}} = A_{I,J} - \omega S_{I,J}, \tag{5.22}$$

and

$$\Xi_{I,J} = \left. \frac{\partial^2 Q^{\mathcal{B}}(\boldsymbol{\epsilon}^+, \boldsymbol{\epsilon}^-; \mathbf{p}^+, \mathbf{p}^-)}{\partial p_I^{+*} \partial p_J^{-*}} \right|_{\substack{\mathbf{p}^\pm = \mathbf{0} \\ \boldsymbol{\epsilon}^\pm = \mathbf{0}}} = B_{I,J}, \tag{5.23}$$

where $A_{I,J}$ and $B_{I,J}$ are given in Eqs. (5.10) and (5.11), and $S_{I,J}$ is the overlap matrix of the first-order wave-function derivatives

$$S_{I,J} = \langle \bar{\Psi}_I | \bar{\Psi}_J \rangle. \tag{5.24}$$

Equation (5.21) also involves the perturbed energy gradient vector

$$V_{j,I} = \left. \frac{\partial^2 Q^{\mathcal{B}}(\boldsymbol{\epsilon}^+, \boldsymbol{\epsilon}^-; \mathbf{p}^+, \mathbf{p}^-)}{\partial \epsilon_j^+ \partial p_I^{+*}} \right|_{\substack{\mathbf{p}^\pm = \mathbf{0} \\ \boldsymbol{\epsilon}^\pm = \mathbf{0}}} = -\langle \bar{\Psi}_I | \hat{d}_j | \Psi_0 \rangle, \tag{5.25}$$

corresponding to transition dipole-moment matrix elements.

Finally, the poles in ω of the dynamic dipole polarizability $\alpha^{\mathcal{B}}_{i,j}(\omega)$ provide M positive excitation energies $\{\omega^{\mathcal{B}}_n\}$ (and M opposite deexcitation energies), which can be found from the following generalized eigenvalue equation (see, e.g., Ref. [34])

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix} = \omega^{\mathcal{B}}_n \begin{pmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & -\mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix}, \tag{5.26}$$

where $(\mathbf{X}_n, \mathbf{Y}_n)$ are eigenvectors. The obtained excitation energies $\{\omega^{\mathcal{B}}_n\}$ include the basis-set correction through the potential $\bar{v}^{\mathcal{B}}[\rho](\mathbf{r})$ in Eq. (5.9) and kernel $\bar{f}^{\mathcal{B}}[\rho](\mathbf{r}, \mathbf{r}')$ in Eqs. (5.12) and (5.13), and may be expected to converge faster to their CBS limit, provided good enough approximations are used for the basis-set correction functional used for $\bar{v}^{\mathcal{B}}[\rho](\mathbf{r})$ and $\bar{f}^{\mathcal{B}}[\rho](\mathbf{r}, \mathbf{r}')$. Obviously, the corresponding basis-set corrected total energy of the n$^{\text{th}}$ excited state is given by

$$E^{\mathcal{B}}_n = E^{\mathcal{B}}_0 + \omega^{\mathcal{B}}_n, \tag{5.27}$$

and could also be expected to converge faster to its CBS limit, if the basis-set correction functional is good enough.

5.2.3 Linear-response equations for configuration-interaction wave functions

We now give the more specific form of the linear-response equations for configuration-interaction (CI) wave functions. Given a set of M orthonormal configurations $\{|\Phi_I\rangle\}$, the CI wave function is parametrized as

$$|\Psi(\mathbf{p})\rangle = \sum_{I=1}^{M} p_I |\Phi_I\rangle. \tag{5.28}$$

The ground-state parameters are assumed to be real valued and are denoted by $p^0_I = c_I$, so that the ground-state wave function is $|\Psi_0\rangle = \sum_{I=1}^{M} c_I |\Phi_I\rangle$. From Eq. (5.2), the first-order and second-order derivatives of the intermediately normalized wave function are found to be

$$|\bar{\Psi}_I\rangle = |\Phi_I\rangle - c_I\, |\Psi_0\rangle, \tag{5.29}$$

and

$$|\bar{\Psi}_{I,J}\rangle = 2c_I c_J\, |\Psi_0\rangle - c_J\, |\Phi_I\rangle - c_I\, |\Phi_J\rangle. \tag{5.30}$$

In a spin-restricted formalism with a set of real-valued orthonormal orbitals $\{\varphi_i\} \subset \mathcal{H}^{\mathcal{B}}$, the linear-response matrices in Eq. (5.26) now become

$$A_{I,J} = \langle \Phi_I | \hat{H}^{\mathcal{B}}_{\text{eff}} - \mathcal{E}^{\mathcal{B}}_0 | \Phi_J \rangle + K_{I,J}, \tag{5.31}$$

$$B_{I,J} = K_{I,J}, \tag{5.32}$$

$$S_{I,J} = \delta_{I,J} - c_I c_J, \tag{5.33}$$

where the kernel contribution takes the form

$$K_{I,J} = \sum_{i,j,k,l} \Delta\gamma^I_{i,j}\, \Delta\gamma^J_{k,l}\, \bar{f}^{\mathcal{B}}_{i,j,k,l}, \tag{5.34}$$

with i, j, k, l referring to spatial orbitals. In Eq. (5.34), we have introduced

$$\Delta\gamma^I_{i,j} = \gamma^I_{i,j} - c_I \gamma_{i,j}, \tag{5.35}$$

Table 5.1: Exact total energies of the first 5 eigenstates of even-parity and singlet symmetries for the 1D two-electron Hooke-type atom with $\omega_0 = 1$, and corresponding excitation energies. All energies are in hartree.

State (n,m)	Total energy $E_{n,m}$	Excitation energy $E_{n,m} - E_{0,0}$
$(0,0)$	1.306746	
$(0,2)$	3.187051	1.880305
$(2,0)$	3.306746	2.000000
$(0,4)$	5.144734	3.837988
$(2,2)$	5.187051	3.880305

where $\gamma_{i,j}$ and $\gamma^I_{i,j}$ are the ground-state and transition density matrices, respectively,

$$\gamma_{i,j} = \langle \Psi_0 | \hat{E}_{i,j} | \Psi_0 \rangle = \sum_{I=1}^{M} \sum_{J=1}^{M} c_I c_J \langle \Phi_I | \hat{E}_{i,j} | \Phi_J \rangle, \tag{5.36}$$

$$\gamma^I_{i,j} = \langle \Phi_I | \hat{E}_{i,j} | \Psi_0 \rangle = \sum_{J=1}^{M} c_J \langle \Phi_I | \hat{E}_{i,j} | \Phi_J \rangle, \tag{5.37}$$

where $\hat{E}_{i,j} = \hat{a}^\dagger_{i\uparrow}\hat{a}_{j\uparrow} + \hat{a}^\dagger_{i\downarrow}\hat{a}_{j\downarrow}$ is the spin-summed one-electron density-matrix operator in second quantization. Finally, in Eq. (5.34), $\bar{f}^{\mathcal{B}}_{i,j,k,l}$ are the matrix elements of the basis-set correction kernel $\bar{f}^{\mathcal{B}}[\rho_{\Psi_0}](\mathbf{r},\mathbf{r}')$ over the spatial orbitals

$$\bar{f}^{\mathcal{B}}_{i,j,k,l} = \int_{\mathbb{R}^3\times\mathbb{R}^3} \bar{f}^{\mathcal{B}}[\rho_{\Psi_0}](\mathbf{r},\mathbf{r}')\, \varphi_i(\mathbf{r})\varphi_j(\mathbf{r})\varphi_k(\mathbf{r}')\varphi_l(\mathbf{r}') \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'. \tag{5.38}$$

5.3 One-dimensional model system

5.3.1 Description of the model and exact solutions

We consider the 1D two-electron Hooke-type atom studied in Refs. [30–32]. We work first in the infinite-dimensional spin-free one-electron Hilbert space $\mathfrak{h} = L^2(\mathbb{R},\mathbb{C})$ and the associated non-antisymmetrized tensor-product two-electron Hilbert space $\mathcal{H} = \mathfrak{h} \otimes \mathfrak{h}$. The Hamiltonian is

$$\hat{H} = -\frac{1}{2}\frac{\partial^2}{\partial x_1^2} - \frac{1}{2}\frac{\partial^2}{\partial x_2^2} + \frac{1}{2}\omega_0^2 x_1^2 + \frac{1}{2}\omega_0^2 x_2^2 + \delta(x_1 - x_2), \tag{5.39}$$

involving a harmonic external potential of curvature ω_0^2 (which will be chosen to 1 throughout this study) and a Dirac-delta two-electron interaction. The latter two-electron interaction generates in 1D the same s-wave electron-electron cusp as the Coulomb interaction does in 3D, and it is thus an appropriate model to study the basis convergence [32]. This 1D two-electron Hooke-type atom can be considered as the 1D analog of the well-known 3D two-electron Hooke atom (see, e.g., Refs. [43, 44]).

