IMPROVEMENT OF APPROXIMATIONS WITH SMOOTH MODEL POTENTIALS

SYSTEMATIC IMPROVEMENT OF APPROXIMATIONS WITH SMOOTH MODELS OF THE COULOMB POTENTIAL

BY

CRISTINA E. GONZÁLEZ ESPINOZA, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF CHEMISTRY & CHEMICAL BIOLOGY

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

© Copyright by Cristina E. González Espinoza, October 2018

All Rights Reserved

Doctor of Philosophy (2018) McMaster University

(Chemistry & Chemical Biology) Hamilton, Ontario, Canada

TITLE:	SYSTEMATIC IMPROVEMENT OF APPROXIMATIONS WITH SMOOTH MODELS OF THE COULOMB POTENTIAL
AUTHOR:	Cristina E. González Espinoza B.Sc., (Science - Chemistry) UAEM, Cuernavaca, Morelos, México
SUPERVISORS:	Dr. Paul W. Ayers Dr. Andreas Savin
NUMBER OF PAGES:	xxii, 149

To my parents, for making the path of life easy to travel.

Abstract

Orbital-based methods for electronic-structure calculations are limited to atoms or molecules with up to about 50 electrons. This limitation comes from the requirement of a long expansion in basis functions to approximate correctly the wave function. Replacing the Coulomb interaction with a smooth model potential has two main consequences: first, the wave function becomes cuspless and the expansion in basis functions converges more rapidly, and second, the smooth potential describes a *weaker* interaction at $r \to 0$, which leads to the loss of accuracy.

This thesis explores whether one can construct models with smooth, non-singular, potentials, but without compromising accuracy. The key idea is to use extrapolation procedures to predict the energy for the Coulomb interaction from a sequence of (cheaper) calculations for smooth potentials.

First, inspired in the works of Savin et al. and Sirbu and King, we constructed an *erfgau* potential to reproduce the spectra of the Hydrogen atom. It was proved to improve the performance of the standard $\mathrm{erf}(\mu r)/r$ potential, not only for Hydrogen, but also for the repulsive interaction in Harmonium and the uniform electron gas (UEG). We then considered correcting the erfgau model in two ways, with short-range potentials and a first-order perturbation theory. While the short-range correction improves the $\mu \to 0$ limit, the perturbation correction accelerates the convergence of

the energy errors as μ increases. By replacing the Coulomb electron-electron interaction with a smooth potential, using the semi-stochastic heat-bath configuration interaction method (SHCI) to select key configurations, and extrapolating to the limiting (non-smoothed) Coulomb potential, we were able to retain the accuracy of full configuration interaction (FCI) calculations, at reduced computational cost.

Acknowledgements

First of all I would like to thank the financial support from CONACYT and the Secretariat of Science and Technology of the State of Morelos for the graduate scholarship. I am also thankful to NSERC and Compute Canada for the computational resources.

Big thanks to the Ayers lab for the countless memories and great experiences we shared. Thank you Ton, Rogelio, Peter, Pawel, Kasia, Marco, Ali, Stijn, Carlos, Ahmed, (big) Paul, Farnaz, Matt, David, Sophie, Caitlin, Derrick, Yilin, Anand, Kumru, Michael, Gaby, Jen, Laurent, Xiao, every one of the amazing, incredibly smart and talented people I crossed paths with during my time in the research group, thank you for sharing your ideas and advice with me, I learned so much from you all.

A very special and deserved thanks to Ana, who suffered (and enjoyed?) as much or even more than me this final step in my life as a graduate student. Thank you for the patience and understanding, and for always being there.

I would like to thank Andreas Savin, for teaching me what is important in life: how to make good ice cream. Seriously, there is no day I don't learn something from such a humble and very knowledgeable person. Thank you for all your help.

Finally, to the person with a soft heart and a brilliant mind, my supervisor Paul Ayers, for all your support, for your advice, your patience, for being ambitious and never giving up on us.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The Schrödinger Equation

Although mathematical expressions and complicated-looking formulas recur throughout this thesis, the main subject is nonetheless *Chemistry* in its most fundamental form: the interaction of electrons with atomic nuclei and, especially, each other. The study of electrons' interactions is outside the scope of classical mechanics, and one needs to invoke quantum mechanical methods. In quantum mechanics, measurable properties, also called *observables*, are mathematically expressed as Hermitian operators; the eigenvalues of these operators define the result of a single measurement. The energy operator is the Hamiltonian operator, $\hat{H}$; the corresponding eigenvalue problem is the fundamental equation in nonrelativistic quantum mechanics: the Schrödinger equation. In its time-independent form, the Schrödinger equation reads

$$\hat{H}\Psi(\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N, \mathbf{R}_1, \mathbf{R}_2, ...\mathbf{R}_M) = E\Psi(\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N, \mathbf{R}_1, \mathbf{R}_2, ...\mathbf{R}_M) \qquad (1.1)$$

The Hamiltonian comprises all the operators for the kinetic and potential energy of the particles in the system. For a system with N electrons and M nuclei, the Hamiltonian in atomic units is:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{A=1}^{M}\frac{1}{M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} \quad (1.2)$$

here, M_A is the mass of the nucleus A divided by the mass of the electron, Z_A is the atomic number of nucleus A (same goes for B), and $R_{AB} = |R_A - R_B|$, $r_{iA} = |r_i - R_A|$ and $r_{ij} = |r_i - r_j|$ are the distance between nuclei A and B, between electron i and nucleus A, and between electrons i and j, respectively. The first two terms in equation 1.2 correspond to the kinetic energy of the electrons and of the nuclei. The Laplacian operator ∇^2 is the sum of the second partial derivatives with respect to each Cartesian coordinate:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (1.3)$$

The last three terms in equation 1.2 are the potential energy operators, specifically: the attraction between electrons and nuclei, the repulsion between nuclei, and the repulsion between electrons.

To determine the energy of the system one needs to solve the Schrödinger equation. Unfortunately, solving the Schrödinger equation for more than two particles is rarely tractable. Approximations to reduce the number of variables are required.

1.1.1 The Born-Oppenheimer approximation

Because the nuclei are orders of magnitude heavier than electrons their movement is much slower and, from the electron's perspective, they can be considered *fixed*.

Therefore, one can describe the motion of the electrons and the nuclei separately:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N, \mathbf{R}_1, \mathbf{R}_2, ...\mathbf{R}_M) = \Psi_{elec}(\{\mathbf{r}_i\})\Phi_{nuc}(\{\mathbf{R}_A\}) \tag{1.4}$$

Focusing on the electrons alone, the repulsion potential between the nuclei is constant. Furthermore, neglecting the kinetic energy of the nuclei leads to the Born-Oppenheimer approximation. The resulting operator for the electronic energy is the electronic Hamiltonian,

$$\hat{H}_{elec} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} \tag{1.5}$$

so the electronic Schrödinger equation is

$$\hat{H}_{elec}\Psi_{elec}(\{\mathbf{r}_i\}) = E_{elec}\Psi_{elec}(\{\mathbf{r}_i\}) \tag{1.6}$$

The total energy is then the sum of the electronic energy and the repulsion energy of the nuclei

$$E(\{\mathbf{R}_A\}) = E_{elec} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{r_{AB}}. \tag{1.7}$$

If one is interested in the motion of the nuclei, the nuclear Schrödinger equation in the Born-Oppenheimer approximation is

$$\left(-\frac{1}{2}\sum_{A=1}^{M}\frac{1}{M_A}\nabla_A^2 + E(\{\mathbf{R}_A\})\right)\Phi_{nuc}(\{\mathbf{R}_A\}) = E_{tot}\Phi_{nuc}(\{\mathbf{R}_A\}) \tag{1.8}$$

In this work we are interested only in the electronic problem and from now on when we use $\hat{H}$ we are referring to $\hat{H}_{elec}$, and $\Psi(\mathbf{r}) = \Psi(\{\mathbf{r}_i\})$.

1.1.2 Orbital Approximations and the Slater Determinant

In Chemistry, the orbital approximation is surprisingly effective for the qualitative explanation of the reactivity of atoms and molecules. This approximation assumes that electrons are moving independently from each other and occupying their own individual one-electron wave functions, called orbitals, denoted $\psi(\mathbf{r})$. To completely describe an electron we need functions that depend on the position and spin coordinates of the electron. Such functions are called *spin-orbitals* $\chi(\mathbf{x})$, and are constructed by multiplying a spatial function $\psi(\mathbf{r})$ and a spin function $\alpha(\omega)$ or $\beta(\omega)$,

$$\chi(\mathbf{x}) = \begin{cases} \psi(\mathbf{r})\alpha(\omega) \\ \\ \psi(\mathbf{r})\beta(\omega) \end{cases} . \tag{1.9}$$

The Pauli exclusion principle states that only one electron can occupy each one-electron state. This principle is closely related to the antisymmetry principle, which states that "a many-electron wave function must be antisymmetric with respect of the interchange of the coordinate $\mathbf{x}$ for any two electrons" [SO12]. Thus, the wave function not only has to be solution of the Schrödinger equation, it also needs to satisfy the antisymmetry principle. A way to impose antisymmetry is to approximate the wave function with *Slater determinants*.

The Slater determinant for the N-electron case is

$$\Phi_N(\{\mathbf{x}_i\}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \tag{1.10}$$

here, $\frac{1}{\sqrt{N!}}$ is the normalization factor (assuming orthogonal and normalized spin-orbitals). For simplicity we will write the normalized Slater determinant as

$$|\Phi_N\rangle = |\chi_1\, \chi_2\, \cdots\, \chi_N\rangle. \tag{1.11}$$

For a given set of spin-orbitals it is possible to construct many different Slater determinants, so how do we get the right wave function for the ground state? The answer is given by the variational principle.

1.1.3 The Variational Principle

The variational principle states that the expectation value of the Hamiltonian is an upper bound to the exact ground-state energy

$$\langle E\rangle \geq E_0 \tag{1.12}$$

The expectation value is defined as

$$\langle E\rangle = \frac{\langle\Psi(\mathbf{r})|\hat{H}|\Psi(\mathbf{r})\rangle}{\langle\Psi(\mathbf{r})|\Psi(\mathbf{r})\rangle} \tag{1.13}$$

where

$$\langle\Psi(\mathbf{r})|\hat{H}|\Psi(\mathbf{r})\rangle = \int \Psi^*(\mathbf{r})\hat{H}\Psi(\mathbf{r})d\mathbf{r} \tag{1.14}$$

In other words, if $\Psi(\mathbf{r})$ corresponds to the ground state then $\langle E\rangle = E_0$. For any other state $\langle E\rangle > E_0$. This last statement is true for any trial wave function. Therefore, the ground-state energy is obtained from the full minimization of $\langle E\rangle$ with

respect to all (allowed) N-electron wave functions:

$$E_0 = \min_{\Psi} \langle E \rangle \tag{1.15}$$

1.2 Hartree-Fock (HF) Method

In a nutshell, the Hartree-Fock method consists in selecting a set of orbitals to describe a single electron configuration and iteratively adjusting the orbitals closer to the ground-state orbitals. More formally speaking, this method results from approximating the N-electron wavefunction with a single Slater determinant, and optimizing the orbitals that constitute the determinant using the variational principle. This optimization is a minimization subject to the orthonormalization constraint for the orbitals:

$$\langle \phi_k | \phi_l \rangle - \delta_{kl} = 0 \tag{1.16}$$

The constrained minimization is solved by the method of Lagrange multipliers. The minimum is found when the Lagrangian is stationary,

$$\delta \left(\langle \Psi(\mathbf{r}) | \hat{H} | \Psi(\mathbf{r}) \rangle - \sum_{kl} \epsilon_{kl} \left(\langle \phi_k | \phi_l \rangle - \delta_{kl} \right) \right) = 0 \tag{1.17}$$

Having a trial wave function, one can manipulate the equation above by separation of variables to obtain equations for each one-electron state (spin-orbital):

$$\hat{f} \chi_i(\mathbf{x}) = \epsilon_i \chi_i(\mathbf{x}) \tag{1.18}$$

These are the Hartree-Fock equations, where $\hat{f}$ is the Fock operator,

$$\hat{f}(\mathbf{r}) = -\frac{1}{2}\nabla_i^2 - \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} + \sum_j \hat{J}_j(\mathbf{r}) - \hat{K}_j(\mathbf{r}) \tag{1.19}$$

$\hat{J}$ is the Coulomb (Hartree) operator,

$$\hat{J}_j(\mathbf{r}) = \int \frac{\chi_j^*(\mathbf{r}')\chi_j(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \tag{1.20}$$

and $\hat{K}$ is the exchange operator,

$$\hat{K}_j(\mathbf{r})\chi_k(\mathbf{r}) = \int \frac{\chi_j^*(\mathbf{r}')\chi_k(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}\chi_j(\mathbf{r}) \tag{1.21}$$

In practice, the spatial orbitals are expanded in a set of basis functions:

$$\psi_i(\mathbf{r}) = \sum_{k=1}^{L} C_{ki}\phi_k(\mathbf{r}) \tag{1.22}$$

Then, solving for the HF molecular orbitals is equivalent to solving for the expansion coefficients C_{ki}. By substituting 1.22 into equation 1.18 we get the matrix equation:

$$\hat{f}(i)\sum_{k=1}^{L} C_{ki}\phi_k(i) = \epsilon_i \sum_{k=1}^{L} C_{ki}\phi_k(i). \tag{1.23}$$

Now, we multiply by $\phi_j^*(\mathbf{r})$ on the left and integrate over $\mathbf{r}$,

$$\sum_k C_{ki} \int \phi_j^*(\mathbf{r})\hat{f}(\mathbf{r})\phi_k(\mathbf{r})d\mathbf{r} = \epsilon \int \phi_j^*(\mathbf{r})\phi_k(\mathbf{r})d\mathbf{r}, \tag{1.24}$$

where,

$$\int \phi_j^*(\mathbf{r})\hat{f}(\mathbf{r})\phi_k(\mathbf{r})d\mathbf{r} = F_{jk} \tag{1.25}$$

is the matrix representation of the Fock operator, better known as the Fock matrix, and

$$\int \phi_j^*(\mathbf{r})\phi_k(\mathbf{r})d\mathbf{r} = S_{jk} \tag{1.26}$$

is the overlap matrix. Rewriting equation 1.23 we have

$$\sum_k F_{jk}\, C_{ki} = \epsilon_i \sum_k S_{jk}\, C_{ki}, \quad i = 1, 2, ..., L \tag{1.27}$$

These set equations are the Roothan-Hall equations and can be written as a single matrix equation:

$$\mathbf{F}\,\mathbf{C} = \mathbf{S}\,\mathbf{C}\,\epsilon \tag{1.28}$$

This is a generalized (unless S is the identity matrix) nonlinear eigenvalue problem. To solve this equation one has to construct the Fock matrix, but at the same time the Fock matrix depends on the spin orbitals, which one only knows after the equation has been solved. Therefore, equation 1.28 has to be solved iteratively, in a *self-consistent* fashion:

1 Start with some initial guess for the coefficients matrix

2 Generate the Fock matrix from this coefficients

3 Solve the HF equation to get a new coefficient matrix

4 Repeat 2 and 3 until the coefficients converge.

The Hartree-Fock method is a mean-field method, *i.e.* it treats the electron-electron interaction in an averaged manner, and there is no simultaneous description of the electrons. In quantum chemistry, the type of electronic interaction that the Hartree-Fock method fails to describe is called *electron correlation*. The correlation energy is thus defined as the difference between the exact (non-relativistic) energy and the HF energy at the basis set limit, $L \to \infty$,

$$E_{corr} = E_{exact} - E_0^{HF} \tag{1.29}$$

1.3 Configuration Interaction

As each Slater determinant represents one electronic configuration, approximating the N-electron wave function with a single Slater determinant limits the wave function to a single electronic configuration. One way to account for electronic correlation is by allowing more than one electronic configuration to approximate the wave function. In the configuration interaction method (CI) the exact wave function is expressed as a linear combination of N-electron trial functions (Slater determinants). The full CI (FCI) wave function contains all possible excitations from a reference determinant, Φ_0, typically chosen to be the Hartree-Fock ground state,

$$|\Psi_{CI}\rangle = c_0|\Phi_0\rangle + \sum_{\substack{i\leq i\leq N \\ N+1\leq a\leq 2L}} c_i^a|\Phi_i^a\rangle + \sum_{\substack{1\leq i<j\leq N \\ N+1\leq a<b\leq 2L}} c_{ij}^{ab}|\Phi_{ij}^{ab}\rangle + \sum_{\substack{1\leq i<j<k\leq N \\ N+1\leq a<b<c\leq 2L}} c_{ijk}^{abc}|\Phi_{ijk}^{abc}\rangle + \cdots \tag{1.30}$$

where the limits ensure each excitation is counted only once. The CI equation is also solved by the variational method, giving an ordinary linear eigenvalue problem. To

see this, start by writing the eigenvalue equation

$$\hat{H}|\Psi_{CI}\rangle = E|\Psi_{CI}\rangle \tag{1.31}$$

$$\sum_n \hat{H}|\Phi_n\rangle c_n = E\sum_n |\Phi_n\rangle c_n \tag{1.32}$$

then multiply by $\langle\Phi_m|$ on the left

$$\sum_n \langle\Phi_m|\hat{H}|\Phi_n\rangle c_n = E\sum_n \langle\Phi_m|\Phi_n\rangle c_n \tag{1.33}$$

$$\sum_n H_{mn}c_n = Ec_m \tag{1.34}$$

here

$$\langle\Phi_m|\Phi_n\rangle = \delta_{mn} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \tag{1.35}$$

Using 1.35 in 1.34 results in a system of equations

$$\sum_n \langle\Phi_m|\hat{H}|\Phi_n\rangle c_n = Ec_m \tag{1.36}$$

As the number of electrons N, and the number of spin-orbitals, increases, the number of possible configurations increases exponentially and solving the FCI equation becomes intractable very quickly. A way to reduce the computational cost is to truncate the CI expansion and consider only some configurations. Truncated versions of CI are named according to the type of excitations considered. CI singles (CIS) includes only one-electron excitations, CI singles and doubles (CISD) for one and two-electron excitations, and so on. More sophisticated techniques to

choose the best selection of configurations employ a variety of algorithms for sampling the configuration space, including stochastic, semistochastic and deterministic algorithms. For instance, the selected plus perturbation theory methods (SCI+PT) select the determinants in an iterative manner by adding the most important elements of the first-order correction to the wave function, and once convergence is reached a second-order perturbation correction is applied to the final variational wave function [HMR73, BP74, SE16, TLT$^+$16]. Even when the SCI+PT method reduces the number of determinants incorporated into the wave function it is computationally demanding because *all* the determinants connected to a reference Hamiltonian matrix must be sampled. The semistochastic heat-bath configuration interaction (HCI) method [HTU16, SHJ$^+$17, HUS17] solves this problem by introducing a threshold ϵ to restrict the sampled determinants. HCI only generates the determinants connected to the reference determinant by Hamiltonian matrix elements that are larger than ϵ. At the beginning all single excitations and only double excitations for which $|H_{ik}c_i| > \epsilon$ are generated. The algorithm consists in three steps: 1) the selected space starts with the Hartree-Fock determinant, 2) all connected determinants outside the selected space (*i.e.* determinants for which $|H_{ik}c_i| > \epsilon$ for at least one determinant in the selected space) are added, and 3) the two first steps are repeated until the number of new determinants is less than 1% of the previously selected determinants. One can also include a second-order Epstein-Nesbet perturbation theory (EN-PT) energy correction,

$$\Delta E^{(2)} = \sum_k \frac{\left(\sum_i H_{ki}c_i\right)^2}{E^{(0)} - H_{kk}} \tag{1.37}$$

using only the elements for which $|H_{ik}c_i|$ is greater than a second parameter ϵ_2. The two threshold values ϵ_1, ϵ_2 define a trade-off of efficiency and accuracy, prescribing a

systematic way to get to the FCI solution.

In principle, FCI produces the exact solution of the N-electron problem. In the case of a complete basis, one would obtain the exact energies of all electronic states of the system. In practice we use a finite set of basis functions and the accuracy of the energy is dependent on the choice of this basis. Furthermore, it is known that the convergence of the correlation energy with the size of the basis is very slow. For example, with the correlation-consistent basis sets cc-pVXZ, the correlation energy follows the law [HKKN97]

$$E_X^{corr} = E_{X\to\infty}^{corr} + AX^{-3} \tag{1.38}$$

This slow convergence comes from the difficulty of modeling the behavior of the electronic wave function when electrons come close together, $|\mathbf{r}_i - \mathbf{r}_j| \to 0$, using one-electron basis functions [Kat57, CSM79, JZS85, MP79, HPKW09, PWD94, DGY$^+$06]. We will discuss this subject in further detail in section 1.6, but for now we leave the complicated wave function methods aside and introduce an alternative strategy to account for the correlation energy, density functional theory.

1.4 Density Functional Theory (DFT)

In density functional theory (DFT) the energy of the system is defined as a functional of the electron density

$$E = E[\rho(\mathbf{r})] \tag{1.39}$$

The electron density, $\rho(\mathbf{r})$, is the probability of finding an electron in a volume element centered at $\mathbf{r}$,

$$\rho(\mathbf{r}) = \int \cdots \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) \left(\sum_k^N \delta(\mathbf{r}_k - \mathbf{r}) \right) \Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) d\mathbf{r}_2 d\mathbf{r}_3 ... d\mathbf{r}_N \quad (1.40)$$

$$= N \int \cdots \int |\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)|^2 d\mathbf{r}_2 d\mathbf{r}_3 ... d\mathbf{r}_N. \quad (1.41)$$

The ground-state electron density has some important properties:

- it is a physical observable that can be measured with x-ray experiments,
- it is a non-negative function of **only three variables** $\mathbf{r} = \{x, y, z\}$,

$$\rho(\mathbf{r}) \geq 0 \quad (1.42)$$

- it integrates to the number of electrons,

$$\int \rho(\mathbf{r}) d\mathbf{r} = N \quad (1.43)$$

- it has a cusp near the atomic nuclei due to the divergence in the electron-nuclear potential, and the pointiness of this cusp defines the effective nuclear charge.

All of these properties also hold for excited states, but the cusp condition is significantly more complicated. The proof of the existence of an exact energy functional of the electron density and a variational principle for the ground-state energy was provided by the Hohenberg and Kohn theorems [HK64].

1.4.1 Hohenberg-Kohn (HK) Theorems

As mentioned before, the electron density determines the effective nuclear charge, so it determines the effective external potential $v(\mathbf{r})$, that binds the electrons to the system. Also, it integrates to the number of electrons N. Knowing N and $v(\mathbf{r})$ is enough to determine completely the molecular Hamiltonian, and thus, the energy and the wave function of a molecule. Hence, one can determine all the properties of a molecule from its ground-state electron density. Hohenberg and Kohn showed that this argument can in fact be generalized to *any* reasonable choice for the external potential and, therefore for any electronic Hamiltonian of the form,

$$\hat{H}_v = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i=1}^{N} v(\mathbf{r}_i) + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.44}$$

(compare Eq. (1.5).) Specifically, the first Hohenberg-Kohn theorem states that "*the external potential is determined, within a trivial additive constant, by the ground-state electron density*". A proof can be found in [PY94].

Due to the first Hohenberg-Kohn theorem, the energy can be written as a functional of the ground-state electron density. It is convenient to decompose the energy functional as a sum of physically different energy contributions

$$E_v[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] + V_{ext}[\rho(\mathbf{r})] \tag{1.45}$$

where T is the kinetic energy functional, V_{ee} the electron-electron repulsion energy functional, and V_{ext} the external potential energy functional. The form of each of

these terms will be addressed in the following sections. For now we define

$$V_{ext}[\rho(\mathbf{r})] = \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \tag{1.46}$$

and the total energy:

$$E_v[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \tag{1.47}$$

The second Hohenberg-Kohn theorem gives a variational principle for the ground-state density analogous to (1.12) for the wave function,

$$E_0 = E_v[\rho(\mathbf{r})] \leq E_v[\tilde{\rho}(\mathbf{r})] \tag{1.48}$$

for any N-electron trial density $\tilde{\rho}(\mathbf{r})$. If $\tilde{\rho}(\mathbf{r})$ determines $\tilde{v}(\mathbf{r})$ and the wave function $\tilde{\Psi}$, using the variational principle for the wave function, we have

$$E_v[\rho(\mathbf{r})] = \langle \Psi_0 | \hat{H}_v | \Psi_0 \rangle \tag{1.49}$$

$$E_v[\rho(\mathbf{r})] \leq \langle \tilde{\Psi} | \hat{H}_v | \tilde{\Psi} \rangle = E_v[\tilde{\rho}(\mathbf{r})] \tag{1.50}$$

About v-representability

We know that "*every ground-state electron density corresponds to at most one external potential*". In other words, the external potential is a functional of the electron density. In particular, if the density is the ground-state of some potential $v(\mathbf{r})$ it is said to be *v-representable*. This property is needed so that there exists the one-to-one mapping between the electron density and the ground-state wave function, which means that the second HK theorem is only valid for v-representable densities.

However, not every density is v-representable. Furthermore, practical conditions for v-representability are unknown.

1.4.2 Levy-Constrained Search

Many wave functions can give the same density, so one needs to find a way to determine whether a wave function that gives the ground-state density is, in fact, a true ground-state wave function Ψ_0. Let Ψ_{ρ_0} denote an N-electron wave function that has the ground-state density ρ_0. From the variational principle we know that

$$\langle \Psi_{\rho_0} | \hat{H} | \Psi_{\rho_0} \rangle \geq \langle \Psi_0 | \hat{H} | \Psi_0 \rangle = E_0 \tag{1.51}$$

$$\langle \Psi_{\rho_0} | \hat{T} + \hat{V}_{ee} | \Psi_{\rho_0} \rangle \geq \langle \Psi_0 | \hat{T} + \hat{V}_{ee} | \Psi_0 \rangle \tag{1.52}$$

the right-hand side is the universal HK functional,

$$F_{HK}[\rho_0] = \langle \Psi_0 | \hat{T} + \hat{V}_{ee} | \Psi_0 \rangle \tag{1.53}$$

This equation is satisfied when $\Psi_{\rho_0} = \Psi_0$, and is obtained by the minimization:

$$F_{Levy}[\rho_0] = \min_{\Psi \to \rho_0} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle = F_{HK}[\rho_0] \tag{1.54}$$

Notice that this variational search for the wave function is constrained to the wave functions that give the density ρ_0. One can extend (1.54) to densities that are not v-representable but which correspond to at least one N-electron wave function. Such densities are said to be *N-representable*, and the associated extension of (1.54) is the

Levy-constrained search formulation of the universal density functional

$$F_{Levy}[\rho] = \min_{\Psi\to\rho}\langle\Psi|\hat{T}+\hat{V}_{ee}|\Psi\rangle \tag{1.55}$$

$F_{Levy}[\rho_0] = F_{HK}[\rho_0]$ only if ρ_0 is v-representable [Lev79, LP85b, LP85a]. To be N-representable, the density has to be a non-negative, normalized, and have bounded kinetic energy:

$$\rho(\mathbf{r}) \geq 0, \quad \int \rho(\mathbf{r})d\mathbf{r} = N, \quad \int |\nabla\rho(\mathbf{r})^{1/2}|^2 d\mathbf{r} < \infty \tag{1.56}$$

The Levy-constrained search therefore allows us to use a softer condition, *N-representability*, to avoid the *v-representability* problem. The ground-state energy can be obtained by a double minimization

$$E_0 = \min_{\rho}\left(\min_{\Psi\to\rho}\left(\langle\Psi|\hat{T}+\hat{V}_{ee}|\Psi\rangle + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r}\right)\right) \tag{1.57}$$

$$= \min_{\rho} F_{Levy}[\rho] + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r}. \tag{1.58}$$

1.4.3 The Kohn-Sham (KS) method

While exact in principle, practical explicit approximations to $F[\rho]$ are not very accurate, the kinetic energy functional being the most important and difficult piece. The kinetic energy is of the same order of magnitude as the total energy of the system, so even a small 1% error in the kinetic energy leads to chemically useless results. Faced with the difficulty of finding an adequate explicit approximation to the kinetic energy functional, Kohn and Sham (KS) introduced orbitals and defined the kinetic energy as an implicit functional of the density. They did this by approximating the

system of interacting electrons as a system of non-interacting electrons, where the total kinetic energy is the sum of the individual energies:

$$T_s[\rho] = -\frac{1}{2}\sum_{i}^{N}\langle\psi_i|\nabla_i^2|\psi_i\rangle = \min_{\Psi\to\rho}\langle\Psi|\hat{H}|\Psi\rangle. \tag{1.59}$$

The Kohn-Sham (KS) approximation describes the energy as the sum of the individual electrons' contributions:

$$\hat{H}_s = \sum_{i}^{N}\left(-\frac{1}{2}\nabla_i^2 + v_s(\mathbf{r}_i)\right). \tag{1.60}$$

Similar to HF, the solution for the Hamiltonian $\hat{H}_s$ is a Slater determinant $|\Psi_s\rangle$. With the introduction of the orbitals it is possible to separate the terms that can be determined exactly for the non-interacting system, namely the first two terms in the r.h.s. of equation 1.61,

$$E_v[\rho] = T_s[\rho] + U[\rho] + E_{xc}[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{1.61}$$

where $T_s[\rho]$ is the exact kinetic energy functional of the non-interacting system, and $U[\rho]$ is the classical repulsion,

$$U[\rho] = \frac{1}{2}\int\int\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}'-\mathbf{r}|}d\mathbf{r}\,d\mathbf{r}' \tag{1.62}$$

The last term, $E_{xc}[\rho]$, is the exchange-correlation functional: it contains all the extra contributions to the energy due to the interaction between electrons,

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - U[\rho] \tag{1.63}$$

The Euler-Lagrange equation associated with the stationary condition of $E_v[\rho]$ gives:

$$\mu = v_s(\mathbf{r}) + \frac{\delta T_s[\rho]}{\delta \rho} \tag{1.64}$$

Here, μ is the Lagrange multiplier for the normalization constraint, v_s is the Kohn-Sham effective potential,

$$v_s(\mathbf{r}) = v(\mathbf{r}) + \frac{\delta U[\rho]}{\delta \rho} + \frac{\delta E_{xc}[\rho]}{\delta \rho} \tag{1.65}$$

where the Coulomb potential is

$$\frac{\delta U[\rho]}{\delta \rho(\mathbf{r})} = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}',$$

and the exchange-correlation potential v_{xc},

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})}$$

As a result, we have the self-consistent set of equations:

$$\left(-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}) - \epsilon_j \right) \psi_j(\mathbf{r}) = 0 \tag{1.66}$$

with

$$\rho(\mathbf{r}) = \sum_{j=1}^{N} |\psi_j(\mathbf{r})|^2$$

Notice that Kohn-Sham equations, eq. (1.66), are almost the same as the Hartree-Fock equations, except that the nonlocal Hartree-Fock exchange operator, $\hat{K}$, is replaced by the local exchange-correlation potential $v_{xc}(\mathbf{r})$. Yet eq. (1.66) is, in principle

exact.

1.4.4 Density-Functional Approximations (DFA)

The work of Kohn and Sham opened the door to the development of approximations for the exchange-correlation density functional. The most basic approximation, the Local-Density approximation (LDA), defines the exchange-correlation energy density at each point in space as a function of only the electronic density at that point. This is the reason it is called a *local* approximation. The exchange and correlation energy density is typically taken from the uniform electron gas (UEG) model. The exchange-correlation energy is given by

$$E_{xc} = \int \rho(\mathbf{r}) \epsilon_{xc} \left(\rho(\mathbf{r})\right) d\mathbf{r} \tag{1.67}$$

One can separate ϵ_{xc} into exchange and correlation components,

$$\epsilon_{xc} = \epsilon_x + \epsilon_c$$

with

$$\epsilon_x = C_x \rho^{1/3} \quad C_x = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}$$

the exchange energy of the uniform electron gas. The Dirac exchange energy, is then

$$E_x[\rho] = \int \rho(\mathbf{r}) \left(C_x \rho^{1/3}(\mathbf{r})\right) d\mathbf{r} = C_x \int \rho^{4/3}(\mathbf{r}) d\mathbf{r} \tag{1.68}$$

In the spin-polarized version, also called the local spin-density approximation (LSDA),

$$E_x[\rho_\alpha, \rho_\beta] = 2^{1/3} C_x \int \left(\rho_\alpha^{4/3}(\mathbf{r}) + \rho_\beta^{4/3}(\mathbf{r}) \right) d\mathbf{r} \tag{1.69}$$

Sometimes, these equations are written in terms of a spin-polarization function

$$f_s(\zeta) = \frac{1}{2}[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}] \tag{1.70}$$

where

$$\zeta(\mathbf{r}) = \frac{\rho_\alpha(\mathbf{r}) - \rho_\beta(\mathbf{r})}{\rho(\mathbf{r})} \tag{1.71}$$

Then, the exchange energy per particle for the polarized case is

$$\epsilon_x = C_x f_s(\zeta) \rho^{1/3} \tag{1.72}$$

The correlation energy, however, is obtained numerically from quantum Monte Carlo calculations. For those benchmark results used in practice, an analytic interpolation formula is needed. The parametrization of Vosko, Wilk and Nusar(VWN) [VWN80] is among those most widely used by commercial electronic structure packages.

