2015

The Development of Efficient Geminal-Based Methods for Quantum Chemical Calculations

Brett Allen Cagg
University of South Carolina

The Development of Efficient Geminal-based Methods for Quantum Chemical Calculations

by

Brett Allen Cagg

Bachelor of Science
Missouri Western State University 2009

Master of Science
University of South Carolina 2011

Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in

Chemistry

College of Arts and Sciences

University of South Carolina

2015

Accepted by:

Vitaly Rassolov, Major Professor

Sophya Garashchuk, Committee Chair

Dmitry Peryshkov, Committee Member

Daniel Dix, Committee Member

Lacy Ford, Senior Vice Provost and Dean of Graduate Studies

Dedication

To my beautiful wife, Nikki: I am incredibly grateful for all the support and love you have shown me the past few years. I am glad that we will weather all of life's storms together.

To my mother, Patricia: I am thankful that you always supported my decision to move to South Carolina in search of my dreams. I will always remember the precious moments I got to spend with you these last few years, and I miss you terribly now that you are gone.

Acknowledgments

I gratefully acknowledge the support from my family and friends. I thank my coworkers in the theory group at the University of South Carolina (Dr. Jim Mazzuca, Bryan Nichols, Dr. Lei Wang, and Bing Gu) for many productive discussions about research throughout the past four years. I also acknowledge my colleagues working in the lab of Dr. Donna Chen (Dr. Samuel Tenney, Dr. Randima Galhenage, Mara Levine, Audrey Duke, and Kangmin Xe) for sharing parts of this journey with me. In addition, my family, both in South Carolina and Missouri, has always been incredibly supportive of my work, and I couldn't have accomplished this without their help. Last, and most importantly, I appreciate the time and patience of both my advisor, Dr. Vitaly Rassolov, and the chair of my doctoral committee, Dr. Sophya Garashchuk. I have learned so much from them, and I attribute all my future success to their careful guidance.

Abstract

Incorporation of electron correlation to improve the accuracy of computations remains a driving force in quantum chemical method development. Traditionally, electron correlation effects are divided into static and dynamic correlation parts, and the inclusion of both results in methods that are impractical for large chemical systems. The goal of this doctoral research is to develop a method that efficiently accounts for both components of electron correlation in a separate but balanced manner. The approach focuses on combining a geminal method, called the antisymmetrized product of singlet-type, strongly orthogonal geminals (SSG), with dynamic correlation by either density functionals or a recently developed, linear, two particle, correlation operator.

The SSG method is a quantum chemical method that groups all electrons in a chemical system into pair functions called geminals. Within geminals, electrons can adapt multiple electronic configurations which allows the method to incorporate most static correlation. However, between geminals, the electron-electron repulsion is mean-field and a strong orthogonality constraint forbids intergeminal electron excitation. Combining SSG with electron correlation from a density functional, implemented in SSG(DFT), is shown to improve optimized properties of main group diatomic molecules at small basis sets. However, overcorrelation due to double counting of dynamic correlation is observed at larger basis sets. The SSG($\hat{C}$) method improves the situation by describing dynamic correlation while negating the double counting error inherent in SSG(DFT). In addition, the SS$_p$G method is developed to relax the strong orthogonality constraint and allow intergeminal electronic excitation. The

method provides a small portion of dynamic correlation energy unaccounted for in the aforementioned methods. The SS_pG method is compared to non-orthogonal geminal methods and is shown to be a viable method to relax the constraint. Performance of all methods is analyzed from applications to representative sets of atoms and small molecules and compared to other electronic structure methods as benchmarks.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The Schrödinger Equation and the Wavefunction

In quantum mechanics, all chemical systems evolve in time according to the Schrödinger equation.

$$i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t} = \hat{H}\Psi(\mathbf{r},t)\,. \tag{1.1}$$

Equation 1.1 is a partial differential equation describing the time evolution of the mathematical representation of any chemical system, which is called a wavefunction and denoted by Ψ. If the wavefunction is known for a system under study, we can determine all of its properties and how they will evolve in time. Unfortunately, outside of a few simple systems, there are no known analytical wavefunctions, and approximations are needed to make use of the Schrödinger equation.

When quantum mechanics is applied to chemistry, theoretical chemists are often interested in the structure of the ground state. The ground state wavefunction is defined as the lowest energy wavefunction for a specific nuclear configuration and number of electrons. In the absence of external forces, the ground state wavefunction will not change with time and will solve the time-independent Schrödinger equation:

$$\hat{H}\Psi(\mathbf{r}) = E\Psi(\mathbf{r})\,, \tag{1.2}$$

$$\hat{H} \equiv \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ee} + \hat{V}_{Ne}\,. \tag{1.3}$$

In equation 1.2, Ψ is now the time-independent wavefunction. The Hamiltonian operator $\hat{H}$ is the sum of kinetic energy operators for nuclei $\hat{T}_n$ and electrons $\hat{T}_e$, as well as, the potential energy operators for interactions between nuclei $\hat{V}_{nn}$, electrons $\hat{V}_{ee}$, and both $\hat{V}_{ne}$. Now, instead of a partial differential equation, the Schrödinger equation takes on the form of an eigenvalue equation. While this simplifies the search for the ground state wavefunction, this equation is still too complicated to be solved due to the positional interdependence of the last three terms in expression 1.3.

A further simplification, the Born-Oppenheimer approximation [6], assumes that electronic and nucleic motion are separable within a chemical system. The Born-Oppenheimer approximation is usually adequate because electrons are much lighter and travel much faster than nuclei, which allows electronic positions to adapt instantaneously to nucleic motion. Thus, in the Born-Oppenheimer approximation, the kinetic energy of the nuclei is assumed to be very small in comparison to the kinetic energy of the electrons, and the first term in equation 1.3 is neglected. Therefore, the Schrödinger equation can be solved for each nuclear configuration which specifies $\hat{V}_{nn}$. The surviving terms form the electronic Hamiltonian operator.

$$\hat{H} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{nn} + \hat{V}_{ee}\,. \tag{1.4}$$

In addition to these simplifications, another widely used assumption is that relativity has little effect on the chemistry of common chemical systems. Relativity, in fact, only starts to strongly affect electronic structure when heavy metals are considered [7]. In the context of this research, we accept all of these approximations and choose to search for the solutions to the time-independent, non-relativistic, electronic form of the Schrödinger equation.

1.2 The Hartree-Fock Method

For multi-electron systems, the aforementioned electronic Schrödinger equation has no analytical solution due to electron-electron repulsion. Thus, further approximations must be made to determine the ground state wavefunctions. Such approximations form the quantum chemical methods implemented in computational chemistry programs. For instance, the Hartree-Fock method, the cornerstone of electronic structure theory, is founded upon the mean-field approximation. The mean-field approximation simplifies the search for the ground state wavefunction from an N-electron problem, where N is the number of electrons, to N one-electron problems, and the resulting Hartree-Fock wavefunction structure is:

$$\Psi_{HF}(\mathbf{r_1}, \mathbf{r_2}...\mathbf{r_N}) = \hat{A}[\phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_2)...\phi_N(\mathbf{r}_N)]\,. \tag{1.5}$$

Ψ_{HF} is an antisymmetrized product of one-electron functions called molecular orbitals and denoted by ϕ_i. The term "Antisymmetric" means that the sign of the wavefunction must change if any two electrons are interchanged. The antisymmetrization operator $\hat{A}$ is a sum over all such permutations, and the antisymmetric wavefunctions satisfy the Pauli exclusion principle. To simplify notation, equation 1.5 can be written as a normalized Slater determinant.

$$\Psi_{HF}(\mathbf{r_1}, \mathbf{r_2}...\mathbf{r_N}) = |\phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_2)...\phi_N(\mathbf{r}_N)|\,. \tag{1.6}$$

The molecular orbitals used to build the Hartree-Fock wavefunction are formed as a linear combination of atomic orbital basis functions:

$$\phi_i(\mathbf{r}) = \sum_{\lambda} C_{i,\lambda}\chi_\lambda(\mathbf{r})\,. \tag{1.7}$$

In equation 1.7 the molecular orbital, denoted as ϕ_i, is a spin up, or α, orbital, and throughout the rest of this dissertation, an overbar over a molecular orbital $\bar{\phi}_i$

denotes a spin down, or β, spin orbital. The Hartree-Fock wavefunction, unless stated otherwise, is built using spin-restricted orbitals, which means that β spin orbitals are restricted to be spatially identical to α spin orbitals. Spin unrestricted orbitals do not have to be identical spatially and often break symmetry if a molecule is stretched to bond breaking distances.

The atomic orbital basis functions, denoted in equation 1.7 as $\chi_i(\mathbf{r})$, used in making the molecular orbitals comprise a basis set which must be chosen when performing electronic structure calculations[8]. Most quantum chemical methods use standard basis sets which combine with the overall accuracy of the method to determine the level of accuracy of the calculation. However, larger basis sets, which produce higher levels of accuracy, are more computationally expensive. Therefore, when choosing a basis set, we want to choose the smallest basis set that will produce desired accuracy. There is a variety of carefully designed basis sets available today with the most popular being either the gaussian basis sets created by Pople [9] or the correlation consistent basis sets created by Dunning[10]. In gaussian basis sets, each basis function is made from a linear combination of gaussian functions of varying widths.

$$\chi_l(\mathbf{r}) = \sum_g B_g \mathbf{r}^l e^{-\alpha_g \mathbf{r}^2} \,. \tag{1.8}$$

It is important to note from equation 1.8 that these gaussian functions can represent s-,p-,d-,f-, or g-type orbitals by changing l, which is the angular momentum quantum number. The widths of the gaussian functions are controlled by changing the size of the α component in the exponential. A larger exponent creates a narrower gaussian. In reality, a variety of gaussian widths are used to closely mimic the more accurate, but less computationally efficient Slater atomic orbitals [11].

The coefficients in equation 1.7, denoted as $C_{i,\lambda}$, are optimized according to the variational principle to give the lowest energy Hartree-Fock wavefunction. The variational principle [11] establishes the Hartree-Fock energy as an upper-bound to the

exact energy.

1.3 Electron Correlation

Traditionally, electron correlation energy is defined as the energy neglected in the Hartree-Fock method within a basis set.

$$E_{corr} = E_{exact} - E_{HF} \,. \tag{1.9}$$

Here, the exact energy E_{exact} is the energy of the exact, non-relativistic, ground state wavefunction within the Born-Oppenheimer approximation using a specific basis set. Whereas, E_{HF} is the energy of the computed Hartree-Fock wavefunction within the same basis set. A long-standing goal of electronic structure is to recover this missing electron correlation energy without incurring impractical computational expense.Thus, it is instructive to consider the failures of the mean-field approximation in Hartree-Fock and the possible corrections to such problems.

First, the Hartree-Fock wavefunction represents one electronic configuration. In comparison, the exact ground state wavefunction Ψ consists of a linear combination of all possible electronic configurations, which is the form of the Full Configuration Interaction method (FCI) wavefunction.

$$\Psi_{FCI}(\mathbf{r_1}, \mathbf{r_2}...\mathbf{r_N}) = \sum_a C_a |\phi_{a1}(\mathbf{r}_1)\phi_{a2}(\mathbf{r}_2)...\phi_{aN}(\mathbf{r}_N)| \,. \tag{1.10}$$

Each determinant contributes to a lowering of energy of the FCI wavefunction, and the CI coefficients C_a are determined that give the lowest FCI energy. Thus, the Hartree-Fock wavefunction, being the lowest energy portion of this wavefunction, is the electronic configuration with the greatest CI coefficient. However, it must be noted that the number of determinants in the FCI wavefunction is exponentially dependent on the number of electrons in the system, and the determination of the CI coefficients increases impractically with system size. Therefore, we cannot use

the FCI method to determine the ground state wavefunction in most cases. On the other hand, it can be surmised that the use of a single configurational wavefunction to describe a chemical system gives rise to a multiconfigurational error.

Correlation energy within a given basis set is usually divided into two types called static and dynamic correlation energy. Correcting the multiconfigurational error accounts for static correlation energy. The classic example of static correlation is observed when stretching the bond in the hydrogen molecule. When a hydrogen molecule is stretched to an infinite bond length, the two hydrogen atoms do not interact. Thus, the energy of an infinitely stretched hydrogen molecule should equal the sum of the energy of two hydrogen atoms. However, when the Hartree-Fock method is used, this is not the case. The Hartree-Fock method is thus not size-consistent, which means the infinitely stretched hydrogen molecules have a spurious, non-physical mutual interaction. This interaction is due to both electrons occupying the bonding orbital in hydrogen even when infinitely stretched. This size consistency can be restored through the use of unrestricted molecular orbitals which will allow orbitals to break spatial symmetry upon stretching. However, the preferred method of accounting for static correlation is through the use of multiple configurations which allows fractional population of both the antibonding and bonding orbital. The result is a lowering of energy, or addition of correlation energy, at long bond distances to recover the atomic limits.

In contrast, dynamic correlation can be thought of as coordinate-based electron correlation. The Hartree-Fock method makes use of the mean-field approximation to simplify the search for an approximate wavefunction. Within the mean-field approximation, each electron experiences the average repulsion from all other electrons. However, each electron should experience repulsion from each individual electron. The type of correlation only increases as electrons are brought closer together. Thus, dynamic correlation energy does play a larger role at closer interelectronic distances

akin to coulombic repulsion.

1.4 Geminal-based Total Electron Correlation Methods

Multiconfigurational methods are meant to recover static electron correlation by including multiple electronic configurations within their wavefunctions. However, the multiconfigurational methods which do incorporate said correlation without incurring FCI expense are still considerably expensive and require careful consideration when building the multiconfigurational portion of the wavefunction. For example, the multiconfigurational self consistent field theory (MCSCF), which is a simple linear combination of more than one electronic configuration, often employs a complete active space simplification to its wavefunction. This simplification uses a subset of orbitals, called active orbitals, to create multiple electronic configurations. These active orbitals are deemed the greatest contributors to molecular bonding and reactivity. The remaining subset of orbitals, which are core orbitals, contribute to all electronic configurations included in the wavefunction. However, there is no well defined method to determine how many and which molecular orbitals are the most "active". Therefore, methods like these must use enough electronic configurations to be able to include sufficient static correlation while also avoiding crippling expense.

As a remedy, a family of well defined, computationally inexpensive, multiconfigurational methods based on two-electron geminals has emerged as an attractive alternative. The advantages of geminal methods can be visualized if we once again consider the Hartree-Fock wavefunction. The Hartree-Fock method builds its wavefunction from the one-electron orbitals. In contrast, geminal theory takes a step beyond the Hartree-Fock method by introducing a wavefunction built of multiconfigurational, two-electron functions called geminals.

$$\psi_A(\mathbf{r_1}, \mathbf{r_2}) = \sum_i D_i |\phi_i(\mathbf{r_1})\bar{\phi}_i(\mathbf{r_2})| \,. \tag{1.11}$$

In equation 1.11, D_i are the geminal expansion coefficients, which are similar to the familiar CI coefficients, and ψ_A represents a single geminal A. These geminals are, in fact, solutions to the two-electron Schrödinger equation as each geminal by itself is an FCI wavefunction for a two-electron system. However, beyond two electrons, the general form of the geminal wavefunction yields problems. First, the most general geminal method, named the antisymmetric product of geminals (APG), which is analogous to the Hartree-Fock method for geminals, is nearly as complicated as an FCI wavefunction.

$$\Psi_{APG}(\mathbf{r_1}, \mathbf{r_2}...\mathbf{r_N}) = |\psi_A(\mathbf{r_1}, \mathbf{r_2})\psi_B(\mathbf{r_3}, \mathbf{r_4})...\psi_{N_g}(\mathbf{r}_{2N_g-1}, \mathbf{r}_{2N_g})|\,. \tag{1.12}$$

This necessitates the use of an approximation to reduce the computational expense. The most popular constraint placed on geminal wavefunctions is called strong orthogonality.

$$\int \psi_a(\mathbf{r}_1, \mathbf{r}_2)\psi_b(\mathbf{r}_1, \mathbf{r}_3)d\mathbf{r}_1 = 0 \qquad a \neq b\,. \tag{1.13}$$

This constraint simplifies the search for the best geminal wavefunction by making each orbital pair "choose" which geminal they populate. Strong orthogonality, thus, eliminates electron correlation between orbitals on different geminals which simplifies normalization. The antisymmetrized product of these strongly orthogonal geminals is appropriately called the antisymmetrized product of strongly orthogonal geminals method (APSG). However, APSG still requires a significant amount of chemical knowledge a priori due to decisions regarding the number of electron configurations included in each geminal. A variant of APSG which optimizes the number of determinants within a geminal, the geminal expansion coefficients, and molecular orbital coefficients was introduced recently as a solution to this quandry [12]. This method, named the antisymmetrized product of Singlet-type Strongly orthogonal Geminals

(SSG), is the multiconfigurational wavefunction of interest in the rest of this dissertation.

While the SSG method is attractive due to its simplicity, it still has problems due to the form of the wavefunction. First, the use of the strong orthogonality constraint reduces computational cost by sacrificing some dynamic correlation. A true description of electron correlation could not hope to be complete without this missing correlation. Second, the mean-field approximation is used intergeminally which causes each geminal to feel the mean-field repulsion of all other geminals. This separate portion of dynamic correlation, the dominant portion, also needs to be recovered in order to improve the accuracy of SSG calculations.

The first part of this dissertation describes how we address the latter question of recovering dynamic correlation energy beyond the mean-field repulsion of geminals. This portion of correlation energy can be recovered either by the formulation of a method-specific perturbation theory, as has already been established for SSG as SSG(EN2)[13], or by using inexpensive standard density correlation functionals. We choose to explore the latter in the SSG(DFT) method described in the next chapter. We continue the study of SSG(DFT) by the exploring the use of more orbital dependent types of correlation with the creation of SSG(PBEα) in chapter three and the SSG($\hat{C}$) method in chapter four. Subsequently, the second part of the dissertation will describe how the dynamic correlation associated with the use of strong orthogonality is recovered. This new method uses a quick, perturbation-like, geminal-at-a-time approach outside of SSG optimization loops and is called SS_pG.