Expressed with the center-of-mass (cm) coordinate $X = (x_1 + x_2)/2$ and the relative (rel) coordinate $x_{12} = x_1 - x_2$, the Hamiltonian is separable [32] and its eigenvalues are

$$E_{n,m} = E^{\mathrm{cm}}_n + E^{\mathrm{rel}}_m \text{ for } n \in \mathbb{N} \text{ and } m \in \mathbb{N}, \tag{5.40}$$

where $E^{\mathrm{cm}}_n = \omega_0(n + 1/2)$ is the center-of-mass contribution and E^{rel}_m is the relative contribution,

$$E^{\mathrm{rel}}_m = \begin{cases} \omega_0(\nu_m + 1/2), & \text{for } m \text{ even}, \\ \omega_0(m + 1/2), & \text{for } m \text{ odd}, \end{cases} \tag{5.41}$$

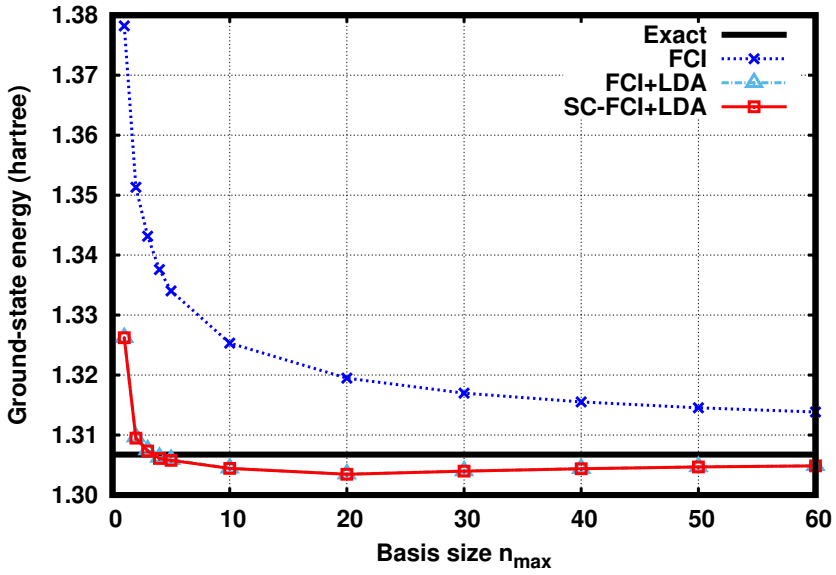

Figure 5.1: Ground-state energy of the 1D two-electron Hooke-type atom with $\omega_0 = 1$ calculated by the standard FCI method, the non-self-consistent FCI+LDA and self-consistent SC-FCI+LDA basis-set corrected methods as a function of the basis size $n_{\max}$.

with the real numbers ν_m being the solutions of the equation [30, 45]

$$2\sqrt{2\omega_0}\,\frac{\Gamma\left(-\frac{\nu_m}{2}+\frac{1}{2}\right)}{\Gamma\left(-\frac{\nu_m}{2}\right)} = -1, \tag{5.42}$$

where Γ is the gamma function. The associated eigenfunctions are

$$\Psi_{n,m}(X, x_{12}) = \psi_n^{\rm cm}(X)\psi_m^{\rm rel}(x_{12}), \tag{5.43}$$

where the center-of-mass eigenfunctions are $\psi_n^{\rm cm}(X) = f_n^{2\omega_0}(X)$ and the relative eigenfunctions are

$$\psi_m^{\rm rel}(x_{12}) = \begin{cases} c_m D_{\nu_m}(\sqrt{\omega_0}|x_{12}|), & \text{for } m \text{ even}, \\ f_m^{\omega_0/2}(x_{12}), & \text{for } m \text{ odd}, \end{cases} \tag{5.44}$$

where D_{ν_m} is the parabolic cylinder function [46] and c_m is a normalization constant. Here, $f_n^{\omega_0}$ designates the Hermite functions (i.e., quantum harmonic-oscillator eigenfunctions) for the frequency ω_0

$$f_n^{\omega_0}(x) = N_n^{\omega_0}\, H_n(\sqrt{\omega_0}x)\, e^{-\omega_0 x^2/2}, \tag{5.45}$$

where H_n are the Hermite polynomials, $N_n^{\omega_0} = (2^n n!)^{-1/2}(\omega_0/\pi)^{1/4}$ is the normalization factor. As announced, for even m, the relative eigenfunctions has the familiar s-wave cusp condition: $\psi_m^{\rm rel}(x_{12}) = \psi_m^{\rm rel}(0)[1 + (1/2)|x_{12}| + O(x_{12}^2)]$ (see Ref. [32]).

We will only consider eigenstates of even-parity symmetry [i.e., invariant under the transformation $(x_1, x_2) \to (-x_1, -x_2)$] and singlet symmetry (i.e., invariant under the exchange $x_1 \leftrightarrow x_2$), corresponding to the eigenstates with both even quantum numbers n and m. The exact total energies of the first 5 of these eigenstates, as well as the corresponding excitation energies, are given in Table 5.1 for $\omega_0 = 1$.

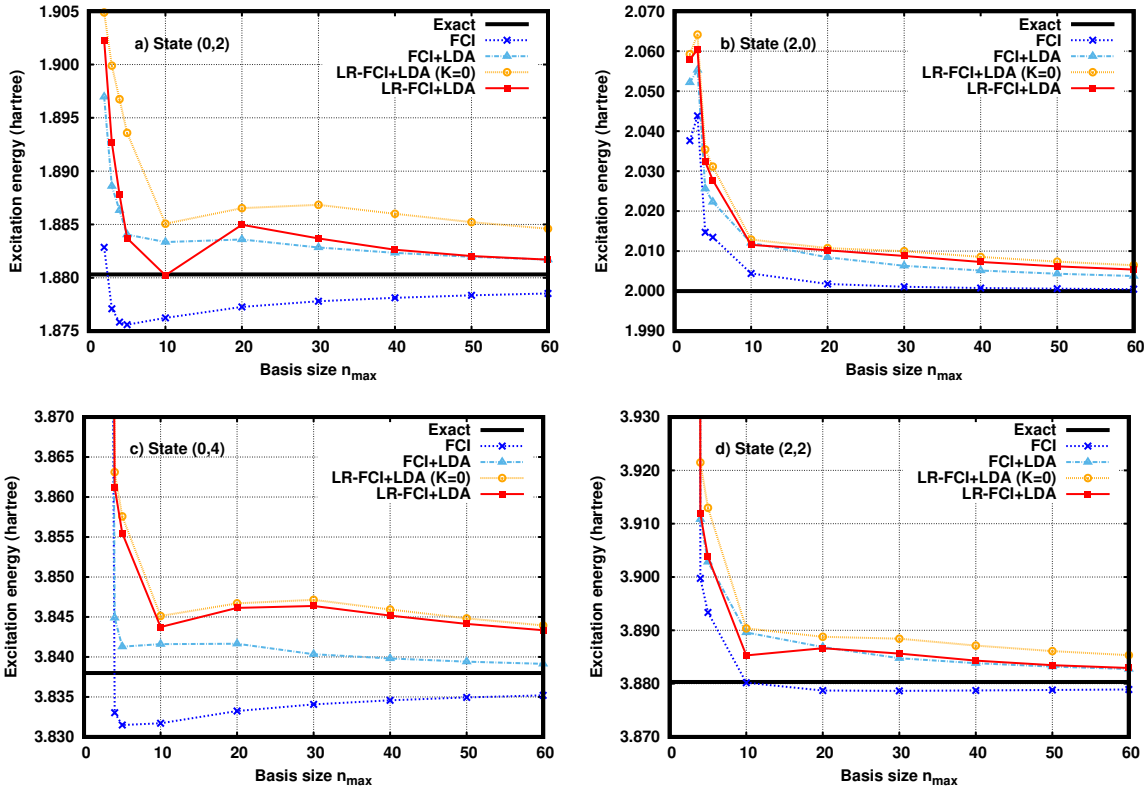

Figure 5.2: Excitation energies of the states $(0,2)$, $(2,0)$, $(0,4)$, and $(2,2)$ of the 1D two-electron Hooke-type atom with $\omega_0 = 1$ calculated by the standard FCI method and the LR-FCI+LDA basis-set corrected method as a function of the basis size $n_{\max}$. For comparison, excitation energies obtained with a zero basis-set correction kernel [LR-FCI+LDA (K=0)] and with a non-self-consistent state-specific approach (FCI+LDA) are also shown.

5.3.2 Full-configuration-interaction calculation in a basis set

We consider finite basis sets of Hermite (or Hermite-Gauss) functions

$$\mathcal{B} = \{f_n^{\omega_0}\}_{n=1,\ldots,n_{\max}}, \tag{5.46}$$

with a fixed parameter $\omega_0 = 1$ and a variable maximal quantum number $n_{\max}$ determining the basis size. The one-electron and two-electron Hilbert spaces corresponding to this basis set are $h^{\mathcal{B}} = \text{span}(\mathcal{B})$ and $\mathcal{H}^{\mathcal{B}} = h^{\mathcal{B}} \otimes h^{\mathcal{B}}$.

For several values of $n_{\max}$, we first perform a Hartree-Fock (HF) calculation to obtain the set of orthonormal HF orbitals $\{\varphi_i\}$, and we then perform a full-configuration-interaction (FCI) calculation for the states of even-parity symmetry. The parameterized FCI wave function is thus $|\Psi_{\text{FCI}}(\mathbf{p})\rangle = \sum_{I=1}^{M} p_I |\Phi_I\rangle$ where $|\Phi_I\rangle = |\varphi_{I_1}\rangle \otimes |\varphi_{I_2}\rangle$, and the orbitals φ_{I_1} and φ_{I_2} are restricted to be of the same parity symmetry. In Fig. 5.1 we report the FCI ground-state energy $E_{0,\text{FCI}} = \langle \Psi_{0,\text{FCI}} | \hat{H} | \Psi_{0,\text{FCI}} \rangle$, where $\Psi_{0,\text{FCI}}$ is the FCI ground-state wave function, as a function of the basis size $n_{\max}$. As expected, the FCI ground-state energy slowly converges toward the exact ground-state energy as $n_{\max}$ increases. The convergence rate is compatible with the theoretical convergence rate of $1/n_{\max}^{1/2}$ determined in Ref. [32].