The next "rung of Jacob's ladder of DFT" are the generalized gradient approximations (GGAs), where the exchange-correlation energy density depends not only on the density, but also on the density at nearby points, as inferred from the density gradient $\nabla\rho(\mathbf{r})$

$$E_{xc}[\rho] = \int \rho(\mathbf{r}) \epsilon_{xc} \left(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}) \right) d\mathbf{r} \tag{1.73}$$

The exchange energy per particle has the general form:

$$\epsilon_x^{GGA} = \epsilon_x^{LSDA}\left(1 + \Delta\epsilon_x\left[\frac{|\nabla\rho(\mathbf{r})|}{\rho^{4/3}(\mathbf{r})}\right]\right) \tag{1.74}$$

Popular GGAs include the BLYP [Bec88, LYP85] and PBE [PBE96] functionals.

The methods that depend also on the kinetic energy density $\tau(\mathbf{r})$ are called Meta-GGAs

$$E_{xc}[\rho] = \int \rho(\mathbf{r})\epsilon_{xc}\left(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}), \tau(\mathbf{r})\right) d\mathbf{r} \tag{1.75}$$

Popular meta-GGAs include the BR and TPSS functionals [BR89, TPSS03].

LDAs, GGAs, and meta-GGAs typically underestimate reaction barriers, bond strengths, polarizabilities, and other properties. These same properties are typically overestimated by the Hartree-Fock method. The hybrid functionals are the affirmative answer to the question: "why not combine DFA and HF methods?". Because the exchange energy is given exactly in HF theory, one can include a fraction of the exact exchange into the energy expression and construct a hybrid functional. The most widely used hybrid functional is B3LYP, which is constructed as follows:

$$E_{xc}^{B3LYP} = (1-a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^{B88} + (1-c)E_c^{LSDA} + cE^{LYP} \tag{1.76}$$

If we replace E_c^{LYP} by E_c^{PW91} we get the hybrid functional B3PW91 [Per91]. Another way to partition the energy contributions is to separate the electronic interaction into the short- and long-range terms

$$\frac{1}{\mathbf{r}_{12}} = \frac{\mathrm{erf}(\omega\mathbf{r}_{12})}{\mathbf{r}_{12}} - \frac{\mathrm{erfc}(\omega\mathbf{r}_{12})}{\mathbf{r}_{12}} \tag{1.77}$$

Here, the error function is used as smoothed step function for the transition between short and long range and ω is the range-separation parameter. Then, the exchange-correlation functional can be written as

$$E_{xc}[\rho] = E_x^{HF,lr}(\omega) + E_x^{DFA,sr}(\omega) + E_c^{DFA} \tag{1.78}$$

Some functionals of this type are: LCωPBE, ωB97XD, and CAMB3LYP [VS06, VHKS06, VS07, CHG08b, CHG08a, YTH04].

1.5 Beyond Kohn-Sham DFT

Even when some correlation energy is recovered through the approximations to the exact functional, the single-reference character of DFA is a practical limitation for the complete description of the electron correlation. In particular, DFAs tend to fail, often catastrophically, to describe strong non-dynamical correlation. In 1995, Savin proposed three strategies to go beyond the single-determinant Kohn-Sham solution, extending the space of search for wave functions that produce the ground-state density [Sav95]. First, the space of wave functions was expanded to include determinants that are *not* eigenfunctions of $\hat{S}^2$. This strategy lowers the energy by allowing determinants that break spin symmetry. The second strategy was to increase the orbital space and improve the calculation systematically. For this, we need to define the universal functional as:

$$\tilde{F}[\rho] = \langle \Psi_\rho | \hat{T} + \hat{V}_{ee} | \Psi_\rho \rangle + \tilde{E}_c[\rho] \tag{1.79}$$

where Ψ_ρ is the wave function in the allowed space that minimizes $\langle \Psi_\rho | \hat{T} + \hat{V}_{ee} | \Psi_\rho \rangle$ and gives the density ρ. Because Ψ_ρ is no longer a single determinant, $\tilde{E}_c$ accounts for

just a portion of the correlation missing in Hartree-Fock. By increasing the orbital space, Ψ_ρ approaches the FCI wave function and $\tilde{E}_c \to 0$. Now, the ground-state energy is the result of the minimization problem:

$$E_0 = \min_\rho \tilde{F}[\rho(\mathbf{r})] + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{1.80}$$

$$= \min_\rho \left(\min_{\substack{\Psi \to \rho \\ \Psi \in subspace}} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle + \tilde{E}_c[\rho(\mathbf{r})] + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \right) \tag{1.81}$$

$$= \langle \Psi_{\rho_{min}} | \hat{T} + \hat{V}_{ee} | \Psi_{\rho_{min}} \rangle + \tilde{E}_c[\rho_{min}(\mathbf{r})] + \int v(\mathbf{r})\rho_{min}(\mathbf{r})d\mathbf{r} \tag{1.82}$$

The third strategy separated the electron-electron interaction into short- and long-range terms,

$$\hat{V}_{ee} = \hat{V}_{ee}^{SR} + \hat{V}_{ee}^{LR} \tag{1.83}$$

using the Yukawa potential as step function:

$$\hat{V}_{ee}^{SR} = \sum_{i,j} v^{SR}(i,j) = \sum_{i<j} \frac{e^{-\mu r_{ij}}}{r_{ij}} \tag{1.84}$$

The long-range portion is then treated with wave function methods and the short-range portion with local (LDA-like) or semilocal (GGA-like) functionals. Similar to equation 1.82, the ground-state energy is given by minimizing over all antisymmetric wave functions Φ:

$$E_0 = \min_\rho F[\rho(\mathbf{r})] + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{1.85}$$

$$= \min_\Phi \left(\langle \Phi | \hat{T} + \hat{V}_{ee}^{LR} | \Phi \rangle + U^{SR}[\rho_\Phi(\mathbf{r})] + E_{xc}^{SR}[\rho_\Phi(\mathbf{r})] + \int v(\mathbf{r})\rho_\Phi(\mathbf{r})d\mathbf{r} \right) \tag{1.86}$$

where

$$U^{SR}[\rho(\mathbf{r})] = \int v^{SR}(1,2)\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \tag{1.87}$$

is the short-range Coulomb (Hartree) energy, and

$$E_{xc}^{SR}[\rho] = F[\rho(\mathbf{r})] - \left(\langle\Phi_{\rho_{min}}|\hat{T} + \hat{V}_{ee}|\Phi_{\rho_{min}}\rangle + U^{SR}[\rho(\mathbf{r})]\right) \tag{1.88}$$

is the short-range exchange-correlation functional. These last two terms were obtained using the Yukawa-like potential in the uniform electron gas. Example calculations with the He atom showed that short-range electron-electron interactions can be transferred from the uniform electron gas to reduce the effort of multi-determinant wave function calculations.

Savin's 1995 work was the first of many others that followed on two main topics: 1) the analysis and development of density functionals using the non-linear adiabatic connection between the KS non-interacting system and the real one [SCT98, CS99, Tou05, GGS06], and 2) combined wave function methods with density functional approximations [LSWS97, PSLS02, TCS05].

1.5.1 DFAs from the non-linear adiabatic connection

An adiabatic connection provides a smooth way to link computationally-easy effective one-particle Hamiltonians to computationally-intractable exact Hamiltonians. The standard adiabatic connection uses a parameter λ to "switch on" the electron interaction. In particular, Harris and Jones defined the adiabatic connection [HJ74b] as:

$$\hat{H}^{\lambda} = \hat{H}^0 + \hat{V}_0 + \lambda(\hat{V}_{ee} + \hat{V}_{Ne} - \hat{V}_0) \tag{1.89}$$

Choosing $\hat{H}^0$ to be the Kohn-Sham Hamiltonian (1.60) we ensure that the density $\rho_0 = \rho$, the exact ground-state density. Then, the correlation energy, at some λ is

$$E_c^\lambda = \langle \Psi^\lambda | \hat{H}^\lambda | \Psi^\lambda \rangle - \langle \Psi^{\lambda=0} | \hat{H}^\lambda | \Psi^{\lambda=0} \rangle \tag{1.90}$$

This expression contains two special cases, when $\lambda = 0$ the correlation energy E_c^λ is zero and when $\lambda = 1$ we get back the standard definition of the correlation energy density functional. Inspection of the correlation energy along the adiabatic connection brings insight into the properties of the exact functional, information that can be used in the development of new functionals or improvements of existing functionals [SCT98, CS99, SCP03]. In practice, most researchers use an alternative, but more complicated, adiabatic connection due to Langreth and Perdew, where the electron density is kept constant along the whole adiabatic connection pathway [LP80].

Alternatively, the separation of the electronic interaction is a possible choice to define a nonlinear adiabatic connection between the non-interacting system and the physical system described with a multi-determinant wave function [TCS05]. In this case the tuning parameter is the range-separation parameter μ. The ground-state energy is the sum of a wave function contribution from the long-range interaction and a density-functional correction for the short-range interaction:

$$E = \langle \Psi^\mu | \hat{T} + \hat{V}_{ee}^{LR} + \hat{V}_{ext} | \Psi^\mu \rangle + U^{\mu,SR}[\rho_\Psi] + E_{xc}^{\mu,SR}[\rho_\Psi], \tag{1.91}$$

where Ψ^μ is the ground-state wave function of the fictitious system with the long-range interaction, the solution of the equation

$$\left(\hat{T} + \hat{V}_{ee}^{\mu} + \hat{V}^{\mu}\right) \Psi^{\mu} = E^{\mu}\Psi^{\mu} \tag{1.92}$$

The potential $\hat{V}^\mu$,

$$\hat{V}^{\mu} = \sum_i v^{\mu}(\mathbf{r}_i) = v(\mathbf{r}_i) + \frac{\delta U^{\mu}[\rho(\mathbf{r}_i)]}{\delta\rho(\mathbf{r}_i)} + \frac{\delta E_{xc}^{\mu}[\rho(\mathbf{r}_i)]}{\delta\rho(\mathbf{r}_i)} \tag{1.93}$$

ensures that the ground-state densities of the fictitious and the real system are the same. This potential can be obtained in a optimized-effective-potential (OEP) fashion, with Lieb maximization:

$$F^{\mu}[\rho] = \sup_{v^{\mu}} \left(E^{\mu}[v^{\mu}] - \int v^{\mu}\rho(\mathbf{r}) d\mathbf{r} \right) \tag{1.94}$$

where $E^\mu[v^\mu] = E^\mu$ from equation 1.92. For two-electron systems the exchange potential is half of the negative Hartree potential $v_x(\mathbf{r}) = -v_h(\mathbf{r})/2$, so the correlation energy can be calculated by integration of the corresponding potential $v_c(\mathbf{r}) = v_{xc}(\mathbf{r}) - v_x(\mathbf{r})$ [CS99, TCS05].

If instead of optimizing the potential one decides to use one of the available approximations for the short-range exchange-correlation density functional, one gets a combined CI-DFA method.

1.5.2 Combined Configuration Interaction and Density functional Approximation (CI-DFA) methods

Given approximate short-range density functionals, the effective external potential can be written as the sum of the physical external potential and the short-range potentials from the density functionals,

$$\hat{V}^{\mu} = v(\mathbf{r}) + \frac{\delta U^{SR,\mu}[\rho_{\Psi^{\mu}}(\mathbf{r})]}{\delta\rho(\mathbf{r})} + \frac{\delta E_{xc}^{SR,\mu}[\rho_{\Psi^{\mu}}(\mathbf{r})]}{\delta\rho(\mathbf{r})} \tag{1.95}$$

The density ρ_{Ψ} is computed self-consistently by solving equation 1.92, and the ground-state energy is given by

$$E_0 = \langle\Psi^{\mu}|\hat{T} + \hat{V}_{ee}|\Psi^{\mu}\rangle + U^{SR}[\rho_{\Psi^{\mu}}(\mathbf{r})] + E_{xc}^{SR}[\rho_{\Psi^{\mu}}(\mathbf{r})] + \int v(\mathbf{r})\rho_{\Psi^{\mu}}(\mathbf{r})d\mathbf{r} \tag{1.96}$$

Examples of methods combining multi-reference CI (MRCI) and LDA functionals for a different separation of the electron interaction were presented by Leininger et al. and Pollet et al. [LSWS97, PSLS02, TCS05]. They use the standard error function for the long-range interaction

$$\hat{V}_{ee}^{LR} = \sum_{i<j} \frac{\mathrm{erf}(\mu\, r_{ij})}{r_{ij}} \tag{1.97}$$

where

$$\mathrm{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2} dt \tag{1.98}$$

In general, the energies computed with the combined method were closer to the experiment, better than the separated methods, and even better than FCI (in a

small basis). Furthermore, it was demonstrated that by removing the cusp and using a short-range functional one can alleviate the slow convergence of wave function methods to basis set limit full-CI. This last observation brings us back to the problem of the correlation cusp, and other approaches to tackle it.

1.6 Dealing with the correlation cusp

The cause of the slow convergence of orbital-based methods is the poor description of the correlation cusp. The true wave function satisfies Kato's cusp theorems [Kat57], which prove that all wave functions are continuous everywhere, and their partial derivatives are all bounded except when the potential diverges, where derivative-discontinuities occur. More specifically, the cusp condition states that at the point of coalescence of two electrons, far from the nuclei and the other electrons, the first partial derivative of the wave function with respect to the distance between the two electrons, averaged over a sphere centered on the coalescence point, is equal to half the value of the wave function at that point

$$\left\langle \left(\frac{\partial \Psi}{\partial r_{ij}} \right)_{r_{ij}=0} \right\rangle_{\theta_{ij}\phi_{ij}} = \frac{1}{2} \Psi_{r_{ij}=0} \tag{1.99}$$

While $\Psi_{r_{ij}=0}$ is almost never zero, the left-hand side of equation 1.99 is zero for any finite expansion of the wave function in Slater determinants. Therefore, the wave function of orbital-based methods does not obey the cusp condition. A direct solution to this problem is to use wave functions that include an explicit dependence on r_{ij}.

1.6.1 Explicitly correlated methods: R12/F12

The importance of using functions that satisfy the cusp conditions was first considered by Slater [Sla28a, Sla28b] and Hylleraas [KBV12, Hyl29, Hyl30]. In particular, with a three-term wave function,

$$\Psi(r_1, r_2, r_3) = N(1 + c_1(r_1 - r_2)^2) + c_2 r_{12})e^{-\alpha(r_1+r_2)} \tag{1.100}$$

where the parameters $\alpha = -1.81607$, $c_1 = 0.130815$ and $c_2 = 0.291786$ were determined variationally, Hylleraas achieved an energy error of 1.3 mHartree [Hyl28]. It is apparent that the exquisite accuracy of this very compact wave function arises from including terms dependent on $r_1 \pm r_2$. Extending Hylleraas idea to N-electrons systems is straightforward. However, his explicitly correlated wave functions are accompanied by high-dimensional integrals, and clever ways to evaluate these integrals efficiently are needed. For example, the transcorrelated method uses a similarity transformation of the Hamiltonian to eliminate the Coulomb singularities analytically. Because this wave function is cuspless it can be represented precisely by Slater determinants [BH69b, BH69a]. Another example are the R12 and F12 methods, which use the resolution-of-the-identity (RI) to reduce three- and four-electron integrals to two-electron integrals [Kut85].

However, one can avoid the treatment of the correlation in the wave function by replacing the singular Coulomb interaction with a smooth potential. The eigenfunction of a smooth Hamiltonian is also smooth, and less effort is needed to approximate this wave function.

1.6.2 Smooth potentials

The cost of Gaussian-type-orbital-based methods can be reduced by replacing the Coulomb electron-electron interaction by a smooth potential. Lloyd-William et al. gained a factor of 32 in the cost of diffusion Monte Carlo (DMC) and CID calculations of the UEG using a pseudopotential fitted to describe correctly the scattering properties of the Coulomb potential [LWNC15] in the UEG. Additionally, a faster convergence of the wave function of a smooth Hamiltonian was observed by Sirbu and King. In their work they showed that one needs a shorter partial wave expansion with the smooth potential, and therefore a smaller angular momentum, to converge the error in the partial wave expansion with respect to the whole wave function [SK02]. In their following paper they used in Rayleigh-Schrödinger perturbation theory, with the smooth long-range potential as reference Hamiltonian and the complementary short-range potential as the perturbation. Sub-μHartree precision was obtained for spherically-symmetric two-electron systems. It was demonstrated that approximations to the reference wave functions with fewer basis functions bring computational benefits with an acceptable loss of precision.

The goal of this thesis is to develop a reduced-cost method for the electronic Schrödinger Equation by using smooth potentials. We aim to leverage calculations for the smoothed potentials (which we can perform at low cost) to make accurate predictions for the true physical system, with the singular Coulomb interaction. Each Chapter of this thesis represents one milestone towards this long-term goal:

- How do we choose the smooth potential? Is an *erfgau* form sufficiently good?
- The use of the model potential has two important effects: 1) some loss of accuracy, but 2) faster convergence of the wave function. How can we improve

the accuracy without losing the cost reduction? Is it possible to define a method to recover the physical solution by improving the approximation systematically?

- Are extrapolation techniques feasible/decent?

Chapter 2

Smooth models for the Coulomb potential

Smooth model potentials with parameters selected to reproduce the spectrum of one-electron atoms are used to approximate the singular Coulomb potential. Even when the potentials do not mimic the Coulomb singularity, much of the spectrum is reproduced within the chemical accuracy. For the Hydrogen atom, the smooth approximations to the Coulomb potential are more accurate for higher angular momentum states. The transferability of the model potentials from an attractive interaction (Hydrogen atom) to a repulsive one (Harmonium and the uniform electron gas) is discussed[1].

[1]This chapter was originally published as [GEAKS16]. Reprinted by permission from: Springer Nature, Theoretical Chemical Accounts, Copyright 2016, Springer-Verlag Berlin Heidelberg (2016)

2.1 Philosophy

How feasible is it to find a model for the Coulomb interaction that is easier to evaluate but still reproduces key properties of the physical interaction? A logical starting point is a system with *no interaction*, as in Kohn-Sham density functional theory (DFT)[KS65]. The Kohn-Sham (KS) approximation starts from a non-interacting system, described as the sum of the individual electrons' contributions to the energy:

$$\hat{H}_s = \sum_i^N \left(-\frac{1}{2}\nabla_i^2 + v_s(r_i) \right). \tag{2.1}$$

In order to ameliorate the effect of omitting the Coulomb repulsion between the electrons, an extra term, the exchange-correlation functional $E_{xc}[\rho]$, is introduced into the energy expression. For a given external potential $v(r)$

$$E_v[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int v(r)\rho(r)dr \tag{2.2}$$

where $T_s[\rho]$ is the kinetic energy functional of the non-interacting system, $J[\rho]$ is the classical repulsion, and $E_{xc}[\rho]$ is the exchange-correlation functional, which must be approximated [PY94]. The Euler-Lagrange equation associated with the stationary condition of $E_v[\rho]$ can be transformed into a self-consistent set of equations:

$$\left(-\frac{1}{2}\nabla^2 + v(r) + \int \frac{\rho(r')}{|r-r'|}dr' + v_{xc}(r) - \epsilon_j \right) \phi_j(r) = 0 \tag{2.3}$$

$$\rho(r) = \sum_{j=1}^{N} |\phi_j(r)|^2$$

$$v_{xc}(r) = \frac{\delta E_{xc}[\rho(r)]}{\delta \rho(r)}.$$

In principle, the KS solutions are exact when E_{xc} is exact, and the KS orbitals yield the exact density of the system with N electrons in the external potential $v(r)$. The accuracy of KS density functional approximations (DFA) depends on the approximation one uses for E_{xc}.The simplest approximation is the local density approximation (LDA)[CA80, PZ81, PW92]. In LDA it is assumed that the exchange-correlation functional is local,

$$E_{xc}[\rho(r)] = \int \epsilon_{xc}(\rho(r))dr, \tag{2.4}$$

where the exchange-correlation energy density $\epsilon_{xc}(\rho(r))$ at r is taken from the uniform electron gas with density $\rho(r)$.

To accurately recover the effect of omitting the interaction between the electrons, one constructs an adiabatic connection that links the KS *non-interacting* system with the physical *interacting* system. Traditionally, this adiabatic connection is written as a function of the strength of the interaction, using a simple multiplicative factor λ [HJ74a, LP75, GL76, Yan98]:

$$\hat{H}_\lambda = \sum_{i=0}^{N} -\frac{\nabla_i^2}{2} + v_\lambda(r_i) + \frac{1}{2}\sum_{j\neq i}\frac{\lambda}{r_{ij}}. \tag{2.5}$$

Computational studies of the adiabatic connection have been performed for few-electron atomic systems, and provide significant insight into the structure of the exact exchange-correlation density functional [Sei03, MSCY06, CMSY07, CS99, PCL$^+$03, TCH09, TCH10]. An alternative to the traditional adiabatic connection is to write the Coulomb interaction as the sum of a short-range piece and a long-range piece. The long-range piece of the potential is usually chosen to be smooth (or at least

nonsingular), so that it is relatively easy to approximate solutions to the Schrödinger equations when only the long-range piece is included.

So far we have reviewed traditional strategies that add density-functional corrections to an "easy" system to approximate the real system. Can we use the real system to construct the model? For example, is it possible to select an interaction potential, different from the Coulomb one, that nonetheless reproduces a certain target property of the system? For example, one might wish to select an interaction potential that preserves the energy spectrum of an atom. This strategy is not new. Valance and Bergeron [VB89] show how to construct analytically solvable pseudopotentials and model potentials, in the framework of supersymmetric quantum mechanics, that reproduce experimental spectra. Starting from a one-electron one-dimensional Hamiltonian H_1 associated with the potential V_1, they found a supersymmetric partner H_2, characterized by a second potential V_2, with almost the same spectrum as H_1. H_2 is missing the ground-state of H_1. A similar approach has been used by Lepage[Lep98] in the field of elementary particle physics, where the Hamiltonian is constructed to reproduce low-energy features of a particular physical system.

In the next section we define an expression for the model potential. We then explain two different model potentials that accurately reproduce the lowest-energy eigenvalues of the Hydrogen atom. In section 4, we use the same models for the Coulomb potential to replace the Coulomb repulsion between the electrons in two-electron Harmonium. Finally, the exchange energy of the uniform electron gas that results from one of the models is compared to the standard approximation.

2.2 Ansatz

Analogous to the inverse problem of finding the Kohn-Sham potential from a given density, where oscillatory potentials and/or shifted potentials can reproduce the exact density numerically [SCP03], finding the potential given the spectrum is not trivial because the solution is not unique. Therefore, we restrict the analytical form of the potential by imposing some constraints. We would like to eliminate the singularity of the Coulomb potential because solving the Schrödinger equation for a singular operator is computationally demanding. We also wish to preserve the long-range asymptotic form of the potential, so that the long-range electrostatics is correct. An interaction potential that satisfies these constraints is:

$$\frac{1}{r} \rightarrow V_\mu(r) = c\, \exp(-\alpha^2 r^2) + \frac{\mathrm{erf}(\mu r)}{r}. \tag{2.6}$$

We prove in this article that despite the simplicity of this *erfgau* [TSF04] type of potential, it is flexible enough for our purposes.

2.3 H atom

2.3.1 Construction of $V_\mu(r)$

To determine the parameters in the model potential we consider what happens when we replace the Coulomb interaction between the nucleus and the electron in the hydrogen atom by the model potential $V_\mu(r)$ (3.4). Thus, we replace the Hamiltonian of the Hydrogen atom

$$\hat{H}_0(\mathbf{r}) = -\frac{1}{2}\nabla^2 - \frac{1}{r} \tag{2.7}$$

by a modified Hamiltonian

$$\hat{H}_\mu(\mathbf{r}) = -\frac{1}{2}\nabla^2 - V_\mu(r) \tag{2.8}$$

with $V_\mu(r)$ defined in such a way that

$$\lim_{\mu\to\infty} \hat{H}_\mu = \hat{H}_0, \quad \text{i.e.} \quad \lim_{\mu\to\infty} V_\mu = \frac{1}{r}. \tag{2.9}$$

As we know, in the case of the long-range term

$$\frac{\mathrm{erf}(\mu\, r)}{r} \underset{\mu\to\infty}{\sim} \frac{1}{r}.$$

Thus, condition (2.9) is fulfilled if $[c\, \exp(-\alpha^2 r^2)] \to 0$ when $\mu \to \infty$. Besides, we would like the spectrum of $\hat{H}_\mu$ to be as close as possible to the spectrum of $\hat{H}_0$. This can be achieved by properly choosing $c = c(\mu)$ and $\alpha = \alpha(\mu)$.

Consider $\hat{H}_0$ as the unperturbed operator, and

$$w_\mu = \hat{H}_\mu - \hat{H}_0 \tag{2.10}$$

as a perturbation. First, we notice that for the bound states of the Hydrogen atom

$$\langle\psi_i|\mathrm{erfc}(\mu\, r)/r|\psi_i\rangle \underset{\mu\to\infty}{\sim} \mu^{-(2l+2)}, \tag{2.11}$$

and

$$\langle\psi_i|c\, \exp(-\alpha^2 r^2)|\psi_i\rangle \underset{\alpha\to\infty}{\sim} c\, \alpha^{-(2l+3)}, \tag{2.12}$$

where l is the angular momentum quantum number, a necessary condition for the

spectra of $\hat{H}_\mu$ and $\hat{H}_0$ to coincide is that these two expectation values have the same asymptotic form. (See Appendix A for more details about the μ-dependence of the wavefunction.) The simplest choice, adopted in this paper, is to take α as a linear function of μ. This implies that c also has to be linear in μ. Thus, we can set

$$c = \gamma\,\mu, \qquad \alpha = \kappa\,\mu, \tag{2.13}$$

where γ and κ are μ-independent parameters.

According to Eqs. (2.11), (2.12) and (2.13), the asymptotic expansion of the expectation value of w_μ may be written as

$$\langle\psi_i|w_\mu|\psi_i\rangle = \sum_{j=2l+2} d_j(\gamma,\kappa)\,\mu^{-j}, \quad \mu >> 1. \tag{2.14}$$

We select γ and κ so that the two leading terms in expansion (2.14) vanish. The lowest order terms ($j = 2$ and $j = 3$) correspond to $l = 0$ states. For angular momenta $l > 0$ the leading terms in Eq. (2.14) are $O(\mu^{-4})$ or smaller. Then, the choice of $c(\mu)$ and $\alpha(\mu)$ is dictated by the requirement that the eigenvalues of the S states are as correct as possible for $\mu \to \infty$. The explicit form of expansion (2.14) for $l = 0$ states reads (see Appendix B)

$$\langle\psi_i|w_\mu|\psi_i\rangle = \left(1 - \frac{\sqrt{\pi}\gamma}{\kappa^3}\right)\mu^{-2} + \left(-\frac{8}{3\sqrt{\pi}} + \frac{4\gamma}{\kappa^4}\right)\mu^{-3} + O(\mu^{-4}). \tag{2.15}$$

Solving equations $d_2 = d_3 = 0$ for γ and κ, we obtain:

$$c = \frac{27}{8\sqrt{\pi}}\,\mu\, = 1.904\,\mu, \quad \alpha = \frac{3}{2}\,\mu\, = 1.500\,\mu. \tag{2.16}$$

In addition to the asymptotic behavior, for practical calculations we need the optimal

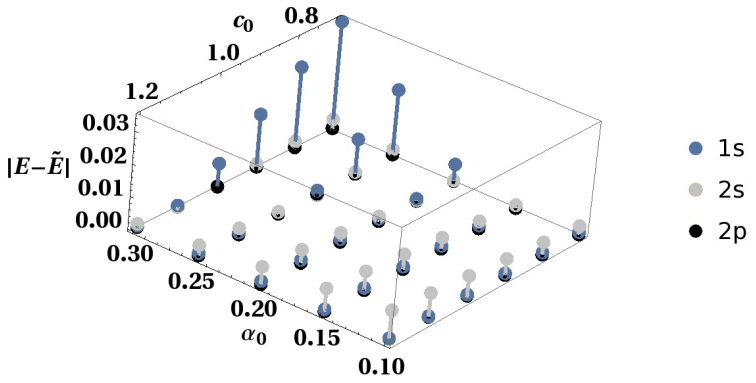

Figure 2.1: The accuracy of the model potential spectrum with respect to the parameters c_0 and α_0. Energy errors (in %), for $\mu = 1$.

parameters c and α at finite values of μ. Taking the linear forms

$$c = \gamma\,\mu + c_0, \qquad \alpha = \kappa\,\mu + \alpha_0 \tag{2.17}$$

with the parameter γ and κ from the previous step we can use c_0 and α_0 to further optimize the spectrum. The condition for the elimination of μ^{-2} term remains the same as before. The equations $d_3 = 0$ and $d_4 = 0$ are given in Appendix B. The behavior of the spectrum of the model potential versus c_0 and α_0 is shown in Figure 2.1, where for fixed $\mu = 1.0$ energies of 1s, 2s and 2p states are displayed. As one can see, there is a range for which pairs of (c_0, α_0) give reasonably small errors of the energy values. The dependence of the relative error on n and l is discussed in section 2.3.2.

To select the best linear forms of c and α we constructed a grid on the intervals $c = [-0.5, 0.0]$ and $\alpha = [1.0, 3.0]$, and then we computed the error in the eigenvalues of the 1s, 2s, 2p, 3s, 3p, and 3d states for $\mu = [0.5, 2.0]$. We defined the best choice for the parameters as the minimax choice: the c, α that minimized the maximum

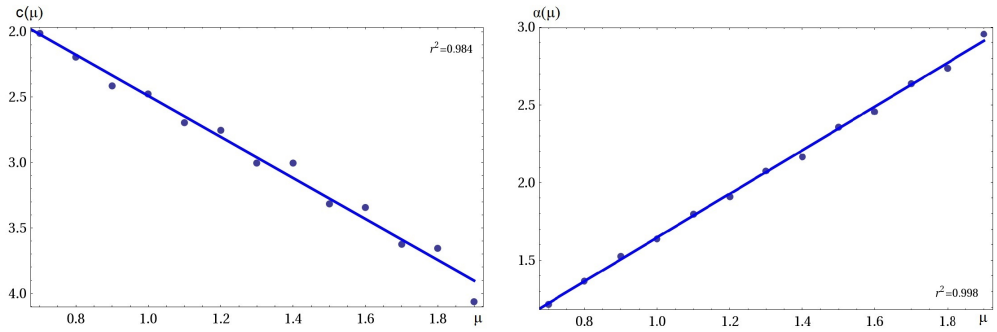

Figure 2.2: Linear regressions for the two parameters $c(\mu)$ and $\alpha(\mu)$ of the simple fit in Eq. (2.19). Here, the dots are the optimal values of the parameters according to Eq. (2.18), the lines are the least-squares linear regressions, and r^2 are their corresponding coefficients of determination.

absolute deviation between the eigenvalues with the model potential and the exact result from the Coulomb interaction,

$$\delta = \min_{c(\mu),\alpha(\mu)} \{ \max_{n,l} |E_{\text{Coulomb}} - E_{\text{model}}| \}. \tag{2.18}$$

As seen in Figure 2.2, the best values of (c, α) can be modeled by a linear function,

$$\begin{aligned} c &= 0.923 + 1.568\,\mu \\ \alpha &= 0.241 + 1.405\,\mu. \end{aligned} \tag{2.19}$$

Note that the resulting fit is very similar to one the linear forms obtained from perturbation theory $\{c_0 = 0.943,\ \gamma = 1.904\,\mu\}$ and $\{\alpha_0 = 0.247,\ \kappa = 1.5\,\mu\}$ (see equation (A.12) of Appendix B). With this fitted form the interaction does not vanish when $\mu = 0$, therefore the spectrum cannot be exact for small values of μ, in contrast to the asymptotic form (3.5), in which both c and α are proportional to μ.