Chapter 2

Density Functional Model of Multireference Systems Based on Geminals[1]

2.1 Abstract

In this chapter, a density functional model based on the variationally optimized, multiconfigurational wavefunction created from strongly orthogonal singlet-type electron geminals, also called SSG, is presented. A rescaled correlation-only PBE functional is used to account for dynamic correlation absent in the geminal wavefunction. The performance of the model is assessed from geometry optimizations on the G2/97 test set diatomics, using two different basis sets. The results presented here are encouraging for the development of a geminal hybrid that approximates total electron correlation in a computationally efficient manner. Some of the results included in this chapter have been reproduced from information in reference [14], and permission for the reproduction of these results is granted by the publisher.

2.2 Introduction

Density Functional Theory and Dynamic Correlation

Density Functional Theory (DFT) is a theory meant to circumvent problems associated with the search for the correct ground state wavefunction for a chemical system.

[1]Cagg, B. A.; Rassolov, V. A. *Chem. Phys. Lett.* 2012, *543* , 205-207. Reprinted here with permission from the publisher.

In fact, instead of searching for the ground state wavefunction, the ground state electron density ρ is the principle quantity of interest in DFT.

$$\rho \equiv \rho(\mathbf{r_1}) = \int \Psi(\mathbf{r_1}, \mathbf{r_2}, ...\mathbf{r_N})^* \Psi(\mathbf{r_1}, \mathbf{r_2}, ...\mathbf{r_N}) \delta\mathbf{r_2} \delta\mathbf{r_3} ... \delta\mathbf{r_N} \tag{2.1}$$

In equation 2.1, the square of the N-electron wavefunction Ψ is integrated over all electron coordinates save one. Usually, the wavefunction used to produce the electron density is assumed to be a Hartree-Fock-like, single-determinant wavefunction to simplify the search. It is important to note here that the electron density is a local variable of three spatial coordinates. In contrast, the wavefunction is a function of 3N spatial coordinates. The search for the ground state density as opposed to the corresponding wavefunction is justified because the classical terms in the energy expression for a ground state wavefunction are actually functionals of the ground state density.

$$E_{DFT} = T[\rho] + V_{ext}[\rho] + J[\rho] + XC[\rho] \tag{2.2}$$

In equation 2.2, The first three terms are the kinetic energy functional, external potential energy functional, and interelectron repulsion functional. The external potential energy V_{ext} actually results from the familiar nuclear interactions with other nuclei and the electrons. Similarly, the interelectron potential $J[\rho]$ and $T[\rho]$ are respectively the V_{ee} and T_e from equation 1.4. Thus, the only unfamiliar term in the DFT energy expression is the last term $XC[\rho]$ which is called the exchange-correlation density functional. The exchange-correlation functional houses all of the energy arising from quantum effects, which include, for example, the effects of the Pauli-exclusion principle. Unfortunately, although it is theorized to exist [15], the correct universal exchange-correlation functional that precisely predicts the ground state density for any chemical system is not known. Therefore, DFT researchers have

created a number of approximate exchange-correlation functionals designed to search for the best approximate ground state density.

The standard exchange-correlation density functionals are often designed as the sum of two separate functionals $X[\rho]$ the exchange functional, and $C[\rho]$ the correlation functional. The exchange functional is designed to mimic the antisymmetric nature of a single-determinant wavefunction. In contrast, the correlation functional attempts to modify the electron density by adding missing dynamic correlation missing. This makes DFT potentially more accurate than the Hartree-Fock method while avoiding some of the expense involved in searching for the ground state wavefunction. However, since the exchange and correlation functionals are only approximate, DFT runs into problems that are not encountered in Hartree-Fock. The most famous example is the spurious self-interaction of electrons treated with DFT. A possible correction and excellent introduction to self-interaction in DFT is provided in reference [16].

Nonetheless, Density Functional Theory is viewed as sufficiently accurate for chrmical properties and has gained widespread use. The popularity of Density Functional Theory in computation of molecular properties [17] arises due to both the qualitative accuracy of the mean-field model in many chemical applications and progress in defining models of electron correlation that are useful to non-specialists [18]. In fact, the approximate correlation functionals, which are derived from these models, have become increasingly dependent on input quantities beyond just the electron density. The popular PBE correlation functional [19], used later in this chapter, uses both the density and the gradient of the density $\nabla\rho$ in its formulation to provide an account for the change in electron correlation with change in the density. This provides information about electron correlation in the "neighborhood" of the electron which is usually termed semi-local correlation information.

Providing Dynamic Correlation to the SSG Wavefunction

As mentioned in the previous chapter, the use of multiconfigurational (MC), or multireference (MR), wavefunctions is complicated by the need to properly account for static correlation efficiently while neglecting as few dynamic correlation effects as possible. Therefore, it is highly desirable to combine a DFT-like description of dynamic correlation effects with a simple well defined MC wavefunction ansatz. Such a combined model must separate correlation effects into two non-overlapping domains: (i) correlation effects arising from explicit interaction of configurations, and (ii) those described by the density-based functional of interest. This can be performed in coordinate space based on range-separation [20, 21, 22], or in configuration space based on orbital classification [23, 24]. In fact, one of the simplest MR wavefunctions in the form of the Generalized Valence Bond using Perfect Pairing (GVB-PP) [25] was used in combination with a Local Spin Density Approximation (LSDA) density functional by Kraka [26] to study dissociation energies of diatomics. The LSDA functional provides correlation based solely on the electron density and thus is the simplest form of DFT correlation functional. It was found that the LSDA functional improves the GVB results only modestly, and this deficiency was attributed to the shortcomings of the GVB wavefunction.

In our first attempt at creating a multireference-DFT hybrid, we combine a SSG and correlation energy from a density functional in a new method called SSG(DFT). For development purposes, we focused solely on studying the unrestricted version of SSG(DFT) called USSG(DFT) but it was formulated in such a way that we could have chosen any form of spin restriction of SSG. Unlike GVB, the spin-unrestricted SSG (USSG) wavefunction describes essentially all static correlation, as indicated by the smallness of perturbative amplitudes when second-order perturbation theory is applied to USSG describing known MR systems [13]. The USSG wavefunction is well defined, rigorously size consistent and, unlike most other MR methods, can

be optimized in the efficient Atomic Orbital (AO) representation. This makes it a suitable candidate for a reference wavefunction to be combined with the DFT description of dynamic correlation.

Our ultimate goal is to formulate the simplest computational model applicable to MR systems with the accuracy and degree of empiricism comparable to mainstream DFT models. Because two-electron integrals are required for the SSG model, the use of exact exchange, as apposed to approximate exchange, does not incur additional computational costs. In single reference DFT, use of pure exact exchange with correlation functionals, while tempting due to the latter being a small fraction of the former, leads to poor accuracy in geometry optimizations [27]. This may not be the case for the SSG reference. Therefore in the present work we explore the use of a correlation-only functional combined with exact SSG energy components. Absence of approximate exchange functionals avoids most of the self-interaction error. For simplicity we use the correlation functional outside of the SCF optimization cycles, and evaluate it at the end of wavefunction optimization. Including the correlation functional in the self-consistent search for the ground state wavefunction is not technically difficult, but the self-consistent search in the similar GVB-LSDA method by Kraka has shown it should make little difference energetically.

In addition, it should be noted that since the SSG wavefunction describes two-electron systems exactly in the complete basis set limit, a suitable correlation functional to be used with SSG need not be derived based on a two-electron system, as is done in the Lee-Yang-Parr LYP functional [28]. Instead, we choose the original Perdew-Burke-Ernzenhof, PBE, functional [19] based on its popularity in chemical applications and our positive experience with the PBEsol version of it [29] in transition metal applications [30].

2.3 The SSG(DFT) Method

The original PBE functional is designed to recover all dynamic correlation energy. The SSG model describes some of it, typically 20 - 30 % in a given basis [31, 32, 12]. Expressed in the language of the correlation hole, the SSG model describes the medium to short range contribution of the correlation hole to the correlation energy. Description of the shortest range depends on the presence of high angular momentum functions in the basis set, and ,therefore, is strongly basis set dependent. This is likely similar to the $O(l^{-4})$ dependence of the atomic correlation energy [33, 34] on the limiting angular momentum quantum number l of the basis set. Description of the long range contribution depends on the degree of localization of geminals. It is probably relatively basis set independent, but does depend on the system, particularly on the bond lengths. Construction of the universal DFT correlation functional compatible with the SSG model is, therefore, challenging. As a simple, exploratory, well defined approach we rescaled PBE correlation energy by a basis set dependent factor, computed by matching the atomic correlation energies for the first and second row atoms [35, 36]. It is important that we do not match the total energies, in order to avoid compensating for the basis set incompleteness on the mean-field level.

Our model is defined as a three step process. First, the SSG wavefunction, Ψ_{SSG}, is computed as in the ab initio SSG model by minimizing the expectation value, E_{SSG}, of the exact Hamiltonian, $\hat{H}$, with respect to variation of all parameters (molecular orbitals, geminal expansion coefficients, and assignments of orbital pairs to individual geminals).

$$E_{SSG} = \frac{\langle \Psi_{SSG} | \hat{H} | \Psi_{SSG} \rangle}{\langle \Psi_{SSG} | \Psi_{SSG} \rangle}, \tag{2.3}$$

Second, the one-electron spin density and dimensionless density gradient are evaluated for Ψ_{SSG} and used to compute E_C^{PBE} according to the traditional DFT procedure

[19]. Finally, the final energy is defined as a sum of the two components, with DFT scaled by a basis set-dependent factor.

$$E_{SSG(PBE)} = E_{SSG} + \lambda_{PBE} E_C^{PBE} \,, \tag{2.4}$$

The rescaling parameter λ_{PBE} is defined to minimize the root-mean-square deviation of atomic correlation energies, as computed by Chakravorty *et al* [36], from our model correlation energy. The model correlation energy is defined as the difference between $E_{SSG(PBE)}$ and spin-unrestricted mean-field energy E_{UHF} evaluated in the same basis. The parameter λ_{PBE} is the only new parameter in the model.

In testing SSG(PBE) we used the popular 6-31G* basis set and a much larger G3MP2large set [37]. The rescaling parameter values were found to be 0.9695 for 6-31G* set and 0.8674 for G3MP2large set.

2.4 Results

One of the most practical indicators of the quality of chemical models describing the molecular ground state is equilibrium geometry. Therefore, we examined the bond lengths of all 28 diatomic molecules from the G2/97 test set [38]. The results are summarized in Table 2.1 for both 6-31G* and G3MP2large basis sets. The resulting bond lengths are compared to experimental values.

The experimental bond length of each molecule is given in the first column and the results of both methods are seen as a difference in bond length between experiment and the method compared. It was noted in earlier studies [12] that the SSG method was superior to MP2, HF, and CCSD if only the diatomics with elements of moderate electronegativity were considered. The worst predictions for molecular geometry came when diatomics of extreme electronegativity such as F_2 and Na_2 were considered. This problem is now resolved by applying the PBE correlation correction to the SSG calculation: Addition of DFT correlation energy makes the SSG(PBE) method

superior to SSG and to other methods when calculating these highly electronegative diatomic molecules at the 6-31G* basis set, as shown by root mean square deviations of computed geometries from the experimental ones in Table 2.2.

Table 2.1 The deviations from experimental bond length [1] for SSG(PBE) and SSG methods using the 6-31G* and G3MP2large basis sets. The ΔR and $\Delta R'$ values correspond to 6-31G* and G3MP2large respectively. The experimental bond lengths and deviations are reported in angstroms.

molecule	R_e(exp)	$\Delta R(SSG(PBE))$	$\Delta R(SSG)$	$\Delta R'(SSG(PBE))$	$\Delta R'(SSG)$
LiH	1.5957	0.035	0.057	-0.008	0.013
BeH	1.3429	0.000	0.009	-0.006	0.001
CH	1.1199	-0.012	0.000	-0.018	-0.010
NH	1.0362	-0.006	0.009	-0.018	-0.008
OH	0.9697	-0.005	0.006	-0.018	-0.007
FH	0.9168	0.000	0.013	-0.016	-0.007
HCl	1.2746	-0.009	0.008	-0.017	-0.005
Li_2	2.6729	0.016	0.058	-0.014	0.030
LiF	1.5639	-0.015	-0.005	-0.012	0.009
CN	1.1718	-0.012	-0.003	-0.024	-0.015
CO	1.1283	-0.012	-0.002	-0.021	-0.012
N_2	1.0977	-0.008	0.003	-0.018	-0.010
NO	1.1508	-0.015	-0.001	-0.025	-0.011
O_2	1.2075	-0.043	0.019	0.048	-0.025
F_2	1.4119	0.021	0.092	0.006	0.079
Na_2	3.0788	0.013	0.088	0.085	0.144
Si_2	2.2460	-0.057	-0.022	-0.070	-0.045
P_2	1.8934	-0.010	0.022	-0.029	-0.004
S_2	1.8892	-0.020	0.030	-0.027	0.019
Cl_2	1.9879	-0.030	0.083	-0.039	0.007
NaCl	2.3608	0.003	0.044	0.004	0.049
SiO	1.5097	-0.026	-0.008	-0.036	-0.022
SC	1.5349	-0.021	0.007	-0.026	-0.014
SO	1.4811	-0.020	-0.008	-0.050	-0.033
ClO	1.5696	-0.001	0.111	-0.003	0.048
FCl	1.6283	-0.001	0.058	-0.035	0.010
H_2	0.7414	-0.001	0.005	-0.003	0.001
HS	1.3400	-0.011	0.007	-0.017	-0.006

Table 2.2 Root mean squared deviation and average bond length deviations for the compared methods using 6-31G* basis and G3MP2large basis sets. All values are in angstroms.

method	RMSD_{6-31G*}	RMSD_{G3MP2L}	$\langle R_e \rangle_{6-31G*}$	$\langle R_e \rangle_{G3MP2L}$
SSG(PBE)	0.0201	0.0313	-0.009	-0.018
SSG	0.0424	0.0372	0.023	0.006
B3LYP	0.0263	0.0136	0.014	0.002
PBE	0.0338	0.0185	0.028	0.014
HF	0.0743	0.1117	0.004	0.007
MP2	0.0289	0.0445	0.019	-0.006
CCSD	0.0332	0.0280	0.020	-0.009

Part of the high accuracy of the new method is fortuitous, as seen from the larger G3MP2large basis set results. The G3MP2large set yields bond lengths that are on average 0.018 Å too short, leading to RMSD that is comparable (but slightly inferior) to CCSD, better than SSG and MP2, and worse than traditional DFT.

We plot the bond length dependence of the CO molecule energy to explain the results. The SSG(PBE) bond energy computed with 6-31G* basis is compared to other methods in Figure 2.1.

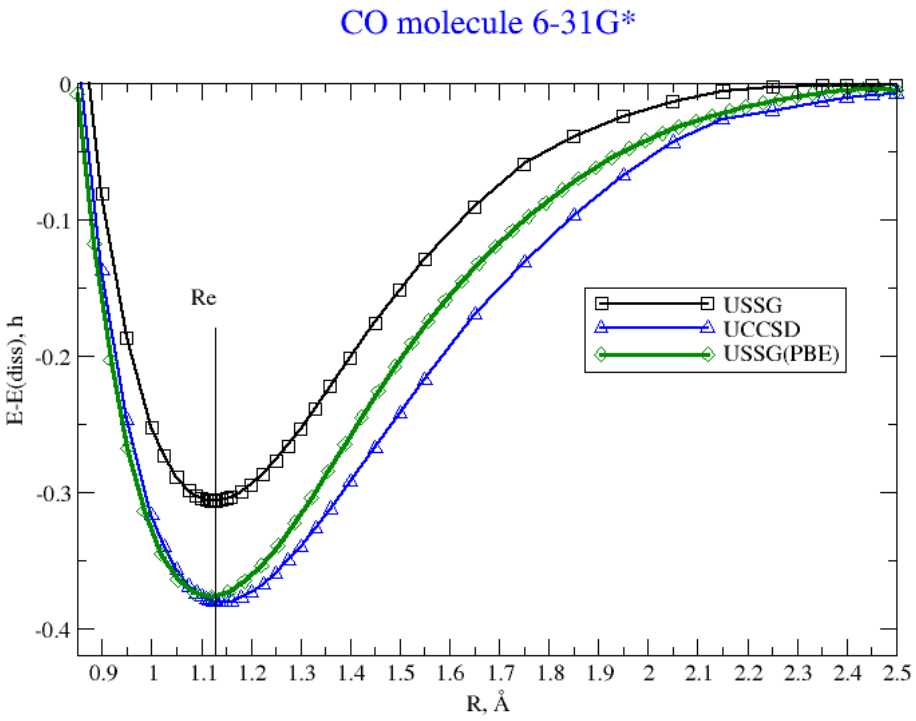


Figure 2.1 Potential energy surface of CO molecule at the 6-31G* basis set. Calculations with SSG(PBE) model are compared with SSG and CCSD. All energies are given relative to dissociated energies and computed with the same methods.

Compared to the coupled-cluster methods, which probably follow the exact correlation energy fairly well near the equilibrium bond distance, the PBE contribution vanishes too fast as the bond is stretched. This causes shortening of the equilibrium bonds. This shortcoming of the model is consistent with the PBE correlation functional being overly sensitive to the medium-range correlation hole, which changes the most rapidly as the bond breaks. It is also consistent with the better performance of our model using a smaller basis set, in which SSG recovers an insignificant fraction of the dynamic correlation. It is likely that a modification of a correlation functional is required to fix the problem.