We construct a local-density approximation (LDA) for the basis-set correction functional $\bar{E}^{\mathcal{B}}[\rho]$

$$\bar{E}^{\mathcal{B}}_{\text{LDA}}[\rho] = \int_{\mathbb{R}} \rho(\mathbf{r}) \bar{\varepsilon}^{\mathcal{B}}(\rho(\mathbf{r})) \, \mathrm{d}\mathbf{r}, \tag{5.47}$$

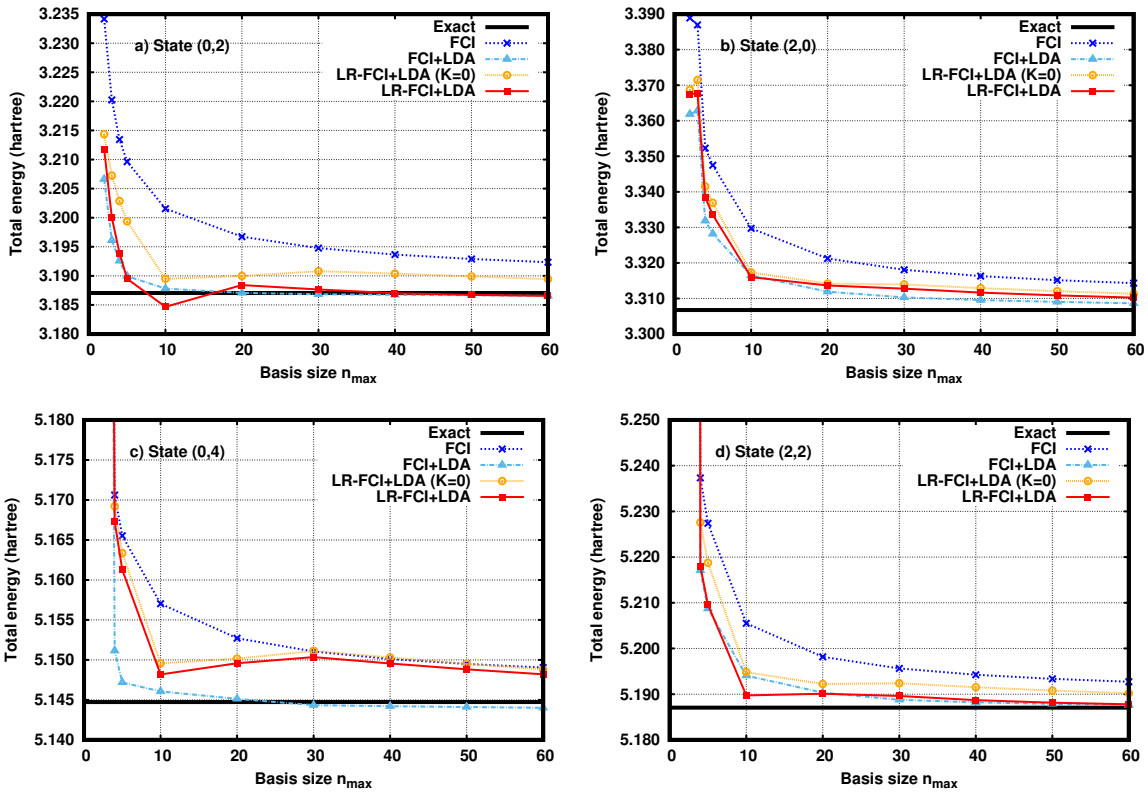

Figure 5.3: Excited-state total energies of the states $(0,2)$, $(2,0)$, $(0,4)$, and $(2,2)$ of the 1D two-electron Hooke-type atom with $\omega_0 = 1$ calculated by the standard FCI method and the LR-FCI+LDA basis-set corrected method as a function of the basis size $n_{\max}$. For comparison, excited-state total energies obtained with a zero basis-set correction kernel [LR-FCI+LDA (K=0)] and with a non-self-consistent state-specific approach (FCI+LDA) are also shown.

where the energy per particle $\bar{\varepsilon}^{\mathcal{B}}(\rho)$ is defined in exactly the same way as in Ref. [32], i.e. as the complementary multideterminant correlation energy per particle of a two-electron finite uniform electron gas with electron-electron interaction projected in the basis set $\mathcal{B}$. For convenience, we fit the numerically calculated energy per particle $\bar{\varepsilon}^{\mathcal{B}}(\rho)$ to a rational fraction

$$\bar{\varepsilon}^{\mathcal{B}}(\rho) \approx \frac{\sum_{i=0}^{4} a_i^{\mathcal{B}} \rho^i}{1 + \sum_{j=1}^{4} b_j^{\mathcal{B}} \rho^j}, \tag{5.48}$$

where the values of the coefficients $a_i^{\mathcal{B}}$ and $b_j^{\mathcal{B}}$ for each basis size $n_{\max}$ are given in the Supplementary Material.

We perform a ground-state FCI calculation including self-consistently the basis-set correction LDA functional according to Eq. (5.5). The required LDA basis-set correction potential is obtained by straightforward differentiation of Eq. (5.47)

$$\bar{v}_{\mathrm{LDA}}^{\mathcal{B}}[\rho](\mathbf{r}) = \bar{\varepsilon}^{\mathcal{B}}(\rho(\mathbf{r})) + \rho(\mathbf{r}) \left. \frac{\mathrm{d}\bar{\varepsilon}^{\mathcal{B}}(\rho)}{\mathrm{d}\rho} \right|_{\rho=\rho(\mathbf{r})}. \tag{5.49}$$

The resulting energy, labelled as SC-FCI+LDA, is reported as a function of the basis size $n_{\max}$ in Fig. 5.1. We see that the basis-set correction LDA functional is very effective in reducing the basis-set incompleteness error, resulting in a fast convergence of the SC-FCI+LDA ground-state energy toward the exact ground-state energy. For comparison, we also show in Fig. 5.1 the non-self-consistent approximation [21, 32], labelled as FCI+LDA,

$$E_{0,\mathrm{FCI+LDA}} = E_{0,\mathrm{FCI}} + \bar{E}_{\mathrm{LDA}}^{\mathcal{B}}[\rho_{\Psi_{0,\mathrm{FCI}}}]. \tag{5.50}$$

On the scale of the plot, it is superimposed with the SC-FCI+LDA energy, showing that the non-self-consistent approximation is an excellent approximation for calculating the ground-state energy of the present system. The same trends have been observed in atomic and molecular systems [27].

We then perform linear-response calculations on-top of the ground-state SC-FCI+LDA calculations according to Eq. (5.26). The required LDA basis-set correction kernel is obtained by differentiation of Eq. (5.49)

$$\bar{f}_{\mathrm{LDA}}^{\mathcal{B}}[\rho](\mathbf{r},\mathbf{r}') = \left[2 \left. \frac{\mathrm{d}\bar{\varepsilon}^{\mathcal{B}}(\rho)}{\mathrm{d}\rho} \right|_{\rho=\rho(\mathbf{r})} + \rho(\mathbf{r}) \left. \frac{\mathrm{d}^2\bar{\varepsilon}^{\mathcal{B}}(\rho)}{\mathrm{d}\rho^2} \right|_{\rho=\rho(\mathbf{r})} \right] \delta(\mathbf{r}-\mathbf{r}'). \tag{5.51}$$

The resulting linear-response basis-set corrected excitation energies, labelled as LR-FCI+LDA, are reported in Fig. 5.2 as a function of the basis set $n_{\max}$ for the four excited states considered in Table 5.1, and compared to the excitation energies obtained by standard FCI. The first thing to note is that the FCI excitation energies have a much faster basis convergence than the FCI ground-state energy. This is somehow expected since the same electron-electron cusp condition applies for both the ground state and the considered excited states, and therefore the short-range correlation effects normally responsible for the slow basis convergence should partially cancel out in the excitation energies. Accelerating the basis convergence of excitation energies is thus a more subtle task than accelerating the basis convergence of ground-state energies. In fact, LR-FCI-LDA does not provide any improvement over standard FCI but instead mostly deteriorates the basis convergence of excitation energies. We may attribute these disappointing results to the limited accuracy of the LDA basis-set correction potential and kernel.

In Fig. 5.2, we also show excitation energies obtained with a basis-set correction kernel set to zero, such that

$$A_{I,J} \simeq \langle \Phi_I | \hat{H}_{\mathrm{eff}}^{\mathcal{B}} - \mathcal{E}_0^{\mathcal{B}} | \Phi_J \rangle, \tag{5.52}$$

and

$$B_{I,J} \simeq 0. \tag{5.53}$$

This approximation is labelled as LR-FCI+LDA (K=0) in the figures of the present paper. It is somewhat consoling to see that the LDA kernel does nevertheless improve the excitation energies, albeit sometimes by a small amount. Finally, Fig. 5.2 also reports the excitation energies obtained by the non-self-consistent state-specific approach of Ref. [29], labelled as FCI+LDA. In this approach, the excited-state energy of the n^{th} excited state is estimated as

$$E_{n,\text{FCI+LDA}} = E_{n,\text{FCI}} + \bar{E}^{\mathcal{B}}_{\text{LDA}}[\rho_{\Psi_{n,\text{FCI}}}], \tag{5.54}$$

where $E_{n,\text{FCI}} = \langle \Psi_{n,\text{FCI}} | \hat{H} | \Psi_{n,\text{FCI}} \rangle$ is the FCI total energy of the n^{th} excited state with wave function $\Psi_{n,\text{FCI}}$. The excitation energy is then given by $E_{n,\text{FCI+LDA}} - E_{0,\text{FCI+LDA}}$. Globally, the state-specific FCI+LDA approach gives excitation energies quite similar to the LR-FCI+LDA method, except for the state $(0,4)$ where FCI+LDA gives clearly excitation energies that more rapidly converge with the basis size. Since the state-specific FCI+LDA approach only involves the energy density functional $\bar{E}^{\mathcal{B}}_{\text{LDA}}[\rho]$ and not its derivatives, it may indicate that LDA is more accurate for the energy than for the potential and kernel.

We discuss now the total excited-state energies which are reported in Fig. 5.3 as a function of the basis set $n_{\max}$ for the four excited states considered. Here, we observe that the FCI excited-state energies exhibit a similar convergence with respect to $n_{\max}$ as the FCI ground-state energy. This is expected since, as mentioned before, the same electron-electron cusp condition applies for both the ground state and the considered excited states. In comparison to the case of the excitation energies, there is no partial cancellation of short-range electron correlation effects, and it is thus an easier task to accelerate the basis convergence of total excited-state energies. Globally, the LR-FCI+LDA excited-state energies [Eq. (5.27)] tend to have less basis-set incompleteness error than the standard FCI excited-state energies, and converge faster with $n_{\max}$ to the exact energies. However, the performance of the basis correction is not uniform over all the states considered. For the state $(0,2)$, the basis-set correction is very effective in reducing the error and accelerating the basis convergence. For the states $(2,0)$ and $(2,2)$, the basis-set correction again effectively reduces the error but does not seem to significantly change the convergence rate for large $n_{\max}$. For the state $(0,4)$, the basis-set correction only reduces the error for small $n_{\max}$ but does not improve the standard FCI energy for $n_{\max} \gtrsim 30$.

Comparison with the total excited-state energies obtained with a zero basis-set correction kernel [LR-FCI+LDA (K=0)] shows again that the LDA kernel improves the basis convergence, even though the effect is small for some of the states. The state-specific FCI+LDA approach gives total excited-state energies very similar to the LR-FCI+LDA ones, except again for the state $(0,4)$ where FCI+LDA gives total excited-state energies that rapidly converge with basis size.

As a final comment, we note that the FCI total energies are of course systematically higher than the exact total energies for the ground and excited states, which makes possible a partial compensation of errors in the FCI excitation energies. By contrast, the basis-set corrected total energies converge to the exact total energies from below for the ground state and from above for the excited states, and thus the basis-set corrected excitation energies do not enjoy any compensation of errors.