The optimum parameters for a Hydrogen-like atom with the nuclear charge Z may be obtained from the ones determined for the case of $Z = 1$ by a simple scaling procedure. Eq. (3.4) becomes

$$\frac{Z}{r} \rightarrow V_{\mu}^{Z}(r) = Z\left[c_Z\, \exp(-\alpha_Z^2 r^2) + \frac{\mathrm{erf}(\mu_Z r)}{r}\right] \tag{2.20}$$

and

$$\hat{H}_{\mu}^{Z}(\boldsymbol{r}) = -\frac{1}{2}\nabla^2 - V_{\mu}^{Z}(\mathbf{r}) = Z^2\, \hat{H}_{\mu}(\boldsymbol{p}), \tag{2.21}$$

where $\boldsymbol{p} = Z\,\boldsymbol{r}$ and

$$c_Z = Z\,c, \quad \alpha_Z = Z\,\alpha, \quad \mu_Z = Z\,\mu. \tag{2.22}$$

2.3.2 Results

Potentials

The model potentials we consider in this paper [$\mathrm{erf}(\mu r)/r$, asymptotic (Eq. (3.5)), and fitted (Eq. (2.19)] corresponding to $\mu = 1$ are compared with the Coulomb potential and with a modified long-range potential, $\mathrm{erf}(3r)/r$, in Figure 2.3. One might suspect that adding an optimized Gaussian term to $\mathrm{erf}(\mu r)/r$ would give a potential that mimics the effect of increasing μ in the long-range term. This is not the case for the potentials we consider in this paper. At first glance the potential $\mathrm{erf}(3r)/r$ (dashed line) seems similar to the asymptotic and to the fitted potentials. But the $\mathrm{erf}(3r)/r$ potential is *always* weaker than the Coulomb potential, while the latter potentials, though in some intervals of r they are also weaker, in other intervals they are stronger than the Coulomb potential. This may explain why the asymptotic

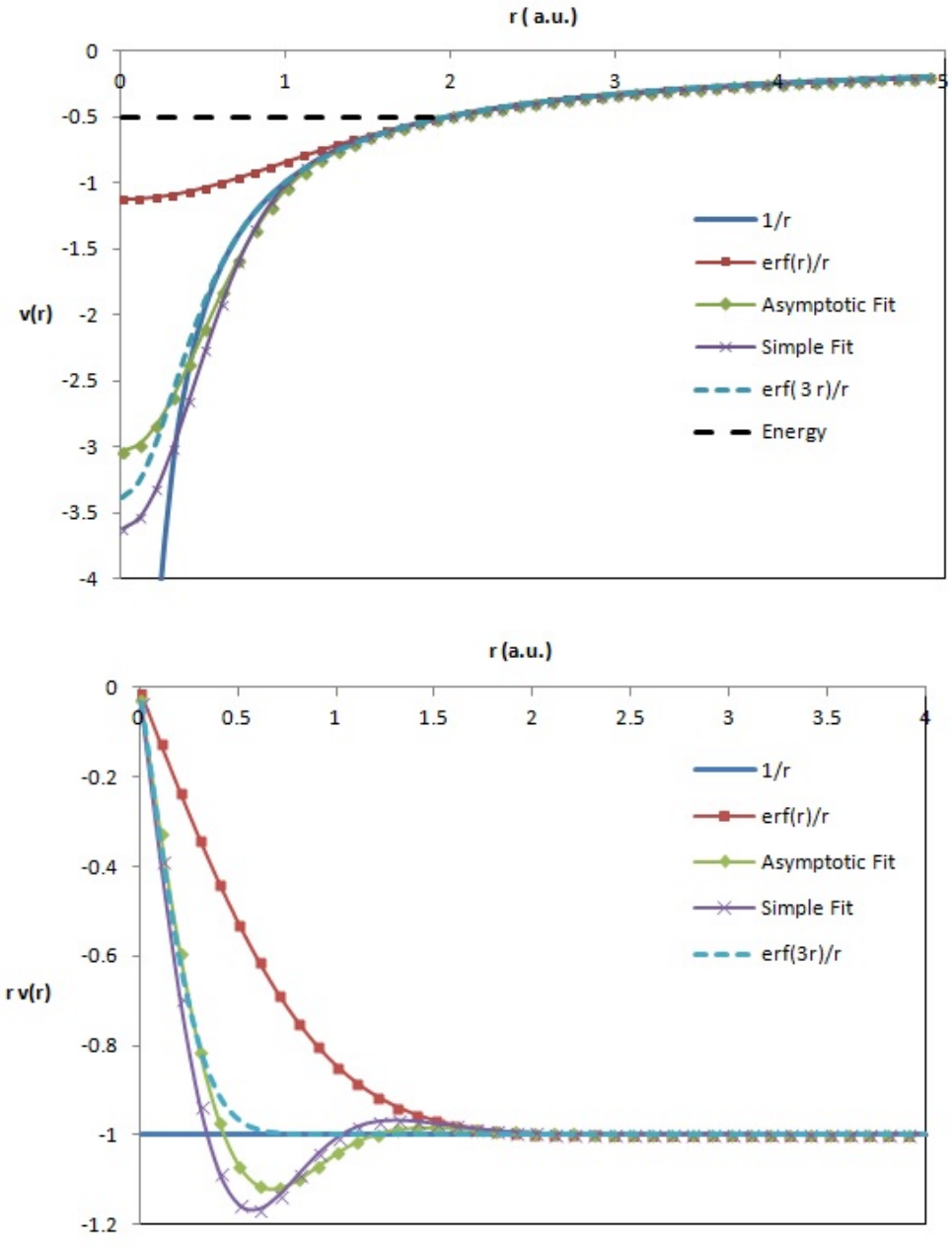

Figure 2.3: Comparison between the Coulomb potential (solid line) and the model potentials: long-range(squares), asymptotic (diamonds), fitted (stars) and modified long-range potential (dashed line).

and fitted potentials reproduce the spectrum much better than the modified long-range potential: the effects of too strong and too weak regions of the model potentials cancel each other, leaving the eigenvalues relatively unchanged.

Eigenvalues

The percentage errors in the eigenvalues of Hydrogen with long-range, asymptotic, and fitted potentials are presented in Figure 2.4. As the quantum number increases, the amplitude of the eigenfunctions near the nucleus decreases, the long-range part of the potential dominates, and the eigenvalues approach the exact ones. A clear improvement is found in the asymptotic and fitted potentials compared with the traditional long-range potential. As expected, the fitted potential produced the smallest errors. However, for $\mu > 1.5$, the difference between the asymptotic and the fitted models is rather small; see Figure 2.5.

The advantage of adding a Gaussian term is clear when we compare against a modified long-range potential, such as $\mathrm{erf}(3\mu r)/r$, Figure 2.6 . The Gaussian term lets us "get away with" a much smaller value of μ, and seems to work better for s-type orbitals than the erf-based potential.

As expected from the asymptotic analysis, better results are obtained for higher angular momentum because the centrifugal barrier, $l(l+1)/2r^2$, pushes the electron away from the nucleus, into a region where the difference between the model potential and the Coulomb potential is negligible (see Figures 2.7 and 2.8). On the other hand, when μ is close to zero, the eigenvalues from the model potential are very poor, because the short-range Gaussian term cannot bind an electron when the angular momentum is too high. When we look at the wavefunction (for example, the 1s and

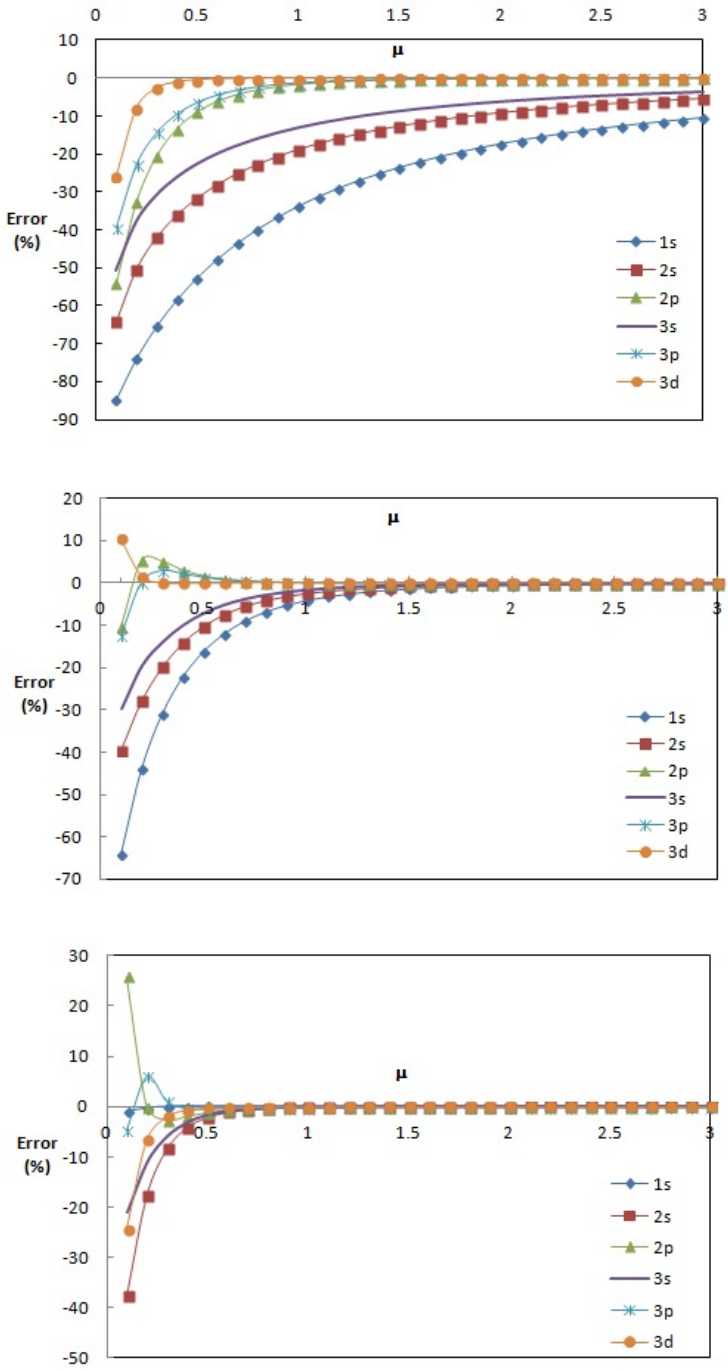

Figure 2.4: Percentage errors in the eigenvalues of Hydrogen. From top to bottom: a)long-range $\mathrm{erf}(\mu r)/r$, b) asymptotic [Eq. (3.5)], and c)fitted [Eq. (2.19)] potentials.

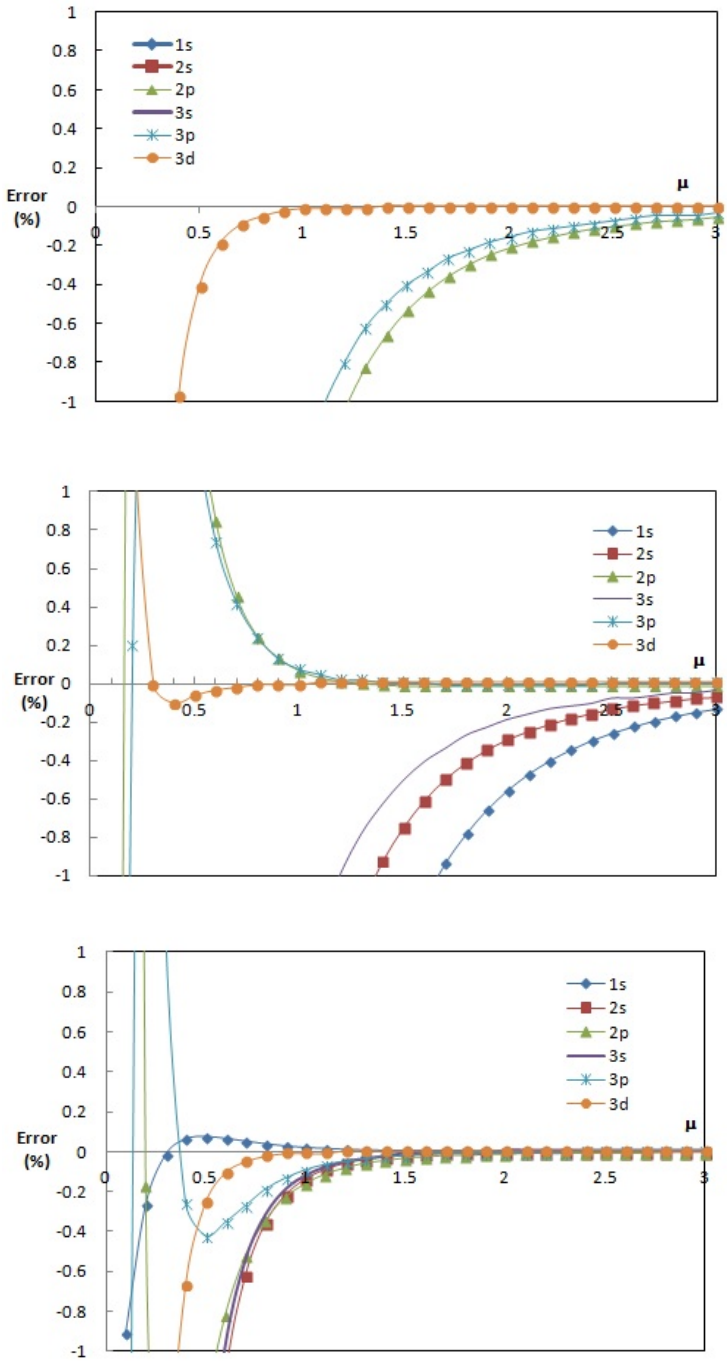

Figure 2.5: Close-up of the percentage error of the eigenvalues of Hydrogen. As in Figure 2.4, the curves are, from top to bottom: a)long-range $\mathrm{erf}(\mu r)/r$, b) asymptotic [Eq. (3.5)], and c)fitted [Eq. (2.19)] potentials.

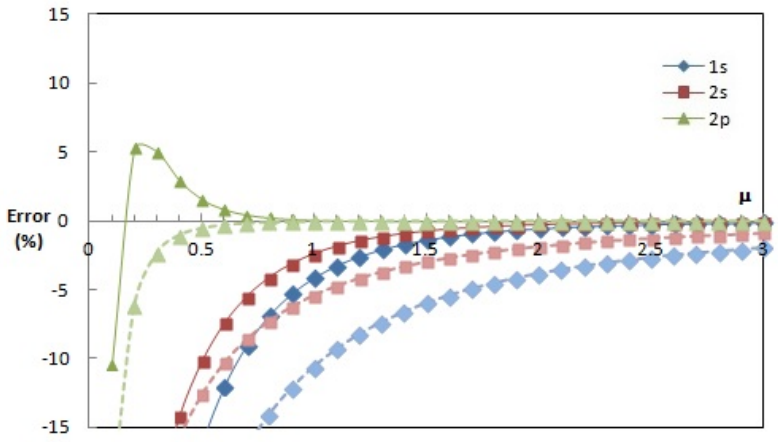

Figure 2.6: Comparison of the percent error in the first three eigenvalues of the Hydrogen atom between the asympotic potential, (thick lines) and $\mathrm{erf}(3\mu)/r$ (dashed lines).

2s orbitals (Fig. 2.9)), we see that even though the eigenvalues are very similar, the eigenfunctions can be quite different.

How important is this difference? Is the perturbation still small if the spectrum is nearly reproduced? Moreover, can we use the same approach for other types of interactions, e.g. a repulsive potential? There are several ways to assess the transferability of our model potentials. Below we use the same replacement for the electron-electron repulsion in two model systems, Harmonium and the uniform electron gas.

2.4 Harmonium

To explore whether the model potentials can be used to describe repulsive interactions, we consider a system of two interacting electrons confined in a harmonic oscillator

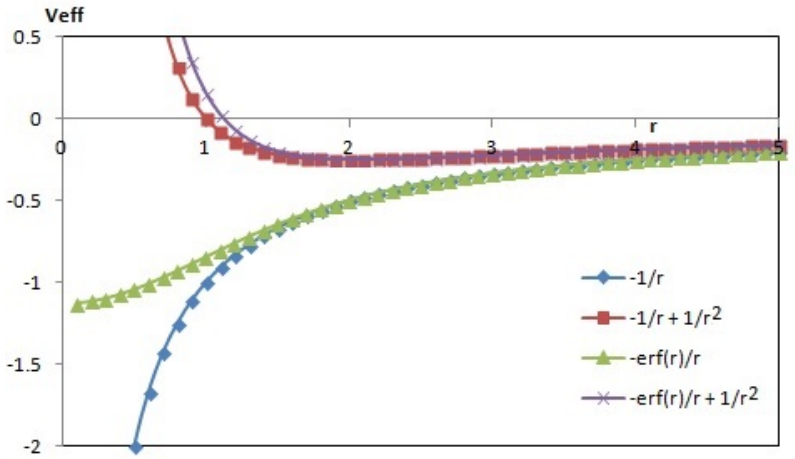

Figure 2.7: Comparison between the radial potentials $-1/r$ and $-\mathrm{erf}(r)/r$ in the Hydrogen atom, when $l = 0$ (diamonds and squares, respectively) and $l = 1$ (triangles and stars).

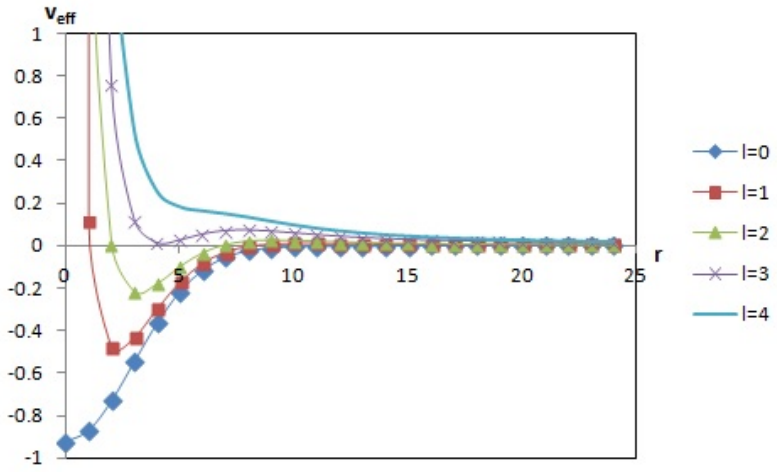

Figure 2.8: The effect of the centrifugal term on the radial potential in the $\mu \to 0$ limit for the fitted erfgau interaction [Eq. (2.19)].

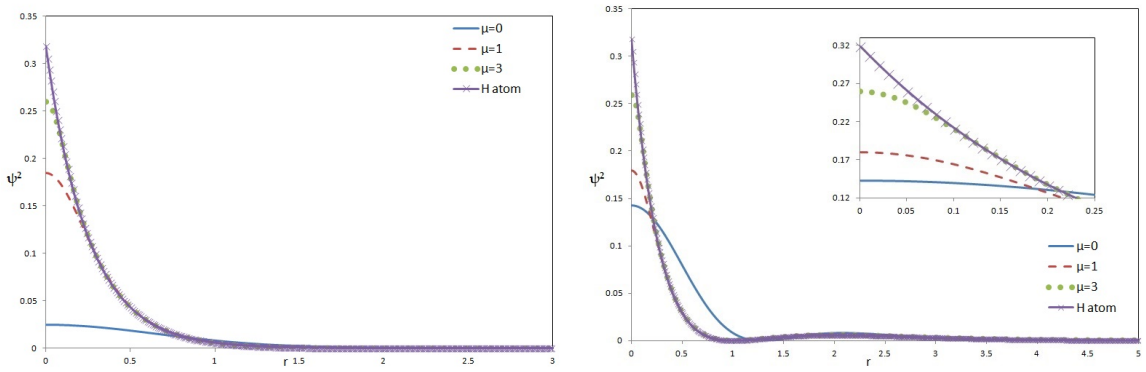

Figure 2.9: The orbital densities $|\psi_{1s}(r)|^2$(left) and $|\psi_{2s}(r)|^2$(right) derived from the model Hamiltonian of Hydrogen, Eq. (2.8) with $c = 0.923 + 1.568\,\mu$, $\alpha = 0.241 + 1.405\,\mu$, using different values of μ.

potential, called Harmonium [Tau93, KC03]. The Hamiltonian of Harmonium is:

$$\hat{H}^{\mathrm{h}}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{1}{2}\nabla_1^2 + \frac{\omega^2\, r_1^2}{2} - \frac{1}{2}\nabla_2^2 + \frac{\omega^2\, r_2^2}{2} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}, \tag{2.23}$$

where the superscript h stands for *harmonium*. This Hamiltonian is separable if one rewrites it in terms of the center of mass and the relative coordinates

$$\mathbf{R} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2), \quad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2. \tag{2.24}$$

In the new coordinates

$$\hat{H}^{\mathrm{h}}(\mathbf{r}_1, \mathbf{r}_2) = \hat{H}^{\mathrm{h}}_{\mathbf{r}}(\mathbf{r}) + \hat{H}^{\mathrm{h}}_{\mathbf{R}}(\mathbf{R}), \tag{2.25}$$

where

$$\hat{H}^{\mathrm{h}}_{\mathbf{r}}(\mathbf{r}) = -\nabla^2_{\mathbf{r}} + \frac{\omega^2\, r^2}{4} + \frac{1}{r}, \tag{2.26}$$

$$\hat{H}^{\mathrm{h}}_{\mathbf{R}}(\mathbf{R}) = -\frac{1}{4}\nabla^2_{\mathbf{R}} + \omega^2\, R^2 \tag{2.27}$$

and the Schrödinger equation separates into two equations:

$$\hat{H}^{\mathrm{h}}_{\mathbf{r}}(\mathbf{r})\, \Phi_{nlm}(\mathbf{r}) = \epsilon_{nl}\, \Phi_{nlm}(\mathbf{r}) \tag{2.28}$$

and

$$\hat{H}^{\mathrm{h}}_{\mathbf{R}}(\mathbf{R})\, \xi_{\nu\lambda\mu}(\mathbf{R}) = \eta_{\nu\lambda}\, \xi_{\nu\lambda\mu}(\mathbf{R}), \tag{2.29}$$

where n, l, m and ν, λ, μ are quantum numbers and the total energy is equal to $E_{\nu\lambda;nl} = \eta_{\nu\lambda} + \epsilon_{nl}$.

In the case of Harmonium we can use the same approximation for the Coulomb potential as we did for the Hydrogen atom but now, instead of the *attractive* Coulomb interaction we have the *repulsive* one. Thus, the modified Hamiltonian for the relative motion of two electrons reads

$$\hat{H}^{\mathrm{h}}_{\mathbf{r};\mu}(\mathbf{r}) = -\nabla^2_{\mathbf{r}} + \frac{\omega^2\, r^2}{4} + V_{\mu}(r). \tag{2.30}$$

Due to the spherical symmetry, it is convenient to express the solutions of Eqs. (2.28) and (2.29) in spherical coordinates. In particular, if we set

$$\Phi_{nlm}(\mathbf{r}) = \frac{1}{r}\, \phi_{nl}(r)\, Y_{lm}(\hat{r}), \tag{2.31}$$

where $Y_{lm}(\hat{r})$ are spherical harmonics, then in the case of Eq. (2.28) with the modified Hamiltonian (2.30) we have

$$\left[-\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + \frac{\omega^2\, r^2}{4} + V_\mu(r)\right] \phi_{nl}(r) = \epsilon_{nl}\, \phi_{nl}(r). \tag{2.32}$$

2.4.1 How Harmonium is computed

In order to assess the model potentials for Harmonium we solved Eq. (2.32) numerically. To this end we discretized this equation on a grid of N equidistant points for $r \in [0, a]$ with the boundary conditions $\phi_{nl}(0) = \phi_{nl}(a) = 0$ and the approximation

$$\frac{d^2}{dr^2}\, \phi_{nl}(r) \approx \frac{1}{h^2}\left[\phi_{nl}(r-h) - 2\phi_{nl}(r) + \ \phi_{nl}(r+h)\right],$$

where $h = a/N$. The discretized equation may be written as

$$\sum_{j=0}^{N} \left(A_{ij} - \epsilon_{nl}\, \delta_{ij}\right) \phi_{nl}(r_j) = 0,$$

where $r_j = j\, h$, $j = 0, 1, \ldots, N$ and

$$A_{ij} = \left(\frac{2}{h^2} + \frac{l(l+1)}{r_j^2} + \frac{\omega^2\, r_j^2}{4} + V_\mu(r_j)\right) \delta_{ij} - \frac{\delta_{i,j+1} + \delta_{i,j-1}}{h^2}$$

It has been solved using standard LAPACK subroutines. In the calculations we set $N = 10\,000$ and $a = 10/\omega$. In order to discard errors due to the numerical procedure, we computed and compared the eigenvalues of two states for which we know the analytic solution:

$$\phi_1 \ \sim\ r^{l+1}\,(1\ +\ \omega\, r)\, \mathrm{e}^{-\omega\, r^2/4},\ \ \omega = \frac{1}{2(l+1)},\quad E_1 = \omega\left(l + \frac{5}{2}\right) \tag{2.33}$$

and

$$\phi_2 \sim r^{l+1}\left[1 + \frac{r\,(1+\omega\,r)}{2(l+1)}\right]\mathrm{e}^{-\omega\,r^2/4}, \tag{2.34}$$
$$\omega = \frac{1}{2(4l+5)}, \quad E_2 = \omega\left(l+\frac{7}{2}\right).$$

For $l = 0$, the percentage errors in the eigenvalues obtained with the numerical integration are $5.76 \times 10^{-6}\%$ and $6.52 \times 10^{-6}\%$, respectively.

2.4.2 Results

In figure 2.10, the eigenvalues of Harmonium with $n = 1$ and $l = 0$ are shown for the long-range, asymptotic, and fitted potentials. Similar to the Hydrogen atom, a deterioration at small μ is observed, but now the asymptotic potential is slightly better than the fit to the hydrogenic spectrum, and both are much better than the uncorrected $\mathrm{erf}(\mu r)/r$ potential. As the harmonic confinement weakens ($\omega \to 0$), the average distance between electrons increases, and the models become more accurate because of their correct $1/r$ asymptotics. For strongly confined electrons ($\omega >> 1$), however, the Gaussian correction factors that were adapted to the hydrogenic spectrum do not seem appropriate for modelling the short-range $1/r$ interaction. In Table 2.1 we collected the smallest value of μ such that the error is always less than 1%, for the different values of ω. It is clear that as the range of the average distance between the electrons decreases ($\omega \to \infty$), we also need to scale the range-separation parameter ($\mu \to 0$).

In order to investigate the interplay between the strength of confinement and the

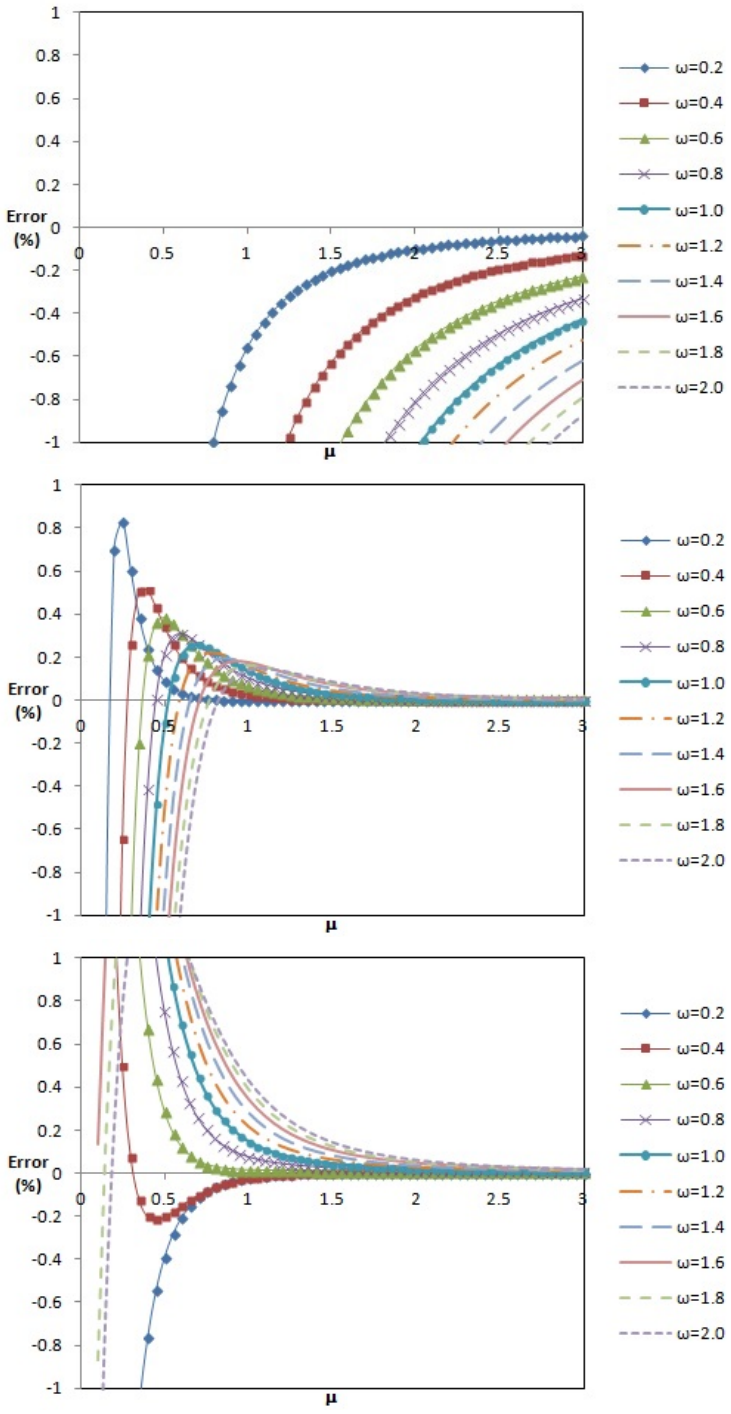

Figure 2.10: Errors (in %) of the Harmonium eigenvalues as function of μ, for different ω. From top to bottom:a)long-range $\mathrm{erf}(\mu r)/r$, b) asymptotic (Eq. 3.5), and c)fitted (Eq. 2.19) potentials.

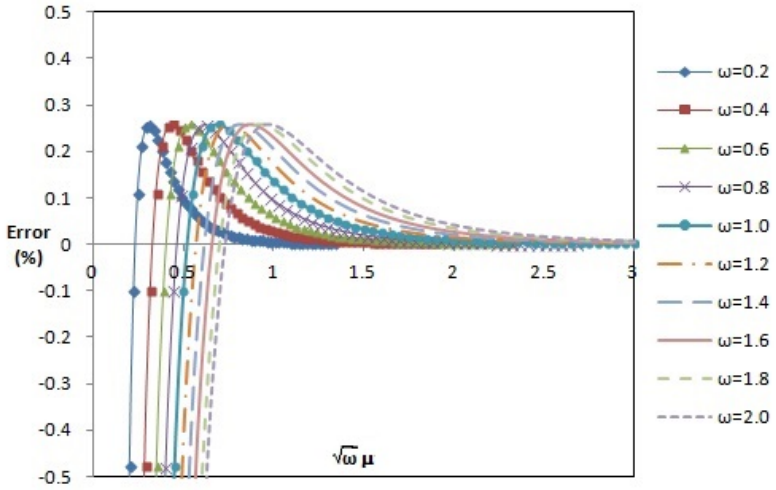

Figure 2.11: Errors (in %) of the Harmonium eigenvalues as function of $\sqrt{\omega}\mu$, for different ω using the scaled asymptotic [Eq. (3.5)] potential.

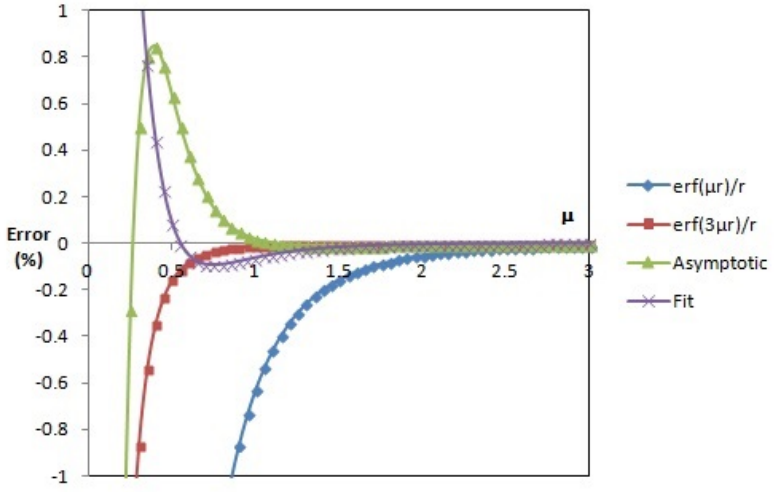

Figure 2.12: Errors (in %) of the excitation energy to the first excited $l = 0$ state of Harmonium, as function of μ, for $\omega = 1$. Long-range $\mathrm{erf}(\mu r)/r$ (diamonds), $\mathrm{erf}(3\mu r)/r$ (squares), asymptotic [Eq. (3.5)] (triangles), and fitted [Eq. (2.19)] (stars) potentials.

Table 2.1: Smallest value of the range-parameter μ needed to obtain a percentage error less than 1% with the model potentials for Harmonium, for a given value of ω.

ω	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
Asymptotic	0.2	0.25	0.35	0.4	0.45	0.5	0.5	0.55	0.6	0.6
Fitted	0.4	0.25	0.4	0.5	0.55	0.6	0.65	0.65	0.65	0.65

parameters of the model potential let us scale the variable in Eq. (2.28) with Hamiltonian (2.30) to reduce the confinement parameter to $\omega = 1$. After the substitution $\rho = \sqrt{\omega}\, r$. We get

$$\left[-\nabla_\rho^2 + \frac{\rho^2}{4} + \frac{1}{\sqrt{\omega}}\left(\frac{\mathrm{erf}(\tilde{\mu}\rho)}{\rho} + \tilde{c}\,\mathrm{e}^{-\tilde{\alpha}^2\rho^2}\right)\right]\phi_{nl}(\rho) = \tilde{\epsilon}_{nl}\phi_{nl}(\rho), \tag{2.35}$$

where

$$\tilde{c} = c/\sqrt{\omega}, \quad \tilde{\alpha} = \alpha/\sqrt{\omega}, \quad \tilde{\mu} = \mu/\sqrt{\omega}, \quad \tilde{\epsilon}_{nl} = \epsilon_{nl}/\omega. \tag{2.36}$$

Thus, to compensate for changing ω we have to properly scale parameters and multiply the potential by $\sqrt{\omega}$. In Fig. 2.11 we can see that by appropriately scaling the parameters of the model potential we get, for all values of ω, exactly the same energies.