2.5 Conclusion and Possible Corrections

Our goal is to formulate a DFT-like practical method applicable to MR chemical systems, and, therefore, based on a MR wavefunction. We argue that the spin-unrestricted geminal model USSG is a promising functional form as a reference, as it is shown to incorporate most static correlation, when applied to MR systems such as ozone or transition metal hydrides [13]. Here, we present a simple approach of using a standard correlation-only density functional with the SSG wavefunction. In the new method, called SSG(DFT), the DFT correlation energy is evaluated outside of wavefunction optimization loops, and scaled to avoid double counting of correlation effects. The resulting model is size-consistent and is largely free from self-interaction error due to use of exact exchange. When used with the widely popular 6-31G* basis set to optimize diatomic geometries, the model is superior to other traditional methods. The use of a larger basis set indicates that this accuracy is partially fortuitous, but the proposed method will likely achieve higher accuracy after modification of standard functionals. The main effect of the PBE correction is to shorten the bonds, and it has the most pronounced effects on the diatomics with extremely electronegative atoms, which are poorly described by the SSG model.

It is clear there are two deficiencies that must be solved if the SSG(DFT) method is to become more accurate. First, correlation energy cannot be cleanly separated as we propose in the introduction. SSG describes some small amount of dynamic correlation which is double counted in SSD(DFT) in all cases. This is partially reduced by the rescaling parameter in SSG(DFT) but must be mitigated more rigorously. Certainly at short distances there is substantial double counting of dynamic correlation within the method that worsens with basis set size. Second, SSG(DFT) must attempt to capture more non-local forms of electron correlation to recover neglected energy highlighted in figure 2.1. The latter correction is the focus of the next chapter.

Chapter 3

SSG(PBEα): Inclusion of Orbital Dependent Dynamic Correlation

3.1 Abstract

An orbital dependent density functional correction to the SSG(DFT) method is presented. The α functional, a custom made density functional based on the α inhomogeneity parameter used in the TPSS and MGGA_MS2 functionals, is mixed with the PBE functional to create the modified SSG(DFT) method we call SSG(PBEα). This is done to reduce an error inherent in the SSG(PBE) method arising due to the semi-local character of PBE. Geometric optimizations are performed on the 28 diatomic molecules from the G2/97 test set at the 6-31G* and G3MP2large levels of theory. The mixing of α and PBE is optimized to give the best results for each basis set, and the results of both basis sets improve from the original SSG(PBE) method.

3.2 Introduction

The (SSG) method has been shown to be a rather good starting approximation to the Full Configuration Interaction (FCI) method based on graded orthogonality arguments [3] due to its recovery of configurational-based, static correlation energy. In the last chapter the accuracy of the SSG method was shown to improve by approximately adding the remaining portion of correlation energy using a correlation density functional. This method, the SSG(DFT) method, added neglected dynamic correlation energy via scaled down contributions from standard DFT correlation functionals out-

side of SSG optimization loops which showed great promise for geometry optimization at small basis sets. However, those promising results deteriorated when calculations were performed at larger basis sets. The large basis set problem was diagnosed to be two-fold. First, the empirical scaling parameter for DFT functionals used to reduce a well-known, correlation double counting error was not sufficient to completely negate the error. At small basis sets, the resulting overcorrelation was small enough that the results were not substantially affected. However, at large basis sets, the residual double counting error resulted in dramatically shortened bond lengths when compared to experiment. A second error became obvious when analyzing the potential energy surface, reproduced in figure 2.1, for the carbon monoxide molecule at 6-31G*. Here, the SSG(DFT) curve displayed a steep loss of correlation energy just outside of equilibrium relative to the highly accurate CCSD method. This showcased a particular failure in PBE, and most density correlation functionals, to account for longer range correlation effects.

In this chapter, we address the latter problem in SSG(PBE) by mixing in the energy calculated from a custom made, orbital dependent functional we call the α functional. The α functional is a functional built on a semi-local building block used for the meta-GGA functionals created by Perdew and coworkers [39, 40] and is chosen due to its apparent sensitivity to bond character within a molecule[41]. The resulting SSG(PBEα) method is then used to geometrically optimize the same 28 diatomic molecules from the G2/97 test set as in the previous chapter at the 6-31G* and G3MP2large levels of theory.

3.3 The PBEα Functional

The orbital dependent density functional α is based on the inhomogeneity parameter used in the well known TPSS functional[39] and MGGA_MS2 [40] meta-GGA functionals. These meta-GGA functionals form the third rung in the Jacob's ladder

metaphor used extensively by Perdew in describing the accuracy of density functionals in comparison to experiment[39]. In the metaphor, the Local Spin Density Approximation is the lowest rung and uses only the electron density described in equation 2.1 in the formulation of density functionals. The second rung uses the Generalized Gradient Approximation (GGA), on which PBE is formulated, and makes use of the density and the gradient of the density. The meta-GGA functionals described here use another semi-local ingredient in the formulation of density functionals called the kinetic energy density τ.

$$\tau(\mathbf{r}) = \sum_{i,\sigma}^{occ} \frac{|\nabla \phi_{i,\sigma}(\mathbf{r})|^2}{2} \,. \tag{3.1}$$

It should be noted, in equation 3.1, that the kinetic energy density depends on the gradient of the molecular orbitals and that the sum runs over occupied molecular orbitals of α and beta spin. In practice, the kinetic energy density, like electron density in DFT, is calculated on a spatial grid. However, like the density gradient $\nabla\rho$, kinetic energy density is a semi-local term. In some applications, as in the PKZB functional [42], it is used in conjunction with the Weiczacker kinetic energy density to reduce the self-interaction error inherent in DFT.

$$\tau^w = \frac{|\nabla\rho|^2}{8\rho} \,. \tag{3.2}$$

However, the kinetic energy density is of interest in this chapter because it is shown to be the main ingredient in an important and sensitive density functional parameter called the α parameter.

$$\alpha = \frac{\tau - \tau^w}{\tau^{unif}} \,. \tag{3.3}$$

In equation 4.4, the τ^{unif} is the kinetic energy density of the uniform electron gas; the fundamental model system upon which all density correlation functionals are based. The α parameter provides a sensitive measure of bond character due to the

comparisons between these three forms of kinetic energy density. The Wieczacker kinetic energy density is the kinetic energy density of a single orbital, and the kinetic energy density of the uniform electron gas most resembles kinetic energy density resulting from metallic bonding. Thus, if the kinetic energy in the bonding region of a molecule is much less than unity, the bond is covalent. If the kinetic energy density in such regions is close to unity, the bonding is more metallic.

As was shown in the SSG(PBE) method, density functionals, when combined with SSG, have an overall inability to recover correlation effects at intermediate bond lengths when bonds are starting to break. We believe that the sensitivity to changes in bond character that the α functional exhibits may be the key to correcting this error within SSG(DFT). However, we keep the computational simplicity of SSG(DFT) by converting the α parameter into a density functional to mix with the density functional chosen in the SSG(DFT) method.

A few changes have been made to the α parameter in order to work within the SSG formulation. First, we modify the kinetic energy density τ to correctly calculate the kinetic energy density of SSG as opposed to a single-determinant method. The original definition of the kinetic energy density is applied to only occupied orbitals and assumes whole occupation of said orbitals. In comparison, the SSG wavefunction partially occupies all orbitals and a sum over all orbitals ϕ_i is required with an inclusion of the geminal expansion coefficients D_i.

$$\tau^G = \sum_{i,\sigma} \frac{|\nabla \phi_{i,\sigma}|^2 D_i^2}{2} \,. \tag{3.4}$$

We call this modified version of kinetic energy density the geminal kinetic energy density and signify it as τ^G. In addition, the Wieczacker kinetic energy density is modified to make a spin dependent version.

$$\tau_{sd}^w = \sum_{\sigma} \frac{|\nabla \rho_\sigma|^2}{8\rho_\sigma} \,. \tag{3.5}$$

We choose this spin dependent τ_{sd}^{w} in our studies due to a better linear fit of the spin dependent α functional to atomic correlation results for open shell atoms. The α parameter is also made dimensionless and multiplied by total density to produce the resulting α functional.

$$\alpha(\rho) = \frac{\tau^{G} - \tau_{sd}^{w}}{\rho^{2/3}} \,. \tag{3.6}$$

The DFT dynamic correlation calculation is performed outside of self consistent loops on the resulting SSG wavefunction with the separate PBE and α functionals. The functionals are both scaled down to reproduce known atomic correlation energies for elements of atomic number 2-17 [35, 36] as was done previously for SSG(PBE) calculations. The scaling is done independently, and the coefficients are basis set specific. The resulting scaled down results are then mixed together using optimized mixing prefactors b and b' with the sum of these prefactors being equal to unity.

$$E_{SSG(PBE\alpha)} = bE_{PBE} + b'E_{\alpha} \,. \tag{3.7}$$

The mixing prefactors are optimized empirically with respect to root mean squared deviation (RMSD) from equilibrium bond length for the 28 diatomic molecules in the G2/97 test set[38]. In this way, the created α functional attempts to provide the best possible correction to SSG(PBE) by adding more non-local correlation at intermediate bond lengths.

3.4 Results and Conclusions

The previous SSG(PBE) calculations showed promise for the small 6-31G* basis set by reducing the RMSD for the G2/97 test set diatomics in half when compared with SSG[14]. These same calculations were performed with new wavefunctions that were re-optimized and the results were nearly the same as can be seen from figure 3.1 with only two molecules, ClO and SC, being significantly different. The RMSD worsens

with these new wavefunctions to 0.0214 from the previously stated 0.0202 but it is still superior to all other methods compared as stated previously. These new re-optimized wavefunctions were used for testing the new SSG(PBEα) calculation with 6-31G*.

The remarkable improvement in RMSD for 6-31G* for SSG(PBE) over SSG was deemed at least partly fortuitous because of the equally remarkable disintegration of RMSD to 0.0313 using the large G3MP2large [37] basis set. The α functional was designed to improve upon SSG(PBE) results by provide missing electron correlation energy at intermediate bond lengths. This lack of correlation energy in SSG(PBE) can be visualized from the steep ascent in the potential energy surface of carbon monoxide relative to CCSD right outside of equilibrium. The SSG(α) curve on the other hand has a much slower rise in energy at the same bond lengths as seen in figure 3.4. However, if the potential energy surface of SSG(α) is analyzed at extended bond lengths it is observed that a spurious long range interaction occurs between the carbon and oxygen atoms. This is not unique to the carbon monoxide molecule and indeed is a feature of the use of the α functional. Thus, we use α as a small correction to SSG(PBE) to add a small amount of intermediate bond length correlation while only allowing a minimal amount of this spurious interaction. Finally, it should be noted that while SSG(α) requires that molecules be stretched significantly farther than standard SSG(DFT) to obtain atomic limits, SSG(α)), as well as SSG(PBEα), retain the size consistency of SSG.

Table 3.1 The deviations from experimental bond length [1] for SSG(PBE) with original and re-optimized wavefunctions using 6-31G* basis set. $\Delta R(SSG(PBE))$ corresponds to calculations with original wavefunctions, while $\Delta R(SSG(PBE))'$ corresponds to corrected deviations from re-optimized wavefunctions. The experimental bond lengths and deviations are reported in angstroms.

molecule	R_e(exp)	$\Delta R(SSG(PBE))$	$\Delta R(SSG(PBE))'$
LiH	1.5957	0.035	0.036
BeH	1.3429	0.000	0.000
CH	1.1199	-0.012	-0.012
NH	1.0362	-0.006	-0.006
OH	0.9697	-0.005	-0.007
FH	0.9168	0.000	0.001
HCl	1.2746	-0.009	-0.008
Li_2	2.6729	0.016	0.016
LiF	1.5639	-0.015	-0.018
CN	1.1718	-0.012	-0.015
CO	1.1283	-0.012	-0.011
N_2	1.0977	-0.008	-0.007
NO	1.1508	-0.015	-0.015
O_2	1.2075	-0.043	-0.042
F_2	1.4119	0.021	0.021
Na_2	3.0788	0.013	0.013
Si_2	2.2460	-0.057	-0.056
P_2	1.8934	-0.010	-0.009
S_2	1.8892	-0.020	-0.021
Cl_2	1.9879	-0.030	-0.030
NaCl	2.3608	0.003	0.004
SiO	1.5097	-0.026	-0.025
SC	1.5349	-0.021	-0.012
SO	1.4811	-0.020	-0.020
ClO	1.5696	-0.001	0.042
FCl	1.6283	-0.001	0.000
H_2	0.7414	-0.001	-0.001
HS	1.3400	-0.011	-0.010

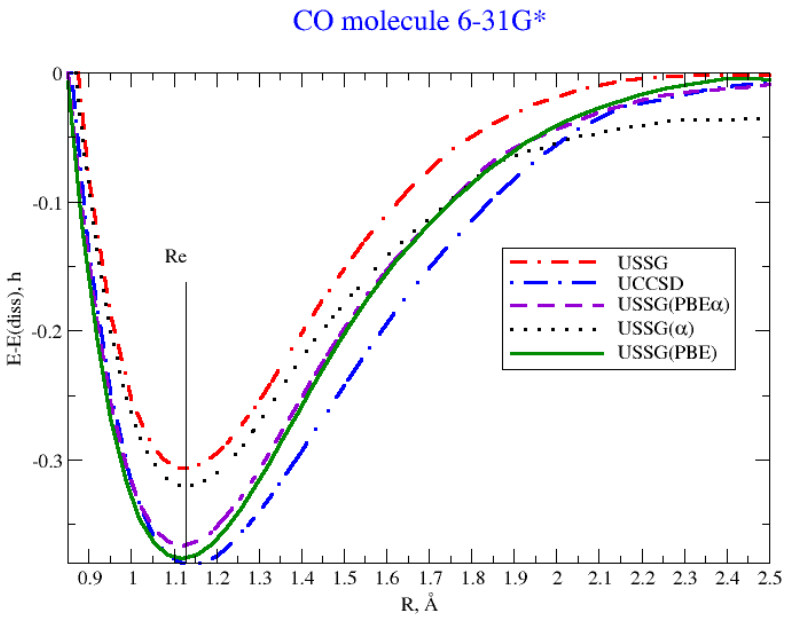


Figure 3.1 Potential energy surface of CO molecule using the 6-31G* basis set. Calculations with SSG(α) SSG, SSG(PBE) and CCSD are shown. All energies are given relative to dissociated energies, computed with the same methods.

Optimizing the mixing prefactors b and b' in SSG(PBEα) produced a slight improvement in the small 6-31G* basis set compared to SSG(PBE) with a mixing of 0.17 α to 0.83 PBE. The RMSD improved from 0.0214 to 0.0200 with notable improvements to LiH, Si_2, P_2, and S_2 as seen in table 3.2. A more significant improvement is observed when the larger G3MP2large basis set is used. When mixing 0.15 α and 0.85 PBE, the RMSD improves from 0.0313 using SSG(PBE) to 0.0270 using SSG(PBEα). It is interesting to point out that the most significant improvement is observed in Na_2 which improved from 0.085 Å in deviation to 0.028 Å. In this case, SSG(PBEα) shortened the bond, which may be due to the weak character of the disodium bond. All save seven molecules either improved slightly or stayed the same in the larger basis set. This improvement makes SSG(PBEα) slightly better than

Table 3.2 The deviations from experimental bond length [1] for SSG(PBEα) and SSG(PBE) using the 6-31G* and G3MP2large bases. The $\Delta R(SSG(PBE\alpha))$ and $\Delta R(SSG(PBE))$ columns are results for the 6-31G* basis sets, while the $\Delta R(SSG(PBE\alpha))'$ and $\Delta R(SSG(PBE))'$ columns are the results for the G3MP2large basis set. The experimental bond lengths and deviations are reported in angstroms.

molecule	R_e(exp)	$\Delta R(SSG(PBE\alpha))$	$\Delta R(SSG(PBE))$	$\Delta R(SSG(PBE\alpha))'$	$\Delta R(SSG(PBE))'$
LiH	1.5957	0.021	0.036	-0.018	-0.008
BeH	1.3429	-0.004	0.000	-0.009	-0.006
CH	1.1199	-0.011	-0.012	-0.018	-0.018
NH	1.0362	-0.003	-0.006	-0.016	-0.018
OH	0.9697	-0.005	-0.007	-0.015	-0.018
FH	0.9168	0.004	0.001	-0.013	-0.016
HCl	1.2746	-0.006	-0.008	-0.015	-0.017
Li_2	2.6729	-0.014	0.016	-0.024	-0.014
LiF	1.5639	-0.031	-0.018	0.033	-0.012
CN	1.1718	-0.012	-0.015	-0.021	-0.024
CO	1.1283	-0.009	-0.011	-0.020	-0.021
N_2	1.0977	-0.005	-0.007	-0.016	-0.018
NO	1.1508	-0.013	-0.015	-0.024	-0.025
O_2	1.2075	-0.039	-0.042	-0.045	-0.048
F_2	1.4119	0.029	0.021	0.012	0.006
Na_2	3.0788	0.013	0.013	0.028	0.085
Si_2	2.2460	-0.047	-0.056	-0.066	-0.070
P_2	1.8934	0.000	-0.009	-0.023	-0.029
S_2	1.8892	-0.014	-0.021	-0.022	-0.027
Cl_2	1.9879	-0.028	-0.030	-0.037	-0.039
NaCl	2.3608	-0.005	0.004	0.011	0.004
SiO	1.5097	-0.022	-0.025	-0.034	-0.036
SC	1.5349	-0.008	-0.012	-0.026	-0.026
SO	1.4811	-0.017	-0.020	-0.049	-0.050
ClO	1.5696	0.048	0.042	0.001	-0.003
FCl	1.6283	0.006	0.000	-0.033	-0.035
H_2	0.7414	0.001	-0.001	-0.009	-0.003
HS	1.3400	-0.008	-0.010	-0.016	-0.018

CCSD in both bases while DFT is superior only at larger basis sets.