5.4 Conclusions

In this work, we have extended the DFT-based basis-set correction method to the linear-response formalism, allowing one to calculate excited-state energies. We have given the general linear-response equations, as well as the more specific equations for configuration-interaction wave functions. As a proof of concept, we have applied this approach to the calculations of excited-state

energies in 1D two-electron model system with harmonic potential and a Dirac-delta electron-electron interaction. The results obtained with FCI wave functions expanded in a basis of Hermite functions and a LDA basis-set correction functional within the adiabatic approximation show that the present linear-response basis-set correction method unfortunately does not help in accelerating the basis convergence of excitation energies. However, it does significantly accelerate the basis convergence of excited-state total energies.

These mixed results should now be checked on real 3D molecular systems. Possibly, for these systems, an important ingredient to add to the basis-set correction functional will be the on-top pair density. The fact that the simple non-self-consistent state-specific basis-set correction approach was found in Ref. [29] to help accelerating the convergence of excitation energies in molecular systems gives us hope that the present linear-response basis-set correction method could be useful as well for these systems.

Supplementary Material

The Supplementary Material contains the coefficients of the LDA basis-set correction energy per particle in Eq. (D.1) for different basis sizes $n_{\max}$.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme Grant agreement No. 810367 (EMC2).

Author Declarations

The authors have no conflicts to disclose.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Bibliography

[1] T. Helgaker, W. Klopper, H. Koch and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[2] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).

[3] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2002).

[4] J. J. Shepherd, A. Grüneis, G. H. Booth, G. Kresse and A. Alavi, Phys. Rev. B **86**, 035111 (2012).

[5] W. Kutzelnigg and J. D. Morgan III, J. Chem. Phys. **96**, 4484 (1992).

[6] T. Kato, Comm. Pure Appl. Math. **10**, 151 (1957).

[7] R. T. Pack and W. Byers-Brown, J. Chem. Phys. **45**, 556 (1966).

[8] M. Kállay and J. Gauss, J. Chem. Phys. **121**, 9257 (2004).

[9] M. A. Watson and G. K.-L. Chan, J. Chem. Theory Comput. **8**, 4013 (2012).

[10] P. Baudin, J. S. Marín, I. G. Cuesta and A. M. J. Sánchez de Merás, J. Chem. Phys. **140**, 104111 (2014).

[11] A. Chrayteh, A. Blondel, P.-F. Loos and D. Jacquemin, J. Chem. Theory Comput. **17**, 416 (2021).

[12] S. Ten-no and J. Noga, WIREs Comput. Mol. Sci. **2**, 114 (2012).

[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] T. Shiozaki and H.-J. Werner, Mol. Phys. **111**, 607 (2013).

[16] H. Fliegl, C. Hättig and W. Klopper, J. Chem. Phys. **124**, 044112 (2006).

[17] C. Neiss, C. Hättig and W. Klopper, J. Chem. Phys. **125**, 064111 (2006).

[18] C. Neiss and C. Hättig, J. Chem. Phys. **126**, 154101 (2007).

[19] A. Köhn, J. Chem. Phys. **130**, 104104 (2009).

[20] M. Hanauer and A. Köhn, J. Chem. Phys. **131**, 124118 (2009).

[21] E. Giner, B. Pradines, A. Ferté, R. Assaraf, A. Savin and J. Toulouse, J. Chem. Phys. **149**, 194301 (2018).

[22] P.-F. Loos, B. Pradines, A. Scemama, J. Toulouse and E. Giner, J. Phys. Chem. Lett. **10**, 2931 (2019).

[23] Y. Yao, E. Giner, J. Li, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **153**, 124117 (2020).

[24] Y. Yao, E. Giner, T. A. Anderson, J. Toulouse and C. J. Umrigar, J. Chem. Phys. **155**, 204104 (2021).

[25] E. Giner, A. Scemama, P.-F. Loos and J. Toulouse, J. Chem. Phys. **152**, 174104 (2020).

[26] P.-F. Loos, B. Pradines, A. Scemama, E. Giner and J. Toulouse, J. Chem. Theory Comput. **16**, 1018 (2020).

[27] E. Giner, D. Traore, B. Pradines and J. Toulouse, J. Chem. Phys. **155**, 044109 (2021).

[28] D. Traore, J. Toulouse and E. Giner, J. Chem. Phys. **156**, 174101 (2022).

[29] E. Giner, A. Scemama, J. Toulouse and P.-F. Loos, J. Chem. Phys. **151**, 144118 (2019).

[30] R. J. Magyar and K. Burke, Phys. Rev. A **70**, 032508 (2004).

[31] N. T. Maitra, F. Zhang, R. J. Cave and K. Burke, J. Chem. Phys. **120**, 5932 (2004).

[32] D. Traore, E. Giner and J. Toulouse, J. Chem. Phys. **156**, 044113 (2022).

[33] B. Mussard, E. Coccia, R. Assaraf, M. Otten, C. J. Umrigar and J. Toulouse, in *Novel Electronic Structure Theory: General Innovations and Strongly Correlated Systems*, edited by P. E. Hoggan (Academic Press, 2018), Advances in Quantum Chemistry Vol. 76, pp. 255–270.

[34] R. McWeeny, *Methods of Molecular Quantum Mechanics. Second Edition* (Academic Press, London, 1992).

[35] O. Christiansen, P. Jørgensen and C. Hättig, Int. J. Quantum Chem. **68**, 1 (1998).

[36] P. Salek, O. Vahtras, T. Helgaker and H. Ågren, J. Chem. Phys. **117**, 9630 (2002).

[37] T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen and K. Ruud, Chem. Rev. **112**, 543 (2012).

[38] E. Fromager, S. Knecht and H. J. A. Jensen, J. Chem. Phys. **138**, 084101 (2013).

[39] P. Norman, K. Ruud and T. Saue, *Principles and Practices of Molecular Properties* (Wiley, Oxford, 2018).

[40] J. Toulouse, *Introduction to the calculation of molecular properties by response theory*, Lecture notes, Sorbonne Université, Paris, 2018, `hal.science/hal-03934866`.

[41] M. E. Casida, in *Recent Developments of Modern Density Functional Theory*, edited by J. M. Seminario (Elsevier, Amsterdam, 1996), p. 391.

[42] R. Bauernschmitt and R. Ahlrichs, Chem. Phys. Lett. **256**, 454 (1996).

[43] M. Taut, Phys. Rev. A. **48**, 3561 (1993).

[44] H. F. King, Theor. Chem. Acc. **94**, 345 (1996).

[45] T. Busch and G. Huyet, J. Phys. B, **36**, 2553 (2003).

[46] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover Publications, New York, 1983).

Conclusion

Originally, the density-based basis-set correction method has been implemented as an a posteriori correction. It only needed a density for the complementary functional and permitted corrections to energies and differences of energies. In this thesis, we have extended the basis-set correction method to other molecular properties.

This work started with the implementation of the self-consistent equation that includes the basis-set correction functional in the optimization of the wave function. In that way, we have access to first-order response molecular properties through expectation values. The dipole moment is the molecular property we chose for a first application of the self-consistent basis-set correction method. We also worked on a non-self consistent procedure for the correction of the coupled-cluster dipole moment using finite differences of energies.

In order to better understand the basis-set correction method, we developed a one-dimensional (1D) model. The latter model helped in the development of a local-density approximation (LDA) for the basis-set correction functional. We also used a similar 1D model for testing the extension of the basis-set correction method to a linear-response formalism.

The work on the implementation of the self-consistent equation and its application to the dipole moment is exposed in Chap. 2. Using expectation values, we computed dipole moments using the eigenvectors of a self-consistent effective Schrödinger equation. Results showed improved basis convergences of the dipole moments in comparison with standard methods.

We also used the implementation to compute ground-state energies with the self-consistent procedure. Due to the comparable results with the a-posteriori approximation, we conclude that the a-posteriori correction is valid for the ground-state energy. Therefore one can access a viable correction at the cost of a single density-functional calculation.

The last conclusion made us wonder if a similar procedure is valid for the dipole moment. In Chap. 3, we show how to compute a correction to the dipole moment using a finite difference of short-range correlation energies at a given density. In addition, we convinced ourselves that corrections computed with the Hartree-Fock density were comparable with the use of the full-configuration-interaction density. The latter conclusion allows the use of a correction at the cost of a Hartree-Fock calculation which can be applied to any correlated method. By getting rid of self-consistent methods, dipole moments can be computed using coupled-cluster methods and the Hartree-Fock-based correction provides enhanced basis-set convergences.

With the purpose to better understand the density-based basis-set correction method and the approximations we used, we developed in Chap. 4 a 1D model with two electrons interacting through a Dirac-delta interaction. Using the latter interaction, we recover systematic and slow convergences of ground-state energies with the number of Hermite-Gauss basis functions. This model led us to quantities that are easily computable and manipulable.

The next step was to develop a LDA-based complementary functional without any range-separated density-functional theory flavor. This development required understanding the behavior of the uniform electron gas model in the presence of the basis-set correction potential which is not invariant under translation.

Finally, in Chap. 5, we used a similar one-dimensional model for the extension of the basis-set correction method to molecular properties computed with the linear-response formalism. In that case, the basis-set correction enters through the second-order derivative of the complementary functional. Our tests on the model system showed that the excitation energies (eigenvalues of the linear-response equation) are not improved using the LDA-based density functional but the basis convergences of the excited-state total energies are accelarated.

Finally, these three years permitted to lay the foundations of what we hope to be a significant field of development: the extension of the density-based basis-set correction method to molecular properties. The logical continuation of this work would be to understand and validate (or not) the use of the second-order derivative of the short-range density functional and the on-top pair density for molecular properties of three-dimensional (3D) systems. This work should start with the extension of the 1D linear-response basis-set correction formalism to 3D systems.

On the other hand, we aim at extending the 1D complementary functional, free of range-separated density functionals, to 3D systems. To this end, we will need to define the properties of the 3D uniform electron gas and build a process to generate a basis-set correction functional that will depend on the basis set.

Appendices

A

The impact of diffuse functions on the basis-set correction of dipole moment

This appendix is related to Chap. 3. In Tab. C.13, we notice that the basis-set correction does not improve the basis convergence of the LiH dipole moment. Here, we study the impact of diffuse functions on the basis-set convergence of the LiH dipole moment. Indeed, the authors in Ref. [1] remind that additional diffuse functions are often needed for the computation of molecular properties and therefore dipole moments.