To show that the same method can be used for excited states, we computed the excitation energy from the ground state to the first excited $l = 0$ state, using the long-range, erf(3μr)/r, asymptotic, and fitted potentials (see Figure 2.12). There is some cancellation of errors (i.e., the energy spacing is better than the absolute energy), but the results are still poor for small values of μ, confirming that the parameters in the model potential should be ω-dependent. We should note, however, that the erfgau potential with fixed parameters is still much better than the raw $\mathrm{erf}(\mu r)/r$ potential.

2.5 Uniform electron gas

We now examine the effect of using a modified potential on the energy of the uniform electron gas. Consider the Hartree-Fock energy of the N-particle spin-unpolarized uniform electron gas confined in the volume Ω with density $\rho = N/\Omega$, in the limit where N and Ω go to infinity at constant ρ. The one-electron reduced density matrix has the well-known form[JM73]

$$\gamma(\mathbf{r},\mathbf{r}') = 3\rho\frac{\sin(x) - x\cos(x)}{x^3}\,, \quad \text{with } x = k_F|\mathbf{r}-\mathbf{r}'|, \tag{2.37}$$

which is not affected when we replace the Coulomb interaction, both attractive and repulsive, with the modified interaction (Eq.(3.4)), because it depends only on the Fermi wavenumber $k_F = (3\pi^2\rho)^{1/3}$. Furthermore, the compensation of the electrostatic terms is maintained (*i.e.* the electrostatic contribution sums up to zero, just as in the standard case). However, the exchange energy is modified to

$$E_x = -\frac{1}{2}\int\int d\mathbf{r}\, d\mathbf{r}'\, \gamma(\mathbf{r},\mathbf{r}')^2 V_\mu(|\mathbf{r}-\mathbf{r}'|). \tag{2.38}$$

Using $\int \rho\, d\mathbf{r} = N$ and transforming the variables of integration, one obtains

$$E_x = -\frac{6}{\pi^2}N\int_0^\infty x^2\left(\frac{\sin(x) - x\cos(x)}{x^3}\right)^2 V_\mu\left(\frac{x}{k_F}\right)dx. \tag{2.39}$$

Replacing V_μ with the erfgau form of interest to us, Eq. (3.4), we can then separate the integral into two terms, the Gaussian function term and the error function term

$$\begin{aligned}
&\int_0^\infty x^2 \left(\frac{\sin(x) - x\cos(x)}{x^3}\right)^2 V_\mu\left(\frac{x}{k_F}\right) dx = \\
&\int_0^\infty x^2 \left(\frac{\sin(x) - x\cos(x)}{x^3}\right)^2 c\, e^{-\alpha^2(\frac{x}{k_F})^2} dx \\
+&\int_0^\infty x^2 \left(\frac{\sin(x) - x\cos(x)}{x^3}\right)^2 \frac{\mathrm{erf}(\mu\,\frac{x}{k_F})}{\frac{x}{k_F}} dx.
\end{aligned} \tag{2.40}$$

Both integrals can be easily evaluated using standard tools for numerical computations such as Mathematica[Mat].

In Figure 2.13 we show the exchange energy per particle $\epsilon_x = E_x/N$, as function of μ and the density parameter $r_s = (3/4\pi\,\rho)^3$, using the asymptotic potential (Eq.(3.5)). We observe that the model works well for large $\mu\, r_s$, but it does not seem possible to correct the interaction at short range. Here, it is important to notice the similarity with Harmonium. As ω controls the distance between the electrons, r_s describes the electron density distribution. A small value of ω translates to short interparticle distances, making the gas "denser", and as consequence, difficult to describe with the smooth potentials. This indicates that the optimal value of μ, just as in Harmonium, depends on the range of the interaction, r_s, so μ should be system-dependent.

2.6 Summary

Is it possible to replace the Coulomb potential with another potential that is computationally more convenient and, if so, how should the approximate potential be constructed? In this work we examine if adding a Gaussian function improves the

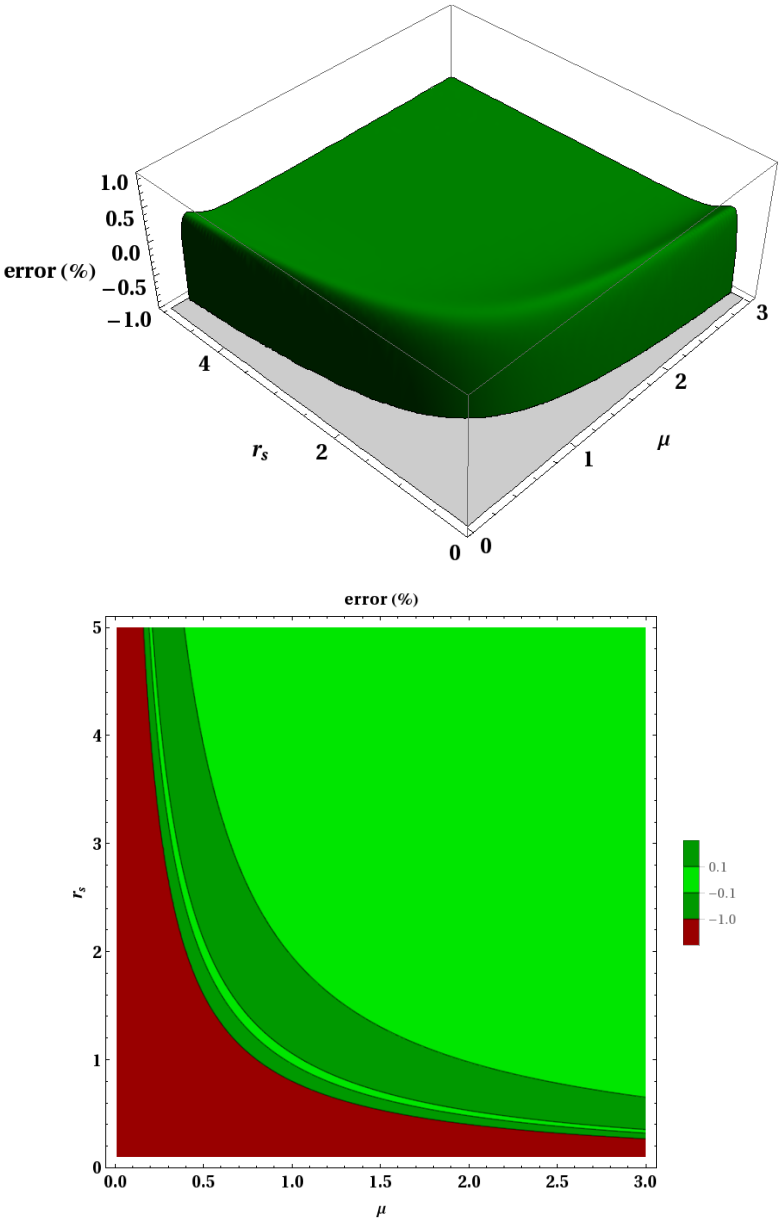

Figure 2.13: Error (in %) in $\epsilon_x(r_s, \mu)$, for the uniform electron gas using the model asymptotic potential (Eq. 3.5).

performance of the traditional $\mathrm{erf}(\mu r)/r$ potential used in range-separated DFT. As the measure of the correctness of the model potential we have chosen the difference between the spectra of the Hydrogen atom calculated using the Coulomb potential and the modified one. It appears that for a reasonable range of parameters defining the new potential not only can the spectrum of the Hydrogen atom be accurately reproduced, but using the same potential to replace the *repulsive* Coulomb potential in the Harmonium atom gives a significant improvement over the uncorrected $\mathrm{erf}(\mu r)/r$ potential.

The remaining question is whether one could somehow correct the residual error in the new potential. One way to do this would be to, as in range-separated DFT, use a correction functional for the neglected short-range contributions to the exchange-correlation energy. However, this biases one's treatment towards the ground-state energy and electron density: a different (and certainly much harder to construct) functional would be needed to correct other properties (e.g., excited-state properties) of the system. There is another way, however: the results of a few calculations at sufficiently large values of μ can be extrapolated to the physical $\mu \to \infty$ limit. This approach is applicable to any property, not just those that are readily accessible from KS DFT. Furthermore, replacing the electron-electron repulsion potential with a smooth function has major computational advantages, as it allows one to use smaller basis sets, with fewer polarization functions.

There are also cases where it may be favorable to replace the Coulombic electron-nuclear interaction with a model potential like those considered in this paper. For example, these smoothed Coulomb potentials could be used, instead of pseudopotentials, for diffusion quantum Monte Carlo and plane-wave DFT calculations. In those

cases, the procedure would be the same: the system would be solved for several choices of the smoothed electron-nuclear interaction, and the results then extrapolated to the physical $\mu \to \infty$ limit.

Chapter 3

Accuracy and efficiency of smooth Coulomb-like model potentials

3.1 Introduction

[1]Orbital-based methods are widely used in many-body electronic structure theory for studies of strongly correlated systems. The disadvantage of these methods is that as the number of orbitals increases the computational cost scales rapidly while the precision improves rather slowly [PWD94, HKKN97, Tru98, DGY$^+$06, HPKW09]. The main reason for this slow improvement is related to the challenge of describing correctly the electron-electron correlation cusp. The correlation cusp results from the divergence in the Coulomb interaction. In practice, there are two ways to deal with this problem: 1) use explicitly-correlated wave functions, or 2) eliminate the singularity from the Hamiltonian. On one hand, explicitly-correlated methods yield very

[1]This Chapter is a manuscript to be submitted for publication, and some details of the introduction may seem redundant.

compact wave functions (fast convergence), but the resulting Hamiltonian is much more complicated, involving three- and four-electron integrals. Although R12/F12 methods reduce the three- and four-electron integrals to two-electron integrals by the resolution-of-identity (RI), the majority of commercial packages for electronic structure calculations are not designed to support these methods and the cost of performing an accurate resolution of the identity is significant. On the other hand, one can avoid the correlation-cusp problem altogether by using smooth potentials or pseudopotentials to describe the electronic interaction. For instance, Lloyd-Williams *et al.* proposed a pseudopotential constructed to reproduce the scattering properties of the Coulomb potential [LWNC15], thereby obtaining a 30-fold decrease in computational cost. The speedup in the calculations is due to the cuspless shape of the resulting wave function and the fact that a smaller basis set is needed to reproduce it. Sirbu and King also demonstrated that one can achieve sub-μHartree accuracy [SK02, SK03] by accelerating the convergence of orbital-based methods. Their strategy was to separate the electronic interaction into long-range and short-range components and use Rayleigh-Schrödinger perturbation theory. The long-range (smooth) term was included in the reference Hamiltonian and the short-range term was treated using (high-order) perturbation theory. While the pseudopotential of reference [LWNC15] is a polynomial expansion, Sirbu and King used a combination of the standard error function and a Gaussian function,

$$\hat{V}_{\mathrm{SK}}^{\gamma} = \frac{\mathrm{erf}(\gamma r)}{r} + \frac{\gamma\left(7-(\gamma r)^2\right)}{4\sqrt{\pi}} e^{-(\gamma r)^2} \tag{3.1}$$

The coefficient and exponent of the Gaussian function were chosen to make certain perturbation terms disappear. This idea of partitioning the electronic interaction has

been used by Savin and co-workers for methods combining configuration interaction (CI) and density functional approximations (DFAs)[LSWS97]. Furthermore, Toulouse *et al.* designed a similar *erfgau* model, the error function plus a Gaussian term, where the parameters of the Gaussian function were selected to make the interaction vanish at $r \to 0$ [Tou04]. It was shown that the use of a soft interaction, together with a short-range energy functional, accelerates the convergence of the wave function method to the basis-set and full-CI limits.

Recently, we constructed a smooth model potential with a similar form. Our approach is based on the idea of replacing the physical Hamiltonian

$$\hat{H} = -\frac{1}{2}\nabla^2 + V_{ne} + V_{ee} \tag{3.2}$$

with a model Hamiltonian of the form

$$\hat{H}^{\mu} = -\frac{1}{2}\nabla^2 + V_{ne} + V_{ee}^{\mu} \tag{3.3}$$

where

$$\hat{V}_{ee}^{\mu} = \sum_{i<j} \frac{\mathrm{erf}(\mu\, \mathrm{r}_{ij})}{\mathrm{r}_{ij}} + c(\mu) e^{-\alpha(\mu)^2 \mathrm{r}_{ij}^2}; \quad c(\mu) = c'\mu, \quad \alpha(\mu) = \alpha'\mu, \tag{3.4}$$

with the parameters c' and α' chosen so that the leading terms of the asymptotic expansion of the difference between the physical Hamiltonian and the model Hamiltonian vanish when $\mu \to \infty$ [GEAKS16]. The values of the parameters are:

$$c(\mu) = \frac{27}{8\sqrt{\pi}}\,\mu \quad \alpha(\mu) = \frac{3}{2}\,\mu. \tag{3.5}$$

It is known that the error in the energy, $E(\mu) - E(\infty)$, decays as μ^{-2} when the

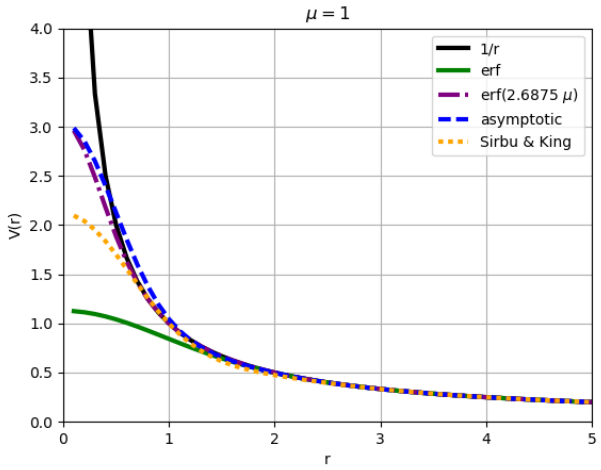

Figure 3.1: The effect of having a Gaussian function to the shape of the potentials. The shape of the potential becomes sharper after the addition of a Gaussian function, becoming closer to the real interaction. The error function that coincides with the asymptotic erfgau potential at $r = 0$ is obtained when $\mu_{\text{erf}} = 2.6857\mu_{\text{asymp}}$.

model potential is the error function [Tou04]. For the asymptotic erfgau potential in equations (3.4) and (3.5), the addition of the Gaussian makes the potential sharper, closer to the Coulomb potential (see figure 3.1), and it is expected to have faster improvement at smaller μ. This was confirmed by examining the energies of some two-electron systems. Electronic energies were computed for H^-, He, Li^+, and Be^{2+}, with the full configuration interaction method (FCI) and Dunning aug-cc-pVXZ basis sets [Dun89, PWD94], X=D,T,Q,5. All the FCI energies were obtained using the PyCI library [refa], the Hartree-Fock one- and two- electron integrals were computed with HORTON 2.0.1 package [refb]. The energy errors are shown in figure 3.1. It is clear that the error of the asymptotic erfgau model converges faster than the error function alone, and the value of μ necessary to obtain a specified accuracy is always

smaller for the asymptotic model. However, the slow convergence with respect to the basis set is still a problem. The local-character of the range-parameter becomes evident when we put together the errors of all the systems (see figure 3.1). This can be partially corrected by scaling the parameter $\mu = \mu Z_{eff}$. Toulouse *et al.* observed that there is no universal value of μ for every system, as it defines the interaction range over which the LDA exchange-correlation effects are transferable from the homogeneous system to the real one [TCS05]. They proposed some choices for an effective local $\mu(\mathbf{r})$, proportional to the density and the gradient of the density of the inhomogeneous system [Tou05] obtaining a considerable correction to the short-range LDA correlation energy. The importance using a local parameter $\mu(\mathbf{r})$ is subject to further investigation but it is outside of the scope of this work.

As the range-separation parameter becomes smaller, the interaction becomes flatter; the curvature of the wave function as $|\mathbf{r}_i - \mathbf{r}_j| \to 0$ becomes flatter too, which makes it easier to approximate. Nevertheless, as we saw from the FCI results, the errors worsen as μ approaches zero. Therefore, to be able to take advantage of the softness of the model interactions one needs to correct for the imperfection of the interaction at short-range. We shall consider two strategies to mitigate the errors when using model potentials: 1) adding a density functional for the short-range part of the electronic interaction (Savin *et al*), and 2) adding a correction by means of perturbation theory (Sirbu and King). In the next section, we revise the short-range corrections from LDA functionals and propose an analogous approach using Hartree-Fock potentials. In section 3 we make use of a first-order perturbation correction. Although the perturbation correction has a similar effect to the short-range correction, with the perturbation-theory correction the error decays more rapidly with respect

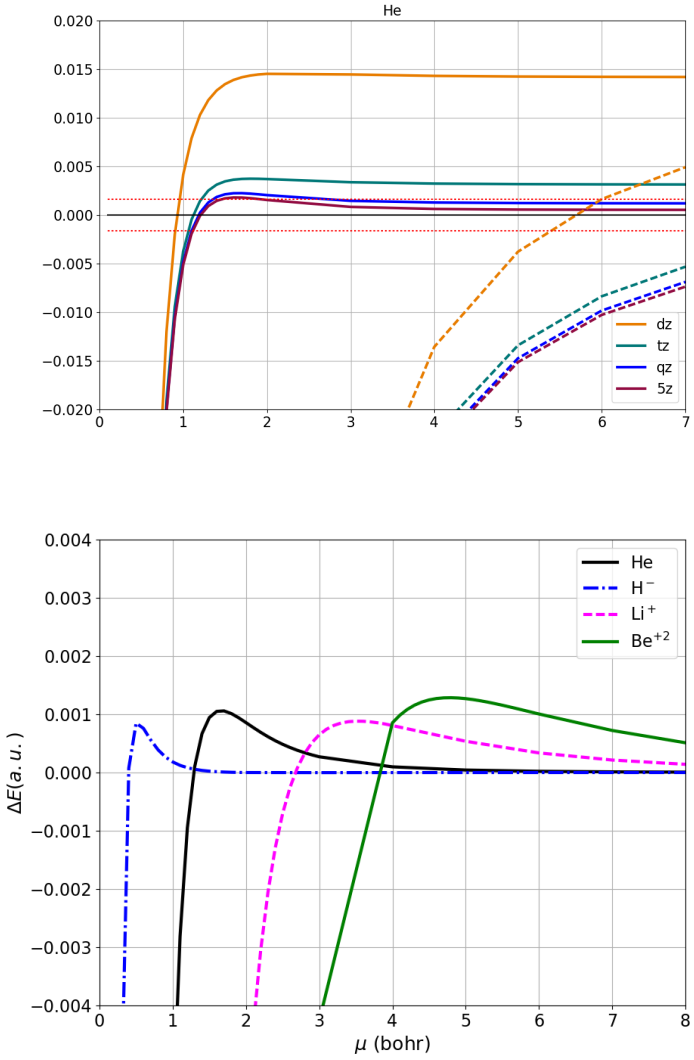

Figure 3.2: Errors in energy ($\Delta E = E^{\mu} - E^{\infty}$) for the asymptotic erfgau potential with two-electron systems. At the top, the basis-set dependence of the FCI energy for He; the dashed lines correspond to the standard error function potential and the solid lines refer to the *erfgau* (asymptotic) potential. At the bottom, the first four elements of the iso-electronic series of He demonstrate the dependence of the asymptotic model to the effective nuclear charge.

to the range-separation parameter than with the short-range correction alone. In section 4, we propose a method to construct model potentials by variational perturbation theory. The meaning of the range-separation parameter, as well as the importance of the shape of the model potentials to the convergence rate of the wave function expansion is addressed in section 5. Finally, conclusions are presented in section 6.

3.2 Short-range corrections

Leinninger *et al.*[LSWS97] showed that it is possible to tackle the convergence problem of wave-function methods by combining them with short-range density functionals. Furthermore, an improvement over the wave-function result was observed. This surprising result can be attributed to the fact that the combined method recovers some of the dynamic correlation associated with the correlation cusp from the uniform electron gas while maintaining the capabilities of the multi-determinant wave function for modeling strong-correlation effects. In CI-DFA combined methods the energy is defined as

$$E^{(\mu,sr-LDA)} = \langle \Psi^\mu | \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^\mu | \Psi^\mu \rangle + U^{sr,\mu}[\rho_{\Psi^\mu}] + E_{xc}^{sr,\mu}[\rho_{\Psi^\mu}] \tag{3.6}$$

where Ψ^μ comes from solving

$$\hat{H}^\mu \Psi^\mu = E^\mu \Psi^\mu \tag{3.7}$$

and $\hat{H}^\mu$ is

$$\hat{H}^\mu = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{lr,\mu} + \frac{\delta U^{sr,\mu}[\rho_\Psi]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}^{sr,\mu}[\rho_\Psi]}{\delta \rho(\mathbf{r})}. \tag{3.8}$$

Equation (3.6) has to be solved self-consistently. The short-range exchange-correlation functional for the error function interaction was taken from the analytic expression proposed by Paziani *et al.* [PMGGB06]. In figure 3.3 we present the energy errors of the corrected model with respect to the reference [DHC91]. It is clear that including the short-range correction greatly improves the energy.

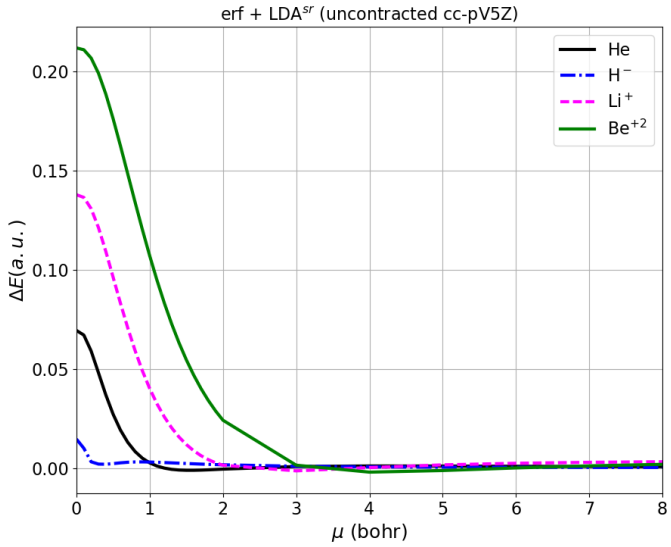

Figure 3.3: Energy errors of the error function with the short-range LDA correction for the iso-electronic series of He with the uncontracted cc-pV5Z basis. An important improvement in the energies comes from the correlation effects transferred from the short-range functional.

One can use the asymptotic model (3.4) instead of the error function. For the asymptotic potential, the correlation energy functional was computed by modifying the uniform electron gas functional of Freeman [Fre77] with the modified *erfgau* potential. The exchange term was obtained by analytic integration of equation 3.36 in

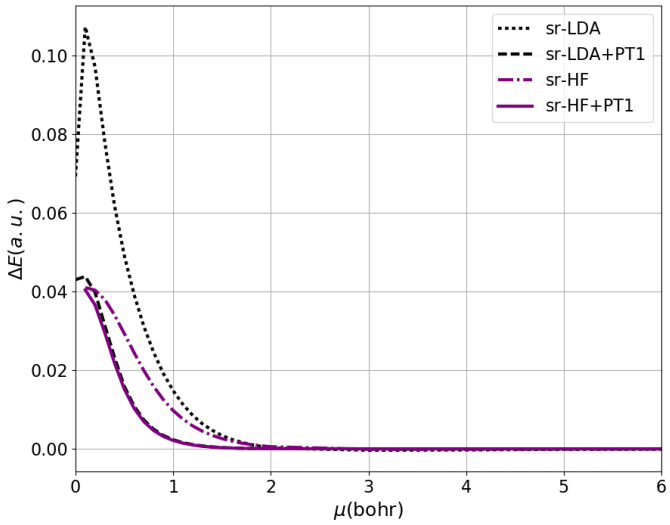

Figure 3.4: Errors in energy of $\Delta E = E^{(\mu,sr-\mathrm{LDA/HF})} - E^{\infty}$ (3.9) (top) and $\Delta E = E^{(\mu,sr-\mathrm{LDA/HF+PT1})} - E^{\infty}$ (3.21), where V_{ee}^{μ} is the asymptotic (ergau) model (3.5), for the He atom with the aug-cc-pVQZ basis. The short-range HF correction its better than the LDA short-range functional correction. The first-order perturbation correction accelerates the convergence of the errors with μ and seems preferable to a correction based on short-range DFT(sr-LDA).

[FW77] with the modified asymptotic erfgau potential. The details of how the short-range functional for the asymptotic model was obtained can be found in appendix B.

Contrary to what we expected, the LDA-corrected asymptotic erfgau model was not a great improvement over the traditional LDA-corrected error function model (fig. 3.4). We believe this can be attributed to the approximate strategy used to calculate the short-range correlation energy, and expect that better results would be obtained using a functional parametrized by Quantum Monte Carlo calculations (as

was done to parameterize the error function sr-LDA functional), instead of coupled-cluster theory.

Alternatively, we could just add the short-range Hartree and exchange potentials from the Hartree-Fock method to $\hat{H}^{\mu}$ instead of the short-range density functional,

$$\hat{H}^{(\mu\,,sr-HF)} = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{lr,\mu} + v_h^{sr,\mu}[\rho_{HF}] + v_x^{sr-HF,\mu}[\rho_{\mathrm{HF}}] \tag{3.9}$$

where $v_h^{sr,\mu}[\rho_{HF}]$ and $v_x^{sr-HF,\mu}[\rho_{\mathrm{HF}}]$ are the short-range Hartree and exchange potentials from the Hartree-Fock density, respectively. This equation (3.9) can be rewritten as:

$$\hat{H}^{(\mu\,,sr-\mathrm{HF})} = \left(\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + v_{ne}(\mathbf{r}_i) + j^{sr,\mu}(\mathbf{r}_i) + k^{sr,\mu}(\mathbf{r}_i)\right) + \hat{V}_{ee}^{lr,\mu}. \tag{3.10}$$

Both terms are added into the one-electron integrals in the AO basis in the form of a Fock matrix:

$$h_{ac} = h_{ac}^{\mathrm{core}} + \sum_d \sum_b P_{db}^{tot}\langle ab|cd\rangle - P_{db}^{\beta}\langle ab|dc\rangle, \tag{3.11}$$

being:

$$j^{sr,\mu} = \sum_d \sum_b P_{db}^{tot}\langle ab|cd\rangle \tag{3.12}$$

$$k^{sr,\mu} = \sum_d \sum_b P_{db}^{\beta}\langle ab|dc\rangle. \tag{3.13}$$

Note that we only work in the closed-shell case so:

$$P_{db}^{\alpha} = P_{db}^{\beta} = \frac{1}{2} P_{db}^{tot}. \tag{3.14}$$

After solving $\hat{H}^{(\mu,sr-x)}\Psi^{(\mu,sr-x)} = E^{(\mu,sr-x)}\Psi^{(\mu,sr-x)}$, where $\hat{H}^{(\mu,sr-x)}$ could be either $\hat{H}^{(\mu,sr-\mathrm{LDA})}$ or $\hat{H}^{(\mu,sr-\mathrm{HF})}$. The final energy is corrected by subtracting the short-range potentials and adding the LDA/HF short-range energies

$$\begin{aligned} E^{(\mu,sr-x)} =& \langle \Psi^{\mu} | \hat{H}^{(\mu,sr-x)} | \Psi^{\mu} \rangle \\ & - \langle \Psi^{\mu} | v_H^{sr,\mu}[\rho_{\Psi^{\mu}/\mathrm{HF}}] + v_x^{sr,\mu}[\rho_{\Psi^{\mu}/\mathrm{HF}}] | \Psi^{\mu} \rangle \\ & (- \langle \Psi^{\mu} | v_c^{sr,\mu}[\rho_{\Psi^{\mu}}] | \Psi^{\mu} \rangle) \\ & + U^{sr,\mu}[\rho_{\Psi^{\mu}}] + E_x^{sr,\mu}[\rho_{\Psi^{\mu}}] \\ & (+ E_c^{sr,\mu}[\rho_{\Psi^{\mu}}]) \end{aligned} \tag{3.15}$$

When $\mu = 0$ we recover the LDA/HF energy. The main difference between these short-range corrections is that the density-functional corrected model introduces extra stabilization from the short-range correlation energy functional. This is missing in HF. Therefore, with the HF correction, the best one can do is the FCI limit, and eq. (3.15) connects the mean-field result and the correlated one. Whether one uses LDA or HF, the error function model always converges more slowly to the $\mu \to \infty$ limit than the asymptotic model. We also notice that the LDA-corrected method has a faster decay than its HF-equivalent. The advantage of the Hartree-Fock-FCI method over the density-functional is practical: it requires only one FCI calculation, while in the LDA-FCI method the FCI equations have to be solved repeatedly until the density used to evaluate the sr-LDA functional is self-consistent with the FCI

density.

We have demonstrated that a considerable improvement is achieved if some short-range interaction is included. If one uses a perturbation theory correction instead, is it comparably accurate? Or can one improve the results even further?

3.3 First-order perturbation correction

Inspired by the strategy of Sirbu and King, we can use the long-range interaction to define a reference Hamiltonian

$$\hat{H}^{(0)} = \hat{H}^{\mu} = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{lr,\mu} \tag{3.16}$$

and perform perturbation theory to correct for the difference between the Hamiltonian with the Coulomb interaction and the reference Hamiltonian

$$\hat{H}^{(1)} = \hat{H}^{sr} = \hat{H} - \hat{H}^{\mu} \tag{3.17}$$

Hence, the first-order corrected energy is

$$E^{PT1} = E^{(0)} + E^{(1)} = \langle \Psi^{\mu} | \hat{H}^{\mu} | \Psi^{\mu} \rangle + \langle \Psi^{\mu} | \hat{H}^{sr} | \Psi^{\mu} \rangle \tag{3.18}$$

$$= \langle \Psi^{\mu} | \hat{H} | \Psi^{\mu} \rangle \tag{3.19}$$

In practice, eq. (3.19) is evaluated using the 1- and 2-reduced density matrices (RDMs),

$$\langle \Psi^{\mu} | \hat{H} | \Psi^{\mu} \rangle = \sum_{pq} \gamma_{pq}^{\mu} h_{pq} + \frac{1}{2} \sum_{pqrs} \Gamma_{pqrs}^{\mu} g_{pqrs} + V_{ext} \tag{3.20}$$

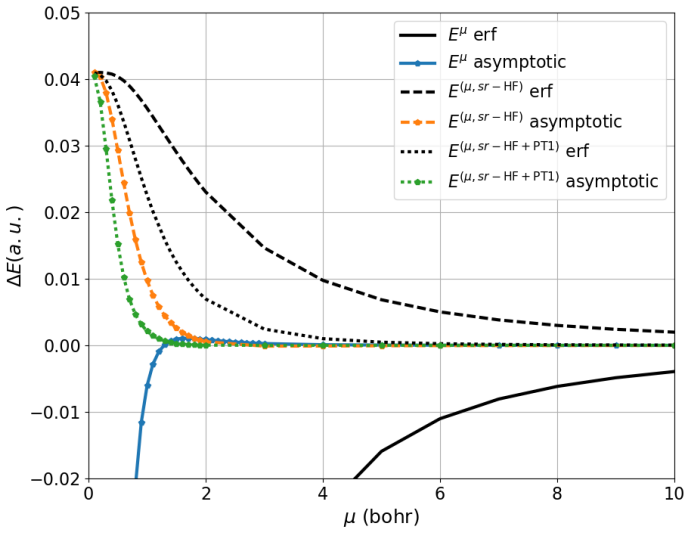

Figure 3.5: Comparison of the errors in energy of the model potentials with the two corrections: $\Delta E = E^{(\mu, sr-HF)} - E^{\infty}$ (3.9) and $\Delta E = E^{(\mu, sr-\mathrm{HF+PT1})} - E^{\infty}$ (3.21) for the He atom with the aug-cc-pVQZ basis, computed with FCI. The short-range HF correction defines the $\mu \to 0$ limit (HF energy), while the first-order perturbation improves the rate of convergence.

where γ^{μ} and Γ^{μ} are the 1- and 2-RDM from Ψ^{μ}, h_{pq} and g_{pqrs} are the 1- and 2-electron integrals, and V_{ext} is the external potential energy.

Notice that because the model wave function is an eigenfunction of the real Hamiltonian, we have an upper bound to the FCI energy. Furthermore, we can include a LDA/HF short-range potential in the model Hamiltonian. This correction is introduced indirectly to the wave function Ψ^{μ} and it's reflected in the evaluation of the expectation value of the real Hamiltonian:

$$E^{(\mu,\, sr-\mathrm{PT1})} = \langle \Psi^{\mu,sr-x} | \hat{H} | \Psi^{\mu,sr-x} \rangle \tag{3.21}$$

According to the results shown in figure 3.4, whether the short-range potential comes from LDA or HF doesn't seem to make a difference after evaluating the perturbation correction. Particularly, in figure 3.7 it is clear how the short-range correction changes limit at $\mu \to 0$ (blue line), and the perturbation correction accelerates the improvement of the results with μ (pink line). Hence, the best results are obtained by using a short-range correction to compute the wave function and first-order perturbation theory to compute the energy.