Although improvement for both basis sets are observed using the α functional, there is still a significant deterioration in RMSD at larger basis sets. This may be due to the the aforementioned double counting error which was not a focus in creation of the SSG(PBEα) method. However, it must be pointed out that the double counting error can never be completely eliminated using standard density functionals due to a lack of wavefunction information included in the electron density. On the other hand, the results show that if these SSG hybrid methods can incorporate some significant amount of orbital dependent information into their formulation we can see a dramatic improvement in results. The α functional does an adequate job at correcting a few of the worst behaving diatomics observed in the G3MP2large basis

for SSG(PBE), but cannot improve beyond the remarkable RMSD of 6-31G* and cannot compete with GGA or hybrid DFT functionals at large bases as seen in table 3.2. It is obvious that using an explicitly orbital dependent functional is useful in regaining some intermediate correlation behavior lost in using standard semi-local density functionals. A more robust, custom-built, non-local functional or operator must be considered if any further significant improvements are to be made. Such an operator is studied in the next chapter.

Chapter 4

SSG($\hat{C}$): A Dynamically Correlated Multireference Method that Completely Avoids the Double Counting Error[1]

4.1 Abstract

SSG is combined with a recently developed, linear, two-particle operator that approximates dynamic correlation in a chemical system. The use of the operator, which we call the correlation operator, or $\hat{C}$, allows the selective inclusion of correlation energy intergeminally where interactions are mean-field. The resulting hybrid method is the first combination of a multireference (MR) wavefunction with DFT-like correlation component that excludes the double counting error that manifests in all other MR-DFT combinations. In addition, the correlation operator is explicitly orbital dependent and recovers correlation effects at intermediate bond lengths much more effectively than previous incarnations of SSG(DFT). The SSG($\hat{C}$) method is tested by predicting the dissociation energies of the twenty-eight diatomic molecules in the G2/97 test set, as well as, the ten, first-row, diatomic, transition metal hydrides at the 6-31G* and G3MP2large basis sets. The results illuminate that the use of the correlation operator outside of optimization loops show promise for a balanced description of main group and transition metal systems. However, work still needs to be done if the correlation operator is to be used in optimizing the wavefunction.

[1]Cagg, B. A.; Rassolov, V. A. To be submitted to *J. Chem. Phys.*

4.2 Introduction

As described in previous chapters, there are two problems that plague SSG(DFT) as a computational method. First, DFT was used in conjunction with SSG in SSG(DFT) to account for needed intergeminal dynamic correlation inexpensively. However, some dynamic correlation is already included intrageminally as SSG accounts for total electron correlation energy for any two-electron system at the complete basis set limit. Due to this inclusion of dynamic correlation within geminals, any SSG(DFT) method using standard density functionals will double count the intrageminal portion of dynamic correlation energy. This double counting of correlation energy cannot be completely excluded without modification to standard density functionals due to the dependence on the electron density which contains no information about the specific geminal population of the molecular orbitals. Second, current standard density correlation functionals rely too heavily on local and semi-local information. Most standard density functionals rely on a three dimensional, local electron density along with the gradient of density and other semi-local quantities in order to make predictions. It is, therefore, difficult to recover dynamic correlation with appreciable accuracy at intermediate bond lengths in molecules. In fact, if we combine a standard density functional with a correction that adds some amount of semi-local orbital dependent correlation in an SSG(DFT), as was done in SSG(PBEα) geometric optimization results improve as seen in the previous chapter.

In this chapter, we continue to make progress towards an efficient dynamic correlation correction for SSG by addressing both deficiencies in SSG(DFT) at once. We do so by combining the SSG method with a recently developed DFT-like approximation to dynamic correlation energy that is explicitly orbital dependent and non-local. This approximation takes the form of a linear, two-electron operator, dubbed the correlation operator or $\hat{C}$. The correlation operator has been shown to provide a balanced description of excited states for a single determinant wavefunction[43] but

has yet to be applied to a multiconfigurational wavefunction. Unlike standard density functionals, the correlation operator utilizes an abundance of wavefunction information to add dynamic correlation to an electronic structure method. In fact, since the operator is not electron density dependent the new method can use information about geminal structure to exclusively add said correlation intergeminally where interactions are mean-field. Thus, this new SSG($\hat{C}$) method promises to provide a computationally inexpensive method to account for total electron correlation while avoiding the double counting of correlation effects. Additionally, the new method is capable of recovering non-local correlation energy due to the orbital dependence of the correlation operator. The results of dissociation energy calculations with $SSG(\hat{C})$ on a set of twenty-eight main group and ten transition metal hydride diatomics are presented at two different basis sets. The results from SSG($\hat{C}$) are then compared with SSG(PBE) and a host of other ab initio and density functional based methods.

4.3 Computational Methods

The Formulation of the Correlation Operator

The correlation operator has been formulated semi-empirically [44], in the harmonic eigenbasis [45], and based on physical arguments [46]. However, it has proved difficult to evaluate the correlation operator matrix elements over atomic centered basis functions. In order to reduce the computational complexity, the current form of the correlation operator takes on the form of a two-electron operator that approximates electron correlation based on the expectation value of the overlap of an adaptive gaussian function with the wavefunction, which produces two electron terms as seen in equation 4.1 [43].

$$\langle\Psi|\hat{C}|\Psi\rangle = \sum_{i,j}^{occupied} \langle\phi_i(\mathbf{r_1})\phi_j(\mathbf{r_2})|\hat{C}|\phi_i(\mathbf{r_1})\phi_j(\mathbf{r_2})\rangle - \langle\phi_i(\mathbf{r_1})\phi_j(\mathbf{r_2})|\hat{C}|\phi_j(\mathbf{r_1})\phi_i(\mathbf{r_2})\rangle\,. \tag{4.1}$$

In principle, in order to calculate the maximum overlap of the operator with any combination of two molecular orbitals, an infinite number of gaussian functions of varying widths would need to be incorporated into the correlation operator. Instead, a gaussian function with an adaptive exponent α_0 is used to adjust the width of the operator gaussian to the varying widths of the atomic basis functions that make up the molecular orbitals. This is done by first reformulating equation 4.1 in terms of atomic orbital basis functions and setting α_0 to be proportional to the average widths of any combination of gaussian primitives, described by equation 1.8, that make up these basis functions.

$$\langle\Psi|\hat{C}|\Psi\rangle = \sum_{\lambda,\mu,\nu,\eta} P_{\lambda,\mu}P_{\nu,\eta}\langle\chi_\lambda(\mathbf{r_1})\chi_\mu(\mathbf{r_2})|\hat{C}|(\chi_\nu(\mathbf{r_1})\chi_\eta(\mathbf{r_2}) - \delta\chi_\eta(\mathbf{r_1})\chi\nu(\mathbf{r_2}))\rangle\,. \tag{4.2}$$

In equation 4.2, the molecular orbitals are expanded in terms of a linear combination of atomic orbitals as proposed in equation 1.7 to produce an expectation value over a combination of four atomic orbitals. The equation has been simplified by the use of one-electron density matrices $P_{\lambda,\mu}$ and $P_{\nu,\eta}$.

$$P_{\lambda,\mu} = \sum_{\sigma} \sum_{i,j}^{occupied} C_{i,\lambda,\sigma}C_{j,\mu,\sigma}\,. \tag{4.3}$$

the density matrices are summed over both spin up and spin down electrons with the spin variable denoted as σ. The delta function in equation 4.2 is zero when spins of electron one and two are different and one when they are the same.

The adaptive width, α_0, can now be set proportional to the average widths of the four basis functions its comparing.

$$\alpha_0 = \frac{1}{2}\sqrt{\frac{\alpha_1\alpha_2\alpha_3\alpha_4}{\alpha_1+\alpha_2+\alpha_3+\alpha_4}(\frac{1}{\alpha_1}+\frac{1}{\alpha_2}+\frac{1}{\alpha_3}+\frac{1}{\alpha_4})}\,. \tag{4.4}$$

In equation 4.4, the exponents α_x, where x=1, 2, 3, 4 are the average gaussian exponents that make up all four gaussian basis functions as explained in the introduction chapter. This makes the matrix elements of the correlation operator simple to calculate and, in fact, the elements are analytically derived for all combinations of basis functions. In addition to the use of an adaptive gaussian operator, the correlation operator is formulated to depend on relative electron coordinates r_{12}, which is the modulus of the distance between electronic coordinates $\mathbf{r_1}$ and $\mathbf{r_2}$. The dependence of the operator on only r_{12} forces the operator to be rotationally and translationally invariant which is a condition met by the universal two-electron correlation operator. In other words, a rotation or shift in cartesian coordinates of a molecule should not change the expectation value of the operator. Thus, a correlation operator which uses an s-type operator gaussian would resemble equation 4.5.

$$\hat{C}_s = \sum_{i,j} C_{i,j}|exp(-\alpha_i r_{12}^2)\rangle\langle exp(-\alpha_j r_{12}^2)|\,. \tag{4.5}$$

The total correlation operator is, in practice, taken to be the sum of the expectation values of the s-type $\hat{C}_s$ and p-type $\hat{C}_p$ gaussian operator.

$$\hat{C} = \hat{C}_s + \hat{C}_p\,. \tag{4.6}$$

In theory, gaussians of even higher angular momentum are used, but the dependence of the correlation energy on the angular momentum quantum number l decreases very rapidly according to studies with tight hookium [45]. Therefore, contributions from the p-type operator gaussian are included as a first order correction to the dominant contribution from the s-type gaussian.

Finally, it should be noted that for both the s-type and p-type operator the is a normalization constant included with the operator. These coefficients are set to reproduce exact correlation energy for the helium and argon atoms.

Combining SSG and $\hat{C}$

It should be noted that in the previous section we developed the correlation operator formalism relative to a single determinant wavefunction. When the correlation operator acts on a single determinant wavefunction, it is designed to approximately account for all dynamic correlation energy. The main goal of this chapter is to combine this complete description of dynamic correlation with SSG just as we have in the previous chapters dealing with the SSG(DFT) methods. However, in contrast to density functional theory, the correlation operator accounts for electron correlation using expectation values over the entire wavefunction instead of the electron density. This gives the correlation operator two advantages which promise an improvement over SSG(DFT). First, unlike the density functionals, the correlation operator is completely non-local and should provide a more accurate description of correlation energy that was missing at intermediate bond lengths in SSG(PBE). Second, the correlation operator can easily distinguish between intergeminal and intrageminal orbitals. If we only add electron correlation between orbitals that are intergeminal to one another, then the double counting error encountered in the previous two chapters can be completely eliminated.

However, combining the correlation operator with SSG requires that we design the correlation operator to work with multiconfigurational wavefunctions. In single determinant wavefunctions, the correlation operator provided electron correlation to only occupied orbitals. In an SSG wavefunction, all molecular orbitals are partially occupied and multireference one-electron density matrices will have to be used.

$$P^{\alpha,A}_{\lambda,\mu} = \sum_{a\in A} D_a^2 C^a_\lambda C^a_\mu \,, \tag{4.7}$$

$$P^{\beta,A}_{\lambda,\mu} = \sum_{a\in A} D_a^2 \bar{C}^a_\lambda \bar{C}^a_\mu \,, \tag{4.8}$$

$$P^{\alpha,I}_{\lambda,\mu} = \sum_{i}^{openshell} C^i_\lambda C^i_\mu \,, \tag{4.9}$$

$$P^{0,T}_{\lambda,\mu} = \sum_{A} [P^{\alpha,A}_{\lambda,\mu} + P^{\beta,A}_{\lambda,\mu}] + P^{\alpha,I}_{\lambda,\mu} \,, \tag{4.10}$$

$$P^{\epsilon,A}_{\lambda,\mu} = \sum_{a\in A} D_a C^a_\lambda \bar{C}^a_\mu \,. \tag{4.11}$$

The first two density matrices, $P^{\alpha,A}$ and $P^{\beta,A}$, sum over all alpha and beta spin molecular orbitals in a geminal respectively. The density matrix $P^{\alpha,I}$ is the familiar one electron density matrix of single reference orbitals applied to the open shell orbitals of geminals, and the sum of the first three matrices over all geminals in a molecule results in the $P^{0,T}$ density matrix. The last density matrix is defined in the seminal SSG article [12] and is used to calculate intrageminal energies. The D^a coefficients in each density matrix are the geminal expansion coefficients and the C^a and C^i are the molecular orbital coefficients. The bars over the coefficients denote the orbital as being a β spin orbital. These density matrices are used to create an analogous correlation energy expression to 4.2 which is a multireference correlation operator energy expression.

$$E_{SSG(\hat{C})} = E_{SSG} + E_{\hat{C}} - \sum_{A} E^A_{\hat{C}} \,. \tag{4.12}$$

The energy expression for SSG($\hat{C}$) in equation 4.12 is the sum of the SSG energy and expectation value of the correlation operator. It should be noted that the correlation provided by the correlation operator will include dynamic correlation of

intrageminal electrons. These electrons are already dynamically correlated and this intrageminal correlation will need to be subtracted. Fortunately, this is easily done by excluding correlation operator contributions from orbitals included within the same geminals. Thus, the true correlation operator contribution to SSG in equation 4.12 is shown as the sum of total correlation operator energy with the sum of intrageminal correlation contributions $E^A_{\hat{C}}$ over all geminals subtracted out.

4.4 Results

In this section, the dissociation energy for the diatomic molecules in the G2/97 test set and the ten, first-row, neutral, transition metal hydride diatomics predicted by the non-iterative, SSG($\hat{C}$), and iterative, SSG$\hat{C}$, combinations of SSG and the correlation operator are presented. The predicted dissociation energies at the 6-31G* and G3MP2large levels of theory are compared with experimental values and the deviation is reported for all molecules. All calculations were performed at experimental geometries provided by [47] for main group molecules and [48] for the transition metal hydrides.

The dissociation energy calculations in this section are performed by first subtracting the molecular energy of the dissociating molecule from the atomic energy of dissociated fragments. An example calculation on the carbon monoxide molecule is shown below in equation 4.13.

$$D_0^{CO} = E^C + E^O - E^{CO}\,. \tag{4.13}$$

Here E^C and E^O are the energies of the carbon and oxygen atoms calculated with the same testing method at the same basis set. The energy of carbon monoxide E^{CO} is then the energy calculated at the experimental bond length with the same method and basis set. If the molecule is bound in the method and basis set tested, then the atomization energy D_0^{CO} will be positive and represents the non-relativistic

energy needed to dissociate the molecule. It should be noted here that the correlation operator in conjunction with SSG carries its size consistency.

The experimental dissociation energy, for comparison, is taken from equation 4.14 where D_0 is the atomization energy[47], E_{ZPE} is the zero-point energy[48, 49] , and E_{SO} is the spin-orbit coupling energy for each molecule.

$$D_e = D_0 - E_{ZPE} - E_{SO}\,. \tag{4.14}$$

For testing purposes, we compare the dissociation energy results of SSG($\hat{C}$) and SSG$\hat{C}$ with those provided from the SSG and SSG(PBE) methods to visualize the improvement of the new methods. In order to gauge the overall success of the method, we also compare to three wavefunction based methods (HF, MP2, and CCSD) as well as two density functional methods (B3LYP and PBE).

The Non-iterative SSG($\hat{C}$) Method

Initially, testing was performed with the correlation operator perturbatively as was done with SSG(DFT). However, the goal of this research is to incorporate the correlation operator in the optimization of the wavefunction where it is then able to improves the geminal structure.

First, we choose to limit our discussion to those of main group dissociation energy results. The results are reported in table 4.1 for the 6-31g* basis set and table 4.2 for the G3MP2large basis set. The numbers shown for each method are the deviation from experimental dissociation energy as calculated in equation 4.14 in kcal/mol. Here, a negative value would mean that the predicted dissociation energy is less than experimental. The numbers at the bottom of each column are the overall root-mean-squared deviation (RMSD) for each method. From both tables we can confirm that SSG(PBE) is certainly a notable improvement over SSG in both basis sets as it halves the RMSD of SSG. However, it is also obvious that these results are not comparable in

accuracy to any of the other wavefunction based or density functional based methods. It is also curious to see that the RMSD for SSG(PBE) does not noticeably change between basis sets.

On the other hand, SSG($\hat{C}$), outperforms SSG(PBE) at both bases and considerably improves at the larger G3MP2large basis. In fact, the new SSG($\hat{C}$) method outperforms the popular wavefunction based CCSD method at both basis sets. Furthermore, SSG($\hat{C}$) also outperforms the PBE at the G3MP2large basis set. However, the B3LYP density functional method and perturbation theory method in the form of MP2 seem to set the pace in accuracy for main group chemistry at larger bases.