In Tab. A.1, we report the convergences of Hartree-Fock and CCSD(T) dipole moments with cc-pCVXZ, aug-cc-pCVXZ, and d-aug-cc-pCVXZ basis sets. We notice that Hartree-Fock and CCSD(T) dipole moments are not converged using the cc-pCVQZ basis set. However, the use of d-aug-cc-pCVXZ basis sets does not improve dipole moments regarding their aug-cc-pCVXZ counterparts. We conclude that aug-cc-pCVXZ basis sets already contain enough diffuse functions.

Finally, no major change occurs while adding diffuse functions (for the same X). The latter remark is consistent with the feature of our basis-set correction functional approximation which should target short-range correlation.

For this study, augmented basis sets have been developed using the recipe provided in Ref. [1] to generate extra-diffuse functions for Y-aug-cc-p(C)VXZ basis sets for Y greater than 1. In this framework, even-tempered parameters α and β are respectively defined as the smallest exponent and the ratio between the two most diffuse functions in the aug-cc-p(C)VXZ set. Finally, each added function will have an exponent $\alpha\beta$ for the d-aug-cc-pCVXZ set, $\alpha\beta^2$ for the t-aug-cc-pCVXZ set, $\alpha\beta^3$ for the q-aug-cc-pCVXZ set, and so on.

Table A.1: The second and third columns are the Hartree-Fock and CCSD(T) dipole moments of LiH in atomic units computed from finite differences of energies using cc-pCVXZ, aug-cc-pCVXZ and d-aug-cc-pCVXZ basis sets. The experimental value is 2.3135 a.u.. The last column corresponds to the correction to the dipole moment computed from a finite difference of the PBEUEG short-range correlation energy at the HF density.

Basis/Method	**HF**	**CCSD(T)**	$\bar{d}^{\mathcal{B}}$ (PBEUEG)
cc-pCVDZ	2.3378	2.2495	0.0050
cc-pCVTZ	2.3497	2.2780	0.0015
cc-pCVQZ	2.3580	2.2880	0.0006
CBS		2.2892	
aug-cc-pCVDZ	2.3667	2.3155	0.0041
aug-cc-pCVTZ	2.3609	2.2960	0.0014
aug-cc-pCVQZ	2.3612	2.2935	0.0006
CBS		2.2915	
d-aug-cc-pCVDZ	2.3643	2.3115	0.0037
d-aug-cc-pCVTZ	2.3613	2.2960	0.0014
d-aug-cc-pCVQZ	2.3613	2.2935	0.0006
CBS		2.2917	
t-aug-cc-pCVDZ	-	-	0.0035
t-aug-cc-pCVTZ	-	-	0.0013
t-aug-cc-pCVQZ	-	-	0.0006
CBS			

Bibliography

[1] D. E. Woon and T. H. Dunning, J. Chem. Phys. **100**, 2975 (1994).

B

Linear-response basis-set correction applied to excited-state energies of the berrylium atom

This appendix is related to Chap. 5. We report the preliminary tests of the linear-response basis-set correction method for the berrylium atom. Here, we use the following variant for the multi-determinantal correlation functional

$$\bar{E}^{\mathcal{B}}_{\text{LDA-UEG}}[n] = \int \mathrm{d}\mathbf{r}\, e^{\text{sr},\mu^{\mathcal{B}}(\mathbf{r}),\text{LDA-UEG}}_{\text{c,md}}(n(\mathbf{r}), \zeta(\mathbf{r}), n_2^{\text{UEG}}(\mathbf{r})), \tag{B.1}$$

where

$$e^{\text{sr},\mu,\text{LDA-UEG}}_{\text{c,md}}(n, \zeta, n_2^{\text{UEG}}) = \frac{e_{\text{c}}^{\text{LDA}}(n, \zeta)}{1 + \beta(n, \zeta, n_2^{\text{UEG}})\mu^3}, \tag{B.2}$$

where

$$\beta(n, \zeta, n_2^{\text{UEG}}) = \frac{3}{2\sqrt{\pi}(1 - \sqrt{2})} \frac{e_{\text{c}}^{\text{LDA}}(n, \zeta)}{n_2^{\text{UEG}}}. \tag{B.3}$$

We use this variant because we have access to the second-order derivative of $e_{\text{c}}^{\text{LDA}}(n, \zeta)$ from Ref. [1].

In Fig. B.1, we show the excitation energies to the 2s2p ^{1}P state on the panel a), and to the $2p^2$ ^{1}D state on the panel b). The FCI excitation energies correspond to the energy differences between the FCI eigenvalues associated with each state without the basis-set correction potential. The LR-FCI-LDA (K=0) curve corresponds to the linear-reponse eigenvalues without the effect of the basis-set correction kernel in Eq. 5.34. Finally, the LR-FCI-LDA curve corresponds to the linear-response eigenvalues with the basis-set correction kernel.

For both excitation energies, we notice that the linear-response equations lead to eigenvalues greater than the FCI eigenvalues. Therefore, the linear-response framework does not correct the FCI estimations of the excited states. Moreover, the kernel contribution tends to deteriorate the results as compared to the LR-FCI-LDA (K=0) excitation energies.

In Fig B.2, we show the total energies of the 2s2p ^{1}P state on the panel a), and of the $2p^2$ ^{1}D state on the panel b). For the FCI energy, we report the eigenvalues of the FCI solution. For the linear-response calculations, the ground-state energy is the one from the self-consistent basis-set correction method.

By contrast with the excitation energies, both LR-FCI-LDA (K=0) and LR-FCI-LDA total energies are closer to the experimental value than the FCI energy. Moreover, the LR-FCI-LDA total energies show better convergences than the LR-FCI-LDA (K=0) results.

Finally, this results are similar with those obtained in Chap. 5 : the linear-response exitation energies do not improve the FCI excitation energies, whereas the total energies are improved. However, this results need to be read with some care: we use cc-pVXZ basis sets due to computational costs that forbid us to use aug-cc-pVXZ basis sets (or cc-pCVXZ). Therefore, this may limit the accuracy.

The preliminary results point to the need of improving the linear-response basis-set correction method by implementing more accurate basis-set correction functionals, in particular depending on the on-top pair density.

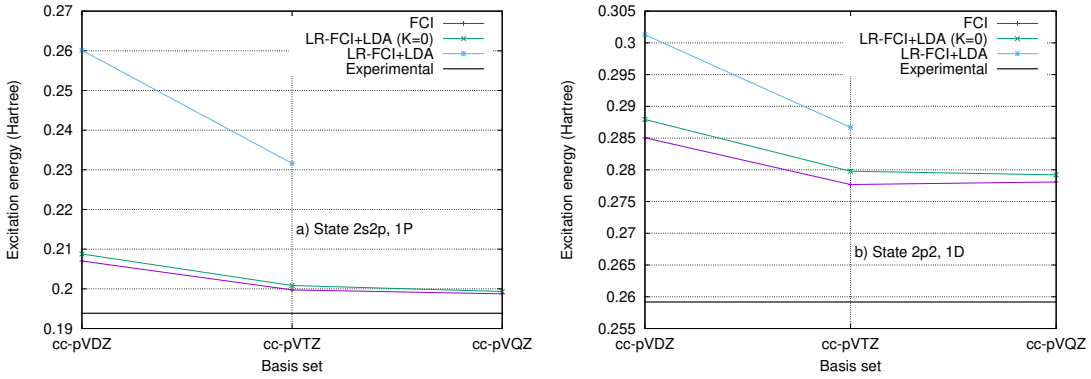

Figure B.1: Excitation energies from the $2s^2 2\ ^1S$ ground state to the a) singlet 2s2p ^1P and b) singlet $2p^2\ ^1D$ states. The experimental values are taken from Ref. [2].

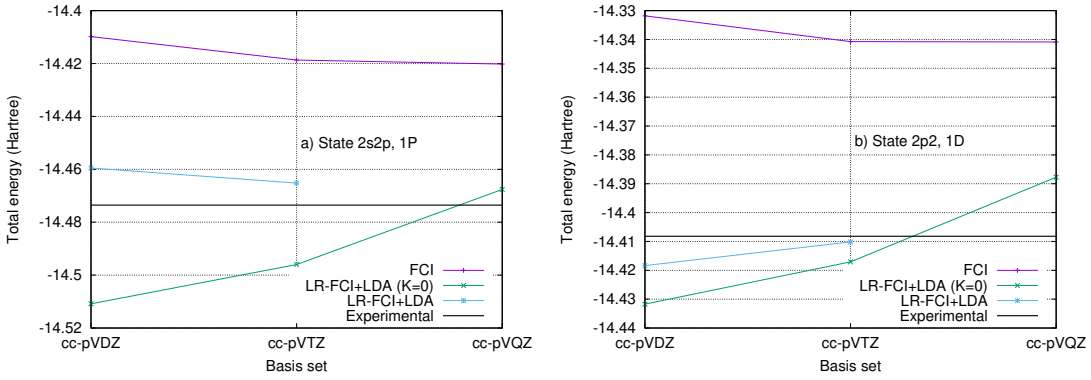

Figure B.2: Total energies of the a) singlet 2s2p ^1P and b) singlet $2p^2\ ^1D$ states. The experimental values are computed from the excitation energies in Ref. [2] and a reference ground-sate energy of -14.66737 Ha from Ref [3].

Bibliography

[1] E. Rebolini, A. Savin and J. Toulouse, Mol. Phys. **111**, 1219 (2013).

[2] A. Kramida, Yu. Ralchenko, J. Reader and and NIST ASD Team, NIST Atomic Spectra Database (ver. 5.10), [Online]. Available: `https://physics.nist.gov/asd` [2023, May 2]. National Institute of Standards and Technology, Gaithersburg, MD. (2022).

[3] K. A. Kabir and A. Halder, in *AIP Conference Proceedings*, AIP Publishing LLC (2015), vol. 1660, p. 090044.

C

Supplementary material: Basis-set correction for coupled-cluster estimation of dipole moments

Geometries

Table C.1: Input xyz geometries (in angström) and their references.