Applying the variational principle to the first-order energy is an alternative way to improve the model interaction. In other words, one can construct the model interaction so that the first-order perturbation correction is minimized. This is the subject of the following section.

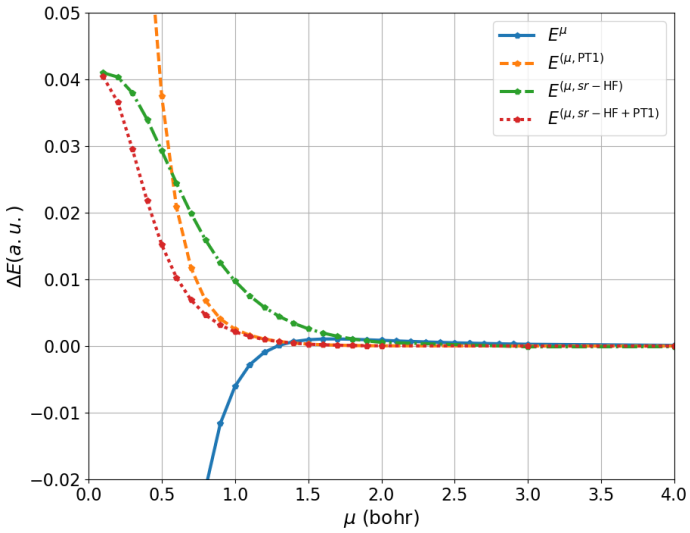

Figure 3.6: Dependence of the energy errors of the different models on the range-parameter μ with respect to the limit $\mu \to \infty$ in the same basis (aug-cc-pVQZ) using the asymptotic erfgau potential as computed with FCI. E^{μ} represent the energies calculated with the model potentials without any extra correction, $E^{(\mu,\, sr-\mathrm{HF})}$ is defined in (3.9), $E^{(\mu,\,\mathrm{PT1})}$ in (3.19), and $E^{(sr-\mathrm{HF}+\mathrm{PT1})}$ in (3.21) using the short-range HF correction.

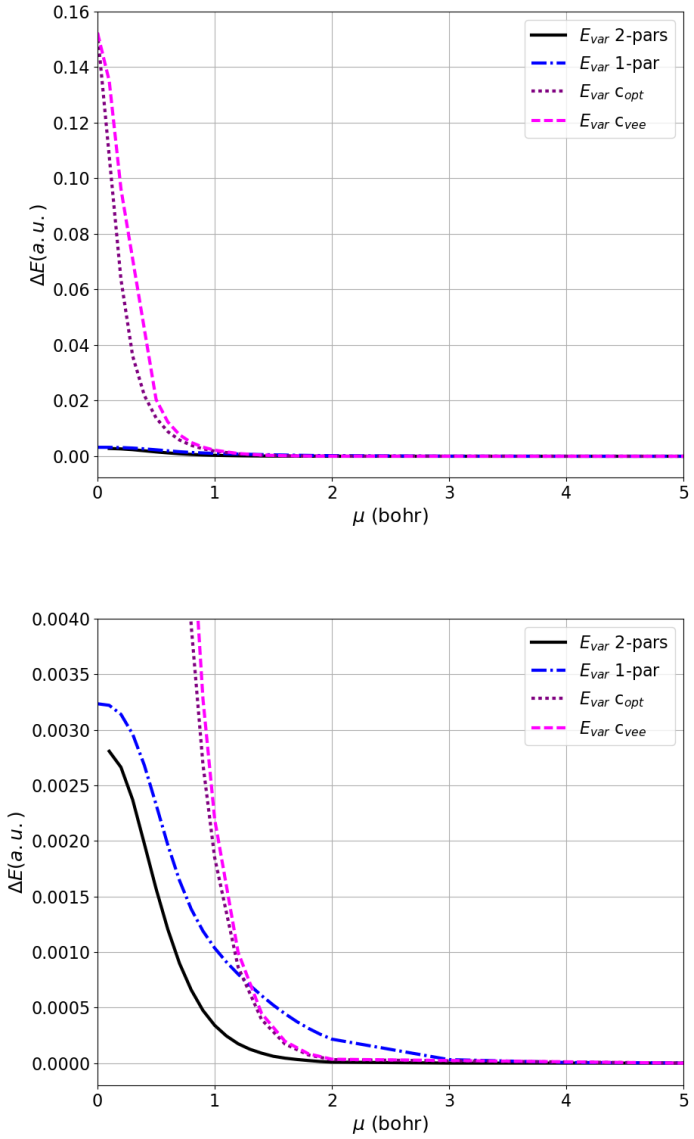

Figure 3.7: Comparison of the errors in the FCI/aug-cc-pVQZ energy with respect to the limit $\mu \to \infty$ with all the variational models for He atom. The models which optimize the exponent of the Gaussian function (equations (3.22) and (3.23) represented as E_{var} 2-pars and 1-par respectively) are much more accurate than the ones that optimize only the coefficient ($E_{var}c_{opt}$ for equation (3.25) and $E_{var}c_{v_{ee}}$ for (3.26)).

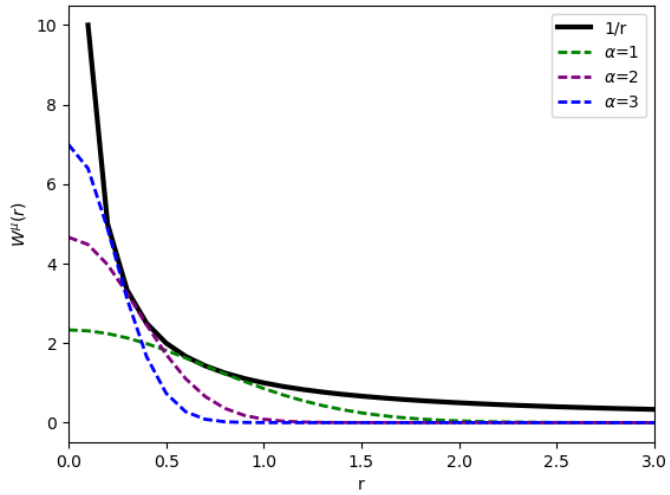

Figure 3.8: Form of the one-parameter model potential (eq. (3.23)) with three different values of α.

3.4 Variational models

Starting from a model Hamiltonian defined in equations (3.3) and (3.4), we would like to choose the Gaussian (*i.e.*, choose c' and α') that minimizes the expectation value of the physical Hamiltonian, E_{var}:

$$E_{\text{var}} = \min_{c',\alpha'} \langle \Psi^{\mu} | \hat{H} | \Psi^{\mu} \rangle. \tag{3.22}$$

where Ψ^{μ} comes from solving (3.7) with FCI.

This minimization is non-linear and contains many local minima. Finding a reasonable choice for α' and c' is easier if we reduce the number of parameters by replacing

Table 3.1: Energy errors in mHartree for the isoelectronic series of He using the one-parameter variational method (3.23) and the aug-cc-pVQZ basis.

μ	H^-	H_2	He	Li^+	Be^{+2}	C^{+4}
0.0	2.911	3.808	3.234	60.167	56.975	57.041
0.1	2.866	3.784	3.220	59.639	56.838	56.819
0.2	2.465	3.813	3.139	59.749	56.865	56.827
0.5	0.633	3.482	2.327	64.957	65.189	55.067
1.0	0.091	2.737	1.035	61.720	63.852	62.542
2.0	0.004	3.293	0.215	58.864	62.326	62.225
5.0	0.004	2.283	0.001	59.449	60.622	57.792

the long-range interaction with a similar model:

$$\hat{W}^{\mu} = \sum_{i<j} \frac{\mathrm{erf}(\mu\, \mathrm{r}_{ij})}{\mathrm{r}_{ij}} + \sqrt{2e}\, \alpha\, e^{-\alpha^2 \mathrm{r}_{ij}^2} \tag{3.23}$$

and variationally optimize only α in (3.22). The expression above is chosen so that the Gaussian functions touches the $1/r$ Coulomb potential from below for a given α:

$$ce^{\alpha^2 r^2} = 1/r \tag{3.24}$$

The interaction (3.23) is compared with the Coulomb potential in figure 3.8.

One can also keep α fixed and optimize $c(\mu)$. If we use the same strategy as before, assigning the same value of the asymptotic potential for the exponent, the energy is

$$E = \min_{c} \langle \Psi^{\mu}(c, \alpha = \alpha_{asymp}) | \hat{H} | \Psi^{\mu}(c, \alpha = \alpha_{asymp}) \rangle \tag{3.25}$$

Another possibility is to choose c so the perturbation correction is zero, *i.e.* the difference between the expectation value of the Coulomb potential and the one of the

Table 3.2: Optimized parameters, in atomic units, for the He atom using the one-parameter (1-par (3.23)), two-parameters (2-pars (3.22)) variational methods, and the methods that only optimized the coefficient of the Gaussian function, c_{opt} (3.25) and $c_{v_{ee}}$ (3.26).

μ	2-pars c'	2-pars α'	1-par α	c_{opt}	$c_{v_{ee}}$
0.1	22.99	10.35	1.34	38.02	1.16
0.2	11.54	5.27	1.39	9.70	1.11
0.5	4.82	2.48	1.73	3.16	1.26
1.0	2.83	1.79	2.56	2.13	1.97
2.0	2.16	1.58	4.51	1.91	3.77
5.0	1.86	1.49	10.38	1.89	9.49

model potential is zero:

$$\langle\Psi^{\mu}(c,\alpha=\alpha_{asymp})|\hat{V}_{ee}|\Psi^{\mu}(c,\alpha=\alpha_{asymp})\rangle-\langle\Psi^{\mu}(c,\alpha=\alpha_{asymp})|\hat{W}^{\mu}|\Psi^{\mu}(c,\alpha=\alpha_{asymp})\rangle=0 \tag{3.26}$$

A change in the exponent has a greater impact on the precision of the model, as it defines the shape of the Gaussian function. Models that optimize the exponent achieve higher precision than the ones optimizing only the coefficient. When we examine the error difference between all the variational models, figure 3.7, we observe how the models where the exponent is optimized have errors that are orders of magnitude smaller than the errors obtained with the models that optimize the coefficient.

The short-range correction can also be included in this case, either during the minimization or *a posteriori*, evaluating E^{sr-PT1} once the parameters for $\hat{V}^{\mu}_{ee}$ in $\hat{H}^{\mu}$ are determined. We will call this energy E^{sr-PT1}_{var}. In figure 3.9 we compare the errors of the models E_{var} from equations (3.22) and (3.23) with and without short-range HF correction.

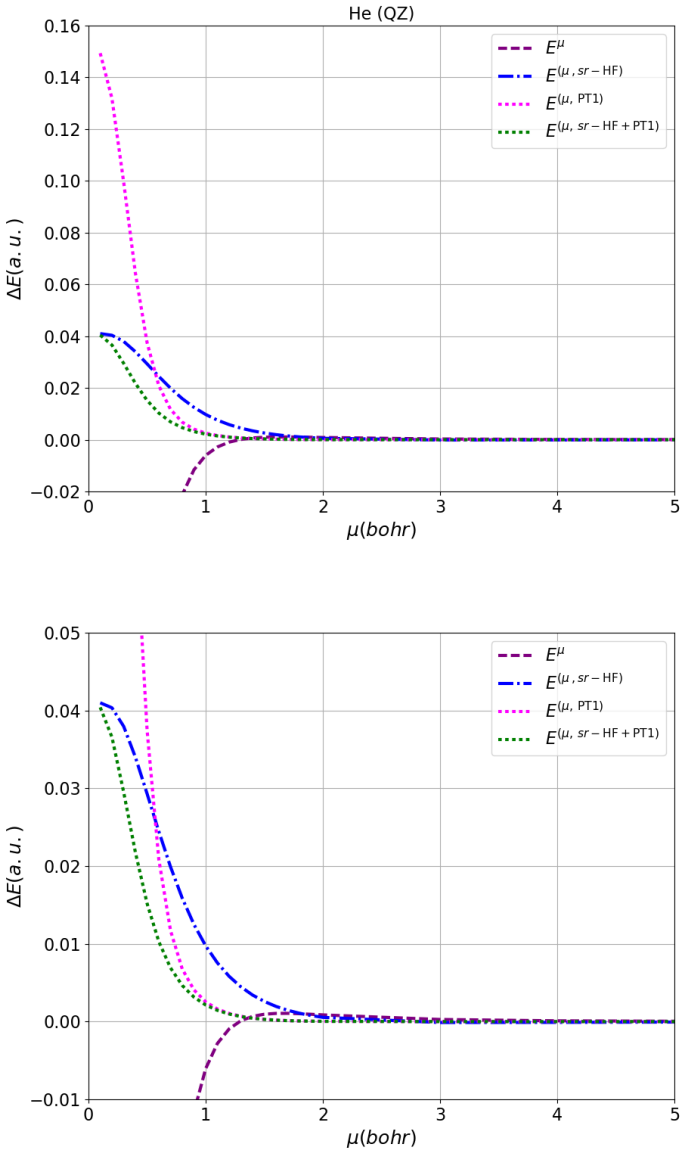

Figure 3.9: Dependence of the energy errors on the range-parameter μ, for He atom, using the variational method (eq. (3.22)) in FCI/aug-cc-pVQZ calculations with two different long-range potentials: 1) the two-parameter model in eq. (3.4), and 2) the one-parameter model in eq. (3.23). Solid lines are used for wavefunctions without short-range correction and dashed lines for wave functions with short-range correction. The second panel is a close-up of the first.

3.5 Effect of the model potentials on the wave function expansion

We now take a closer look into the effect of using the model potentials and corrections to the wave function. For instance, we already mentioned that the eigenfunction of a smooth Hamiltonian is also smooth; it doesn't have a cusp. This clearly accelerates the convergence with respect to the 1-electron basis set. Does it also accelerate the convergence of the CI expansion?

We can start by looking at the number of configurations with coefficients larger than some arbitrary value, for example $\epsilon = 10^{-3}$. We observe how as μ decreases, the number of "important" configurations also decreases (see figure 3.10.) It is apparent that the shape of the potential has an impact on the CI expansion. If one wants to reproduce the energy of the the optimized 2-parameter model, for example, one needs a good portion of the FCI expansion (figure 3.11). One can make a connection between the flatness of the interaction potential and the number of important electronic configurations. This makes more sense if we think of what is happening to the interaction between electrons: as $\mu \to 0$ the interaction becomes weaker, until at $\mu = 0$ just one configuration is needed to describe the positions of the electrons relative to the nuclei. This feature of smooth potentials can be relevant to algorithms for selecting configurations in order to reduce the computational cost. We use a Heat-bath CI, a semi-stochastic algorithm developed by Holmes et al. [HTU16, SHJ$^+$17] to examine the effect of the smooth models in the selection of determinants.

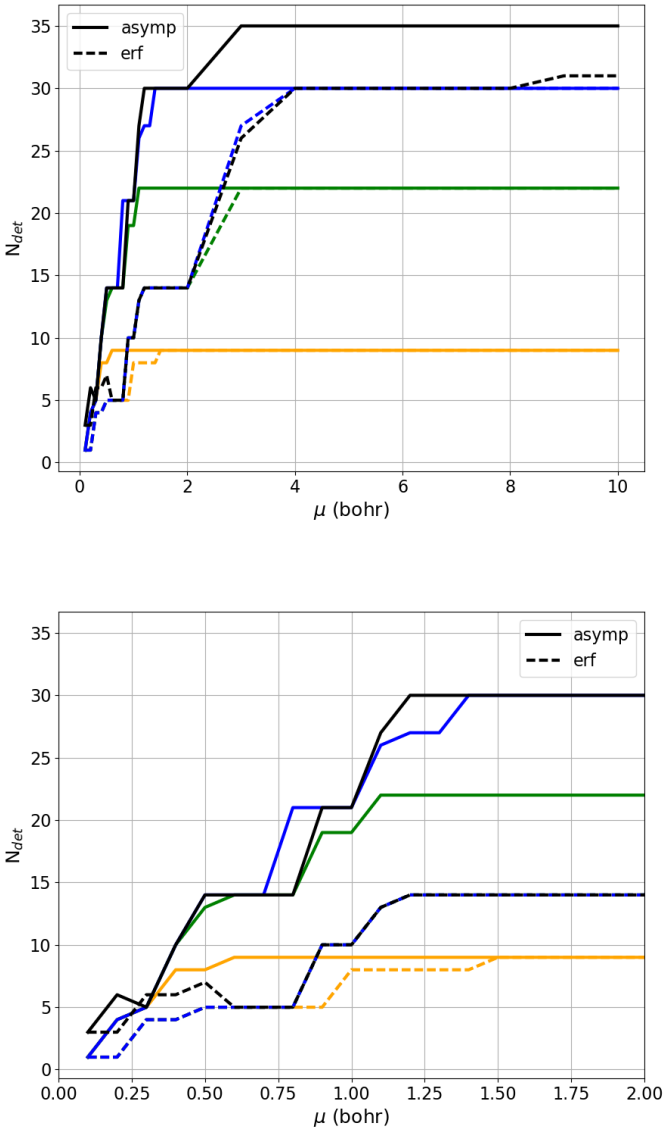

Figure 3.10: Number of determinants in the FCI expansion of He atom, in the natural orbital basis, that have coefficients greater than 10^{-3} using the asymptotic potential and the traditional error function potential. Different basis sets of the family aug-cc-pVXZ were used: DZ in yellow, TZ in green, QZ in blue and 5Z in black. The figure at the bottom is a close-up of the top panel.

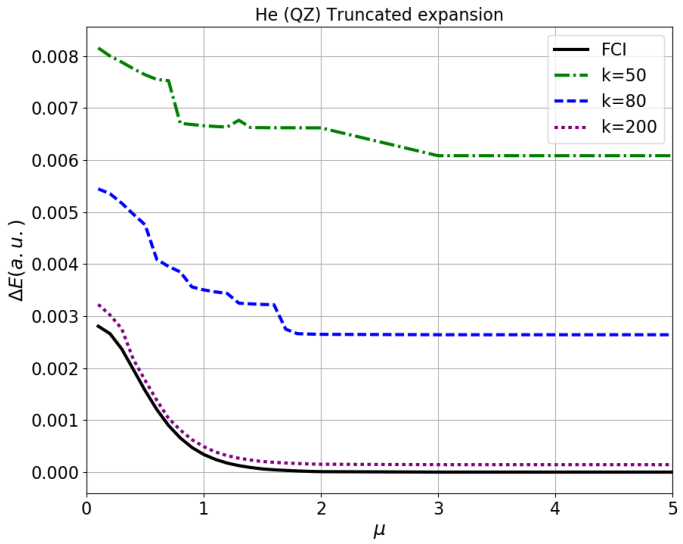

Figure 3.11: Energy errors of truncated CI expansions of the wave function for the He atom, using the variational optimized 2-parameter model E_{var} in equation (3.22). The model wave function is constructed after a FCI calculation with the aug-cc-pVQZ basis, using the first k determinants with larger coefficients. The energy is evaluated as the expectation value of this truncated expansion. A large number of determinants is needed to reproduce the wave function of the 2-parameter model.

Table 3.3: Energy errors (Hartree) relative to FCI and number of determinants used in SHCI calculations with the error function potential for different values of ϵ_1 for He atom with the aug-cc-pVQZ basis. Although a considerable reduction of the CI expansion is obtained when μ is small, there is a substantial loss of accuracy. The SHCI calculations converge to mHartree accuracy already for $\epsilon_1 = 0.01$.

μ	$\epsilon_1 = 10^{-2}$		$\epsilon_1 = 10^{-4}$		$\epsilon_1 = 10^{-8}$	
	ΔE	N_{det}	ΔE	N_{det}	ΔE	N_{det}
0.1	-0.9846	11	-0.9848	24	-0.9848	134
0.2	-0.8756	11	-0.8757	35	-0.8757	349
0.5	-0.5916	14	-0.5918	103	-0.5918	561
1.0	-0.2980	25	-0.2982	225	-0.2982	561
1.5	-0.1613	35	-0.1617	289	-0.1617	561
2.0	-0.0967	61	-0.0971	340	-0.0971	561
5.0	-0.0155	93	-0.0159	428	-0.0159	561
10.0	-0.0035	95	-0.0039	450	-0.0039	561

3.5.1 Combination with Semi-stochastic Heath-Bath Configuration Interaction method (SHCI)

The SHCI algorithm generates only the determinants connected to the reference determinant by Hamiltonian matrix elements that are larger than some threshold ϵ_1. A second-order Epstein-Nesbet perturbation energy correction is added using only the elements for which $|H_{ik}c_i|$ is greater than a second parameter ϵ_2.

We studied the convergence of energy and the size of the CI expansion first with the He atom using the best models obtained in previous sections. We first look at the effect of replacing the Coulomb potential with the smooth potentials, without any extra correction. In the interest of studying only on the models and corrections proposed in the present work, we set the parameter $\epsilon_2 = 0.1$ to not include the EN-PT correction. Tables 3.3 and 3.4 collect energy errors and number of determinants

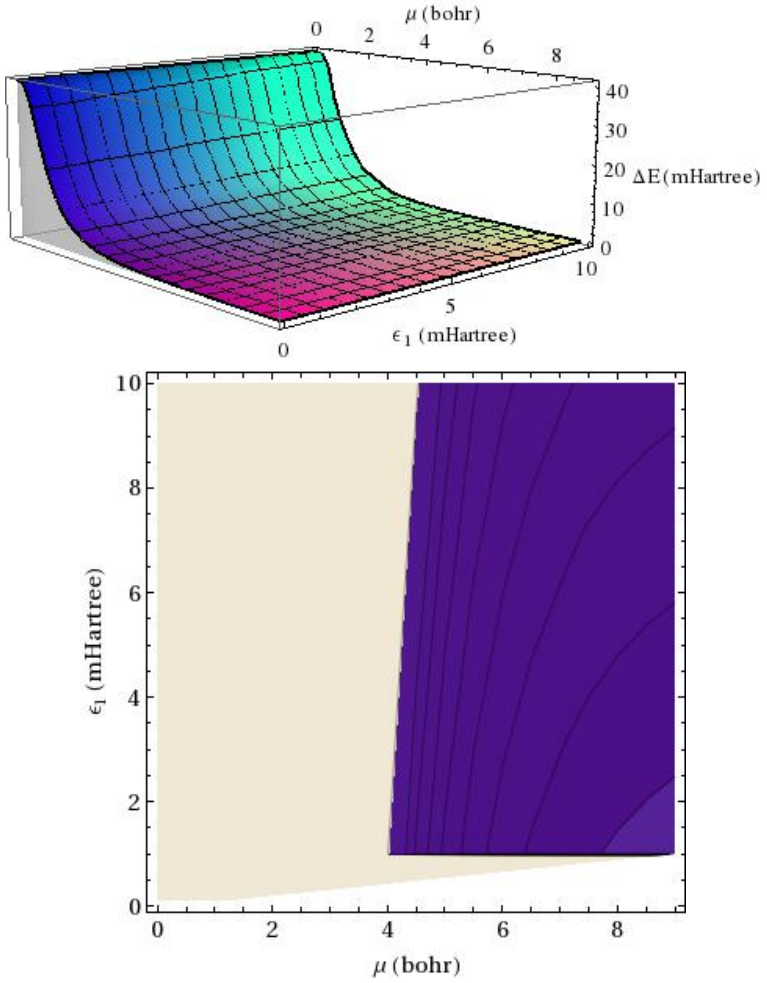

Figure 3.12: Dependence of the energy error on both parameters μ and ϵ_1 for the short-range plus perturbation $E^{(sr-\mathrm{HF+PT1})}$ corrected error function model. The contour levels of the plot at the bottom are spaced by 0.1 mHartree. The model energy of the He atom is compared with the FCI result at the same basis set (aug-cc-pVQZ). High precision (1 mHartree) is obtained only for $\mu > 4$.

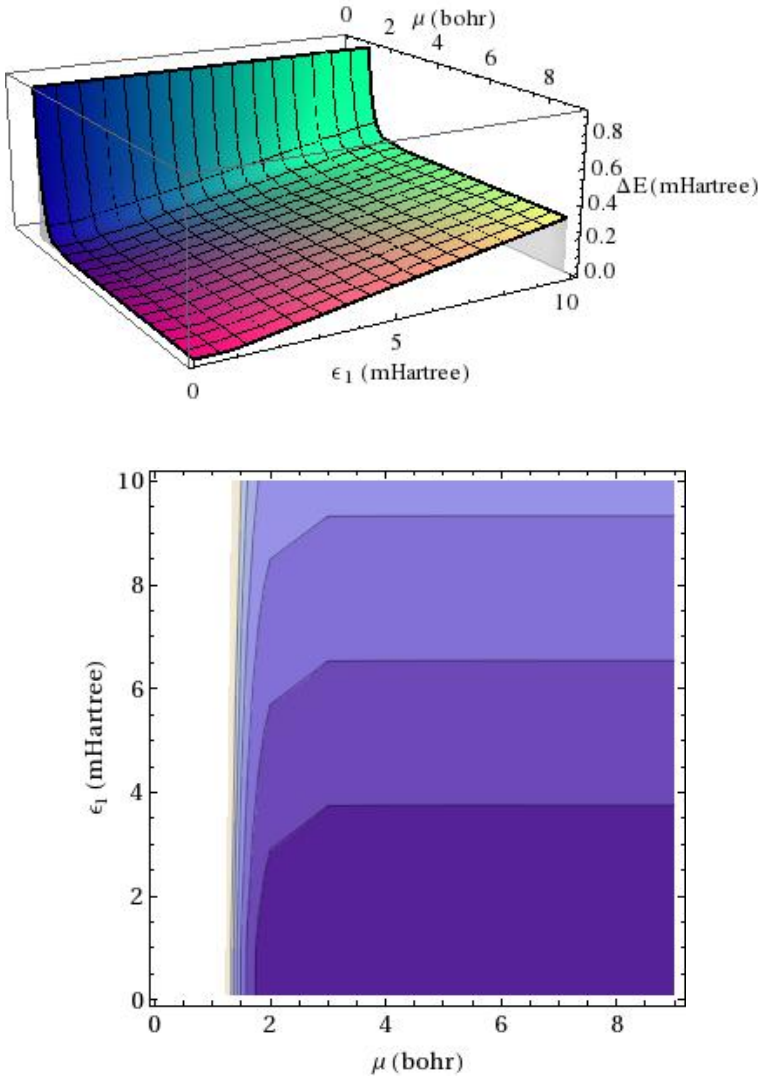

Figure 3.13: Dependence of the energy error on both parameters μ and ϵ_1 for the short-range plus perturbation $E^{(sr-\text{HF-PT1})}$ corrected asymptotic model. The contour levels of the plot at the bottom are spaced by 0.1 mHartree. The model energy of the He atom is compared with the FCI result with the same basis set (aug-cc-pVQZ). Compared with the error function corrected model 3.12, the asymptotic model improves much faster, in both ϵ_1 and μ.

Table 3.4: Energy errors (Hartree) relative to FCI and number of determinants used in SHCI calculations with the asymptotic potential for different values of ϵ_1 for He atom with the aug-cc-pVQZ basis. The FCI calculation contains 561 number of determinants but (normally) fewer determinants suffice at small μ and moderately small ϵ_1.

μ	$\epsilon_1 = 10^{-2}$		$\epsilon_1 = 10^{-4}$		$\epsilon_1 = 10^{-8}$	
	ΔE	N_{det}	ΔE	N_{det}	ΔE	N_{det}
0.1	-0.8005	11	-0.8006	38	-0.8006	302
0.2	-0.5409	14	-0.5411	86	-0.5411	561
0.5	-0.1175	35	-0.1178	286	-0.1178	561
1.0	-0.0057	84	-0.0060	431	-0.0060	561
1.5	-0.0012	89	-0.0009	453	-0.0009	561
2.0	-0.0012	88	-0.0008	456	-0.0008	561
5.0	-0.0004	88	$-4.25\text{x}10^{-5}$	456	$-4.25\text{x}10^{-5}$	561
10.0	-0.0003	88	$-3.02\text{x}10^{-6}$	455	$-3.02\text{x}10^{-6}$	561

obtained with the error function potential and the asymptotic erfgau potential at several values of μ. For both methods, energetic convergence is achieved at $\epsilon = 10^{-4}$, and the asymptotic erfgau model is more precise. In figures 3.12 and 3.13 one can see that the number of determinants required by the asymptotic erfgau potential increases faster with respect to both μ and ϵ_1. These results confirm what we discussed previously: there is a compromise between accuracy and computational cost that depends on the shape of the model potential, and hence, on how closely the model resembles the physical interaction. Now, let us consider what happens when a short-range correction is added. In this case, the limit of $\mu = 0$ becomes the mean-field limit, instead of the noninteracting limit, which already makes a big difference, improving the results over the uncorrected model not only in terms of energy but reducing the number of determinants needed. Figure 3.14 shows how the length of the CI expansion depends on the choice of the model electronic-interaction. We used the He atom to

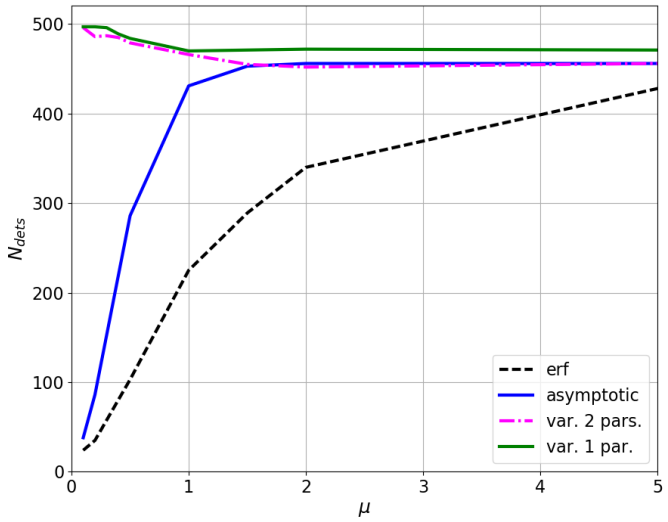

Figure 3.14: The number of determinants used by SHCI using $\epsilon_1 = 100\mu$Hartree using 1)the error function long-range interaction, 2) the asymptotic long-range interaction, 3) the variational two-parameter model (3.22), and 4) the variational one-parameter model 3.23. Although the variational optimized models are very accurate they are just as expensive as the calculation with the physical (Coulomb) interaction.

compare the two soft long-range interactions obtained by variationally optimizing the Gaussian functions. As expected these are the most accurate model potentials, but they are also by far the most expensive computationally because the variationally optimized Gaussian is very pointy (c.f. Table 3.2).

3.5.2 C_2 dissociation

The effect of the model potentials and the correction depends also on the atom or molecule studied. We would like to be able to describe correctly larger atoms and molecules. In particular, for multi-determinant methods, it is important to describe

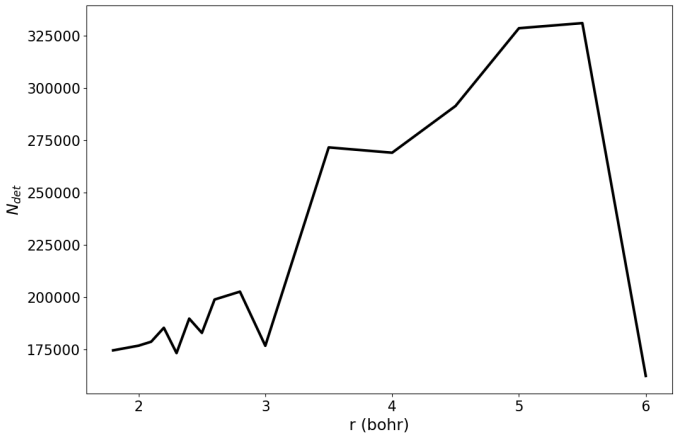

Figure 3.15: The number of determinants used in SHCI with ϵ_1 =100μHartree and the cc-pVDZ basis along the dissociation of C_2, using the physical Coulomb potential. The number of determinant increases substantially as the bond breaks. Once the atoms are dissociated the number of determinants drops again.

correctly the correlation energy, especially the strong correlation. In this section we assess the accuracy and cost of the soft potentials with the Carbon dimer. The most difficult calculations are at bond-breaking distances. These are also the most expensive calculations (see figure 3.15). We can take one of these distances, for example $r = 4.5$ to compare the erf and erfgau models. When we looked at the CI expansion of the short-range corrected models, we observed instability of the erf-srLDA model when combined with SHCI, particularly at small μ. This instability shows up also in the number of determinants used by the SHCI method. In addition, we found that the growth of the determinant expansion of carbon dimer, with respect to μ, has a similar behavior to the one of the He atom (see figure 3.16). The rate of increase in determinants is much faster for the asymptotic model, explaining the fast convergence of the energies. According to figure 3.17 even if we try to represent the

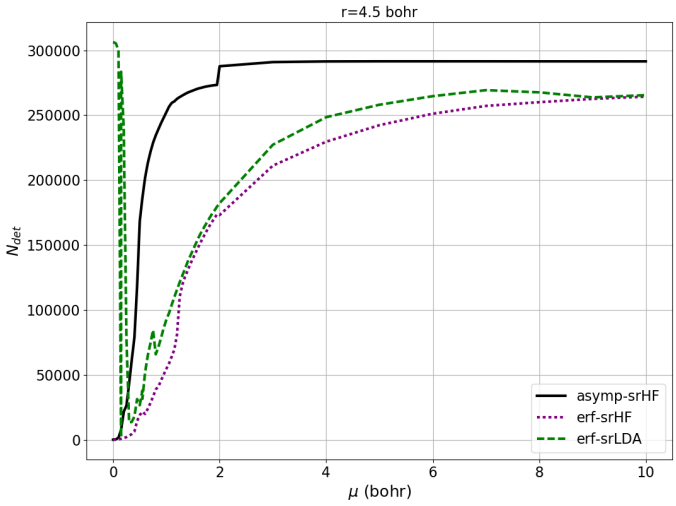

Figure 3.16: The number of determinants used in SHCI with $\epsilon_1 = 100\mu$Hartree and the cc-pVDZ basis, for C_2 at a bond-breaking distance, 4.5 bohr. Dotted lines represent the error function long-range interaction with short-range LDA correction, and the solid lines are used for the asymptotic long-range interaction with short-range HF correction. For small μ erf-srLDA had problems achieving self-consistency, which is why were needed a large number of determinants for $\mu \approx 0.3$.