Table 4.1 Deviation from experimental dissociation energies for SSG, SSG(PBE) and SSG($\hat{C}$) with five other methods using the 6-31g* basis set. All dissociation energies are calculated with respect to experimental ground state geometries, and all values are in kcal/mol. Root mean squared deviation is provided for each method at the bottom of the table.

molecule	$D_e Expt.$	$\Delta D_e(SSG)$	$\Delta D_e(SSG(PBE))$	$\Delta D_e(SSG(\hat{C}))$	$\Delta D_e(HF)$	$\Delta D_e(MP2)$	$\Delta D_e(CCSD)$	$\Delta D_e(B3LYP)$	$\Delta D_e(PBE)$
LiH	57.81	-12.11	3.73	-18.55	-25.77	-16.28	-12.01	-0.97	-5.8
BeH	49.64	-12.74	-6.11	-1.61	1.89	-1.49	-8.43	8.29	5.97
CH	83.89	-12.44	8.51	-0.42	-30.07	-16.33	-13.38	-0.73	-0.99
NH	83.51	-23.07	7.22	-1.39	-36.07	-18.41	-16.45	0.96	1.61
OH	106.63	-28.72	-3.77	-20.5	-43.91	-16.95	-17.62	-4.7	-2.94
FH	141.14	-39.61	-14.52	-12.2	-53.49	-17.79	-21.58	-13.08	-10.3
HCl	107.11	-29.02	-7.08	-4.42	-35.74	-18.1	-17.99	-7.3	-5.68
Li_2	24.43	-3.42	9.73	-17.33	-22.15	-10.1	-3.06	-4.21	-4.97
LiF	139.3	-40.2	-17.64	-14.77	-52.79	-8.42	-15.17	-6.91	-4.04
CN	179.22	-69.04	-19.32	-8.82	-92.4	-27.63	-23.47	-2.91	16.04
CO	259.63	-66.84	-24.21	-14.44	-88.61	-3.24	-20.82	-7.16	7.45
N_2	228.62	-78.76	-10.07	-26.39	120.55	-15.6	-34.21	-6	8.77
NO	153.17	-75.95	-23.92	-29.07	-104.72	-13.61	-27.47	-0.73	17.17
O_2	121.01	-83.57	-48.21	-38.56	-91.52	-2.73	-20.25	3.7	24.81
F_2	39.25	-36	-17.63	6.86	-74.45	-1.16	-10.08	2.86	19.47
Na_2	16.8	-1.86	9.65	15.76	-17.68	-6.38	-1.18	0.21	0.98
Si_2	75.58	-29.38	-5.1	22.21	-54.02	-14.64	-19.22	-4.89	1.99
P_2	117.26	-55.29	-9.76	17.28	-90.18	-23.57	-33.26	-8.91	-2.47
S_2	102.86	-64.67	-33.21	17.12	-62.34	-17.85	-25.9	-7.46	4.41
Cl_2	59.65	-46.94	-24.66	13.34	-49.05	-17.71	-21.87	-11.84	-2.43
NaCl	98.8	-27.85	-5.27	5.41	-29.74	-10.62	-13.39	-8.46	-6.78
SiO	192.95	-66.9	-30.12	-23.05	-91.02	-8.29	-25.72	-12.57	-2.9
SC	171.97	-68.2	-29.95	-0.72	-79.49	-10.35	-22.83	-10.26	3.57
SO	129.81	-84.65	-51.14	-21.14	-87.34	-17.2	-30.984	-13.42	1.44
ClO	65.43	-57.94	-30.41	-12	-66.18	-19.1	-22.12	-6.41	8.2
FCl	62.69	-45.23	-23.37	6.01	-58.57	-7.89	-15.18	-6.4	5.15
H_2	109.24	-11.84	7.27	-11.84	-27.48	-16.58	-11.85	0.53	-4.42
HS	88.16	-21.74	-0.2	-0.81	-31.59	-18.85	-17.04	-3.65	-3.54
RMSD		48.27	21.57	16.54	65.02	14.97	20.29	7.1	8.78

Table 4.2 Deviation from experimental dissociation energies for SSG, SSG(PBE), and SSG($\hat{C}$) and five other methods using the G3MP2large basis basis set. All dissociation energies are calculated with respect to experimental ground state geometries, and all values are in kcal/mol. Root mean squared deviation is provided for each method at the bottom of the table.

molecule	$D_e Expt.$	$\Delta D_e(SSG)$	$\Delta D_e(SSG(PBE))$	$\Delta D_e(SSG(\hat{C}))$	$\Delta D_e(HF)$	$\Delta D_e(MP2)$	$\Delta D_e(CCSD)$	$\Delta D_e(B3LYP)$	$\Delta D_e(PBE)$
LiH	57.81	-2.46	12.33	-0.82	-23.8	-7.16	-1.56	0.5	-4.51
BeH	49.64	-5.83	-0.01	5.37	0.72	2.91	-1.23	8.2	5.72
CH	83.89	-2.38	16.47	9.5	-27.15	-6.11	-3.31	1.43	0.71
NH	83.51	-15.41	14.08	-2.62	-32.45	-7.16	-5.08	4.28	4.67
OH	106.63	-21.31	0.86	2.6	-38.4	-3.03	-4.6	1.28	2.89
FH	141.14	-30.82	-8.9	-3.61	-44.33	1.1	-4.98	-2.26	0.38
HCl	107.11	-22.25	-1.88	2.19	-30.25	-3.05	-5.12	-2.19	-0.91
Li_2	24.43	-2.34	9.74	0.06	-20.7	-7.04	-0.6	-3.45	-4.03
LiF	139.3	-38.39	-17.99	-4.47	-48.72	2.17	-6.87	-2.8	-0.79
CN	179.22	-64.3	-16.94	-4.08	-99.6	-7.25	-16.05	-0.21	17.95
CO	259.63	-65.16	-26.52	0.75	-85.32	7.74	-13.24	-4.57	8.95
N_2	228.62	-70.41	-3.93	-16.44	-113.42	2.23	-18.76	0.44	14.28
NO	153.17	-72.82	-23.32	-8.58	-100.01	0.18	-16.22	1.6	18.6
O_2	121.01	-84.05	-51.61	-9	-88.09	6.13	-13.28	2.41	22.37
F_2	39.25	-42.63	-24.91	1.69	-77.16	0.53	-11.64	-2.98	12.85
Na_2	16.8	2.89	13.28	-6.86	-17.09	-3.43	0.24	0.69	1.56
Si_2	75.58	-16.52	5.9	20.35	-48.31	-5.68	-13.3	-0.8	5.65
P_2	117.26	-53.9	-12.92	3.64	-79.15	-6.43	-18.91	-1.47	4.15
S_2	102.86	-58.3	-28.68	20.37	-51.95	-1.19	-13.28	-0.19	11.56
Cl_2	59.65	-40.96	-18.65	20.91	-41.59	-1.07	-10.23	-5.12	4.98
NaCl	98.8	-24.83	-4.71	6.39	-28.26	0.35	-4.98	-5.79	-4.11
SiO	192.95	-56.73	-23.46	0.97	-81.89	6.34	-14.2	-4.96	3.27
SC	171.97	-62.96	-27.43	4.18	-74.74	2.26	-15.37	-5.86	7.28
SO	129.81	-75.48	-44.89	6.86	-74.74	-0.45	-16.87	-3.68	10.46
ClO	65.43	-53.58	-27.92	8.03	-60.39	-6.44	-12.22	0.09	15.08
FCl	62.69	-45.96	-25.55	5.32	-55.64	0.48	-10.46	-3.43	7.85
H_2	109.24	-1.79	15.55	1.79	-25.55	-6.86	-1.79	0.89	-4.79
HS	88.16	-13.63	5.78	4.69	-27.2	-6.46	-5.33	-0.04	-0.32
RMSD		45.56	21.17	8.80	61.11	4.79	10.98	3.3	9.32

From the results, there are two groups of molecules that can be highlighted as problematic. First, from the 6-31G* data it is obvious that the SSG($\hat{C}$) are the worst for oxygen containing molecules with the worst being that of O_2. The second group, with contributions to the error in both bases, comes with respect to homonuclear diatomic molecules. We will hold off discussing these problems and a possible solution until the next section.

In switching our focus from main group diatomics to transition metal hydrides, the are some aspects of the data analysis that must be noted. First, we are interested in testing the SSG methods on transition metal systems because these systems need a multiconfigurational description[50]. The other methods we use for comparison are not multireference and shouldn't fare as well on such systems. Second, the results for 6-31G* are reported here mostly for completeness. It is known that Hartree-Fock results are inaccurate for transition metal atoms using 6-31G* due to the diffuseness of the atomic basis functions [51]. These basis set inaccuracies manifest in SSG, as they do in Hartree-Fock calculations, as incorrect ground state energies for the first row transition metals. Therefore, the results are shown here to give some qualitative insight on the error in the respective methods. Third, the experimental data on transition metal systems varies within the literature. We take experimental values from another computational study on these transition metal hydride systems as the experimental values we use [48].

From the RMSD of 6-31G* data reported in table 4.3, it can be assumed that, like main group molecules, SSG($\hat{C}$) is an improvement upon both SSG and SSG(PBE) and may even be competitive with the most accurate methods compared in this study. However, it must be noted that the basis set inaccuracies certainly muddle the conclusions at this basis as SSG error in dissociation energy prediction should be lower than the Hartree-Fock method.

More representative conclusions can be drawn from the G3MP2large basis set re-

Table 4.3 Root mean squared deviation from experimental dissociation energies for the first row diatomic transition metal hydrides using SSG, SSG(PBE) and SSG($\hat{C}$) at the 6-31G* basis set. Results for wavefunction and density functional based methods are included for comparison. All RMSD are calculated with respect to experimental ground state geometries, and all values are in kcal/mol.

method	RMSD_{6-31G*}
SSG	24.31
SSG(PBE)	22.91
SSG($\hat{C}$)	11.31
B3LYP	10.6
PBE	11.45
HF	20.11
MP2	22.06
CCSD	18.64

sults that do not suffer from these same inaccuracies. The results for the G3MP2large basis are shown in table 4.4. Here the RMSD for SSG(PBE) shows a substantial improvement more in line with the improvement made by SSG($\hat{C}$). We believe this is somewhat due to the fair accuracy of the parent SSG method with respect to such calculations. Nonetheless, both hybrid methods are also comparable to the PBE functional and perturbation theory. SSG($\hat{C}$) is slightly more accurate and presents a slight improvement over SSG(PBE). In fact, SSG($\hat{C}$) is just two kcal/mol worse on average than B3LYP for these simple transition metal systems. It is also important to note that the accuracy of SSG($\hat{C}$) is relatively unchanged in the G3MP2large basis between calculations on transition metal and main group diatomics.

Table 4.4 Deviation from experimental dissociation energies for the first row diatomic transition metal hydrides using SSG, SSG(PBE) and SSG($\hat{C}$) at the G3MP2large basis set. Results for five other methods are included for comparison. All dissociation energies are calculated with respect to experimental ground state geometries, and all values are in kcal/mol. Root mean squared deviation is provided for each method at the bottom of the table.

molecule	$D_e Expt.$	$\Delta D_e(SSG)$	$\Delta D_e(SSG(PBE))$	$\Delta D_e(SSG(\hat{C}))$	$\Delta D_e(HF)$	$\Delta D_e(MP2)$	$\Delta D_e(CCSD)$	$\Delta D_e(B3LYP)$	$\Delta D_e(PBE)$
ScH	50.39	-7.3	0.43	0.44	-14.03	-9.05	-7.03	3.94	4.37
TiH	50.55	-5.66	5.29	11.77	-9.35	-3.99	0.26	8.33	14.16
VH	51.88	-5.85	5.42	14.93	-9.25	-1.59	4.47	11.22	21.98
CrH	46.94	-3.7	7.42	1.87	-22.38	-17.77	2.02	8.27	5.7
MnH	41.14	-2.66	6.49	13.69	-6.15	-4.52	-1.87	-1.78	4.09
FeH	39.63	-13.03	-1.4	10.07	-15.5	-10.66	1.1	15.85	16.98
CoH	48.91	-24.97	20.48	-1.99	-27	-18.98	-1.32	9.99	13.02
NiH	62.72	-39.72	-26.4	-16.84	-40.4	-29.36	-0.69	3.3	4.57
CuH	63.17	-16.5	0.23	-2.02	-31.34	-2.15	-3.3	0.06	1.8
ZnH	22.02	-0.79	10.01	17.06	-2.06	-1.59	1.35	1.35	1.07
RMSD		16.68	11.72	11.12	21.22	13.14	3.05	8.04	11.11

The Iterative USSG$\hat{C}$ and RUSSG$\hat{C}$ Methods

Calculations with iterative use of the correlation operator to optimize geminal structure were also performed and are shown in table 4.5 and 4.6. In the preceding subsection it was mentioned that there are two classes of molecules that need an improved description. First, the class of homonuclear diatomic molecules in the G2/97 test set seems to cause problems in both basis sets. Upon further examination, the core set of geminals in these molecules were on average too delocalized. When a molecule is stretched, the valence geminals will change drastically, but the core geminals will be remain unchanged and should resemble atomic geminals even at equilibrium geometry. This is not the case for these delocalized core geminals. The problem is exacerbated by the sensitivity of the correlation operator. Since the electrons in core geminals at equilibrium geometry are spread out over the entire molecule, the correlation operator will overcorrelate these core geminals and thus predict a larger dissociation energy than experimental. This is corrected in our iterative calculations by forcing the core geminals to be Hartree-Fock-like single determinant geminals and providing them intrageminal correlation via the correlation operator. This reduces the amount of correlation these core geminals can add to the molecular dissociation energies. This remedy is shown to improve these results in table 4.5.

Second, the class of oxygen containing molecules presents a problem. The iterative correlation operator method is used with respect to two spin restrictions placed upon the geminal wavefunctions called USSG and RUSSG . The first is the same unrestricted formulation discussed in chapter one. We have used this formulation in SSG(DFT) which could easily be called USSG(DFT) in all research discussed so far. USSG requires the use of α and β molecular orbitals that can be different spatially. The RUSSG formalism uses a partially restricted formulation that is only applicable to pair based wavefunctions like SSG. In RUSSG, intergeminal α and β molecular orbitals are restricted to be spatially orthogonal. The advantage of using a par-

tially restricted formulation of SSG can be surmised from comparing the USSG$\hat{C}$ and RUSSG$\hat{C}$ at 6-31G* in table 4.5. The nearly 5.5 kcal/mol improvement in RMSD stems from the improvement in oxygen containing molecules. This is due to the ability of RUSSG to optimize molecules and atoms to chemically accurate states. USSG, while accurate in most cases, can sometimes allow incorrect states by introducing spin contamination into the wavefunction. This spin contamination allows the lowering of energy by the interaction of unrestricted intergeminal orbitals. The partial spin restriction reduces the amount of spin contamination which allows the chemically correct states to be predicted more frequently. Despite this improvement, there is not an similarly impressive improvement for G3MP2large due to some of these same oxygen containing molecules.

In transition metal hydrides, the iterative correlation problems seems to have great difficulty with the later transition metal hydrides as shown in table 4.6. The NiH and CoH molecules seem to be the most difficult molecules to describe. The problem is not corrected by the use of either USSG$\hat{C}$ or RUSSG$\hat{C}$ and warrants further study.

Table 4.5 Deviation from experimental dissociation energies using USSG($\hat{C}$), USSG$\hat{C}$, and RUSSG$\hat{C}$ for the G2/97 test set diatomic molecules at 6-31G* (left) and G3MP2large (right) basis sets. All dissociation energies are calculated with respect to experimental ground state geometries, and all values are in kcal/mol. Root mean squared deviation is provided for each method at the bottom of the table.

molecule	$D_e Expt.$	$\Delta D_e(USSG(\hat{C}))$	$\Delta D_e(USSG\hat{C})$	$\Delta D_e(RUSSG\hat{C})$	$\Delta D_e(USSG(\hat{C}))$	$\Delta D_e(USSG\hat{C})$	$\Delta D_e(RUSSG\hat{C})$
LiH	57.81	-18.55	-11.24	-11.24	-0.82	-1.02	-1.01
BeH	49.64	-1.61	-4.12	-6.23	5.37	2.95	0.18
CH	83.89	-0.42	-4.32	-5.29	9.5	5.77	4.57
NH	83.51	-1.39	-0.30	-6.11	-2.62	-11.49	3.30
OH	106.63	-20.5	-16.33	-7.64	2.6	-1.98	0.57
FH	141.14	-12.2	-19.02	-15.60	-3.61	-10.07	-5.09
HCl	107.11	-4.42	-12.25	-6.83	2.19	-4.55	-8.68
Li_2	24.43	-17.33	-2.93	-2.93	0.06	-0.63	-0.13
LiF	139.3	-14.77	-14.60	-11.18	-4.47	-11.85	-6.87
CN	179.22	-8.82	-5.75	-11.29	-4.08	0.33	-4.02
CO	259.63	-14.44	-23.58	-10.71	0.75	-17.07	-6.51
N_2	228.62	-26.39	-9.10	-7.05	-16.44	3.64	7.43
NO	153.17	-29.07	-19.78	-12.38	-8.58	-7.35	-6.26
O_2	121.01	-38.56	-35.93	-23.55	-9.00	-25.17	-30.71
F_2	39.25	6.86	-8.69	-1.85	1.69	-14.47	-4.50
Na_2	16.8	15.76	0.04	0.05	-6.86	1.59	1.59
Si_2	75.58	22.21	-6.82	-8.04	20.35	2.54	10.57
P_2	117.26	17.28	2.99	6.29	3.64	-12.42	-13.07
S_2	102.86	17.12	-0.25	0.47	20.37	3.88	-8.00
Cl_2	59.65	13.34	-9.70	1.14	20.91	-0.85	-9.10
NaCl	98.8	5.41	0.28	5.70	6.39	1.13	-3.00
SiO	192.95	-23.05	-34.66	-18.69	0.97	-19.57	-4.40
SC	171.97	-0.72	-18.75	-11.00	4.18	-11.67	-7.23
SO	129.81	-21.14	-30.14	-21.71	6.86	-10.55	-22.77
ClO	65.43	-12	-23.58	-11.36	8.03	-11.61	-16.08
FCl	62.69	6.01	-14.29	-5.45	5.32	-14.50	-13.64
H_2	109.24	-11.84	-11.84	-11.84	-1.79	-1.83	-1.79
HS	88.16	-0.81	-6.55	-3.73	4.69	-0.05	-2.18
RMSD		16.54	15.98	10.54	8.80	10.01	9.96

Table 4.6 Deviation from experimental dissociation energies using USSG($\hat{C}$), USSG$\hat{C}$, and RUSSG$\hat{C}$ for the test set of transition metal hydrides at 6-31G* (left) and G3MP2large (right) basis sets. All dissociation energies are calculated with respect to experimental ground state geometries and all values are in kcal/mol. Root mean squared deviation is provided for each method at the bottom of the table.

molecule	$D_e Expt.$	$\Delta D_e(USSG(\hat{C}))$	$\Delta D_e(USSG\hat{C})$	$\Delta D_e(RUSSG\hat{C})$	$\Delta D_e(USSG(\hat{C}))$	$\Delta D_e(USSG\hat{C})$	$\Delta D_e(RUSSG\hat{C})$
ScH	50.39	-2.45	-8.29	0.08	0.44	3.20	4.32
TiH	50.55	0.33	7.89	-5.67	11.77	2.92	2.74
VH	51.88	-1.86	-9.43	-9.95	14.93	2.89	3.04
CrH	46.94	12.06	11.31	7.53	1.87	3.57	4.11
MnH	41.14	6.21	-9.88	-13.67	13.69	2.26	-1.49
FeH	39.63	-4.12	-15.21	-14.95	10.07	-1.45	-5.37
CoH	48.91	-23.29	-20.62	-27.55	-1.99	-21.45	-20.36
NiH	62.72	-21.71	-33.52	-42.23	-16.84	-34.65	-30.68
CuH	63.17	-0.51	2.60	2.72	-2.02	-33.17	-15.91
ZnH	22.02	7.32	1.70	-5.30	17.06	3.06	0.28
RMSD		11.31	14.95	17.82	11.12	16.78	13.01

4.5 Conclusions

The results reported for the new hybrid geminal method SSG($\hat{C}$) and its iterative analogs USSG$\hat{C}$ and RUSSG$\hat{C}$ warrants further study for two reasons. First, they are the first dynamic correlation corrected multireference methods to completely avoid the double counting error and account for longer range dynamic correlation effects while keeping a Hartree-Fock-like expense. The improvement of the correlation operator can be visualized from the consistently more accurate results provided by SSG($\hat{C}$) over the flawed SSG(PBE) method. Second, the consistent accuracy between main group molecules and transition metal molecules is a welcome development in response to the known inconsistencies with respect to density functional theory. In fact, in transition metal systems the accuracy of the SSG($\hat{C}$) method rivals the accuracy of the immensely popular B3LYP DFT method. A further improvement could be realized with the use of the iterative combinations USSG$\hat{C}$ and RUSSG$\hat{C}$ as they correct two of the main problems for the perturbative method at small basis sets. However, improvement is still needed in describing open shell molecules and atoms.