Molecules			
CO [1]			
C	0.0000	0.0000	0.0000
O	0.0000	0.0000	1.1282
BeH [1]			
Be	0.0000	0.0000	0.0000
H	0.0000	0.0000	1.3426
BF [1]			
B	0.0000	0.0000	0.0000
F	0.00000	0.0000	1.2669
BH [2]			
B	0.0000	0.0000	0.0000
H	0.0000	0.0000	1.2324
CH [1]			
C	0.0000	0.0000	0.0000
H	0.0000	0.0000	1.1199
NH [1]			
N	0.0000	0.0000	0.0000
H	0.0000	0.0000	1.0362
CH_2 (singlet) [1]			
C	0.0000	-0.0000	0.1734
H	0.0000	-0.8623	-0.5202
H	0.0000	-0.8623	-0.5202
FH [2]			

F	0.0000	0.0000	0.0000
H	0.0000	0.0000	0.9170
H_2O [1]			
O	0.0000	0.0000	0.1173
H	0.0000	0.7572	-0.4692
H	0.0000	-0.7572	-0.4692
BN [1]			
B	0.0000	0.0000	0.0000
N	0.0000	0.0000	1.3250
BO [1]			
B	0.0000	0.0000	0.0000
O	0.0000	0.0000	1.2045
LiH [1]			
Li	0.0000	0.0000	0.0000
H	0.0000	0.0000	1.5949
LiF [1]			
Li	0.0000	0.0000	0.0000
F	0.0000	0.0000	1.5639
LiN [1]			
Li	0.0000	0.0000	0.0000
N	0.0000	0.0000	1.8690

Detailed results for the molecules

Table C.2: CO

Basis/Method	**HF**	**CCSD(T)**	**CCSD(T)+PBEUEG**
aug-cc-pVDZ	-0.10199	0.05550	0.04398
aug-cc-pVTZ	-0.10499	0.05000	0.04414
aug-cc-pVQZ	-0.10433	0.04600	0.04273
aug-cc-pV5Z	-0.10421	0.04500	0.04360
CBS		0.04485	

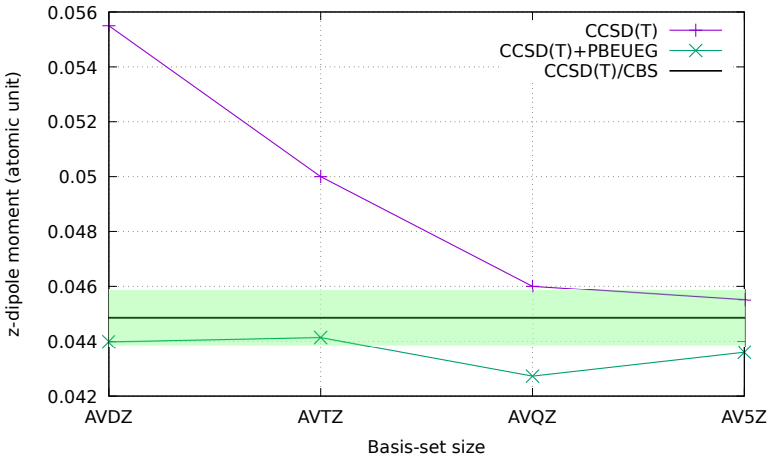

Figure C.1: CO

Table C.3: BeH

Basis/Method	**ROHF**	**ROCCSD(T)**	**ROCCSD(T)+PBEUEG**
aug-cc-pVDZ	0.11017	0.09550	0.08416
aug-cc-pVTZ	0.11076	0.09100	0.08746
aug-cc-pVQZ	0.11199	0.09050	0.08941
aug-cc-pV5Z	0.11218	0.09050	0.08980
CBS		0.09030	

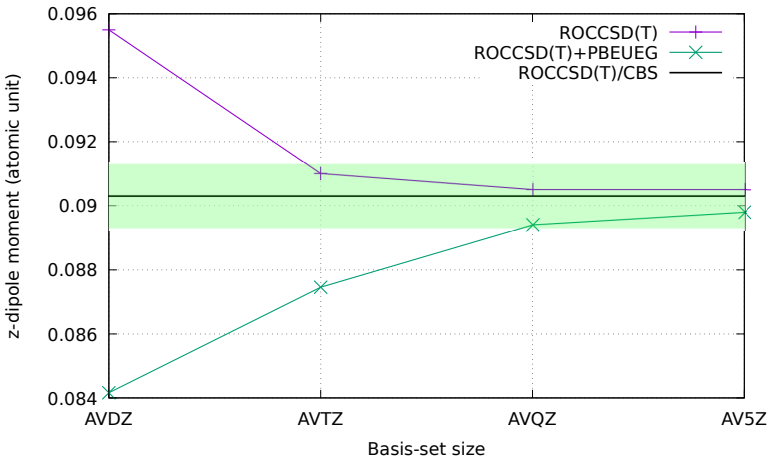

Figure C.2: BeH

Table C.4: BF

Basis/Method	**HF**	**CCSD(T)**	**CCSD(T)+PBEUEG**
aug-cc-pVDZ	0.34436	0.34100	0.33287
aug-cc-pVTZ	0.33390	0.32700	0.32351
aug-cc-pVQZ	0.33314	0.32300	0.32082
aug-cc-pV5Z	0.33328	0.32200	0.32068
CBS		0.32081	

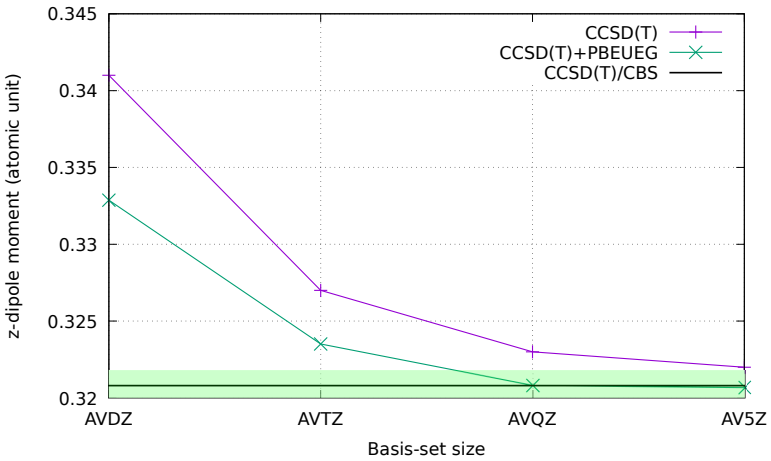

Figure C.3: BF

Table C.5: BH

Basis/ Method	**HF**	**CCSD(T)**	**CCSD(T)+PBEUEG**
aug-cc-pVDZ	0.68796	0.52950	0.54162
aug-cc-pVTZ	0.68649	0.54500	0.55002
aug-cc-pVQZ	0.68493	0.54750	0.54986
aug-cc-pV5Z	0.68496	0.54850	0.54980
CBS		0.54953	

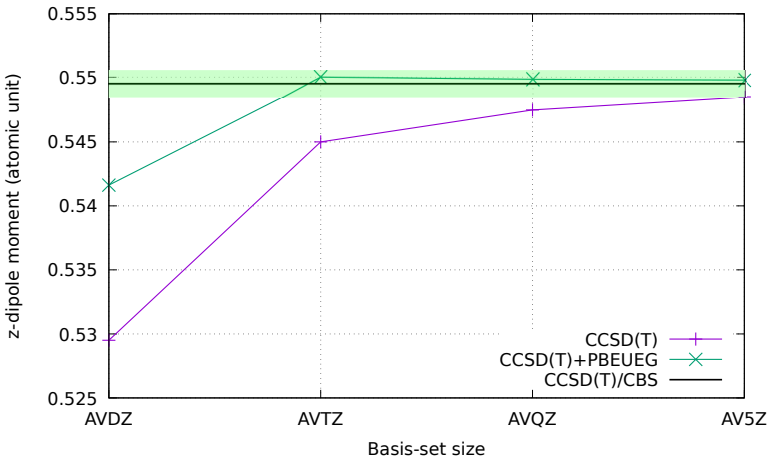

Figure C.4: BH

Table C.6: CH

Basis/Method	**ROHF**	**ROCCSD(T)**	**ROCCSD(T)+PBEUEG**
aug-cc-pVDZ	0.62348	0.54150	0.55427
aug-cc-pVTZ	0.62000	0.54950	0.55481
aug-cc-pVQZ	0.61871	0.55150	0.55405
aug-cc-pV5Z	0.61858	0.55250	0.55396
CBS		0.55368	

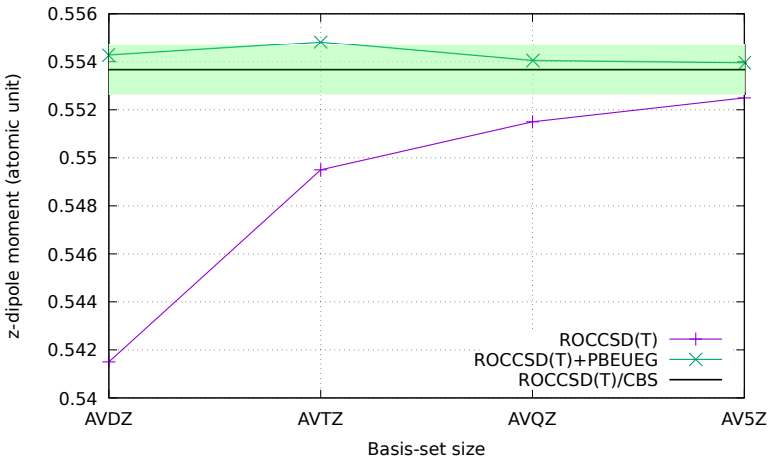

Figure C.5: CH

Table C.7: NH

Basis/Method	**HF**	**CCSD(T)**	**CCSD(T)+PBEUEG**
aug-cc-pVDZ	0.63850	0.59350	0.60792
aug-cc-pVTZ	0.63505	0.59950	0.60519
aug-cc-pVQZ	0.63381	0.60200	0.60464
aug-cc-pV5Z	0.63384	0.60350	0.60506
CBS		0.60504	