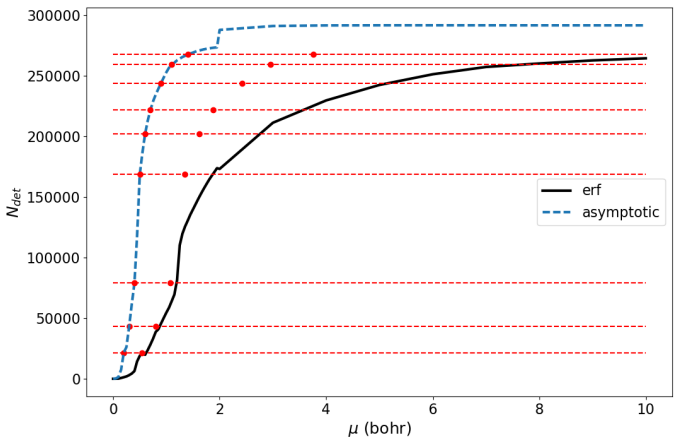

Figure 3.17: The number of determinants used by SHCI using $\epsilon_1 = 100\mu$Hartree and cc-pVDZ basis, for C_2 at a bond-breaking distance: 4.5 bohr. Comparison between the error function and the asymptotic long-range interactions with short-range HF correction. The red dots represent the expected number of determinants of $\text{erf}(\mu' r)/r$ where $\mu' = c_a \mu$, so that $\text{erf}(\mu' r)/r = \text{erf}(\mu r)/r + c_{asymp}\mu \exp(-\alpha^2_{asymp}\mu^2 r^2)$ at $r = 0$. The prediction works well for $\mu < 1.0$. The contribution of the Gaussian function to the asymptotic potential makes a difference at larger μ, where the length of the CI expansion increases much more slowly for the error function model potential than for the asymptotic erfgau potential.

asymptotic model with an erf that has the same value at $r = 0$ ($\mu_{\text{erf}} = 2.6875\mu_{\text{erfgau}}$), the CI expansion from the error function is always shorter than the equivalent of the asymptotic model. The same happens with the energy, the convergence of an erfgau model is always faster than the erf model.

3.6 Summary

The goal of using smooth models to approximate the (singular) Coulomb interaction between electrons is to reduce the computational cost of approximately solving the electronic Schrdinger equation by decreasing the length of the configuration interaction expansion and increasing the rate of convergence with respect to the one-electron basis set. In this chapter we demonstrate that these expected advantages do accrue, not only for the conventional error-function potential but also for the asymptotic erfgau potential we developed in chapter 2. However, when we tried to optimize the form of the Gaussian to minimize the expectation value of the energy, a very pointy Gaussian was obtained, and no significant cost savings were found. We therefore conclude that we should focus our attention on the traditional error-function smoothed Coulomb interaction and the asymptotic erfgau form.

Although replacing the Coulomb potential with a smooth model potential reduces the computational cost of (approximately) solving the Schrdinger equation, it also decreases the accuracy: one must somehow correct for the difference between the smooth electron-electron interaction one solved for and the physical electron-electron interaction one wishes to model. We considered several ways of doing this, including correcting for the neglected short-range potential using density functional theory (at the level of the local density approximation) and Hartree-Fock theory. In terms of

accuracy, including a short-range Hartree-Fock correction is always comparable to a short-range LDA correction, but the short-range Hartree-Fock correction is noniterative, making it faster and more robust. (We sometimes have problems converging the short-range LDA correction.)

With both the error function and the asymptotic erfgau interaction, a (first-order) perturbative correction inspired by Sirbu and King is helpful. That is, use the wavefunction computed by solving the smooth model Hamiltonian including short-range Hartree-Fock (or LDA) corrections to evaluate the true physical Hamiltonian. This approach gives impressive results for 2-electron atoms and the Carbon dimer. Further study is needed, however, in order to select between the error-function and the asymptotic-erfgau forms for the smoothed potential. The asymptotic erfgau method gives more accurate energies, but requires more determinants. It is unclear, for now, whether the increased accuracy of the asymptotic erfgau method is worth its increased cost. Since the energy from the asymptotic erfgau potential approaches the exact $\mu \to \infty$ limit as μ^{-4}, while the energy from the error function potential approaches the exact energy only as μ^{-2}, it is reasonable to hypothesize that, at least in application settings where results near the high-accuracy limit are required, the asymptotic erfgau potential will be preferable.

Chapter 4

Extrapolation

In the previous chapters we explored the possibility of replacing the Coulomb electron-electron interaction with soft models. We showed that by using these soft models one can shorten the configuration expansion in selected CI methods, particularly when the range-separation parameter μ takes very small values. In this chapter we propose a method to approximate the FCI energy by extrapolating the results of these cheaper calculations. The reduction in cost comes from combining the soft model potentials and the stochastic heat-bath configuration interaction method (SHCI) [HolTubUmr16]. The accuracy and the cost of the calculations depend on two parameters, the SHCI parameter ϵ_1 and the range-separation parameter μ. We can extrapolate in one dimension, ϵ_1 or μ, and then the other, or do a two-dimensional extrapolation. In this chapter we explore the first strategy, extrapolating $\epsilon_1 \to 0$ first.

4.1 Extrapolating $\epsilon_1 \to 0$

In the semi-stochastic heat-bath configuration interaction (SHCI) method the parameter ϵ_1 determines the accuracy of the variational energy and ϵ_2 determines the accuracy of the second-order Epstein-Nesbet perturbation correction (EN-PT2). More than one approach has been used to extrapolate the total energy to the FCI limit. For example, one can set ϵ_2 to zero and extrapolate the variational energy using a rational function of $\sqrt{\epsilon_1}$. Another option is to evaluate the total energy with several values of ϵ_1 and ϵ_2, then use a rational polynomial in ϵ_2 to extrapolate to $\epsilon_2 = 0$ and, finally, fit a rational polynomial to extrapolate in ϵ_1. As shown in figure 5 of reference [HolTubUmr16], both strategies lead to the same extrapolated energy. More recently, Holmes *et al* proved that it is sensible to extrapolate to the limit where the perturbation correction ΔE_2 is zero. In that case fewer calculations are needed and a simpler (linear or quadratic) fit is sufficient to produce values with μHartree precision [HolUmrShar17]. We chose this last strategy to perform the extrapolation $\epsilon \to 0$. More details about the extrapolation can be found in section 4.3.

4.2 Extrapolating $\mu \to \infty$

In this case, we want to generate a function to model the energy $E(\mu)$. Rational functions and hypergeometric functions are good candidates for our model as they are very flexible, in the sense that they can be used to approximate a wide range of functions. Recently, Mera and co-workers developed a resummation approach based on Meijer-G functions to approximate the Borel sum of divergent series in quantum

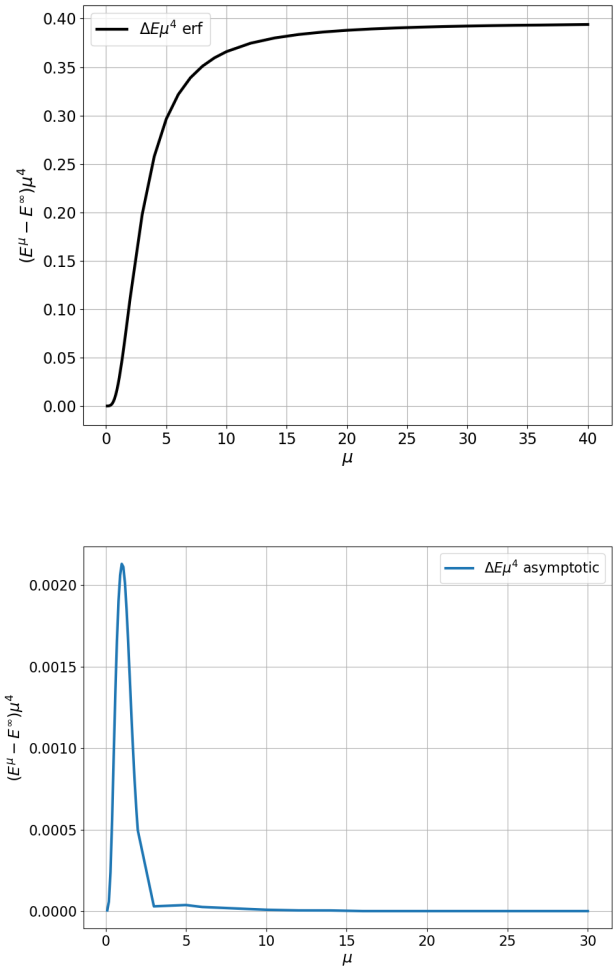

Figure 4.1: The FCI/aug-cc-pVQZ energy of He calculated with the error function (top) and the asymptotic (bottom) model potentials with the short-range Hartree-Fock correction and the first-order perturbation correction. In both cases the energy is observed to approach the true (Coulomb interaction) energy as μ^{-4}. The asymptotic model shows a much faster convergence.

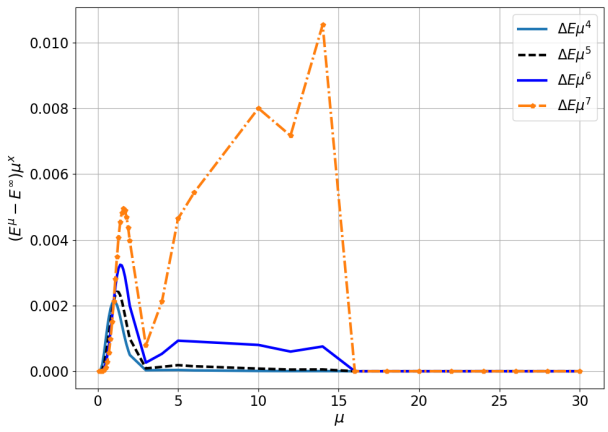

Figure 4.2: The decay of the FCI/aug-cc-pVQZ energy of He calculated using the asymptotic erfgau model potential with the short-range Hartree-Fock correction and the first-order perturbation correction. The energy is observed to fully converge to the true (Coulomb interaction) energy for $\mu > 15$.

mechanics and quantum field theory (QFT). The Meijer-G resummation is philosophically similar to the standard Padé approximants but uses the more general and flexible hypergeometric functions. The algorithm to construct a Meijer-G approximant is the following:

1. We assume that the energy, a function of μ, can be approximated by a Euler-Maclaurin series at the origin,

$$E(\mu) = \sum_{k=0}^{\infty} z_n \mu^k \tag{4.1}$$

   where z_n are normalized, $z_0 = 1$.

2. If we only know the first N coefficients (N being an odd number), z_0, z_1, ..., z_N, we compute the Borel-transformed coefficients, $b_n = z_n/n!$ and the N ratios of

consecutive Borel-transformed coefficients, $r(n) = b_{n+1}/b_n$.

3. The consecutive ratios are fitted with a rational function of n,

$$r_N(n) = \frac{\sum_{m=0}^{l} p_m n^m}{1 + \sum_{m=1}^{l} q_m n^m}, \tag{4.2}$$

here $l = (N-1)/2$.

4. *Hypergeometric vectors*, $(x_1, \ldots, x_l)$ and $(y_1, \ldots, y_l)$, are obtained by finding the roots of the polynomials in the numerator and denominator, *i.e.*, solving

$$\begin{aligned} \sum_{m=0}^{l} p_m x^m &= 0, \\ 1 + \sum_{m=1}^{l} q_m y^m &= 0, \end{aligned} \tag{4.3}$$

These hypergeometric vectors uniquely determine the hypergeometric function

$$B_N(\tau) \equiv {}_{l+1}F_l\left((1, -x_1, \ldots, -x_l), (-y_1, \ldots, -y_l), \frac{p_l}{q_l}\tau\right) \tag{4.4}$$

where ${}_{l+1}F_l$ is a generalized hypergeometric function [ref52].

5. Finally, we perform a Laplace transform

$$E_{B,N}(\mu) \equiv \int_0^\infty e^{-\tau} B_N(\mu\tau) d\tau \tag{4.5}$$

to restore the $n!$ removed from the expansion coefficients with the Borel transform. This expression can be rewritten as

$$E_{B,N}(\mu) = \frac{\Pi_{i=1}^{1}\Gamma(-y_i)}{\Pi_{i=1}^{1}\Gamma(-x_i)} G_{l+1,l+2}^{l+2,1}\left(\begin{matrix} 1, -y_1, \ldots, -y_l \\ 1, 1, -x, \ldots, -x_l \end{matrix}\middle| \frac{q_l}{p_l \mu}\right) \tag{4.6}$$

where $\Gamma(x)$ is the Euler's Gamma function, and

$$G_{p,q}^{m,n}\left(\begin{matrix} a_1, \ldots, a_n; a_{n+1} \ldots a_p \\ b_1, \ldots, b_m; b_{m+1} \ldots b_q \end{matrix}\middle| z\right) = \frac{1}{2\pi i}\int_L \frac{\prod_{j=1}^{m}\Gamma(b_j + s)\prod_{j=1}^{n}\Gamma(1 - a_j - s)}{\prod_{j=n+1}^{p}\Gamma(a_j + s)\prod_{j=m+1}^{q}\Gamma(1 - b_j - s)} z^{-s} ds \tag{4.7}$$

is the Meijer G-function.

The result of this algorithm is a table of Meijer G-functions that approximate the Borel sum of $E(\mu)$. In our particular case, we do not have high order derivatives of $E(\mu)$, we only have some energy values at certain small μ. In order to use the Meijer-G approximation we need to find another function to model the energy that we can differentiate to high order and approximate with an Euler-Maclaurin series. Frequently, people find it useful to constrain the form of the Padé approximant in order to reproduce the behavior of the function it is modeling. For example, we have observed that the energy error of the asymptotic model converges with μ^{-4} (see figure 4.1). We can make use of this empirical knowledge and force the energy model to decay as μ^{-4}. A Padé approximant with this asymptotic behavior is:

$$E_{(n,n)}(\mu) = \frac{a_0 + a_1\mu + a_2\mu^2 + \ldots + a_{n-4}\mu^{n-4} + a_n\mu^n}{1 + b_1\mu + b_2\mu^2 + \ldots + b_n\mu^n} \tag{4.8}$$

With (4.8), we can easily calculate the derivatives of the energy at $\mu = 0$, and

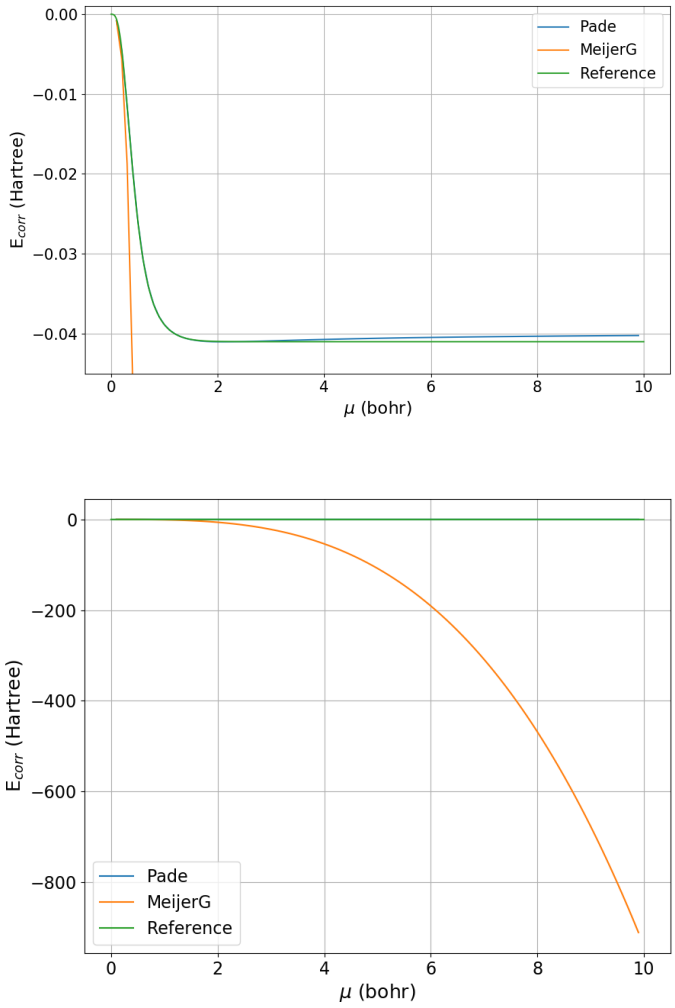

Figure 4.3: He correlation FCI/aug-cc-pVQZ energy modeled with a Meijer-G approximant and a Padé approximant using the asymptotic sr-HF+PT1 corrected and SHCI extrapolated energies, with $\mu_{max} = 1.0$. The Meijer-G function fits the points near zero accurately. However, it does not have explicit information about the asymptotic decay and tends to diverge asymptotically.

therefore, the coefficients of the Maclaurin expansion of $E(\mu)$. Once we have the coefficients of the expansion the rest of the algorithm to construct the Meijer-G approximant is as before. Because we used an approximate function to build the Meijer-G approximant, we need one final optimization. We took the form obtained using the algorithm explained above applied to the Euler-Maclaurin series for $E_{(n,n)}(\mu)$, as an initial guess, and we then optimized the parameters $(x_1, \ldots, x_l)$, $(y_1, \ldots, y_l)$, p_l, and q_l, by nonlinear least squares in order to reproduce the reference energies. Although the method seemed promising, the results were disappointing. As shown in figure 4.3, The Meijer-G approximation seems to reproduce the energies at μ very close to zero, which makes sense since we only gave information about the function and its derivatives at that $\mu = 0$. Additionally, it was pointed out in ref [MeraPedNiko18] that very high-accuracy input data is needed to achieve accurate results with Meijer-G approximants, while we only have the derivatives of an approximation to $E(\mu)$. At the same time, it was clear that the Padé approximant does a great job modeling the energy with only 6 data points. Therefore, we decided to leave the Meijer-G approximant aside and use a Padé approximant as our main energy model.

There are a couple of subtle points about the Padé model. For instance, the coefficients in eq. (4.8) are badly scaled. The magnitude of a_j and b_j are very different for each j, depending on the sizes of μ_k. This occurs because if $\mu_k > 1$, the last terms in the numerator and denominator are enormous compared with the earlier terms. On the other side, if $\mu_k < 1$, the last terms are tiny compared to the earlier terms. This can be mitigated by using orthogonal polynomials in the rational functions. Suppose all the data we have is between $-\mu_{\max} < \mu_k < \mu_{\max}$. Then, it is

sensible to use scaled Chebyshev polynomials,

$$E_{(n,n)}(\mu) = \frac{\sum_{j=0}^{n} a_j T_j\left(\mu/\mu_{\max}\right)}{\sum_{j=0}^{n} b_j T_j\left(\mu/\mu_{\max}\right)} \tag{4.9}$$

To preserve the asymptotics one needs to impose that the coefficients of μ^{n-3}, μ^{n-2}, and μ^{n-1} to be zero. This is done by imposing the constraints:

$$a_{n-1} = 0$$

$$a_{n-3} = 0$$

$$a_{n-2} - n a_n = 0$$

The resulting model is:

$$E_{(n,n)}(\mu) = \frac{\sum_{j=0}^{n-4} a_j T_j\left(\mu/\mu_{\max}\right) + a_n\left(n T_{n-2}\left(\mu/\mu_{\max}\right) + T_n\left(\mu/\mu_{\max}\right)\right)}{\sum_{j=0}^{n} b_j T_j\left(\mu/\mu_{\max}\right)} \tag{4.10}$$

We still have an arbitrary degree of freedom inasmuch as we can multiply both the numerator and the denominator by an arbitrary constant. We resolve this by setting the coefficient of $T_0\left(\mu/\mu_{\max}\right)$ in the denominator equal to one:

$$E_{(n,n)}(\mu) = \frac{\sum_{j=0}^{n-4} a_j T_j\left(\mu/\mu_{\max}\right) + a_n\left(n T_{n-2}\left(\mu/\mu_{\max}\right) + T_n\left(\mu/\mu_{\max}\right)\right)}{1 + \sum_{j=1}^{n} b_j T_j\left(\mu/\mu_{\max}\right)} \tag{4.11}$$

If n gets large, we almost certainly will encounter overfitting problems.

4.2.1 Avoidance of Overfitting

Overfitting occurs when minimizing the error of the model on the training data set no longer benefits the model's performance for data outside the training set. This happens when the model is so tightly fit to specific data points that it is also modeling the underlying noise. It can also happen when additional data leads to a complicated functional form for the model energy that models the true unknown energy function less accurately than a simpler form. Some examples of methods to check for overfitting are: cross-validation (CV), bootstrap, and stepwise regression.

In stepwise regression one starts by fitting all the available data. then one discards each of the parameters in the model one-by-one and recomputes the error in the model, choosing the parameter that, when discarded, gives the smallest least-squares error. An alternative is to use the LASSO method to minimize the errors in the least-squares equations,

$$\frac{1}{\epsilon_k^2}\Bigg[E_{\text{calc}}(\mu_k)\left(1+\sum_{j=1}^{n} b_j T_j\left(\mu/\mu_{\max}\right)\right) \tag{4.12}$$

$$-\left(\sum_{j=0}^{n-4} a_j T_j\left(\mu/\mu_{\max}\right)+a_n\left(nT_{n-2}\left(\mu/\mu_{\max}\right)+T_n\left(\mu/\mu_{\max}\right)\right)\right)\Bigg]^2 \tag{4.13}$$

subject to the constraint that the l_1-norm of the unknown coefficients is not too large,

$$\eta \geq \sum_{j=0}^{n-4}|a_j|+|a_n|+\sum_{j=1}^{n}|b_j| \tag{4.14}$$

Here $E_{calc}(\mu_k)$ are the energies obtained from the extrapolation to $\epsilon \to 0$, and ϵ_k are the precision of the extrapolated values. In practice this minimization is solved using

Lagrange multipliers,

$$\min_{\substack{a_0,a1,\ldots,a_{n-4},a_n \\ b_1,b_2,\ldots,b_n}} \Bigg[\sum_{k=0}^{K} \frac{1}{\epsilon_k^2} \Bigg[E_{calc}(\mu_k) \Bigg(1 + \sum_{j=1}^{n} b_j T_j \left(\mu/\mu_{\max}\right) \Bigg] \tag{4.15}$$

$$- \Bigg(\sum_{j=0}^{n-4} a_j T_j \left(\mu/\mu_{\max}\right) + a_n \left(n T_{n-2} \left(\mu/\mu_{\max}\right) + T_n \left(\mu/\mu_{\max}\right) \right) \Bigg) \Bigg]^2 \tag{4.16}$$

$$+ \alpha \Bigg(\sum_{j=0}^{n-3} |a_j| + |a_n| + \sum_{j=1}^{n} |b_j| \Bigg) \Bigg] \tag{4.17}$$

The selection of the parameter α determines the allowed magnitudes of the parameters of the model. At each value of α one gets a set of parameters, (hopefully) eliminating some of the least important parameters.

The parameters obtained with LASSO are used as guesses to optimize the least-squares error

$$e_{rms}(\alpha) \equiv \min_{\{\text{nonzero } a_j, b_j\}} \sqrt{\frac{1}{K+1} \sum_{k=0}^{K} \left(\frac{E_{(n,n)}(\mu_k) - E_{\text{calc}}(\mu_k)}{\epsilon_{prec,k}} \right)^2} \tag{4.18}$$

where ϵ_{prec} is the estimated accuracy of each computed energy provided to the extrapolation model.

4.2.2 Selection of points

We would like to minimize the cost/benefit ratio. The cost is proportional to the number of determinants used in the SHCI calculation. To predict the number of determinants of a point we have not computed yet, we fit a function dependent on μ and ϵ, $N_{det} = f(\epsilon, \mu)$,

$$N_{det}(\mu) = N_{\text{FCI}} \exp^{-\frac{a\epsilon^n}{1+b\mu^m}}, \tag{4.19}$$

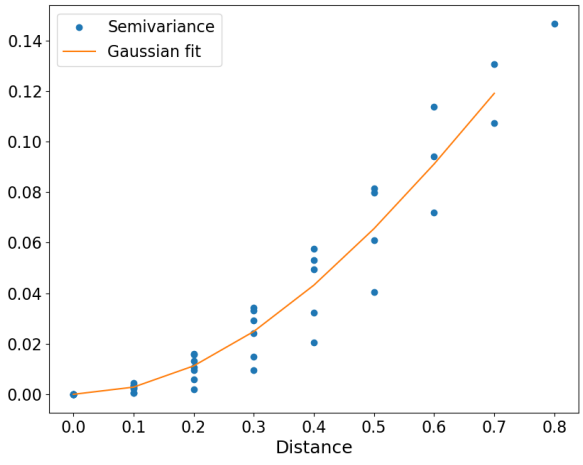

Figure 4.4: The semivariogram model in (4.24). The Gaussian model in orange and the distance in the errors of the model at μ_k and μ_l: $\frac{1}{2}(\Delta\epsilon_{kl})^2$ in blue. This model is used to approximate the statistical error of the data used to build the energy extrapolation model.

if we first extrapolate to $\epsilon \to 0$, then (4.19) could be expressed as

$$N_{det}(\mu) = N_{\text{FCI}} \exp^{-\frac{a}{b+\mu^m}} \tag{4.20}$$

here b is obtained from the limit at $\mu = 0$, where we only have one determinant, so $1 = N_{\text{FCI}} \exp^{-\frac{a}{b}}$.

The error in the model is a function of a) the sensitivity of the extrapolated value to the point $(\epsilon_{\text{calc}}, \mu_{\text{calc}})$ we are about to calculate and b) the amount of knowledge we already have about that point.

One can estimate the sensitivity of the final solution to adding a data-point. We start by assuming that the error in the fit can be modeled as a random process around the model function. For this to be true it is important that the model is roughly of the same accuracy for all μ, but also that it is not overfitting the data. First, we take all pairs of data points available (previously computed), $\{\mu_k, E_k\}$ and compute

$$\Delta\mu_{kl} = \mu_k - \mu_l \tag{4.21}$$

$$\Delta\epsilon_{kl} = \epsilon(\mu_k) - \epsilon(\mu_l) \tag{4.22}$$

where

$$\epsilon(\mu_k) = E_{\text{calc}}(\mu_k) - E_{\text{model}}(\mu_k) \tag{4.23}$$

We use this data to fit a simple Gaussian semivariogram model,

$$\gamma_{s_0,a}(|\Delta\mu|) = s_0 \left[1 - \exp(-a(\Delta\mu)^2)\right] \tag{4.24}$$

to the data $\frac{1}{2}(\Delta\epsilon_{kl})^2$ vs. $|\Delta\mu_{kl}|$. Now we build a statistical model for the error in the

data. The error can be computed using Gaussian processes. The covariance between different computations can be approximated as

$$c_{s_0,a}(|\Delta\mu|) = s_0 \exp(-a(\Delta\mu)^2) \tag{4.25}$$

Suppose that we have n data points that are known and we are going to add two new points. We can estimate the data with the existing points and the other unknown point $\mu_?$. Defining

$$c_{ij} = c_{s_0,a}\left(|\mu_i - \mu_j|\right) = s_0 \exp(-a(\mu_i - \mu_j)^2)$$

the expected error in one of the unknown values is proportional to

$$\sigma(\mu_?) = \sqrt{\begin{bmatrix} w_0 & w_1 & \dots & w_n \end{bmatrix} \begin{bmatrix} c_{0,0} & c_{0,1} & \dots & c_{0,n} \\ c_{1,0} & c_{1,1} & \dots & c_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,0} & c_{n,1} & \dots & c_{n,n} \end{bmatrix} \begin{bmatrix} w_o \\ w_1 \\ \vdots \\ w_n \end{bmatrix} + 2 \begin{bmatrix} c_{0,?} & c_{1,?} & \dots & c_{n,?} \end{bmatrix} \begin{bmatrix} w_o \\ w_1 \\ \vdots \\ w_n \end{bmatrix} + s_0} \tag{4.26}$$

where $c_{k,?} = s_0 \exp(-a(\mu_k - \mu_?)^2)$, and $\mathbf{w}$ is obtained by solving a linear system of equations,

$$\begin{bmatrix} c_{0,0} & c_{0,1} & \dots & c_{0,n} & 1 \\ c_{1,0} & c_{1,1} & \dots & c_{1,n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{n,0} & c_{n,1} \dots & c_{n,n} & 1 & \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \\ m \end{bmatrix} = \begin{bmatrix} c_{0,?} \\ c_{1,?} \\ \vdots \\ c_{n,?} \\ 1 \end{bmatrix} \tag{4.27}$$

The argument of the square root in equation (4.26) is never negative because the

covariance matrix is nonnegative semideninite, so we can rewrite it as:

$$\sigma(\mu_?) = \sqrt{\begin{bmatrix} w_0 & w_1 & \dots & w_n & -1 \end{bmatrix} \begin{bmatrix} c_{0,0} & c_{0,1} & \dots & c_{0,n} & c_{0,?} \\ c_{1,0} & c_{1,1} & \dots & c_{1,n} & c_{1,?} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{n,0} & c_{n,1} & \dots & c_{n,n} & c_{n,?} \\ c_{?,0} & c_{?,1} & \dots & c_{?,n} & c_{?,?} \end{bmatrix} \begin{bmatrix} w_o \\ w_1 \\ \vdots \\ w_n \\ -1 \end{bmatrix}} \tag{4.28}$$

Now, we deal with the issue of defining how do we determine the "importance" of the points to be added. Suppose we have a model for the energy with the form

$$E_{(n,n)}(\mu) = f_{n_{\text{calc}}}\left(\mu | \{\mu_k, E_k\}_{k=1}^{n_{\text{calc}}}\right) \tag{4.29}$$

based on the previous calculations $\{\mu_k, E_k\}_{k=1}^{n_{\text{calc}}}$. Now we pretend to add one additional calculation, which we actually approximated with our current model. That is, we assume the model is correct in which case the computed energy would be predicted by the model (4.29),

$$E_{n_{\text{calc}+1}} = f_{n_{\text{calc}}}\left(\mu_{n+1} | \{\mu_k, E_k\}_{k=1}^{n_{\text{calc}}}\right) \tag{4.30}$$

Then we construct a new model using that point,

$$f_{n_{\text{calc}+1}}\left(\mu | \{\mu_k, E_k\}_{k=1}^{n_{\text{calc}+1}}\right) \tag{4.31}$$

We want to know how much a small change in the value of this point, corresponding to an error in our current model $f_{n_{\text{calc}}}(\mu)$, would affect our final prediction. Therefore,

we define the *importance* of this point,

$$\left| \frac{\partial f_{n_{\text{calc}+1}} \left(\mu = \infty | \{\mu_k, E_k\}_{k=1}^{n_{\text{calc}+1}} \right)}{\partial E_{n_{\text{calc}+1}}} \right| \tag{4.32}$$

If we include the prediction of the cost of that calculation (4.20) and the estimation of the model error at $\mu_{n_{\text{calc}+1}}$ from (4.26), the point we should add is the one for which

$$\frac{\left| \frac{\partial f_{n_{\text{calc}+1}} \left(\mu = \infty | \{\mu_k, E_k\}_{k=1}^{n_{\text{calc}+1}} \right)}{\partial E_{n_{\text{calc}+1}}} \right| \sigma(\mu)}{N_{dets}(\mu)} \tag{4.33}$$

is maximized. The intuition behind this equation is simple: we wish to solve the Schrödinger equation for values of μ that the energy value at $\mu = \infty$ is sensitive to, where the uncertainty in the energy is high, and where the cost of calculation is low.

4.2.3 Termination

At the beginning we do not know the right answer, so how do we decide when to stop computing new points? That is, how do we determine if the extrapolation model is good enough? Because the energy has to converge at some point, one way to evaluate if the extrapolating model has achieved a desired form is to assess the deviation of the predicted values at large values of μ. For this, we make a grid of points $\mu_{max} < \mu_l < \infty$ and compare the maximum distance between the prediction of the current model and the model of the previous iteration,

$$\max_{\mu_{max} < \mu < \infty} |E^m(\mu) - E^{m-1}(\mu)| \tag{4.34}$$

and if the maximum difference is below the desired precision of the prediction then we stop. If the target accuracy is never achieved, then we stop when additional data stops improving the model.