Chapter 5

SSpG: A Strongly Orthogonal Geminal Method with Relaxed Strong Orthogonality[1]

5.1 Abstract

In this final chapter, the second correlation deficiency discussed in chapter 1 related to the strong orthogonality condition placed on SSG wavefunctions is studied. Strong orthogonality is an important constraint placed on geminal wavefunctions in order to make variational minimization of the wavefunction tractable. However, strong orthogonality prevents certain, possibly important, excited configurations from contributing to the ground state description of chemical systems. The method presented in this chapter lifts the strong orthogonality constraint from a geminal wavefunction by computing a perturbative-like correction to each geminal independently from the corrections to all other geminals. Comparisons of this new SS_pG method are made to the non-orthogonal AP1roG and the unconstrained GMFCI method using small atomic and molecular systems. The correction is also compared to DMRG calculations performed on long polyene chains in order to assess its scalability and applicability to large strongly correlated systems. The results of these comparisons demonstrate that although the perturbative correction is small, it may be a necessary first step in the systematic improvement of any strongly orthogonal geminal method. Some of the results reproduced in this chapter can be found in [52] and permission

[1]Cagg, B. A.; Rassolov, V. A. *J. Chem. Phys.* 2014, *141* , 164112. Reprinted here with permission from the publisher.

to reproduce them is granted by the publisher.

5.2 Introduction and Motivation

Geminal formalism

Geminal theory is a conceptually simple and intuitive model that can describe many chemical phenomena which are difficult, if not impossible, to represent using single reference methods such as DFT. In its most general form, the geminal wavefunction is written as an antisymmetrized product of two-electron functions, or geminals. This antisymmetrized product wavefunction may also include any number of single electron orbitals in order to describe open-shell species. In addition, individual geminals may be explicitly correlated by inclusion of a functional dependence on relative electron coordinates, $\mathbf{r}_k - \mathbf{r}_m$, [53] or have an implicit correlation through the expansion in products of one-electron orbitals [54]. The geminal energy is then often computed by the variational minimization of the expectation value of a Hamiltonian applied to a geminal wavefunction with respect to the parameters of the geminals.

Unfortunately, the variational minimization of the most general geminal wavefunction, the APG wavefunction shown in 1.12 is a computationally intractable problem, as it is likely to have an exponential dependence on the system size. However, as was mentioned in the first chapter, in small model systems where an exact (in a given basis) Full Configuration Interaction (FCI) wavefunction can be feasibly determined, the geminal wavefunction often yields energies and dipole moments that are tantalizingly close to FCI results [3]. Therefore, utilizing approximations to simplify the variational search for the best possible geminal wavefunction is desirable.

The two most widely used geminal methods, which are comparable in computational expense to the simplest forms of the Multiconfigurational Self-Consistent Field (MCSCF) methods, are based on different approximations that constrain the con-

struction of geminals. The Antisymmetrized Geminal Power (AGP) assumes that all geminals are functionally identical. The Antisymmetrized Product of Strongly Orthogonal Geminals (APSG), on the other hand, assumes all geminals to be as different as possible by imposing strong orthogonality between geminals as in equation 1.13. It has been shown by Arai [55] that enforcing strong orthogonality is equivalent to requiring different geminals to be expanded in mutually orthogonal one-electron orbital sets.

Both of the aforementioned geminal methods have their particular strengths and weaknesses. The AGP method is particularly useful for the description of delocalized electron pairs, as in the BCS model of superconductivity [56]. However, it is not size-consistent, meaning that AGP energy of two non-interacting subsystems is not equal to the sum of their individual AGP energies. The APSG wavefunction, on the other hand, is rigorously size-consistent provided that only the opposite-spin orbitals are coupled into geminals and that the orbitals themselves are spin-unrestricted and fully optimized [12]. Moreover, the APSG wavefunction optimization can be performed in the Atomic Orbital (AO) basis, making the APSG wavefunction optimization problem to scale with the number of basis functions N better than $O(N^5)$.

Intra-Geminal Correlation and Strong Orthogonality

It is convenient to subdivide all electron correlation into an intra-geminal part and the rest, which we define as inter-geminal correlation. The SSG model omits all inter-geminal correlation, and describes all intra-geminal correlation exactly, subject to the following discussion.

Consider the case of helium gas in a strong external field that may perturb individual atoms. In the case of non-overlapping wavefunctions of individual atoms, all electron correlation effects in this system separate into intra-atomic correlations and van der Waals-type inter-atomic correlation. All other electron correlation effects

vanish due to zero overlap between atomic wavefunctions. The SSG model describes each helium atom in such a system exactly, and omits all inter-atomic interactions beyond the mean-field. For this system, the intra-geminal and intra-atomic correlations are the same.

Now let us consider an atom of beryllium, which is one of the simplest chemical systems with strong static correlation. The mean-field, Hartee-Fock, description of the ground state is a $1s^2 2s^2$ configuration. The dominant correlation effects arise due to strong coupling of $1s^2 2p^2$ configurations to the ground state. The APSG wavefunction describes this correlation via a geminal $\psi_{2s} = D_0^{2s}\phi_{2s}\bar{\phi}_{2s} + D_1^{2s}\phi_{2px}\bar{\phi}_{2px} + \ldots$, where D are variationally optimized geminal expansion coefficients. With a very large basis set, this geminal will also include many other terms corresponding to excitation of $2s$ electrons to high energy orbitals. Such excitations collectively describe dynamic correlation of the $2s$ electrons. At the same time, the other geminal describes the $1s$ electrons, and in a sufficiently large basis it contains some highly excited orbitals a, which describe core correlation: $\psi_{1s} = D_0^{1s}\phi_{1s}\bar{\phi}_{1s} + D_1^{1s}\phi_a\bar{\phi}_a + \ldots$. The APSG wavefunction for beryllium is $\Psi = \hat{A}[\psi_{1s}\psi_{2s}]$. Compared to the exact wavefunction, the configurations corresponding to excitations of $1s$ electrons into $2p$ orbitals are missing from it. Considering each geminal as a correlated two-electron object, the $1s \rightarrow 2p$ excitations are intergeminal and, therefore, are not part of the exact description of the intra-geminal correlation in ψ_{1s}. However, from the chemical point of view, the $1s$ electron pair correlation should include all important excitations, including those into $2p$ orbitals. Therefore, the APSG description of the $1s$ electrons in beryllium is not exact.

The goal of the method presented here is to account for such excitations, as was mentioned in the previous example, in a computationally efficient manner. Such nominally inter-geminal excitations violate strong orthogonality, thus we test our method in comparison to two other geminal methods that provide corrections to the strong

orthogonality approximation. The first is the Geminal Mean-Field Configuration Interaction (GMFCI) method [57] developed by Cassam-Chenaï. GMFCI is based on the graded orthogonality concept[58], with a variationally optimized wavefunction written as a product of non-strongly orthogonal geminals. The other is the AP1roG method developed in the Ayers group [2]. It relies on identifying one orbital pair in each geminal as a reference pair, and allows for the coupled-cluster-like excitations of this pair to all other *non-reference* orbital pairs in the system. We also provide a comparison to the calculations of all-trans-polyenes based on the Density Matrix Renormalization Group (DMRG) as developed in the Chan group [4]. The method is applicable to any APSG wavefunction, including RUSSG. We investigate its performance in application to RSSG and USSG.

5.3 SS$_p$G Formulation

Consider a converged SSG calculation for a system of n_α electrons, that are spin up, and n_β electrons, that are spin down (we assume $n_\alpha \geq n_\beta$), and uses N one-electron basis functions. This system contains n_β geminals and $n_\alpha - n_\beta$ open-shell uncorrelated orbitals. In the SSG model, each of the converged molecular orbitals ϕ_i, $i = 1 \ldots N$ ends up in one and only one geminal, with the exception of the $n_\alpha - n_\beta$ spin down orbitals $\bar{\phi}_k$ that are counterparts to the occupied open-shell ϕ_k spin up orbitals, which are left unoccupied. Each geminal, here labeled with A, can then be expanded as

$$\psi_A = \sum_{i \in A} D_i |\phi_i \bar{\phi}_i|, \;\; \sum_{i \in A} D_i^2 = 1 \,. \tag{5.1}$$

In order to simplify the notation and not to consider the open-shell and correlated geminal orbitals as separate cases with separate energy expressions, we note that a product of two uncorrelated orbitals $\phi_k \bar{\phi}_k$ can be viewed as a geminal with $D_k = 1$. Furthermore, setting to zero all energy matrix elements that involve $\bar{\phi}_k$ makes the energetics of this orbital product identical to the energetics of the single orbital ϕ_k.

Therefore, to simplify the coding and the formula derivation, each open-shell orbital ϕ_k is considered to be part of a special uncorrelated geminal with single expansion coefficient $D_k = 1$ and with all energy matrix elements involving $\bar{\phi}_k$ being zero. This converts the SSG wavefunction from the product of n_β geminals and $n_\alpha - n_\beta$ open-shell orbitals, to the product of n_α geminals (with "special" uncorrelated $n_\alpha - n_\beta$ single-product geminals, each containing β-spin orbitals not contributing to the energy). Using geminal wavefunction Eq. 1.12 with each geminal given by Eq. 5.1, we get for the SSG energy

$$\begin{aligned} E_{SSG} &= \langle \Psi_{SSG} | \hat{H} | \Psi_{SSG} \rangle = \sum_n D_n^2 h_{nn} + \sum_{A>B} \sum_{a \in A, b \in B} D_a^2 D_b^2 < ab||ab > \\ &+ \sum_A \sum_{m,n \in A} D_m D_n < m\bar{m}||n\bar{n} > , \end{aligned} \tag{5.2}$$

where the first sum runs over all spin orbitals, the second sum runs over all pairs of geminals (including the unpaired orbitals), h_{kk} is the expectation value of the one-electron part of the hamiltonian over spin orbitals ϕ_k, and the last term has a summation over the spatial orbitals m and n, with spins given explicitly in the term by the bars labeling beta spin orbitals. Here $< ij||kl >= \int \phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_2)\frac{1}{r_{12}}\phi_k(\mathbf{r}_1)\phi_l(\mathbf{r}_2) - \int \phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_2)\frac{1}{r_{12}}\phi_l(\mathbf{r}_1)\phi_k(\mathbf{r}_2)$, with integration running over both the spatial and the spin coordinates.

Now let us have one geminal, Z, break the strong orthogonality condition by allowing its expansion in all orbitals of the system, while keeping all other geminals frozen

$$\begin{aligned} \tilde{\psi}_Z &= \sum_{i,j=1}^{N} \tilde{D}_{i,j} |\phi_i \bar{\phi}_j| \\ \tilde{\Psi}^Z &= \hat{A}[\psi_A \ldots \tilde{\psi}_Z \ldots] \\ \tilde{E}^Z &= \langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle \\ \Delta E^Z &= \tilde{E}^Z - E_{SSG} \end{aligned} \tag{5.3}$$

For normalized $\tilde{\Psi}^Z$ the energy expression with respect to the coefficients $\tilde{D}_{i,j}$ is bi-linear, so the variational energy minimization is straightforward. The challenge is

in computation of the Hamiltonian matrix elements for this minimization. The excitation of geminal Z into orbital spaces of other geminals breaks the mean-field approximation, and individual terms in the interacting geminals must be considered explicitly. Details of these computations are given in the Appendix.

The essence of the proposed method is to consider breaking of strong orthogonality by all geminals independently from each other, in the spirit of the Independent Electron Pair Approximation (IEPA)[59, 60]. The total energy is defined as

$$E_{SS_pG} = E_{SSG} + \sum_{Z=1}^{n_\alpha} \Delta E^Z$$

The density matrix of the system is also defined as the SSG density matrix plus the sum of all geminal corrections, with each correction defined as the difference between the reference density and the density of the wavefunction from Eq. 5.3. In the present work, we limit our investigation of the new method to energies and energy differences.

Note that $\tilde{\Psi}^Z$ is normalized to unity. It is also possible to define it in intermediate normalization of $\langle\tilde{\Psi}^Z|\Psi_{SSG}\rangle = 1$. We investigate both normalizations, and for the test systems studied here we see no qualitative difference in results. Both forms of $\tilde{\Psi}^Z$ lead to fully size-consistent energy. This is in contrast to widely used *ab initio* formulations such as coupled cluster theory or Møller-Plesset perturbation theory, where intermediate normalization is required. Size consistency of the method follows from localization of geminals on the non-interactive fragments, leading to a size-consistent description of strong orthogonality correction in each geminal.

In the new method the correction to each geminal is determined variationally, but the corrections of different geminals are not coupled to each other. Also, the excitations from a corrected geminal Z to different geminals are not coupled to each other, as explained in the Appendix. In addition, during the optimization of the geminal Z all other geminals remain unmodified, apart from elimination of doubly occupied single orbitals in each configuration due to antisymmetrization. Therefore, the proposed method is fully variational with respect to the correction of a single geminal, but is

perturbative with respect to interaction between the geminals. Therefore, we label the new method SS_pG for (antisymmetrized product) of the Singlet-type Strongly orthogonal-*p*erturbatively corrected Geminals.

We emphasize the following differences with IEPA: (i) The SS_pG method uses only n_α independent excitations, in contrast to $(n_\alpha + n_\beta)((n_\alpha + n_\beta - 1)/2$ excitations in IEPA. Thus, the SS_pG method accounts for a relatively small fraction of the total correlation energy that corresponds to a correlation within each localized electron pair. (ii) The excitations in the SS_pG arise from *correlated* and *localized* electron pairs, making the approximation of their independence less severe than in IEPA.

The SS_pG method is designed to correct for a specific deficiency in the APSG wavefunction related to strong orthogonality, in a computationally efficient manner with better than or close to $O(N^5)$ effective scaling with the number of basis functions. To account for all electron correlation missing in APSG, a number of approaches have been developed over the years, see recent discussion [61]. Here we want to emphasize two approaches. Our group has developed a perturbative correction that includes *all* double excitations from the APSG reference state [13]. The number of configurations that have to be considered leads to approximately $O(N^6)$ scaling, with significant scratch space requirements to store contracted integrals. It is possible to reduce the number of excitations further by using an effective one-body Hamiltonian in the formulation of perturbative equations, as was done by Rosta and Surján [32]. The have also included all intra-geminal two-body terms into the reference Hamiltonian, leading to much more accurate results at the expense of the demanding scaling requirements with the system size, if geminal spaces are large.

5.4 Results

The effectiveness of the SS_pG correction is evaluated by comparison to two geminal methods, AP1roG and Geminal Mean Field Full Configuration Interaction (GMFCI),

in which strong orthogonality is not imposed, for several small atomic and molecular systems. GMFCI is a product of fully unconstrained geminals that has an exponential scaling with the system size. AP1roG [2], on the other hand, is a method with polynomial scaling. It is based on the subdivision of the orbital pairs within each geminal into a single reference pair, constrained to be occupied in a single geminal, and all other orbital pairs, which can have non-zero occupations in all geminals, thus lifting a strong orthogonality constraint. By comparing SS_pG with AP1roG and GMFCI we can examine the significance of the strong orthogonality constraint and compare two different approaches of approximate relaxation of this constraint. All calculations are performed using a modified version of Q-Chem [62] and all systems are studied using the same basis sets and geometries used in the comparative literature.

Comparison to AP1roG

SS_pG calculations are performed on three H_{10} chains with different intermolecular distances. The results are shown in Table 5.1. Since all methods used in this section are, in principle, bound by the RSSG energy from above, we compare the difference between the energy of a given method, and RSSG. AP1roG and SSG are nearly equivalent for these simple systems with the biggest energy difference of only 0.75 mh. Since the minimal STO-6G basis has very few virtual orbital pairs in a single-determinant reference ground state, and since the difference between SSG and AP1roG is mainly in the treatment of these reference virtual orbitals, the discrepancy in this basis is small. The negligible difference is further highlighted by the small difference between the SSG and the Full CI energy.

The strong orthogonality correction yields considerable improvement in the energy for these hydrogen chains. At an intermolecular distance of 2.5 bohr, over nine mh is regained which is over a third of the energy error in SSG. SS_pG becomes even more accurate as intermolecular distance increases. Conversely, AP1roG recovers a

Table 5.1 Total energy of alternating H_{10} chain, relative to SSG, given in millihartrees. The distances between H_2 units are given in bohrs, and the distances within H_2 are constrained at 2 bohrs. Calculations are performed at the STO-6G basis set, and the methods are discussed in the text.