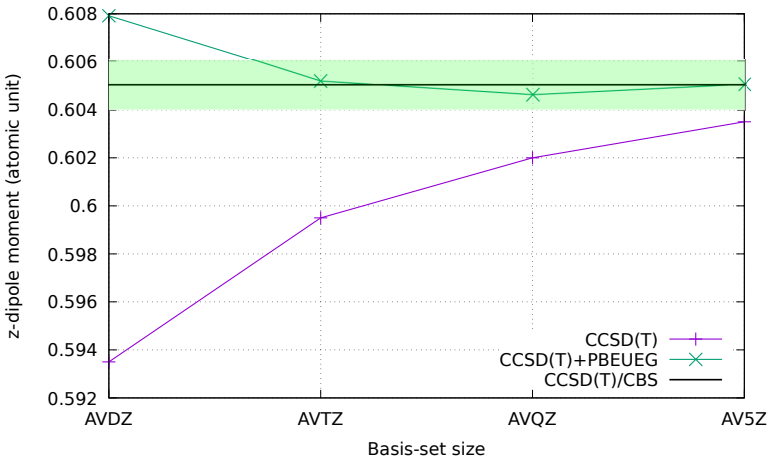

Figure C.6: NH

Table C.8: CH_2

Basis/Method	**HF**	**CCSD(T)**	**CCSD(T)+PBEUEG**
aug-cc-pVDZ	0.74877	0.65600	0.66666
aug-cc-pVTZ	0.74477	0.66000	0.66455
aug-cc-pVQZ	0.74355	0.66200	0.66420
aug-cc-pV5Z	0.74353	0.66350	0.66478
CBS		0.66510	

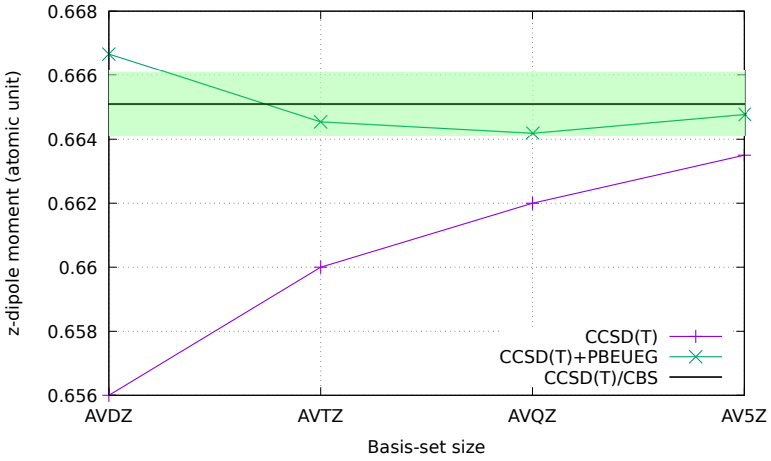

Figure C.7: CH_2

Table C.9: FH

Basis/Method	**HF**	**CCSD(T)**	**CCSD(T)+PBEUEG**
aug-cc-pVDZ	0.75976	0.70350	0.71371
aug-cc-pVTZ	0.75751	0.70450	0.70903
aug-cc-pVQZ	0.75634	0.70700	0.70946
aug-cc-pV5Z	0.75617	0.70750	0.70900
CBS		0.70820	

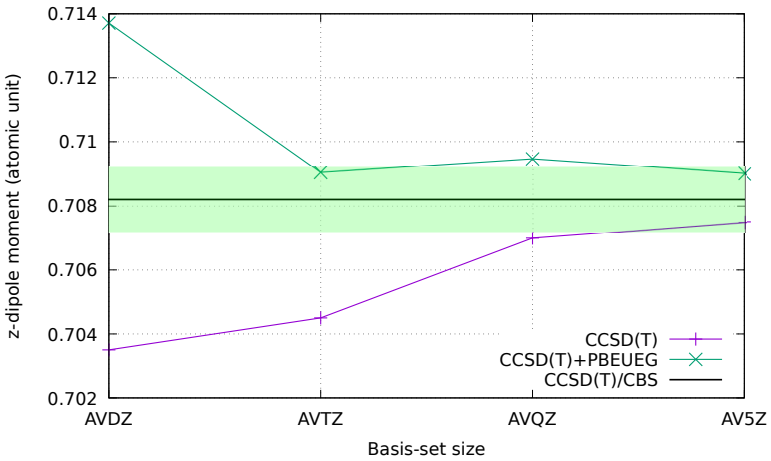

Figure C.8: FH

Table C.10: H_2O

Basis/Method	**HF**	**CCSD(T)**	**CCSD(T)+PBEUEG**
aug-cc-pVDZ	0.78671	0.72700	0.73891
aug-cc-pVTZ	0.78039	0.72400	0.72930
aug-cc-pVQZ	0.77956	0.72650	0.72912
aug-cc-pV5Z	0.77956	0.72800	0.72920
CBS		0.72957	

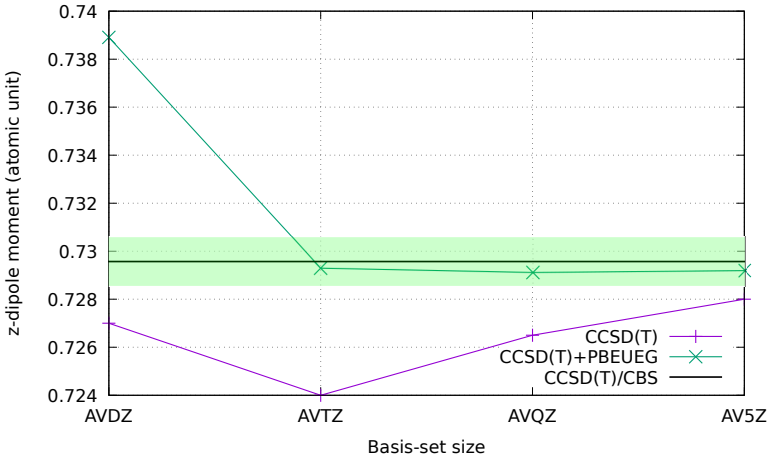

Figure C.9: H_2O

Table C.11: BN

Basis/Method	**ROHF**	**ROCCSD(T)**	**ROCCSD(T)+PBEUEG**
aug-cc-pVDZ	1.13451	0.76250	0.77517
aug-cc-pVTZ	1.13862	0.77550	0.78145
aug-cc-pVQZ	1.13831	0.78400	0.78756
aug-cc-pV5Z	1.13840	0.78650	0.78846
CBS		0.78902	

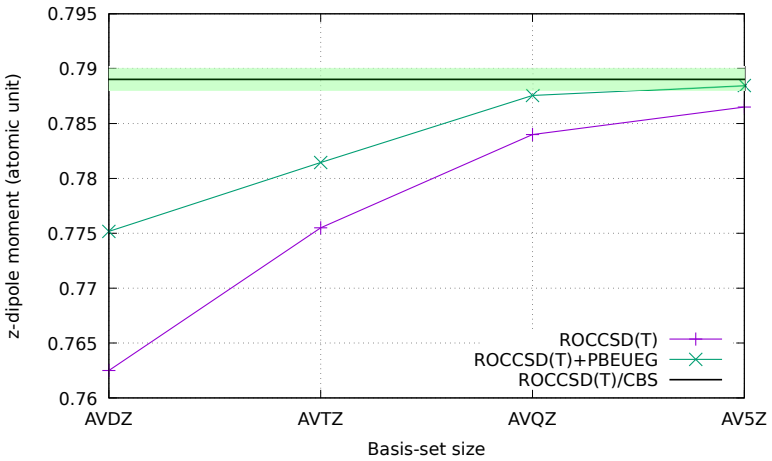

Figure C.10: BN

Table C.12: BO

Basis/Method	**ROHF**	**ROCCSD(T)**	**ROCCSD(T)+PBEUEG**
aug-cc-pVDZ	1.17803	0.88300	0.89417
aug-cc-pVTZ	1.18533	0.89550	0.90153
aug-cc-pVQZ	1.18527	0.90250	0.90622
aug-cc-pV5Z	1.18539	0.90450	0.90698
CBS		0.90647	

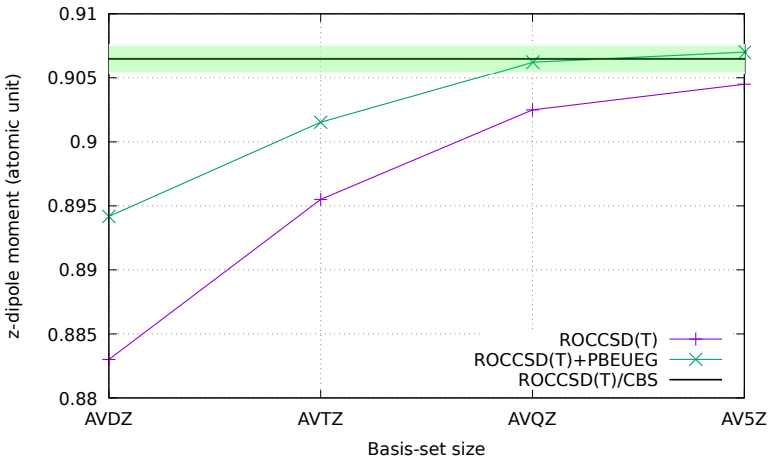

Figure C.11: BO

Table C.13: LiH

Basis/Method	**HF**	**CCSD(T)**	**CCSD(T)+PBEUEG**
aug-cc-pVDZ	2.37055	2.32500	2.32501
aug-cc-pVTZ	2.36235	2.31000	2.30965
aug-cc-pVQZ	2.36153	2.30800	2.30795
aug-cc-pV5Z	2.36129	2.30800	2.30802
CBS		2.30825	

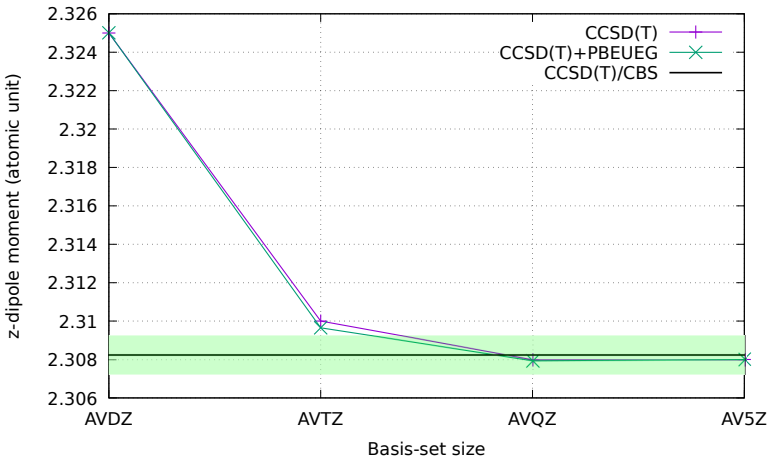

Figure C.12: LiH

Table C.14: LiF

Basis/Method	**HF**	**CCSD(T)**	**CCSD(T)+PBEUEG**
aug-cc-pVDZ	2.56111	2.50400	2.50942
aug-cc-pVTZ	2.54103	2.48300	2.48542
aug-cc-pVQZ	2.53949	2.48250	2.48367
aug-cc-pV5Z	2.53905	2.48250	2.48321
CBS		2.48297	