4.3 Results

In this section we demonstrate how the optimization method works. We took the correlation energy of He atom as example. The energies used for the extrapolation were calculated with the best model found in previous chapter, the asymptotic erfgau potential with short-range HF and first-order perturbation corrections,

$$E^{sr-\mathrm{HF}+\mathrm{PT1}} = \langle \Psi^{(\mu,sr-\mathrm{HF})} | \hat{H} | \Psi^{(\mu,sr-\mathrm{HF})} \rangle \tag{4.35}$$

where $\Psi^{(\mu,sr-\mathrm{HF})}$, comes from solving $\hat{H}^{(\mu,sr-\mathrm{HF})}\Psi^{(\mu,sr-\mathrm{HF})} = E^{(\mu,sr-\mathrm{HF})}\Psi^{(\mu,sr-\mathrm{HF})}$, and

$$\hat{H}^{(\mu,sr-\mathrm{HF})} = \left(\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + v_{ne}(\mathbf{r}_i) + j^{sr,\mu}(\mathbf{r}_i) + k^{sr,\mu}(\mathbf{r}_i) \right) + \hat{V}_{ee}^{lr,\mu} \tag{4.36}$$

The energies are calculated with the SHCI method, setting $\epsilon_2 = 10^{-8}$ and using 3 different values of ϵ_1: 50, 100 and 500 μHartree. We use the total energies and the perturbation correction to extrapolate $\Delta E_2 \to 0$ using a simple rational function dependent on ΔE_2, $\frac{a+b\Delta E_2}{1+c\Delta E_2}$. In most cases the rational function became a linear function (see figure 4.5).

For the extrapolation along the μ axis, we must first decide what is the most expensive calculation we are willing to perform, in terms of which μ to use. In the previous chapter we learned that it is convenient to use values smaller than $\mu = 1.0$

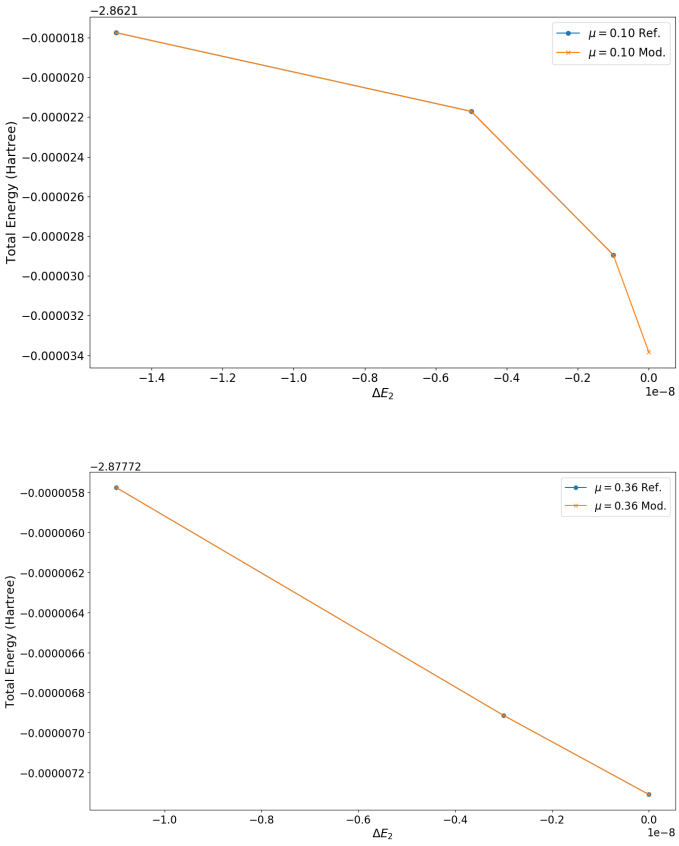

Figure 4.5: Extrapolation in the ϵ-axis of the total energies, variational SHCI energy plus the EN-PT2 perturbation correction ΔE_2, of the He atom with the aug-cc-pVQZ basis. The variational energies were calculated using the asymptotic model with sr-HF+PT1 corrections at two different values of μ=0.1 (top) and 0.3 (bottom). We set $\epsilon_2 = 10^{-8}$ and used three different values of ϵ_1 =50, 100 and 500μHartree to fit a simple rational function of ΔE_2. The blue line labeled as "Ref." represents the energies from SHCI, and the orange line, labeled as "Mod." represents the prediction of the fit. At the top-left corner of each plot it is indicated the first digits of the total energies and in the $y-$axis the numbers shown are the last few digits. Similarly, the numbers in the $x-$axis have a magnitude of 10^{-8}.

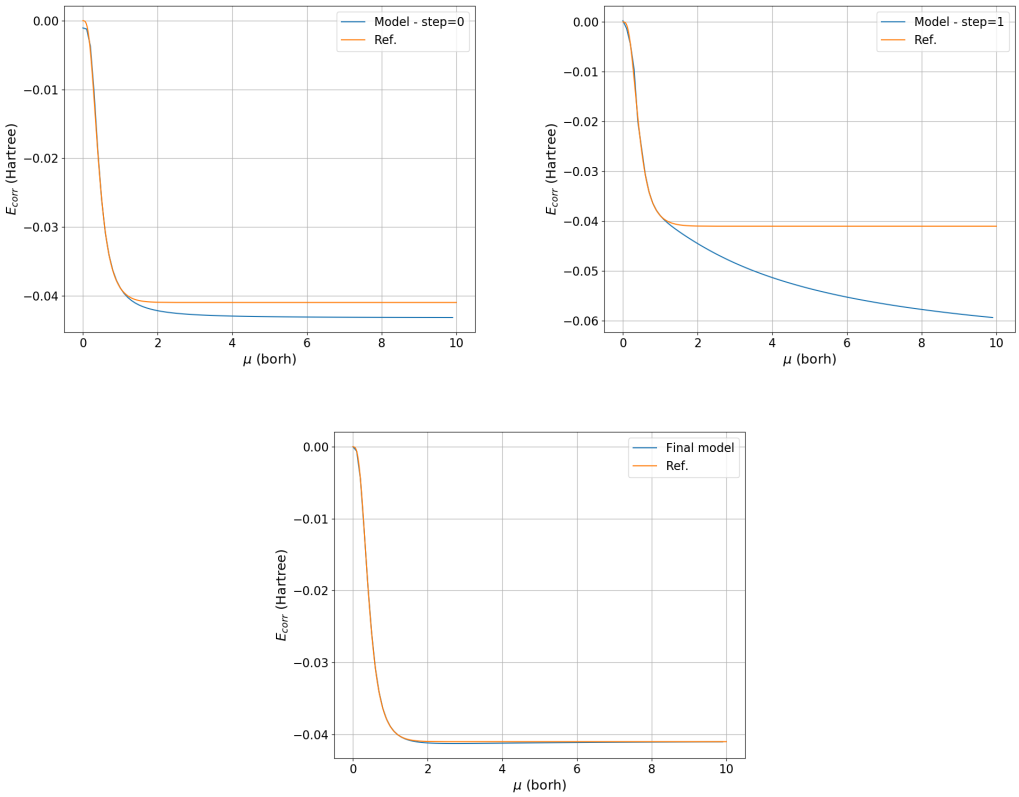

Figure 4.6: Performance of the model during the optimization of the extrapolation model with $\mu_{max} = 1.0$. The figures on the top show the intermediate results after the reduction of parameters (4.44), and (4.46). At the bottom, the final prediction of the optimized extrapolation model after 4 steps and 14 energy calculations.

to have a reduction on the determinant expansion. Therefore, we set $\mu_{max} = 1.0$ as our limit. In order to build the smallest approximant of the form (4.11) we need to start with 6 points. We select 4 points equally spaced between $\mu = 0$ and μ_{max}, for example $\mu_m = \{0.0, 0.24, 0.48, 0.72, 0.96, 1.0\}$. With these information the initial model is,

$$E_{(4,4)}(\mu) = \frac{a_0 T_0\left(\mu/\mu_{\max}\right) + a_4\left(4T_2\left(\mu/\mu_{\max}\right) + T_4\left(\mu/\mu_{\max}\right)\right)}{1 + b_1 T_1\left(\mu/\mu_{\max}\right) + b_2 T_2\left(\mu/\mu_{\max}\right) + b_3 T_3\left(\mu/\mu_{\max}\right) + b_4 T_4\left(\mu/\mu_{\max}\right)} \tag{4.37}$$

To make it easier to read, from now on $T_{n,k}$ will represent $T_n(\mu_k/\mu_{max})$. The linear system of equations to solve are:

$$E_{(4,4)}(\mu_k) = a_0 T_{0,k} + 4a_4 T_{2,k} + a_4 T_{4,k} - E_{(4,4)}(\mu)\left(b_1 T_{1,k} + b_2 T_{2,k} + b_3 T_{3,k} + b_4 T_{4,k}\right) \tag{4.38}$$

that can be rewritten in a matrix form,

$$\begin{pmatrix} T_{0,1} & 4T_{2,1}+T_{4,1} & -E_{(4,4)}(\mu_1)T_{1,1} & \dots & -E_{(4,4)}(\mu_1)T_{4,1} \\ T_{0,2} & 4T_{2,2}+T_{4,2} & -E_{(4,4)}(\mu_2)T_{1,2} & \dots & -E_{(4,4)}(\mu_2)T_{4,2} \\ T_{0,3} & 4T_{2,3}+T_{4,3} & -E_{(4,4)}(\mu_3)T_{1,3} & \dots & -E_{(4,4)}(\mu_3)T_{4,3} \\ \vdots & \vdots & \ddots & \dots & \vdots \\ T_{0,6} & 4T_{2,6}+T_{4,6} & -E_{(4,4)}(\mu_6)T_{1,6} & \dots & -E_{(4,4)}(\mu_6)T_{4,6} \end{pmatrix} \begin{pmatrix} a_0 \\ a_4 \\ b1 \\ \vdots \\ b_4 \end{pmatrix} = \begin{pmatrix} E_{(4,4)}(\mu_1) \\ E_{(4,4)}(\mu_2) \\ E_{(4,4)}(\mu_3) \\ \vdots \\ E_{(4,4)}(\mu_6) \end{pmatrix} \tag{4.39}$$

Solving eq. (4.39) gives us the first set of parameters $\{a_0, a_4, b_1, b_2, b_3, b_4\}$.

The next step is to determine whether a reduction of parameters is needed and/or possible without affecting the precision of the model. For that we ensure the RMS error of the reduced model stays below 1.0. The reduced model is constructed using the Least Absolute Shrinkage and Selection Operator (LASSO) method, setting α of

equation (4.17) to 0.00001. We utilize the LASSO LARS algorithm , so a whole path of values for α is generated. If in the first step, at the smallest α, all the parameters are kept, then we use the next value in the path. We repeat this procedure until some parameters are dropped out. Additionally, in order to conserve the right asymptotics of the energy, we force the model to always keep the highest order terms. This is achieved with an inelegant trick: we multiply the highest order terms by a factor ϕ,

$$E_{(n,n)}(\mu_k) = \frac{\sum_{j=0}^{n-4} a_j T_{j,k} + \phi a_n \ (nT_{n-2,k} + T_{n,k})}{1 + \sum_{j=1}^{n-1} b_j T_{j,k} + \phi \, b_n T_{n,k}} \tag{4.40}$$

so that the coefficients a_n and b_n become small and they do not get eliminated by the LASSO method.

After the first two iterations with LASSO, two parameters were taken off: b_1 and b_3, resulting in a scaled RMS error of 0.73 (c.f. eq. (4.18)). When the third parameter was taken from the model the error increased to an unacceptably large value of 6.28, so the third parameter was retained. After *cleaning-up* the model, the remaining parameters are optimized, by minimizing the RMS error. We also impose penalties to help ensure that the correlation energy is always negative and it converges

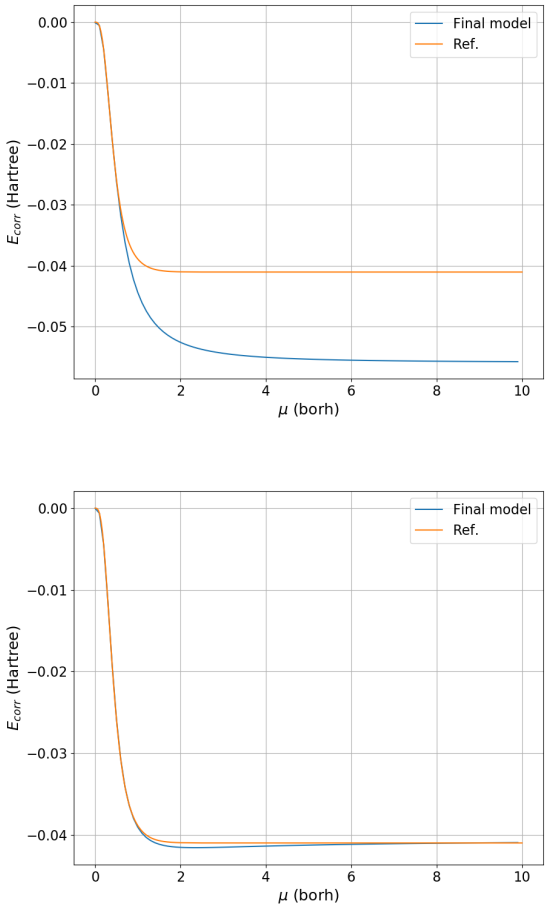

Figure 4.7: Performance of the extrapolation model of the correlation energy of He atom with the asymptotic potential and sr-HF+PT1 corrections. When $\mu_{max} = 0.4$ (top figure) the model does not have enough information about the FCI limit, but it preserves the asymptotic formed imposed. Setting $\mu_{max} = 0.6$ (bottom figure) is sufficient to achieve tolerable precision.

monotonically. Finally, we minimize

$$\min_{a_n,b_n\neq 0}\Bigg[\sum_{p=1}^{N_{\rm data}}\left(\frac{E(\mu_p)-E_{\rm model}(\mu_p)}{\epsilon_{prec;\,p}}\right)^p \tag{4.41}$$

$$+\sum_{t=1}^{N_{\rm test}}\begin{cases}\left(E_{\rm model}(\mu)-E(\mu=0)\right)^2 & E(\mu_t)-E(0)>0\\ 0 & E(\mu_t)-E(0)<0\end{cases} \tag{4.42}$$

$$+\sum_{t=1}^{N_{\rm test}}\begin{cases}\left(E'_{\rm model}(\mu_t)\right)^2 & E'(\mu_t)>0\\ 0 & E'(\mu_t)<0\end{cases}\Bigg] \tag{4.43}$$

where $\mu_{max} < \mu_t < \infty$ are grid points we use to test the monotonic decay of the model at long range. The model after the first reduction of parameters has the form,

$$E_{(4,4)}(\mu_k)=\frac{a_0T_{0,k}+4a_4T_{2,k}+a_4T_{4,k}}{1+b_2T_{2,k}+b_4T_{4,k}} \tag{4.44}$$

and overestimates the energy at $\mu=\infty$ (-0.043235) by approximately 1mHartree. The next step is to add new points. Because we have to reproduce the right asymptotics of the energy, we need to add two points at a time, adding one term to the numerator and one to the denominator of the Padé, of the same order. The two data points, $\mu_?$ and $\mu_{??}$ we should add to the model are the ones that maximize

$$\frac{\left|\frac{\partial f_{n_{\rm calc}+1}\left(\mu=\infty|\{\mu_k,E_k\}_{k=1}^{n_{\rm calc}},\mu_?,E_?\right)}{\partial E_?}\right|\sigma(\mu_?)}{N_{dets}(\mu_?)}+\frac{\left|\frac{\partial f_{n_{\rm calc}+1}\left(\mu=\infty|\{\mu_k,E_k\}_{k=1}^{n_{\rm calc}},\mu_{??},E_{??}\right)}{\partial E_{??}}\right|\sigma(\mu_{??})}{N_{dets}(\mu_{??})} \tag{4.45}$$

If the current number of parameters is not sufficient to reproduce the reference values, then we increase the number of parameters, adding the next term in both the numerator and the denominator. In the present example, the RMS error of the

reduced model was greater than 1, so the parameters a_1, a_5, b_5 where added and a_4 was eliminated. The new system of equations is,

$$E_{(4,4)}(\mu_k) = a_0 T_{0,k} + a_1 T_{1,k} + 5a_5 T_{3,k} + a_5 T_{5,k} - E_{(4,4)}(\mu)\left(b_2 T_{2,k} + b_4 T_{4,k} + b_5 T_{5,k}\right) \tag{4.46}$$

Two extra points where added to the same model and the prediction of the energy got worse (see top-right plot in 4.6). In the next step two more terms were added to the model. Once the approximant was cleaned by removing the least important parameters with LASSO, the model improved again. The final energy model was:

$$E_{(4,4)}(\mu_k) = \frac{a_0 T_{0,k} + a_2 T_{2,k} + 6a_6 T_{4,k} + a_6 T_{6,k}}{1 + b_2 T_{2,k} + b_4 T_{4,k} + b_5 T_{5,k} + b_6 T_{6,k}} \tag{4.47}$$

The extrapolation model predicted the E_{corr} =-0.040869 Hartree, which has a 0.14mHartree error with respect to the FCI limit with the same basis set (aug-cc-pVQZ), $E_{corr,\mathrm{FCI}} =$ 0.0410116, and a 1mHartree of error with respect to the reference energy taken from [DavHagChaUmaFis91] (-0.04204 Hartree). This is an important result taking into account that we were able to approximate the energy at $\mu \to \infty$ knowing only the energy at $\mu < 1$ and the asymptotic erfgau form with which the energy decays.

4.4 Discussion

On the one hand we should select the smallest μ_{max} as possible, because the smaller the μ_{max}, the cheaper the calculation is. However, using a very small μ_{max} means that we are extrapolating further, compromising the accuracy of the results. The smallest

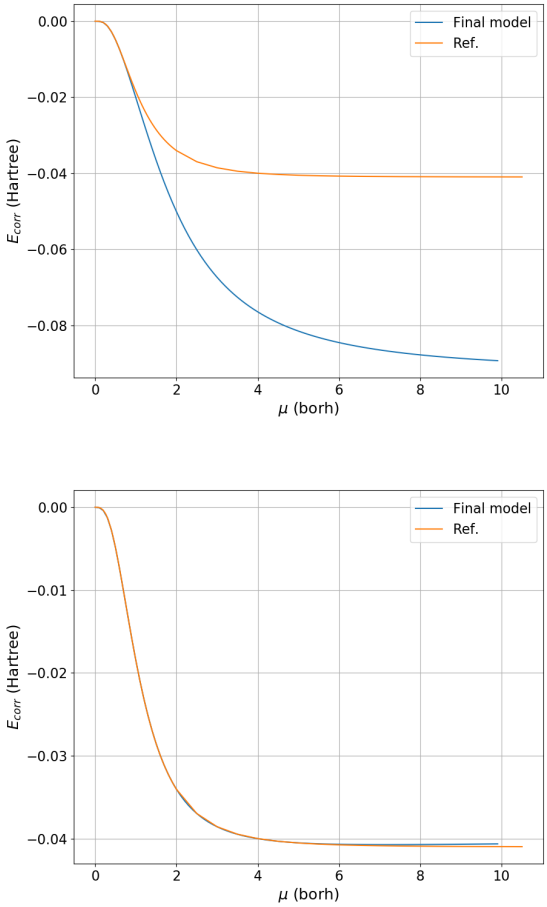

Figure 4.8: Performance of the extrapolation model of the erf potential with short-range and first-order perturbation correction. A much slower convergence of the energy is observed for the standard long-range potential compared with the $erfgau$ form. With $\mu_{max} = 0.6$ (top figure) the model energies are still very far from the FCI value. At least $\mu_{max} = 2.5$ (bottom figure) is needed to reach mHartree precision.

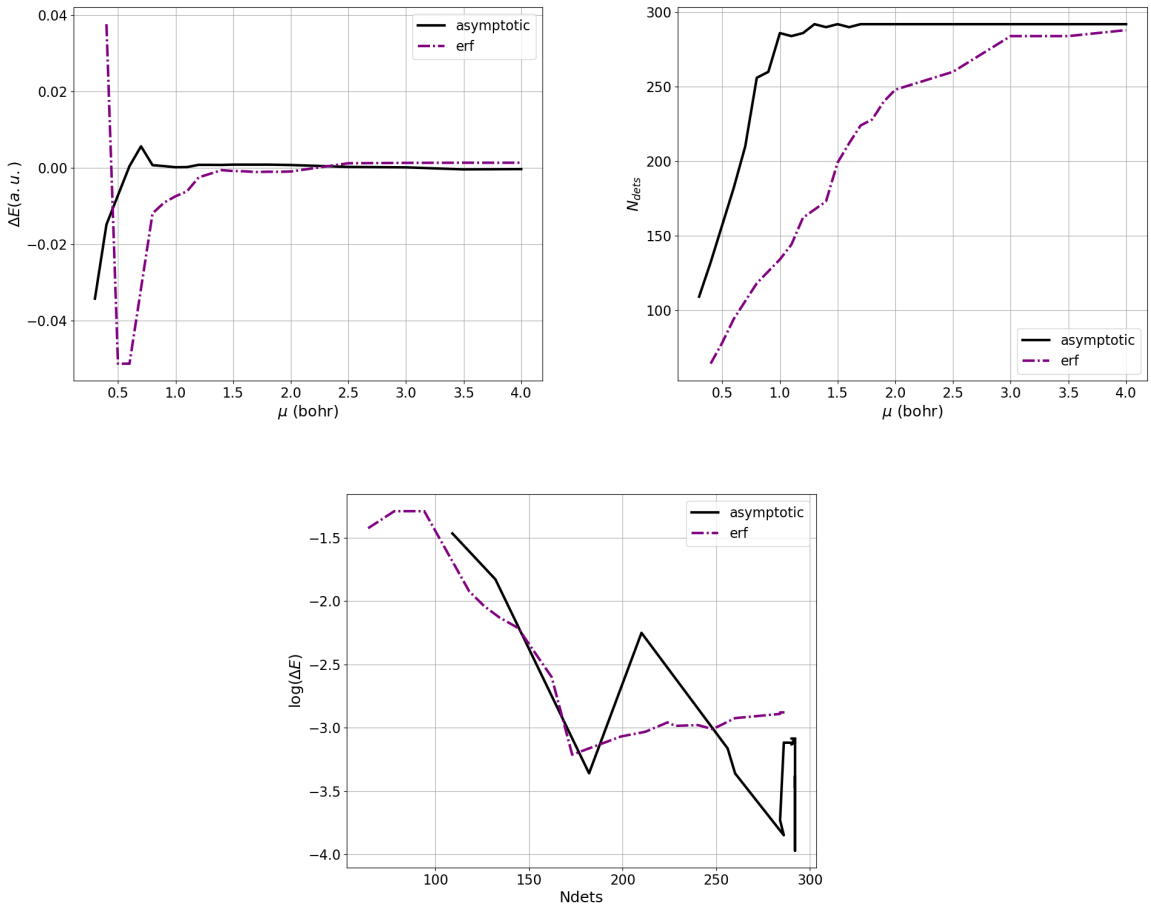

Figure 4.9: Top: Dependence of the energy error and the number of determinants on μ. Bottom:Precision/cost relation of the extrapolation models with the erf and asymptotic potentials. The energy error with respect to the FCI limit with the Coulomb potential, against the number of determinants of the SHCI expansion with the smallest ϵ_1 used: 10μHartree. Note that only determinants with coefficients greater than 10^{-4} where taken into account. Even when it seems to be cheaper to use the erf model, both methods required approximately the same size of the SHCI expansion achieve mHartree accuracy.

μ one can use and still get an accuracy of a few mHartree is $\mu = 0.6$ (figure 4.7). In the same figure we can see how for $\mu = 0.4$ the approximation starts failing to recover the right energies. Note that even when the model predicts a wrong answer, it preserves the asymptotic form imposed.

4.4.1 erf vs erfgau

The same procedure was performed with the error function potential. When we look at the results (figure 4.8), we see that at small μ the extrapolation model can not predict a proper energy because the model potential is very different from the physical interaction. With the erf model we needed to go up to $\mu = 2.5$ to get mHartree precision. In the previous section we showed how the determinant expansion grows more rapidly for the asymptotic erfgau potential, in parallel with the increase of precision. On the contrary, with the erf model one can get CI expansions with very few determinants, albeit with a low precision. Moreover, to reach mHartree precision one needs to use also a very large determinant expansion, almost the same as the asymptotic erfgau model. Even when both methods required about the same determinant expansion to reach high level of precision, the asymptotic erfgau model was slightly more accurate at large μ.

4.5 Conclusions

We proposed a method to extrapolate the model energies obtained by using the SHCI method to accurate estimate the correlation energy with smooth potentials. The extrapolation method was illustrated for the correlation energy of the He atom.

We reproduced the correlation energy at the FCI limit within 140μHartree precision extrapolating from $\mu \leq 1.0$. Even when the maximum value for μ was set to $\mu_{max} = 0.6$ a precision of 5mHartree was achieved. Comparison of the extrapolation model of the error function potential and the asymptotic erfgau potential showed that the cost and accuracy of both methods was equivalent. These proof-of-principle results merely prove that this strategy can work, it should be tested with larger molecules to demonstrate a real improvement in the efficiency of CI calculations.

Chapter 5

Conclusions

5.1 Summary and conclusions

In the field of method-development, a new method or approximation is judged based on two criteria: efficiency, so that the method is fast enough to study interesting systems, and accuracy, so the numbers are reliable enough to be useful. In the second chapter we examined the possibility of constructing a model for the electron-electron interaction to replace the Coulomb potential with a smooth, "easy-to-handle" potential. Specifically an *erfgau* potential was constructed that reproduced a portion of the spectrum of the Hydrogen atom. The new potential showed an improvement in the accuracy over the uncorrected $\mathrm{erf}(\mu r)/r$ potential not only for the attractive Coulomb interaction in the Hydrogen atom, but also for the repulsive electron-electron interaction in Harmonium and the exchange energy of the uniform electron gas.

In chapter three we assessed the *erfgau* model with two-electron systems, and compared it to the standard error function long-range potential. We explored strategies to correct the performance at short-range. Two main approaches were used:1)

short-range correction through an effective potential, using LDA or HF methods; 2) energy corrections from first-order perturbation theory. Additionally, we examined a different way to construct a model potential with an *erfgau* form by variationally optimizing the coefficients of the Gaussian function in order to make the first-order perturbation correction vanish. The variationally optimized models, particularly the ones where the exponent of the Gaussian was optimized, were very accurate, but also very expensive. The potential, while still smooth, became very pointy and, therefore, there was no significant reduction in the computational cost. The combination of the model potentials and the semistochastic heat-bath CI (SHCI) method was presented as a more efficient computational protocol. The FCI result can be achieved when the SHCI parameter $\epsilon_1 \to 0$ and the range-separation parameter $\mu \to \infty$. The dependence on these two parameters defines a way to correct the approximations systematically.

Finally, in the fourth chapter we proposed a method to extrapolate the model energies to the FCI limit, $\epsilon_1 \to 0$, and the limit of the physical interaction, $\mu \to \infty$. We compared the extrapolation model of the standard error function long-range potential and our asymptotic erfgau potential. Our results are inconclusive on whether to use the erf or erfgau potentials.

5.2 Outlook and future work

We have proposed a method to decrease the computational effort of wave function calculations by extrapolating energies computed with more accessible calculations like SHCI calculations with smooth model potentials. Many things are left to explore, for example:

- Obtaining an analytic form for the exchange-correlation energy of the asymptotic erfgau model in the UEG, by fitting to data from quantum Monte Carlo (QMC) calculations. Because this is a nonstandard potential, this would require performing these QMC calculations.

- Extending this method to use a local-μ so that the electron-electron repulsion is treated differently close to and far from the nuclei. This could be done by interpolating between a lower-limit and an upper-limit, for instance

$$\hat{W}^{\mu} \equiv \frac{\operatorname{erf}\left(\left[\mu_{\text{low}} + (\lambda(\mathbf{r}_1, \mathbf{r}_2))(\mu_{\text{high}} - \mu_{\text{low}})\right] r_{12} \right)}{r_{12}} \tag{5.1}$$

  and then approximating this function with a simpler function that is easier to deal with,

$$\hat{W}^{\mu} \equiv \frac{\operatorname{erf}\left(\mu_{\text{low}} r_{12}\right)}{r_1 2} + \lambda(\mathbf{r}_1, \mathbf{r}_2) \left(a \exp(-b r_{12}^2)\right) \tag{5.2}$$

  where the Gaussian function would add some sharpness to the potential, defining the upper-limit, and λ regulates the transition between the two limits. Pursuing the local-μ approach seems less appealing now that we know the consequences of changing the form of the potential, but this was the original goal of this project.

- Instead of using HF as a reference, use the seniority-zero Hamiltonian. It would also be reasonable to treat same-spin and opposite-spin electron pairs differently, perhaps always choosing $\mu \to \infty$ for same spin electrons.

- For the extrapolation approach, sometimes larger μ_{max} may be required to

achieve certain accuracy. One way to compute energies at larger μ while maintaining low computational cost is to use larger ϵ_1. In that case one would need to extrapolate in both dimensions, $\mu \to \infty$ and $\epsilon \to 0$ at the same time. This is possible with a small modification in the form of the model,

$$E_{(KL,KL)}(\epsilon,\mu) \equiv \frac{\sum_{l=0}^{L}\sum_{k=0}^{L-4} a_{lk} T_{l,\epsilon} T_{k,\mu} + \phi \sum_{l=0}^{L} a_{lk} \left(K T_{K-2,\mu} + T_{K,\mu}\right) T_{l,\epsilon}}{1 + b_{01} T_{1,\mu} + b_{10} T_{1,\epsilon} + \sum_{l=1}^{L}\sum_{k=1}^{K-1} b_{lk} T_{l,\epsilon} T_{k,\mu} + \phi \sum_{l=1}^{L} b_{lK} T_{l,\epsilon} T_{K,\mu}} \tag{5.3}$$

where $T_{k,\mu} = T_k\left(\mu/\mu_{max}\right)$, and $T_{k,\epsilon} = T_k\left(\epsilon/\epsilon_{max}\right)$.

Appendix A

Asymptotic model for Hydrogen atom

A.1 Hydrogenic atoms: μ dependence and integrals' scaling.

The leading term in the hydrogenic radial function is

$$R_{nl}(r)^2 \sim Z^3(Zr)^{2l}\mathrm{e}^{-2Zr/n}, \tag{A.1}$$

therefore

$$R_{nl}(r)^2 r^2 dr \sim \rho^{2l+2}\mathrm{e}^{-2\rho/n}d\rho \tag{A.2}$$

where $\rho = Zr$. Now, let

$$I(\mu) = \int_0^\infty f(\mu r)R_{nl}(r)^2 r^2 dr \tag{A.3}$$

then

$$\begin{aligned} I(\mu) &\sim \int_0^\infty f\left(\frac{\mu\rho}{Z}\right)\rho^{2l+2}\mathrm{e}^{-2\rho/n}d\rho \\ &= \left(\frac{Z}{\mu}\right)^{2l+3}\int_0^\infty f(x)x^{2l+2}\mathrm{e}^{-2Zx/\mu n}dx \\ &= \left(\frac{Z}{\mu}\right)^{2l+3}\sum_{i=0}^{\infty}a_i\left(\frac{Z}{\mu}\right)^i, \end{aligned} \tag{A.4}$$

where $x = \mu\rho/Z$ and

$$a_i = \frac{(-1)^i}{i!}\left(\frac{2}{n}\right)^i\int_0^\infty f(x)x^{2l+i+2}dx. \tag{A.5}$$

For $f(\mu\, r/Z) = \mathrm{erf}(\mu\, r)/Zr$, $f(x) = (\mu/Z)\mathrm{erf}(x)/x$ and the power of the asymptotic term is $(2l+2)$.

A.2 Model Hamiltonian from first-order perturbation theory.

We define the model Hamiltonian as

$$\hat{H}_\mu = -\frac{1}{2}\nabla^2 - \left[c\,\exp(-\alpha^2 r^2) + \frac{\mathrm{erf}(\mu\, r)}{r}\right], \tag{A.6}$$

where the parameters c and α should be chosen so that the eigenvalues of this operator are as close as possible to the ones of the physical operator,

$$\hat{H}_0 = -\frac{1}{2}\nabla^2 - \frac{1}{r}. \tag{A.7}$$

The difference between the two operators

$$w_\mu = \hat{H}_\mu - \hat{H}_0 = \frac{\mathrm{erfc}(\mu\, \mathrm{r})}{r} - c \exp(-\alpha^2 r^2) \tag{A.8}$$

is treated as a perturbation. We want the perturbation to vanish as $\mu \to \infty$. Moreover, we would like to keep a single parameter, μ, and make c and α functions of μ. As we mentioned in section 3.1, one way to produce $\langle \psi_i | w_\mu | \psi_i \rangle = 0$ is to choose α increasing with μ. When we expand the expectation value of w_μ for large values of α and μ we obtain, for $l = 0$ states, the integrals of the hydrogenic functions $\psi_{n0} = \{4e^{-r},\ (2-r)e^{-r/2},\ \frac{4}{729}(27 - 18r + 2r^2)e^{-r/3}\}$ are

$$\langle \psi_{n0} | w_\mu | \psi_{n0} \rangle = n^{-3} \left[A(\alpha, c) + B(\mu) \right] + O(\mu^{-5}), \tag{A.9}$$

where

$$A(\alpha, c) = -\frac{c\sqrt{\pi}}{\alpha^3} + \frac{4c}{\alpha^4}$$

comes from $\langle \psi_{n0} | c \exp(-\alpha^2 r^2) | \psi_{n0} \rangle$ and

$$B(\mu) = \frac{1}{\mu^2} - \frac{8}{3\sqrt{\pi}\mu^3} + \frac{3}{2\mu^4},$$

from $\langle \psi_{n0} | \mathrm{erfc}(\mu r)/r | \psi_{n0} \rangle$. Note that we used arbitrary multiplicative factors (4, 1 and $\frac{4}{729}$ respectively) to simplify the expressions. We can use this trick because we want to equate all expressions to 0.