R_{inter}	FCI	AP1roG	SS_pG
2.5	-25.6522	-0.7591	-9.2977
3	-10.9849	-0.1833	-4.7898
4	-1.9876	0.1850	-1.0841

smaller and smaller fraction of FCI energy and becomes nearly equivalent to SSG at 4 bohr intermolecular distance. The slight energy error in AP1roG is likely due to incomplete orbital optimization.

In Table 5.2, we compare the total energy calculated for four other chemical systems with AP1roG, RSSG, and RSS_pG. Neon was studied with three basis sets to investigate the basis set dependence of the proposed method. There is much more interaction between orbital pairs in these systems than in hydrogen chains with the minimal basis. Therefore, we expect a greater differentiation between the methods. Indeed, both geminal methods are significantly worse than FCI, and provide only modest improvement over SSG. The SS_pG is closer to FCI than AP1roG, and its relative performance improves as the basis size is increased.

Table 5.2 Total energy, relative to SSG, given in millihartrees. He-He distance is at 4 bohrs, as in Ref. [2]. The methane geometry is optimized at the Hartree-Fock level.

System	Basis	FCI	AP1roG	SS_pG
Ne	6-31G	-73.8302	-1.2142	-1.1384
Ne	6-311G*	-161.4802	-13.4660	-17.0642
Ne	cc-PVTZ	-206.5669	-13.1737	-20.0516
CH_4	6-311G*	-94.521	-4.4535	-13.4397
He_2	6-31G*	-0.3227	-0.0004	-0.2215
Be	6-31G	-0.5663	-0.05188	-0.2672

Finally in Figures 5.1 and 5.2 we show the potential energy surfaces for the symmetric stretch of both O-H bonds in H_2O as well as the stretching of the triple bond in N_2. Due to the spin restriction used in the previously published graph of AP1roG and its comparable methods[5], our SS_pG method can only be compared to these other methods near equilibrium, where there is no spin polarization in USSG. It can be observed in both surfaces that near equilibrium SS_pG is consistently lower in energy than AP1roG at all points along the potential energy surface. However, the difference between SS_pG and SSG near equilibrium is only $\tilde{1}0$ mh. Interestingly, the SS_pG surfaces are smooth, despite different sizes of orbital subspaces in geminals. This is likely due to the (i) the SSG potential energy surface being continuous by construction [12], (ii) the re-assignment of orbital pairs between the geminals in practice only involve very weakly occupied orbitals, and (iii) the SS_pG wavefunction being less dependent on the definition of the geminal orbital subspaces than the SSG wavefunction, due to relaxed strong orthogonality in the former.

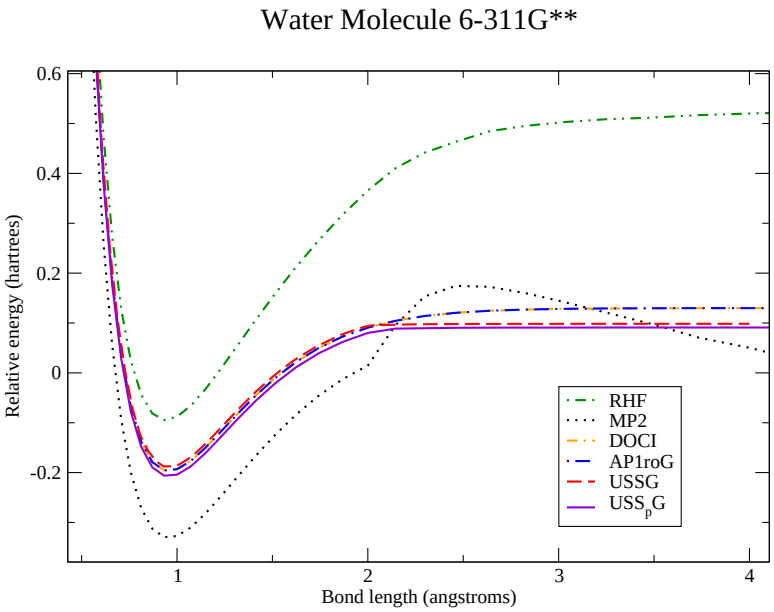


Figure 5.1 Potential energy surface of the symmetric stretching of water, computed with 6-311G** basis. The energy is reported relative to FCI atomic fragments and is measured in hartrees. The molecular angle is set to 104.6 degrees throughout and the bond distances are reported in Å. The lines are drawn through the computed points without interpolation, so sufficiently dense grid is used for smooth appearance.

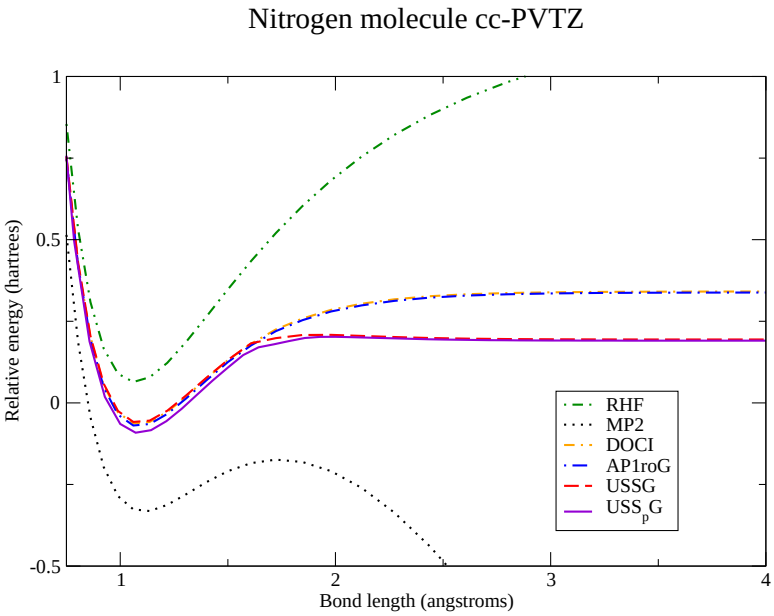


Figure 5.2 Potential energy surface of N_2, computed with cc-pVTZ basis. The energy is reported in hartrees relative to the atomic fragments calculated by NEVPT2 and taken from [5].The lines are drawn through the computed points without interpolation, so sufficiently dense grid is used for smooth appearance.

Comparison to the Completely Non-orthogonal GMFCI Method

In table 5.3, equidistant hydrogen chains are used as a benchmark to compare RSS$_p$G with the single-step GMFCI method. The GMFCI is based on fully unconstrained geminals. With the single-step version, strong orthogonality is relaxed for only one geminal. In the smallest 6 atom chain considered, SS$_p$G recovers slightly less correlation energy than the single-step GMFCI method (both methods recover about 25% of the energy missing in SSG). In larger chains the SS$_p$G energy is below single-step GMFCI. The unpublished GMFCI data on this systems indicates that optimization of more than one geminal in GMFCI drives its energy below that of the SS$_p$G, as expected.

Table 5.3 Total energy, relative to SSG, given in millihartrees. All calculations use the STO-3G basis and the bond-lengths are taken from Ref. [3]: R(Li-Li)=2.673, R(Li-H)=1.5957, R(Be-H)=1.340, R(B-H)=1.2324, R(Be-Be)=2.460, R(H-H)=1.000, in Å. The GMFCI calculations on the hydrogen chains relax strong orthogonality in a single geminal only, as described in Ref. [3].

System	ΔE_{FCI}	ΔE_{GMFCI}	ΔE_{SS_pG}
H_{10}	-61.3694	-7.0264	-13.8799
H_8	-45.5595	-7.0575	-10.3601
H_6	-30.0830	-8.1250	-7.8346
LiH	-0.1894	-0.1694	-0.0852
Be	-0.0254	-0.0254	-0.0238
Li_2	-0.7560	-0.5300	-0.1463
BeH_2	-6.2310	-6.0850	-1.8852
BH	-2.0362	-2.0292	-0.7927
Be_2	-22.8749	-21.7419	-3.4660

The full GMFCI and RSS$_p$G energies become substantially different in the systems where different electron pairs have larger differential overlap, allowing for stronger interaction between them. This is observed in the six molecules at the bottom of Table 5.3. In some systems, such as the beryllium atom, both methods give nearly exact energy. In the beryllium dimer, on the other hand, GMFCI is relatively close to being exact, while SS$_p$G recovers only about 15 % of the energy missing in SSG. Such systems as the beryllium dimer are unlikely to be described with high accuracy by a computationally cheap general purpose method such as SS$_p$G.

Table 5.4 Energies for all-*trans* polyene chains at fixed geometry. All calculations use STO-3G basis, with geometries from [4]. Except for LDMRG, all calculations correlate all orbitals. The last column shows SS_pG energies based on the intermediate normalization of corrected geminals (see text). All energies are in hartrees.

Molecule	E_{RHF}	ΔE_{MP2}	ΔE_{DMRG}	ΔE_{USSG}	ΔE_{USS_pG}	ΔE_{RSSG}	ΔE_{RSS_pG}	$\Delta E'_{RSS_pG}$
C_8H_{10}	-304.88939	-0.486896	-0.177127	-0.354404	-0.379816	-0.353744	-0.380439	-0.380455
$C_{40}H_{42}$	-1519.962544	-2.414667	-0.857935	-1.646457	-1.804361	-1.626526	-1.815888	-1.816225
$C_{48}H_{50}$	-1823.730875	-2.896671	-1.028113	-1.981566	-2.161978	-1.944683	-2.174951	-2.175365

Comparison to DMRG

Finally, we use long all-*trans* polyene chains, with geometries used by Hachmann *et al* [4] as test systems. The minimal STO-3G basis is used in all polyene calculations. Since DMRG results are based on a π-electron active space, while SSG correlates all electrons, direct energy comparison between the methods is less informative. The results for both are nonetheless provided in Table 5.4 for general comparison.

It's important to compare the energy gained by RSS_pG and USS_pG on top of their respective uncorrected counterparts to the reference MP2 calculations provided. From Table 5.4 it is observed that uncorrected USSG calculations yield more correlation energy to the polyene systems than their RSSG counterparts. The perturbative corrections for both spin treatments recovers a small fraction, 20% , of the remaining dynamic correlation, which compares well with similar systems of extended hydrogen chains. However, it is interesting to note that the RSS_pG correction brings the total energy much closer to USS_pG.

The last column in Table 5.4 presents RSS_pG results based on the intermediate normalization of geminal corrections, as described in Sec. 5.3. All other SS_pG results in this work are based on normalized geminals. Comparison of last two columns shows that the difference in results is marginal, because each geminal correction is quite small. It is consistent with our observation that, in general, strong orthogonality constraint is a relatively mild approximation, provided that the orbitals are fully optimized.

As a note on numerical cost, for the longest $C_{48}H_{50}$ chain studied the computation of the energy correction of the single geminal took ca. 300 s running on a single core, slightly longer but comparable to the single SCF iteration in the optimization of the SSG wavefunction. It is significantly shorter than a ca. 5000 s required for the transformation of the integrals from atomic to molecular orbitals.

5.5 Conclusions

We have presented in this chapter a perturbative-like correction to the computationally inexpensive SSG method that lifts the strong orthogonality constraint placed on the wavefunction. The resulting approximation to the energy of a nonorthogonal geminal method is called SS_pG. In every system compared, SS_pG energy is shown to be superior to that of the nonorthogonal AP1roG method. It's also shown that the energy neglected in the strong orthogonality constraint is usually relatively small compared to total dynamic correlation in systems. Comparisons with GMFCI calculation demonstrate that lifting the strong orthogonality constraint alone is not sufficient to obtain highly accurate energies in some systems. Analysis of extended polyene systems shows that improvements to conjugated π-bonded systems are similar to improvements on extended hydrogen chains. It is also shown that the proposed method scales well with the system size, and is applicable to large strongly correlated systems. Finally, we conclude that the new SS_pG method provides a consistent improvement over SSG as it provides a leading correction to the strong orthogonality approximation in SSG. Further work in our group focuses on computationally cheap ways of including dynamic correlation energy into SS_pG.

Bibliography

[1] Huber, K.; Herzberg, G. *Molecular Spectra and Molecular Structure 4. Constants of Diatomic Molecules*; Van Nostrand: Princeton, 1979.

[2] Limacher, P. A.; Ayers, P. W.; Johnson, P. A.; Baerdemacker, S. D.; Neck, D. V.; Bultinck, P. *J. Chem. Theory Comput.* **2013**, *9*, 1394–1401.

[3] Cassam-Chenaï, P.; Rassolov, V. *Chem. Phys. Lett.* **2010**, *487*, 147–152.

[4] Hachmann, J.; Cardoen, W.; Chan, G. K.-L. *J. Chem. Phys.* **2006**, *125*, –.

[5] Limacher, P. A.; Ayers, P. W.; Johnson, P. A.; Baerdemacker, S. D.; Neck, D. V.; Bultinck, P. *Phys. Chem. Chem. Phys.* **2014**, *16*, 5061–5065.

[6] Born, M.; Oppenheimer, J. R. *Annalen der Physik* **1927**, *389*, 457–484.

[7] Pitzer, K. S. *Acc. Chem. Res.* **1979**, *12*, 271–276.

[8] Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

[9] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

[10] Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

[11] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry*, 1st ed.; McGraw-Hill: New York, NY, 1989.

[12] Rassolov, V. A. *J. Chem. Phys.* **2002**, *117*, 5978–5987.

[13] Rassolov, V. A.; Xu, F.; Garashchuk, S. *J. Chem. Phys.* **2004**, *120*, 10385–10394.

[14] Cagg, B. A.; Rassolov, V. A. *Chem. Phys. Lett.* **2012**, *543*, 205–207.

[15] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, 864–871.

[16] Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048–5079.

[17] Kohn, W. *Rev. Mod. Phys.* **1999**, *71*, 1253–1266.

[18] Pople, J. A. *Rev. Mod. Phys.* **1999**, *71*, 1267–1274.

[19] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[20] Gill, P. M. W.; Adamson, R. D.; Pople, J. A. *Mol. Phys.* **1996**, *88*, 1005–1009.

[21] Savin, A. In *Recent Developments and Applications of Modern Density Functional Theory*; Seminario, J. M., Ed.; Elsevier Science B.V.: Amsterdam, 1996; Vol. 4; pp 327–357.

[22] Fromager, E.; Cimiraglia, R.; Jensen, H. J. A. *Phys. Rev. A* **2010**, *81*, 024502.

[23] Gräfenstein, J.; Cremer, D. *Mol. Phys.* **2005**, *103*, 279–308.

[24] Pérez-Jiménez, A. J.; Pérez-Jordá, J. M. *Phys. Rev. A* **2007**, *75*, 012503.

[25] Bobrowicz, F. W.; Goddard III, W. A. In *Methods of Electronic Structure Theory*; Schaefer, H. F., Ed.; Plenum: New York, 1977; Vol. 3; pp 79–127.

[26] Kraka, E. *Chem. Phys.* **1992**, *161*, 149–153.

[27] Rassolov, V. A.; Ratner, M. A.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 4014–4019.

[28] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

[29] Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. *Phys. Rev. Lett.* **2008**, *100*, 136406.

[30] Adams, R. D.; Rassolov, V.; Zhang, Q. *Organometallics* **2012**, *31*, 2961–2964.

[31] Mehler, E. L.; Ruedenberg, K.; Silver, D. M. *J. Chem. Phys.* **1970**, *52*, 1181–1205.

[32] Rosta, E.; Surján, P. R. *J. Chem. Phys.* **2002**, *116*, 878–890.

[33] Schwartz, C. *Phys. Rev.* **1962**, *126*, 1015–1019.

[34] Morell, G. O.; Parr, R. G. *J. Chem. Phys.* **1979**, *71*, 4139–4141.

[35] Davidson, E. R.; Hagstrom, S. A.; Chakravorty, S. J.; Umar, V. M.; Fischer, C. F. *Phys. Rev. A* **1991**, *44*, 7071–7083.

[36] Chakravorty, S. J.; Gwaltney, S. R.; Davidson, E. R.; Parpia, F. A.; Fischer, C. F. *Phys. Rev. A* **1993**, *47*, 3649–3670.

[37] Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Rassolov, V. A.; Pople, J. A. *J. Chem. Phys.* **1999**, *110*, 4703–4709.

[38] Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 42–55.

[39] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

[40] Sun, J.; Haunschild, R.; Xiao, B.; Bulik, I. W.; Scuseria, G. E.; Perdew, J. P. *J. Phys. Chem.* **2013**, *138*, 044113.

[41] Sun, J.; Xiao, B.; Fang, Y.; Haunschild, R.; Hao, P.; Ruzsinszky, A.; Csonka, G. I.; Scuseria, G. E.; Perdew, J. P. *Phys. Rev. Lett.* **2013**, *111*, 106401.

[42] Perdew, J.; Kurth, S.; Zupan, A.; Blaha, P. *Phys. Rev. Lett.* **1999**, *82*, 2544–2547.

[43] Nichols, B.; Rassolov, V. A. *J. Chem.Phys.* **2013**, *139*, 104111.

[44] Rassolov, V. A. *J. Chem. Phys.* **2009**, *131*, 204102.

[45] Rassolov, V. A. *J. Chem. Phys.* **2011**, *135*, 034111.

[46] Rassolov, V. A. *J. Chem. Phys.* **1999**, *110*, 3672–3677.

[47] Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Phys. Chem.* **1991**, *94*, 7221–7230.

[48] Mayhall, N. J.; Raghavachari, K.; Redfern, P. C.; Curtiss, L. A.; Rassolov, V. *J. Chem. Phys.* **2008**, *128*, 144122.

[49] Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V. A.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

[50] Langhoff, S. R.; Bauschlicher, C. W. *Ann. Rev. Phys. Chem.* **1988**, *39*, 181–212.

[51] Mitin, A. V.; Baker, J.; Pulay, P. *J. Chem. Phys.* **2003**, *118*, 7775.

[52] Cagg, B.; Rassolov, V. *J. Chem. Phys.* **2014**, *141*, 164112.