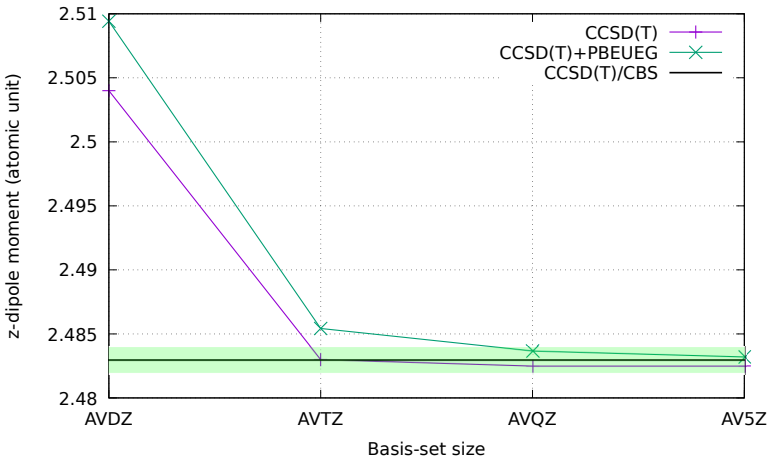

Figure C.13: LiF

Table C.15: LiN

Basis/Method	**ROHF**	**ROCCSD(T)**	**ROCCSD(T)+PBEUEG**
aug-cc-pVDZ	2.90309	2.74200	2.75215
aug-cc-pVTZ	2.90379	2.77300	2.77714
aug-cc-pVQZ	2.90372	2.78250	2.78464
aug-cc-pV5Z	2.90317	2.78450	2.78583
CBS		2.78718	

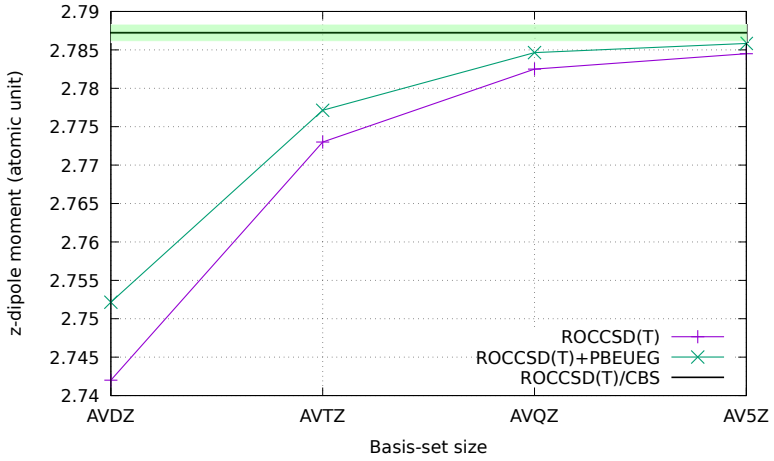

Figure C.14: LiN

Table of errors

Table C.16: Errors in dipole moments (in atomic units) with respect to CCSD(T)/CBS for frozen-core calculations for the entire set of molecules. For each Dunning basis set (AVXZ=aug-cc-pVXZ, $X \in$\{D,T,Q,5\}), the first column contains the errors of the CCSD(T) dipole moments and the second column contains the errors after adding the PBEUEG basis-set correction.

	AVDZ		AVTZ		AVQZ		AV5Z	
	CCSD(T)	+PBEUEG	CCSD(T)	+PBEUEG	CCSD(T)	+PBEUEG	CCSD(T)	+PBEUEG
CO	0.01065	-0.00087	0.00515	-0.00071	0.00115	-0.00212	0.00065	-0.00125
BeH	0.00520	-0.00614	0.00070	-0.00284	0.00020	-0.00089	0.00020	-0.00050
BF	0.02019	0.01207	0.00619	0.00270	0.00219	0.00001	0.00119	-0.00013
BH	0.02003	0.00791	0.00453	-0.00049	0.00203	-0.00033	0.00103	-0.00027
CH	0.01218	-0.00059	0.00418	-0.00112	0.00218	-0.00036	0.00118	-0.00027
NH	0.01154	-0.00288	0.00554	-0.00015	0.00304	0.00041	0.00154	-0.00002
CH_2 (singlet)	0.00910	-0.00156	0.00510	0.00055	0.00310	0.00090	0.00160	0.00032
FH	0.00470	-0.00551	0.00370	-0.00083	0.00120	-0.00126	0.00070	-0.00080
H_2O	0.00257	-0.00934	0.00557	0.00027	0.00307	0.00045	0.00157	0.00037
BN	0.02652	0.013855	0.01352	0.00757	0.00502	0.00146	0.00252	0.00056
BO	0.02347	0.01230	0.01097	0.00494	0.00397	0.00025	0.00197	-0.00051
LiH	0.01675	0.01676	0.00175	0.00140	-0.00025	-0.00030	-0.00025	-0.00023
LiF	-0.02103	-0.02645	-0.00003	-0.00246	0.00047	-0.00070	0.00047	-0.00024
LiN	0.04518	0.03504	0.01418	0.01004	0.00468	0.00254	0.00268	0.00136

Table C.17: Relative errors (%) in dipole moments with respect to CCSD(T)/CBS for frozen-core calculations for the entire set of molecules. For each Dunning basis set (AVXZ=aug-cc-pVXZ, $X \in$\{D,T,Q,5\}), the first column contains the errors of the CCSD(T) dipole moments and the second column contains the errors after adding the PBEUEG basis-set correction.

	AVDZ		AVTZ		AVQZ		AV5Z	
	CCSD(T)	+PBEUEG	CCSD(T)	+PBEUEG	CCSD(T)	+PBEUEG	CCSD(T)	+PBEUEG
CO	23.75	-1.95	11.48	-1.59	2.57	-4.74	1.45	-2.80
BeH	5.76	-6.79	0.78	-3.14	0.22	-0.99	0.22	-0.56
BF	6.29	3.76	1.93	0.84	0.68	0.00	0.37	-0.04
BH	-3.64	-1.44	-0.82	0.09	-0.37	0.06	-0.19	0.05
CH	-2.20	0.11	-0.76	0.20	-0.39	0.07	-0.21	0.05
NH	-1.91	0.48	-0.92	0.02	-0.50	-0.07	-0.25	0.00
CH_2 (singlet)	-1.37	0.23	-0.77	-0.08	-0.47	-0.14	-0.24	-0.05
FH	-0.66	0.78	-0.52	0.12	-0.17	0.18	-0.10	0.11
H_2O	-0.35	1.28	-0.76	-0.04	-0.42	-0.06	-0.22	-0.05
BN	-3.36	-1.76	-1.71	-0.96	-0.64	-0.19	-0.32	-0.07
BO	-2.59	-1.36	-1.21	-0.55	-0.44	-0.03	-0.22	0.06
LiH	0.73	0.73	0.08	0.06	-0.01	-0.01	-0.01	-0.01
LiF	0.85	1.07	0.00	0.10	-0.02	0.03	-0.02	0.01
LiN	-1.62	-1.26	-0.51	-0.36	-0.17	-0.09	-0.10	-0.05

Bibliography

[1] D. Hait and M. Head-Gordon, J. Chem. Theory. Comput. **14**, 1969 (2018).

[2] A. Halkier, W. Klopper, T. Helgaker and P. Jørgensen, J. Chem. Phys. **111**, 4424 (1999).

D

Supplementary material: Basis-set correction based on density-functional theory: Linear-response formalism for excited-state energies

For each basis set $\mathcal{B}$ of size $n_{\max}$, the LDA basis-set correction energy per particle is fitted to a rational fraction

$$\bar{\varepsilon}^{\mathcal{B}}(\rho) \approx \frac{\sum_{i=0}^{4} a_i^{\mathcal{B}} \rho^i}{1 + \sum_{j=1}^{4} b_j^{\mathcal{B}} \rho^j}, \tag{D.1}$$

where the coefficients are given in Tab. D.1.

Table D.1: Coefficients for the rational faction of Eq. (D.1). The coefficients have been found using Mathematica `NonlinearModelFit` function.

$n_{\max}$	$a_0^{\mathcal{B}}$	$a_1^{\mathcal{B}}$	$a_2^{\mathcal{B}}$	$a_3^{\mathcal{B}}$	$a_4^{\mathcal{B}}$	$b_1^{\mathcal{B}}$	$b_2^{\mathcal{B}}$	$b_3^{\mathcal{B}}$	$b_4^{\mathcal{B}}$
1	-0.000071	-0.198377	-4.150011	14.507548	-21.949820	20.296972	139.231729	-321.276975	536.780361
2	-0.000594	-0.173014	0.384433	-0.244367	-0.365982	5.401147	-7.446688	6.256052	8.969527
3	0.000444	-0.330837	0.297792	0.012509	-0.652327	24.748701	-18.057894	4.452590	15.604104
4	0.000795	-0.483374	0.382824	-0.346615	-0.485469	50.965472	-37.484558	17.394304	11.199767
5	0.000879	-0.637201	0.012849	0.251293	-0.793912	82.954134	-50.063365	9.142699	18.248454
10	-0.005058	0.025411	-0.076797	0.085669	-0.034448	-2.385101	3.352381	-2.446328	0.874405
20	-0.003546	0.018777	-0.073303	0.062172	-0.018124	-0.362275	1.663291	-1.437415	0.436582
30	-0.001524	-0.001298	-0.003401	0.004565	-0.001330	-1.064542	0.542581	-0.183654	0.036448
40	-0.001315	-0.001846	-0.000559	0.001925	-0.000579	-1.045314	0.447844	-0.113278	0.017638
50	-0.001259	-0.001685	0.000160	0.001096	-0.000349	-1.036763	0.413128	-0.088704	0.011676
60	-0.001212	-0.001615	0.000544	0.000651	-0.000229	-1.015071	0.382952	-0.073090	0.008439
70	-0.002015	0.026366	-0.069143	-0.167102	-0.010058	-10.127259	52.123001	-7.934000	0.982991