In order to eliminate terms of order μ^{-2}, we set $\sqrt{\pi}\, c\, \alpha^{-3} = \mu^{-2}$. This means that $c = \gamma\, \mu$, $\alpha = \kappa\, \mu$ and $\gamma = \kappa^3/\sqrt{\pi}$. Substituting these values into the expansion (A.9) corresponding to $n = 1$ we see that the coefficient of μ^{-3} vanishes if $\gamma = (2\,\kappa^4)/(3\sqrt{\pi})$.

From the last two equations we get

$$\gamma = \frac{27}{8\sqrt{\pi}}, \quad \kappa = \frac{3}{2}$$

for which the energy of the S states is correct up to μ^{-4}. To eliminate the error for μ^{-4}, we would need to consider corrections from second-order perturbation theory. For $l = 1$ the expansion of the expectation value of w_μ starts with terms proportional to μ^{-4} and, in general, for an arbitrary l, the leading term of the expansion is proportional to $\mu^{-(2l+2)}$.

Now let us consider the expansion of the expectation values of w_μ for the linear forms $\alpha = \kappa\mu + \alpha_0$ and $c = \gamma\mu + c_0$ with the parameter γ and κ from the previous step. The condition for the elimination of the μ^{-2} term remains the same as above. The coefficient of μ^{-3} is equal to

$$d_3 = -\frac{8\, c_0\, \sqrt{\pi}}{27} + 2\, \alpha_0 \tag{A.10}$$

and is the same for all $l = 0$ states. The coefficient of μ^{-4} for $l = 0$ states is n-dependent. For $n = 1$ it is equal to

$$d_4 = \frac{1}{6} + \frac{64\, c_0}{81} - \frac{64\, \alpha_0}{9\, \sqrt{\pi}} + \frac{16}{27}\, \sqrt{\pi}\, c_0\, \alpha_0 - \frac{8\, \alpha_0^2}{3} \tag{A.11}$$

Solving equations $d_3 = d_4 = 0$ for c_0 and α_0 we obtain:

$$c_0 = \frac{27\, \alpha_0}{4\, \sqrt{\pi}}, \qquad \alpha_0 = \frac{8 \pm \sqrt{64 - 18\pi}}{12\sqrt{\pi}}. \tag{A.12}$$

This gives two possible solutions: $\{c_0 = 0.94364, \alpha_0 = 0.24778\}$ and $\{c_0 = 1.92115, \alpha_0 =$

0.50446}. When we use either of these sets of parameters and expand the first-order correction to fourth order in $1/\mu$, we have:

$$1s : 0, \quad 2s : -\frac{1}{48\mu^4}, \quad 3s : -\frac{2}{81\mu^4}, \quad 2p : \frac{1}{6\mu^4},$$

$$3p : \frac{1}{9\mu^4}, \quad 3d : 0.$$

Thus the error of the first-order correction to the eigenvalues is proportional to μ^{-4}.

Appendix B

Implementation of the Modified potentials

B.1 Modified two-electron integrals

Here we use the notation by Ahlrichs [Ahl06]. We start from unnormalized atom-centered GTOs, with orbital exponent α, at the center $\mathbf{A}$, and quantum numbers $\mathbf{a} = (a_x, a_y, a_z)$

$$|\mathbf{a}\rangle = \phi_a(\mathrm{r}) = (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} e^{-\alpha|\mathrm{r}-\mathbf{A}|^2} \tag{B.1}$$

Defining the functions ϕ_b,ϕ_c and ϕ_d in the same manner, the two-electron integral is defined as

$$I(\mathbf{ab}|\mathbf{cd}) = \int \phi_a(\mathrm{r}_1)\phi_b(\mathrm{r}_2)g(\mathrm{r}_{12}\phi_c(\mathrm{r}_1)\phi_d(\mathrm{r}_2) \tag{B.2}$$

where g is an interaction potential that we can define in different ways (including our

erf function), as I will show later.

$$\zeta = \alpha + \beta, \quad \eta = \gamma + \delta, \quad \rho = \frac{\zeta\eta}{\zeta + \eta}, \tag{B.3}$$

$$\mathbf{P} = \frac{\alpha\mathbf{A} + \beta\mathbf{B}}{\zeta}, \quad \mathbf{Q} = \frac{\gamma\mathbf{C} + \delta\mathbf{D}}{\eta}, \tag{B.4}$$

$$T = \rho|\mathbf{P} - \mathbf{Q}|^2, \tag{B.5}$$

$$S_{ab} = e^{\frac{\alpha\beta}{\zeta}|\mathbf{A}-\mathbf{B}|^2}, \quad S_{cd} = e^{\frac{\gamma\delta}{\eta}|\mathbf{C}-\mathbf{D}|^2}. \tag{B.6}$$

We define the basic integral

$$\mathbf{I_0} = \mathbf{I}(00|00) = \left(\frac{\pi}{\gamma_P + \gamma_Q}\right)^{3/2} S_{ab}S_{cd}G_0(\rho, T) \tag{B.7}$$

were

$$G_0(\rho, T) = \int e^{-\rho|\mathbf{r}-\mathbf{P}+\mathbf{Q}|^2} g(\mathrm{r}) d^3\mathbf{r}, \tag{B.8}$$

where $\mathbf{r} = (r_1 - r_2)$. For integrals with higher angular momentum, we use:

$$\hat{D} = \frac{\partial}{2\alpha\partial A_x}, \tag{B.9}$$

$$|\mathbf{a} + \mathbf{1}_x\rangle = \hat{D}|\mathbf{a}\rangle + \frac{a_x}{2\alpha}|\mathbf{a} - \mathbf{1}_x\rangle, \tag{B.10}$$

$$I((\mathbf{a} + \mathbf{1}_x)\mathbf{b}|\mathbf{cd}) = \hat{D}I(\mathbf{ab}|\mathbf{cd}) + \frac{a_x}{2\alpha}I((\mathbf{a} - \mathbf{1}_x)\mathbf{b}|\mathbf{cd}), \tag{B.11}$$

with $\mathbf{1}_i = (\delta_{ix}\delta_{iy}\delta_{iz})$ for $i = x, y, z$. Finally, we need to define $G_n(\rho, T)$:

$$G_n(\rho, T) = \left(-\frac{\partial}{\partial T}\right)^n G_0(\rho, T) \tag{B.12}$$

and

$$\hat{D}S_{ab} = (P_x - A_x)S_a b, \tag{B.13}$$

$$\hat{D}G_k(\rho, T) = -\frac{\rho}{\gamma_p}(P_x - Q_x)G_{k+1}(\rho, T), \tag{B.14}$$

$$\hat{D}(P_x - A_x) = -\frac{\beta}{1\alpha\gamma_p}, \quad \hat{D}(P_x - Q_x) = \frac{1}{2\gamma_P}. \tag{B.15}$$

.

B.1.1 G_0 from different interaction potentials $g(\mathrm{r})$

From equation (52) of Ahlrichs, we obtain the expression to evaluate the error function integrals:

$$g(\mathrm{r}) = \frac{\mathrm{erf}(\mu \mathrm{r})}{\mathrm{r}} : G_0(\rho, T) = \frac{2\pi}{\rho}\frac{\mu}{\sqrt{\mu^2 + \rho}}F_0\left(\frac{\mu^2 T}{\mu^2 + \rho}\right), \tag{B.16}$$

here I will define c_m

$$c_m = \frac{\mu^2}{\mu^2 + \rho} \tag{B.17}$$

then, equation (22) becomes

$$G_0 = \frac{2\pi}{\rho}\sqrt{c_m}F(c_m T) \tag{B.18}$$

and

$$F_n(c_m T) = \int_0^1 (c_m)^n t^{2n} e^{-c_m T t^2} dt. \tag{B.19}$$

This is equivalent to:

$$F_n(T) = \int_0^{\frac{\mu}{\sqrt{\mu^2+\rho}}} \sqrt{cm} t^{2n} e^{-T t^2} dt. \tag{B.20}$$

For the Gaussian function repulsion integrals we have:

$$g(r) = c e^{-\alpha^2 r^2} : G_0 = (\frac{\pi}{\rho+\alpha})^{3/2} c e^{-T f(\alpha)} \tag{B.21}$$

where:

$$f(\alpha) = \frac{\alpha}{\rho+\alpha}. \tag{B.22}$$

And the derivatives are:

$$G_n = (\frac{\pi}{\rho+\alpha})^{3/2} f(\alpha)^n c e^{-T f(\alpha)} \tag{B.23}$$

B.2 Short-range correlation functional for the asymptotic erfgau interaction

The correlation energy is obtained with a modified version of the Freeman [Fre77] fit with the model long-range potential. The short-range correlation energy per particle is computed as:

$$\epsilon_c^{sr,\mu}(r_s) = \epsilon_c^{VWN}(r_s) \left(1 - \frac{\epsilon_c^{lr,\mu}(r_s)}{\epsilon_c^{lr,\mu\to\infty}(r_s)}\right) \tag{B.24}$$

were $\epsilon_c^{VWN}(r_s)$ is the correlation energy per particle with the Coulomb interaction from the parametrization of Vosko, Wilk and Nusar(VWN) [VWN80]. Then, we use a grid with $\epsilon_c^{sr}(r_s)$ to construct an interpolant. For the interpolation we use a change of variables:

$$f^{\mu}(r_s) = \ln(-\epsilon_c^{\mu}(r_s)) \tag{B.25}$$

so the function interpolate on $r_s \in (0, \infty)$ using cubic Splines with natural boundary condition. From the interpolant we can also get the derivative of the function $f^{\mu}(r_s)$:

$$\frac{\partial \ln(-\epsilon_c(\mathbf{r}))}{\partial r_s(\mathbf{r})} = \frac{\frac{\partial \epsilon_c(\mathbf{r})}{\partial \rho(\mathbf{r})}\frac{\partial \rho(\mathbf{r})}{\partial r_s(\mathbf{r})}}{\epsilon_c(\mathbf{r})} \tag{B.26}$$

and

$$\begin{aligned}
\frac{\partial \epsilon_c(\mathbf{r})}{\partial \rho(\mathbf{r})} &= \frac{\epsilon_c(\mathbf{r})\frac{\partial \ln(-\epsilon_c(\mathbf{r}))}{\partial r_s(\mathbf{r})}}{\frac{\partial \rho(\mathbf{r})}{\partial r_s(\mathbf{r})}} \\
&= \frac{\epsilon_c(\mathbf{r}) f'(r_s)}{\frac{\partial \rho}{\partial r_s}} \\
&= \frac{\epsilon_c(\mathbf{r}) f'(r_s)}{-\frac{9}{4\pi r_s^4}}
\end{aligned}$$

Therefore, the short-range correlation potential is

$$v_c(\rho(\mathbf{r})) = \epsilon_c(\mathbf{r}) + \rho(\mathbf{r})\frac{\epsilon_c(\mathbf{r})f'(r_s)}{-\frac{9}{4\pi r_s^4}} \tag{B.27}$$

$$= \epsilon_c(\mathbf{r}) + \rho(\mathbf{r})\frac{\epsilon_c(\mathbf{r})f'(r_s)}{-\frac{9}{4\pi(\frac{3}{4\pi\rho})^{4/3}}} \tag{B.28}$$

$$= \epsilon_c(\mathbf{r}) + \rho(\mathbf{r})\frac{\epsilon_c(\mathbf{r})f'(r_s)}{-\frac{6^{2/3}\pi^{1/3}}{(\frac{1}{\rho})^{4/3}}} \tag{B.29}$$

$$= \epsilon_c(\mathbf{r}) + \rho(\mathbf{r})\frac{\epsilon_c(\mathbf{r})f'(r_s)}{-6^{2/3}\pi^{1/3}\rho^{4/3}} \tag{B.30}$$

$$= \epsilon_c(\mathbf{r}) + \rho(\mathbf{r})\frac{\epsilon_c(\mathbf{r})f'(r_s)}{-\rho(\mathbf{r})(36\pi\rho(\mathbf{r}))^{1/3}} \tag{B.31}$$

$$= \epsilon_c(\mathbf{r}) - \frac{\epsilon_c(\mathbf{r})f'(r_s)}{(36\pi\rho(\mathbf{r}))^{1/3}} \tag{B.32}$$

B.3 Short-range exchange functional for the asymptotic erfgau interaction

The exchange energy of the uniform electron gas per particle is

$$\epsilon_x(\mathbf{r}) = C_x\rho(\mathbf{r})^{1/3} \tag{B.33}$$

and the energy functional

$$E_x[\rho(\mathbf{r})] = C_x\int \rho(\mathbf{r})^{4/3}d\mathbf{r} = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\int \rho(\mathbf{r})^{4/3}d\mathbf{r}, \tag{B.34}$$

so the potential $\delta E_x/\delta\rho$ is

$$v_x(\mathbf{r}) = C_x\frac{4}{3}\rho(\mathbf{r})^{1/3}. \tag{B.35}$$

The expression used in Mathematica [Mat] to compute the exchange energy per particle is (this comes from equation (3.36) of Fetter-Walecka)

$$\epsilon_x(\mathbf{r}) = -4\pi \frac{1}{(k_F)^3/3\pi^2}(2\pi)^6 \frac{4}{3}\pi(k_F)^3 2k_F \int_0^1 4\pi \left(1 - \frac{3}{2}x + \frac{1}{2}x^3\right) dx \tag{B.36}$$

substituting $k_F = 1/\alpha r_s$ and $\alpha = \left(\frac{4}{9\pi}\right)^{1/3}$ and solving the integral we get:

$$\epsilon_x(\mathbf{r}) = -\frac{3}{4r_s}\left(\frac{3}{2\pi}\right)^{2/3} \tag{B.37}$$

which is equal to equation B.33. Now, if we use the model potential $\tilde{V}(\mathbf{r}) = \frac{\mathrm{erf}(\mu\mathbf{r})}{\mathbf{r}} + c(\mu)e^{-\alpha(\mu)^2\mathbf{r}^2}$

$$\epsilon_x(\mathbf{r}) = -4\pi \frac{1}{(k_F)^3/3\pi^2}(2\pi)^6 \frac{4}{3}\pi(k_F)^3 2k_F \int_0^1 4\pi \left(1 - \frac{3}{2}x + \frac{1}{2}x^3\right) \tag{B.38}$$

$$\left(e^{-\frac{(2k_F x)^2}{4\mu^2}} + \frac{(2k_F x)^2}{4\pi} c(\mu)\pi^{3/2} a(\mu)^{-3} e^{-\frac{(2k_F x)^2}{4a(\mu)^2}}\right) dx \tag{B.39}$$

After evaluating the integral in (B.39) and the change of variables: $v_1 = (3\pi^2\rho)^{1/3}$ and $v_2 = v_1^2$, the exchange energy per particle of the asymptotic erfgau potential is:

$$\epsilon_x(\mathbf{r}) = \left(\frac{3^{2/3}a(\mu)c(\mu)}{2\pi^{7/6}} - \frac{a(\mu)c(\mu)e^{-v_2/a(\mu)^2}}{2\times 3^{1/3}\pi^{7/6}} + \frac{3^{2/3}\mu^2}{2\pi^{5/2}} - \frac{e^{-v2/\mu^2}\mu^2}{3^{1/3}\pi^{5/3}}\right)\left(\frac{1}{\rho}\right)^{1/3} \tag{B.40}$$

$$+ \left(-\frac{a(\mu)^3 c(\mu)}{3\pi^{5/2}} + \frac{a(\mu)^3 c(\mu)e^{-v_2/a(\mu)^2}}{3\pi^{5/2}} - \frac{\mu^4}{6\pi^3} + \frac{e^{-v_2/\mu^2}\mu^4}{6\pi^3}\right)\left(\frac{1}{\rho}\right) \tag{B.41}$$

$$- \frac{1}{2}c(\mu)\mathrm{erf}(v_1/a(\mu)) - \frac{\mu\ \mathrm{erf}(v_1/\mu)}{\sqrt{\pi}} \tag{B.42}$$

The short-range exchange potential,

$$v_x = \frac{\delta E_x[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} = \epsilon_x(\mathbf{r}) + \rho(\mathbf{r})\frac{\partial\epsilon_x(\mathbf{r})}{\partial\rho(\mathbf{r})}, \tag{B.43}$$

where

$$\begin{aligned}\frac{\partial\epsilon_x(\mathbf{r})}{\partial\rho(\mathbf{r})} = &\left(-\frac{a(\mu)c(\mu)}{2\times 3^{1/3}\pi^{7/6}} - \frac{a(\mu)c(\mu)e^{-v_2/a(\mu)^2}}{2\times 3^{1/3}\pi^{7/6}} - \frac{\mu^2}{2\times 3^{1/3}\pi^{5/2}}\right)\left(\frac{1}{\rho}\right)^{4/3} \\ &+ \left(\frac{a(\mu)^3c(\mu)}{3\pi^{5/2}} - \frac{a(\mu)^3c(\mu)e^{-v_2/a(\mu)^2}}{3\pi^{5/2}} + \frac{\mu^4}{6\pi^3} - \frac{e^{-v_2/\mu^2}\mu^4}{6\pi^3}\right)\left(\frac{1}{\rho}\right)^2\end{aligned}$$

is then

$$\begin{aligned}v_x(\mathbf{r}) = \epsilon_x(\mathbf{r}) &+ \left(-\frac{a(\mu)c(\mu)}{2\times 3^{1/3}\pi^{7/6}} - \frac{a(\mu)c(\mu)e^{-v_2/a(\mu)^2}}{2\times 3^{1/3}\pi^{7/6}} - \frac{\mu^2}{2\times 3^{1/3}\pi^{5/2}}\right)\left(\frac{1}{\rho}\right)^{1/3} \\ &+ \left(\frac{a(\mu)^3c(\mu)}{3\pi^{5/2}} - \frac{a(\mu)^3c(\mu)e^{-v_2/a(\mu)^2}}{3\pi^{5/2}} + \frac{\mu^4}{6\pi^3} - \frac{e^{-v_2/\mu^2}\mu^4}{6\pi^3}\right)\left(\frac{1}{\rho}\right) \qquad \text{(B.44)} \\ &= \left(\frac{a(\mu)\,c(\mu)}{\pi^{7/6}} - \frac{a(\mu)c(\mu)e^{-v2/a^2}}{\pi^{7/6}} + \frac{\mu^2}{\pi^{5/3}} - \frac{e^{-v2/\mu^2}}{\pi^{5/3}}\right)\left(\frac{1}{3\rho}\right)^{1/3} \\ &- \frac{1}{2}c(\mu)\mathrm{erf}(v_1/a(\mu)) - \frac{\mu\ \mathrm{erf}(v_1/\mu)}{\sqrt{\pi}} \qquad \text{(B.45)}\end{aligned}$$

Bibliography

[Ahl06] R. Ahlrichs. A simple algebraic derivation of the obara-saika scheme for general two-electron interaction potentials. *Phys. Chem. Chem. Phys.*, 8:3072–3077, 2006.

[Bec88] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, 1988.

[BH69a] S. F. Boys and N. C. Handy. A calculation for the energies and wave-functions for states of neon with full electronic correlation accuracy. *Proc. R. Soc. London Ser. A*, 1500(310):43, 1969.

[BH69b] S. F. Boys and N. C. Handy. A condition to remove the indeterminacy in interelectronic correlation functions. *Proc. R. Soc. London Ser. A*, (309):209, 1969.

[BP74] R. J. Buenker and S. D. Peyerimhoff. Individualized configuration selection in ci calculations with subsequent energy extrapolation. *Theor. Chim. Acta*, 35:33–58, 1974.

[BR89] A. D. Becke and M. R. Roussel. Exchange holes in inhomogeneous systems: A coordinate-space model. *Phys. Rev. A*, 39:3761–3767, 1989.

[CA80] D. M. Ceperley and B. J. Alder. *Phys. Rev. Lett.*, 45:566, 1980.

[CHG08a] J.-D. Chai and M. Head-Gordon. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.*, 10:6615–6620, 2008.

[CHG08b] J.-D. Chai and M. Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.*, 128:084106, 2008.

[CMSY07] A. J. Cohen, P. Mori-Sánchez, and W. Yang. *J. Chem. Phys.*, 127:034101, 2007.

[CS99] F. Colonna and A. Savin. *J. Chem. Phys.*, 110:2828, 1999.

[CSM79] D. P. Carroll, H.J. Silverstone, and R.M. Metzger. Piecewise polynomial configuration interaction natural orbital study of 1s2 helium. *J. Chem. Phys.*, 71:4142–4163, 1979.

[DGY+06] N. J. DeYonker, T. Grimes, S. Yockel, A. Dinescu, B. Mintz, T. R. Cundari, and A. K. Wilson. The correlation-consistent composite approach: Application to the g3/99 test set. *J. Chem. Phys.*, 125:104111, 2006.

[DHC91] E. R. Davidson, S. A. Hangstrom, and S. J. Chakravorty. Groun-state correlation energies for two- to ten-electron atomic ions. *Phys. Rev. A*, 44(11):7071–7083, 1991.

[Dun89] T. H. Dunning. Gaussian basis set for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90:1007–1023, 1989.

[Fre77] D. L. Freeman. Coupled-cluster expansion applied to the electron gas: Inclusion of rinf and exchange effects. *Phys. Rev. B*, 15(12):5512–5521, 1977.

[FW77] A. L. Fetter and J. D. Walecka. *Quantum Theory of Many-Particle Systems*. L. I. Schiff, Third Edition, McGraw-Hill Book Company, 1977.

[GEAKS16] C. E. González-Espinoza, P. W. Ayers, J. Karwowski, and A. Savin. Smooth models for the coulomb potential. *Theor. Chem. Acc.*, (135):256, 2016.

[GGS06] P. Gori-Giorgi and A. Savin. Properties of short-range and long-range correlation energy density functionals from electron-electron coalescence. *Phys. Rev. A*, 73:032506, 2006.

[GL76] O. Gunnarsson and B. I. Lundqvist. *Phys. Rev. B*, 13:4274–4298, 1976.

[HJ74a] J. Harris and R. O. Jones. *PJ. Phys. F: Metal Phys.*, 4:1170–1186, 1974.

[HJ74b] J. Harris and R. O. Jones. The surface energy of a bounded electron gas. *J. Phys. F: Metal Phys.*, 4:1170–1186, 1974.

[HK64] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phy. Rev.*, 136(3B):B864–B871, 1964.

[HKKN97] T. Helgaker, W. Klopper, H. Koch, and J. Noga. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.*, 106:9639–9646, 1997.

[HMR73] B. Huron, J. Malrieu, and P. Rancurel. *J. Chem. Phys.*, 58:5745–5759, 1973.

[HPKW09] J. G. Hill, K. A. Peterson, G. Knizia, and H.-J. Werner. Extrapolating mp2 and ccsd explicitly correlated correlation energies to the complete basis set limit with first and second row correlation consistent basis sets. *J. Chem. Phys.*, 131(19):194105, 2009.

[HTU16] A. A. Holmes, N. M. Tubman, and C. J. Umrigar. Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling. *J. Chem. Thoery Comput.*, 12(8):3674–3680, 2016.

[HUS17] A. A. Holmes, C. J. Umrigar, and S. Sharma. Excited states using semistochastic heat-bath configuration interaction. *J. Chem. Phys.*, 147:164111, 2017.

[Hyl28] E. A. Hylleraas. *Z. Phys.*, 48:469, 1928.

[Hyl29] E. A. Hylleraas. *Z. Phys.*, 53:347–366, 1929.

[Hyl30] E. A. Hylleraas. *Z. Phys.*, 65:209, 1930.

[JM73] W. Jones and N. H. March. *Theoretical Solid State Physics*, volume 2. Dover, New York, 1973.

[JZS85] K. Jankowski, D. W. Zaharevitz, and H. J. Silverstone. On definitions of l convergence of atomic correlation energies. *J. Chem. Phys.*, 82:1969–1972, 1985.

[Kat57] T. Kato. On the eigenfunctions of many-particle systems in quantum mechanics. *Commun. Pure Appl. Math.*, 10:151–177, 1957.

[KBV12] L. Kong, F. A. Bischoff, and E. F. Valeev. Explicitly correlated r12/f12 methods for electronic structure. *Chem. Rev.*, 112:75–107, 2012.

[KC03] J. Karwowski and L. Cyrnek. *Computational Methods in Science and Technology*, 9(1-2):67–78, 2003.

[KS65] W. Kohn and L. J. Sham. *Phys. Rev.*, 140:A1133, 1965.

[Kut85] W. Kutzelnigg. r12-dependent terms in the wave function as closed sums of partial wave amplitudes for large l. *Theor. Chim. Acta*, 68(6):445–469, 1985.

[Lep98] G. P. Lepage. *How to renormalize the Scrödinger equation.* J.C.A. Barata, A.P.C. Malbouisson, S.F. Novaes (Eds.), Particles and Fields, Proceedings of the Ninth J.A. Swieca Summer School, World Scientific, Singapore, 1998.

[Lev79] M. Levy. Universal variational functional of electron densities, first-order density matrices, and natural spin-orbitals and solution of the u-representability problem. *Proc. Natl Acad. Sci.*, 76:6062–6065, 1979.

[LP75] D. C. Langreth and J. P. Perdew. *Solid State Commun.*, 17:1425–1429, 1975.

[LP80] D. C. Langreth and J. P. Perdew. *Phys. Rev. B*, 21(12):2828, 1980.

[LP85a] M. Levy and J. P. Perdew. *The constrained search formulation of density functional theory. In Density Functional Methods in Physics.* Dreizler, R.M. and Providencia, J. da, eds.). New York: Plenum, 1985.

[LP85b] M. Levy and J. P. Perdew. Hellmann-feynman, virial and scaling requisites for the exact universal density functional. shape of the correlation potential and diamagnetic susceptibility for atoms. *Phys. Rev. A*, 32:2010–2021, 1985.

[LSWS97] T. Leininger, H. Stoll, H.-J. Werner, and A. Savin. Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.*, 275:151–160, 1997.

[LWNC15] J. H. Lloyd-Williams, R. J. Needs, and G. J. Conduit. Pseudopotential for the electron-electorn interaction. *Phys. Rev. B*, 92:075106, 2015.

[LYP85] M. Levy, W. Yang, and R.G Parr. A new functional with homogeneous coordinate scaling in density functional theory: F[p, a]. *Chem. Phys.*, 83:2334–2336, 1985.

[Mat] Mathematica, Wolfram Research, Inc., Mathematica, Version 11.0, Champaign, IL (2016).

[MP79] G. O. Morrell and R.G Parr. Estimates of helium l-limit energies. *J. Chem. Phys.*, 71:4139–4141, 1979.

[MSCY06] P. Mori-Sánchez, A. J. Cohen, and W. Yang. *J. Chem. Phys.*, 124:091102, 2006.

[PBE96] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approimation made simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.

[PCL+03] R. Pollet, F. Colonna, T. Leininger, H. Stoll, H.-J. Werner, and A. Savin. *Int. J. Quantum Chem.*, 91(2):84–93, 2003.

[Per91] J. P. Perdew. Unified theory of exchange and correlation beyond the local density approximation. *Electronic Stricture of solids '91, edited by P. Ziesche and H. Eschrig (Akademiew Verlag, Berlin)*, pages 11–20, 1991.

[PMGGB06] S. Paziani, S. Moroni, P. Gori-Giorgi, and G. B. Bachelet. *Phys. Rev. B*, 73:155111, 2006.

[PSLS02] R. Pollet, A. Savin, T. Leininger, and H. Stoll. Combining multideterminantal wave functions with density functionals to handle near-degeneracy in atoms and molecules. *J. Chem. Phys.*, 116(4):1250–1258, 2002.

[PW92] J. P. Perdew and Y. Wang. *Phys. Rev. B*, 45:13244, 1992.

[PWD94] K. A. Peterson, D. E. Woon, and T.H. Dunning. Benchmark calculations with correlated molecular wave-functions. 4. the classical barrier height of the h+h2-¿h2+h reaction. *J. Chem. Phys.*, 100:7410–7415, 1994.

[PY94] R. G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford: Oxford University Press, 1994.

[PZ81] J. P. Perdew and A. Zunger. *Phys. Rev. B*, 23:5048, 1981.

[refa] "M. Richer, F. Heidar-Zadeh, G. Sánchez-Díaz, M. Chan (in preparation).".

[refb] "T. Verstraelen, P. Tecmer, F. Heidar-Zadeh, K. Boguslawski, M. Chan, Y. Zhao, T.D. Kim, S. Vandenbrande, D. Yang, C.E. González-Espinoza, S. Fias, P.A. Limacher, D. Berrocal, A. Malek, P.W. Ayers. HORTON 2.0.1, http://theochem.github.com/horton/, 2016.".

[Sav95] Andreas Savin. *Advances in Density Functional Methods*, volume 1. D. P. Chong Editor, World Scientific 1996, 1995.

[SCP03] A. Savin, F. Collona, and R. Pollet. Adiabatic connection approach to density functional theory of electronic systems. *Int. J. Quant. Chem.*, 93:166–190, 2003.

[SCT98] A. Savin, F. Colonna, and J.-M. Teuler. *Electronic Density Functional Theory.* J. F. Dobson et al. (eds.), Springer Science+Business Media New York 1998, 1998.

[SE16] J.B. Schriber and F. A. Evangelista. *J. Chem. Phys.*, 144:161106, 2016.

[Sei03] M. Seidl. *J. Quantum Chem.*, 91:145, 2003.

[SHJ+17] S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar. Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory. *J. Chem. Theory Comput.*, 13:1595–1604, 2017.

[SK02] I. Sirbu and H. F. King. Perturbation theory of the electron correlation cusp based on a partitioning of the electron-electron interaction into

long- and short-range components. *J. Chem. Phys.*, 117(14):6411–6416, 2002.

[SK03] I. Sirbu and H. F. King. Perturbational approach to the electron correlation cusp applied to helium-like atoms. *Int. J. Quant. Chem.*, 92:433–450, 2003.

[Sla28a] J. C. Slater. Central fields and rydberg formulas in wave mechanics. *Phys. Rev.*, 31:333, 1928.

[Sla28b] J. C. Slater. The normal state of helium. *Phys. Rev.*, 32:349, 1928.

[SO12] A. Szabo and N. Ostlund. *Modern quantum chemistry.* Dover, New York, 2012.

[Tau93] M. Taut. *Phys. Rev. A*, 48:3561–3566, 1993.

[TCH09] A. M. Teale, S. Coriani, and J. Helgaker. *J. Chem. Phys.*, 130:104111, 2009.

[TCH10] A. M. Teale, S. Coriani, and J. Helgaker. *J. Chem. Phys.*, 133:164112, 2010.

[TCS05] J. Toulouse, F. Colonna, and A. Savin. Exchange-correlation potentials and local energies per particle along nonlinear adiabatic connections. *Mol. Phys.*, 103:2725–2734, 2005.

[TLT+16] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley. A deterministic alternative to the full configuration interaction quantum monte carlo method. *J. Chem. Phys.*, 145:044112, 2016.

[Tou04] Long-range-short-range separation of the electron-electron interaction in density-functional theory. *Phys. Rev. A*, 70:062505, 2004.

[Tou05] Short-range exchange and correlation energy density functionals: Beyond the local-density approximation. *J. Chem. Phys.*, 134(122):014110, 2005.

[TPSS03] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria. Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.*, 91:146401, 2003.

[Tru98] D. G. Truhlar. Basis-set extrapolation. *Chem. Phys. Lett.*, 294:45–48, 1998.

[TSF04] J. Toulouse, A. Savin, and H.-J. Flad. *Int. J. Quantum Chem.*, 100:1047–1056, 2004.

[VB89] A. Valance and H Bergeron. *J. Phys. B: Mol. Opt. Phys.*, 22:L65–L69, 1989.

[VHKS06] O. A. Vydrov, J. Heyd, A. Krukau, and G. E. Scuseria. Importance of short-range versus long-range hartree-fock exchange for the performance of hybrid density functionals. *J. Chem. Phys.*, 125:074106, 2006.

[VS06] O. A. Vydrov and G. E. Scuseria. Assessment of a long range corrected hybrid functional. *J. Chem. Phys.*, 125:234109, 2006.

[VS07] O. A. Vydrov and G. E. Scuseria. Tests of functionals for systems with fractional electron number. *J. Chem. Phys.*, 126:154109, 2007.

[VWN80] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys*, 58(8):1200, 1980.

[Yan98] W. Yang. *J. Chem. Phys.*, 109:10107–10110, 1998.

[YTH04] T. Yanai, D. Tew, and N. Handy. A new hybrid exchange-correlation functional using the coulomb-attenuating method (cam-b3lyp). *Chem. Phys. Lett.*, 393:51–57, 2004.