[53] King, H. F. *J. Chem. Phys.* **1967**, *46*, 705.

[54] Hurley, A. C.; Lennard-Jones, J. E.; Pople, J. A. *Proc. R. Soc. London, Ser. A* **1953**, *220*, 446–455.

[55] Arai, T. *J. Chem. Phys.* **1960**, *33*, 95–98.

[56] Bardeen, J.; Cooper, L. N.; Schrieffer, J. R. *Phys. Rev.* **1957**, *108*, 1175.

[57] Cassam-Chenaï, P. *J. Chem. Phys.* **2006**, *124*, 194109.

[58] Wilson, S. *J. Chem. Phys.* **1976**, *64*, 1692–1696.

[59] Sinanoğlu, O. *Adv. Chem. Phys.*; John Wiley & Sons, Inc., 1964; pp 315–412.

[60] Nesbet, R. K. *Adv. Chem. Phys.*; John Wiley & Sons, Inc., 1965; pp 321–363.

[61] Jeszenszki, P.; Nagy, P. R.; Zoboki, T.; Szabados, Á.; Surján, P. R. *Int. J. Quantum Chem.* **2014**, *114*, 1048.

[62] Shao, Y. et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

Appendix A

The Computation of SSpG Matrix Elements

In this Appendix, we briefly introduce the computation of matrix elements described in the SS_pG formulation section of chapter five. To streamline the notation, we use indices a, b, c, d to label the orbitals of geminal A, z, x, y to label the orbitals of geminal Z, and p, q, s to label the orbitals of geminal P, i is the running index over the orbitals of geminal A, j is the running index over the orbitals of geminal B, and k runs over the orbitals of all other geminals. No more than three geminals need to be considered.

In computation of the Hamiltonian matrix elements, it is convenient to normalize each wavefunction in the summation in Eq. 5.3. First, let us consider a specific term in the geminal Z: $\tilde{\psi}_Z^{a,\bar{b}} = |\phi_a \bar{\phi}_b|$ corresponding to the overall wavefunction $\tilde{\Psi}^{Z,a,\bar{b}} = \hat{A}[\psi_A \ldots \tilde{\psi}_Z^{a,\bar{b}} \ldots]$. Geminal A is expanded in a set of its orbitals (see Eq 5.1), but the antisymmetrizing operator eliminates the terms in this expansion that have the orbital(s) ϕ_a or $\bar{\phi}_b$ occupied, keeping all other terms intact. Therefore, one can view $\tilde{\Psi}^{Z,a,\bar{b}}$ as a wavefunction based on strongly orthogonal geminals, with the geminal Z being uncorrelated and built out of a product $|\phi_a \bar{\phi}_b|$, and geminal A written as $\sum_{i \in A, i \neq a,b} D_i^{a,\bar{b}} |\phi_i \bar{\phi}_i|$, with coefficients $D_i^{a,\bar{b}} = D_i / \sqrt{1 - \sum_{i \in A, i \neq a,b} D_i^2}$. Thus, the standard APSG equations apply to the calculation of the diagonal matrix elements.

For the off-diagonal elements, definitions of both A and Z geminals are different in $\langle bra|$ and $|ket\rangle$ wavefunctions, and one has to examine each term in the summation of the terms in the modified geminal A explicitly. The one-particle density matrix

element between two terms different by one orbital is

$$\langle\tilde{\Psi}^{Z,a,\bar{b}}|\hat{\Gamma}(1)|\tilde{\Psi}^{Z,a,\bar{c}}\rangle = \Gamma_{\bar{b}\bar{c}}\tilde{S}^{A} - \Gamma_{bc}D_{c}^{a,\bar{b}}D_{b}^{a,\bar{c}}\,, \tag{A.1}$$

where Γ_{bc} is a one-particle matrix element between α-spinorbitals ϕ_b and ϕ_c, $\Gamma_{\bar{b}\bar{c}}$ is the matrix element between β-spinorbitals, and $\tilde{S}^{A} = \sum_{i\neq a,b,c} D_{i}^{a,\bar{b}}D_{i}^{a,\bar{c}}$ is the overlap of geminal A, different from unity due to difference in $\langle bra|$ and $|ket\rangle$ wavefunctions, and dependent on the excitation orbitals a, b, c. Note the negative sign in the last term in Eq. A.1. The two-particle matrix element is

$$\begin{aligned}\langle\tilde{\Psi}^{Z,a,\bar{b}}|\hat{\Gamma}(1,2)|\tilde{\Psi}^{Z,a,\bar{c}}\rangle &= \sum_{k\notin A,Z} D_{k}^{2}\Big((\Gamma_{\bar{b}\bar{c},kk}+\Gamma_{\bar{b}\bar{c},\bar{k}\bar{k}})\tilde{S}^{A} - (\Gamma_{bc,kk}+\Gamma_{bc,\bar{k}\bar{k}})D_{c}^{a,\bar{b}}D_{b}^{a,\bar{c}}\Big) \\ &+ \sum_{k\in A,k\neq a,b,c}(\Gamma_{\bar{b}\bar{c},kk}+\Gamma_{\bar{b}\bar{c},\bar{k}\bar{k}})D_{k}^{a,\bar{b}}D_{k}^{a,\bar{c}} - \sum_{k\in A,k\neq a,b}\Gamma_{\bar{c}\bar{k},bk}D_{k}^{a,\bar{b}}D_{b}^{a,\bar{c}} \\ &- \sum_{k\in A,k\neq a,c}\Gamma_{\bar{b}\bar{k},ck}D_{k}^{a,\bar{c}}D_{c}^{a,\bar{b}} + \Gamma_{\bar{b}\bar{c},aa}\tilde{S}^{A} - \Gamma_{bc,aa}D_{c}^{a,\bar{b}}D_{b}^{a,\bar{c}}\,,\end{aligned} \tag{A.2}$$

where antisymmetrization is implied in the 2-particle density matrix terms Γ containing all four orbitals of the same spin.

Other matrix elements in the case of excitations to a single geminal A are

$$\langle\tilde{\Psi}^{Z,a,\bar{b}}|\hat{\Gamma}(1,2)|\tilde{\Psi}^{Z,b,\bar{c}}\rangle = \tilde{S}^{A}\Gamma_{ab,\bar{b}\bar{c}} + D_{c}^{a,\bar{b}}D_{a}^{b,\bar{c}}\Gamma_{cb,\bar{b}\bar{a}} \tag{A.3}$$

$$\begin{aligned}\langle\tilde{\Psi}^{Z,a,\bar{b}}|\hat{\Gamma}(1,2)|\tilde{\Psi}^{Z,d,\bar{c}}\rangle &= \tilde{S}^{A}\Gamma_{ad,\bar{b}\bar{c}} + D_{c}^{a,\bar{b}}D_{a}^{c,\bar{z}}\Gamma_{\bar{b}\bar{a},cd} - D_{c}^{a,\bar{b}}D_{b}^{d,\bar{c}}\Gamma_{da,cb} \\ &+ D_{d}^{a,\bar{b}}D_{b}^{d,\bar{c}}\Gamma_{\bar{d}\bar{c},ba} - D_{d}^{a,\bar{b}}D_{a}^{d,\bar{c}}\Gamma_{\bar{d}\bar{a},\bar{b}\bar{c}}\end{aligned} \tag{A.4}$$

All other cases involving two geminals can be reduced to these equations, either by interchanging spins (for the case $\langle\tilde{\Psi}^{Z,b,\bar{a}}|\hat{\Gamma}(1,2)|\tilde{\Psi}^{Z,c,\bar{a}}\rangle$ to reduce to Eq. A.2), or by considering the orbitals from geminal Z belonging to geminal A with expansion coefficients $D = 0$. For example, $\langle\tilde{\Psi}^{Z,a,\bar{x}}|\hat{\Gamma}(1)|\tilde{\Psi}^{Z,a,\bar{y}}\rangle$ is reduced to Eq. A.1 by substituting $b = x$ and $c = y$ and assuming $D_{y}^{a,\bar{x}} = D_{x}^{a,\bar{y}} = 0$.

The unique matrix elements that involve excitations from Z to two different geminals A and P are

$$\langle\tilde{\Psi}^{Z,a,\bar{p}}|\hat{\Gamma}(1)|\tilde{\Psi}^{Z,a,\bar{q}}\rangle = \tilde{S}^{A}(\tilde{S}^{P}\Gamma_{\bar{q}\bar{p}} - D_{q}^{a,\bar{p}}D_{p}^{a,\bar{q}}\Gamma_{qp})\,, \tag{A.5}$$

$$
\begin{aligned}
\langle\tilde{\Psi}^{Z,a,\bar{p}}|\hat{\Gamma}(1,2)|\tilde{\Psi}^{Z,a,\bar{q}}\rangle &= \sum_{k\notin A,P,Z} D_k^2\Big[\tilde{S}^A\tilde{S}^P(\Gamma_{\bar{q}\bar{p},kk}+\Gamma_{\bar{q}\bar{p},\bar{k}\bar{k}}) \\
&- \tilde{S}^A D_q^{a,\bar{p}} D_p^{a,\bar{q}}(\Gamma_{qp,kk}+\Gamma_{qp,\bar{k}\bar{k}})\Big] \\
&+ \tilde{S}^A(\tilde{S}^P\Gamma_{aa,\bar{q}\bar{p}} - D_q^{a,\bar{p}} D_p^{a,\bar{q}}\Gamma_{qp,aa}) \\
&+ \sum_{i\in A\neq a} \tilde{S}^P D_i^{a,\bar{p}} D_i^{a,\bar{q}}(\Gamma_{ii,\bar{q}\bar{p}}+\Gamma_{\bar{i}\bar{i},\bar{q}\bar{p}}) \\
&+ \sum_{j\in B\neq p,q} \tilde{S}^A D_j^{a,\bar{p}} D_j^{a,\bar{q}}(\Gamma_{jj,\bar{q}\bar{p}}+\Gamma_{\bar{j}\bar{j},\bar{q}\bar{p}}) \\
&- \sum_{i\in A\neq a} D_i^{a,\bar{p}} D_q^{a,\bar{p}} D_i^{a,\bar{q}} D_p^{a,\bar{q}}(\Gamma_{qp,ii}+\Gamma_{qp,\bar{i}\bar{i}}) \\
&- \tilde{S}^A \sum_{j\in B\neq p,q} (D_q^{a,\bar{p}} D_j^{a,\bar{q}}\Gamma_{\bar{p}\bar{j},qj} + D_j^{a,\bar{p}} D_p^{a,\bar{q}}\Gamma_{pj,\bar{q}\bar{j}}) \\
&- \tilde{S}^A D_q^{a,\bar{p}} D_p^{a,\bar{q}}(\Gamma_{qp,\bar{p}\bar{p}}+\Gamma_{qp,\bar{q}\bar{q}}) \ , \qquad \text{(A.6)}
\end{aligned}
$$

$$
\begin{aligned}
\langle\tilde{\Psi}^{Z,a,\bar{p}}|\hat{\Gamma}(1,2)|\tilde{\Psi}^{Z,b,\bar{q}}\rangle &= \tilde{S}^A\tilde{S}^P\Gamma_{ba,\bar{q}\bar{p}} + D_b^{a,\bar{p}} D_q^{a,\bar{p}} D_a^{b,\bar{q}} D_p^{b,\bar{q}}\Gamma_{\bar{b}\bar{a},qp} \\
&- \tilde{S}^A D_q^{a,\bar{p}} D_p^{b,\bar{q}}\Gamma_{ba,qp} - \tilde{S}^P D_b^{a,\bar{p}} D_a^{b,\bar{q}}\Gamma_{\bar{b}\bar{a},\bar{q}\bar{p}} \ . \qquad \text{(A.7)}
\end{aligned}
$$

Overall, there are $O(N^4)$ matrix elements for each geminal Z, leading to $O(N^6)$ scaling of the computational cost based on matrix diagonalizations, with N being the number of one-electron orbitals in the system. To make the model computationally cheaper and the debugging easier, we set most of the matrix elements to zero based on the geminals involved. Specifically, we assume that a matrix element between $\tilde{\Psi}^{Z,a,\bar{b}}$ and $\tilde{\Psi}^{Z,c,\bar{d}}$ is zero unless either both orbitals in the pair (ϕ_a,ϕ_c) belong to the same geminal, or at least one of them is in the orbital subspace of geminal Z. The same criterion applies to the orbital pair $(\bar{\phi}_b,\bar{\phi}_d)$. Physically, this is equivalent to uncoupling of excitations from Z to two different geminals, with excitations of α and β spin electrons considered separate. This approximation becomes exact in the limit of fully localized geminals, with zero differential overlap between the orbitals of two different geminal subspaces. In practice, geminal orbitals have substantially smaller

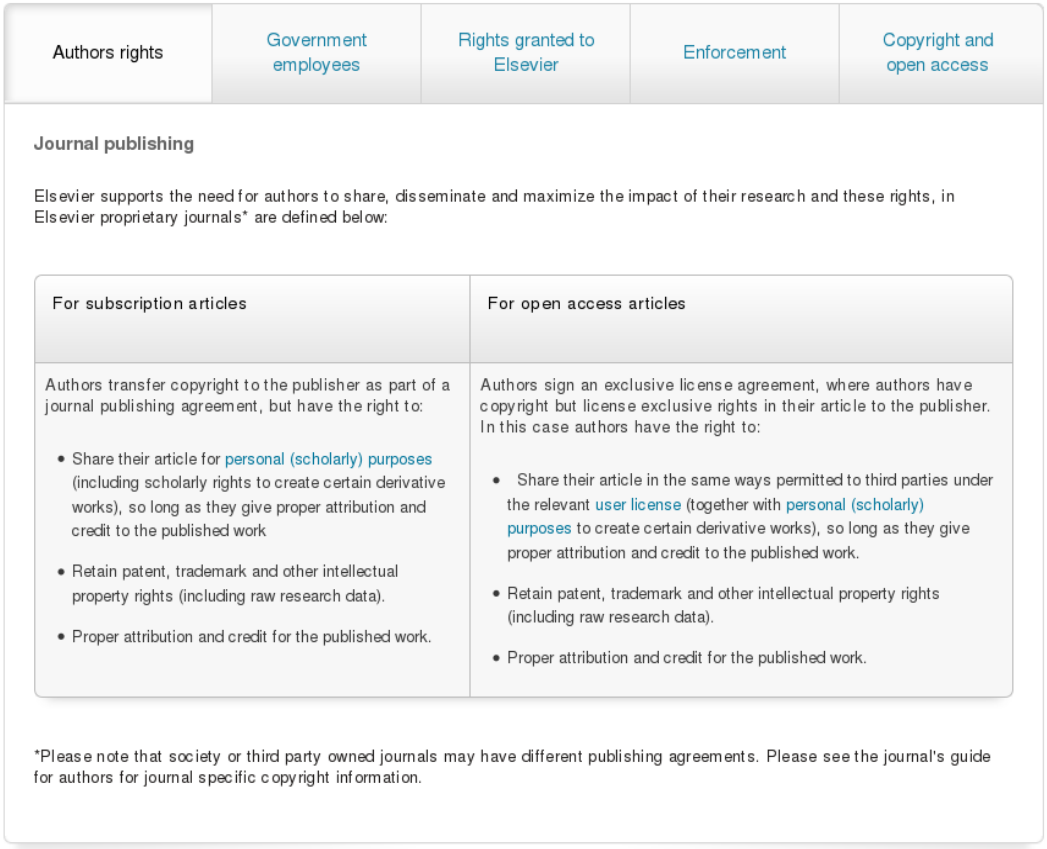
differential overlaps than canonical orbitals due to their spatial localization. The numerical cost of matrix computation and diagonalization is reduced to under $O(N^5)$ for large systems. In practice, the timing and storage requirements are dominated by $O(N^5)$ atomic to molecular orbitals 2-particle integral transformation.

Appendix B

Permission to Reprint

B.1 Chapter 2: Density Functional Model of Multireference Systems Based on Geminals

Permission is granted to all authors to include published work in a dissertation or thesis; as seen on: http://www.elsevier.com/about/policies/author-agreement#authors-rights accessed on 4/21/15.



Authors rights | Government employees | Rights granted to Elsevier | Enforcement | Copyright and open access

Journal publishing

Elsevier supports the need for authors to share, disseminate and maximize the impact of their research and these rights, in Elsevier proprietary journals* are defined below:

For subscription articles	For open access articles
Authors transfer copyright to the publisher as part of a journal publishing agreement, but have the right to: • Share their article for personal (scholarly) purposes (including scholarly rights to create certain derivative works), so long as they give proper attribution and credit to the published work • Retain patent, trademark and other intellectual property rights (including raw research data). • Proper attribution and credit for the published work.	Authors sign an exclusive license agreement, where authors have copyright but license exclusive rights in their article to the publisher. In this case authors have the right to: • Share their article in the same ways permitted to third parties under the relevant user license (together with personal (scholarly) purposes to create certain derivative works), so long as they give proper attribution and credit to the published work. • Retain patent, trademark and other intellectual property rights (including raw research data). • Proper attribution and credit for the published work.

*Please note that society or third party owned journals may have different publishing agreements. Please see the journal's guide for authors for journal specific copyright information.

Personal (scholarly) purposes

Authors can use their articles, in full or in part, for a wide range of scholarly, non-commercial purposes as outlined below:

- Share copies of the article and distribute them via email to colleagues for their research use (also known as 'scholarly sharing').
- Share the article for personal use or for the author's own classroom teaching.
- Use the article at a conference, meeting or for teaching purposes.
- Allow the author's employers to use the article for other internal purposes (such as training).
- Include the article in a printed compilation of the author's works, such as collected writings and lecture notes.
- Inclusion the article in a thesis or dissertation
- Use the article in full or in part to prepare other derivative works, including expanding the article to book-length form, with each work to include full acknowledgement of the article' original publication.

These rights apply for all Elsevier authors who publish their article as either a subscription article or an open access article. In all cases we require that all Elsevier authors always include a full acknowledgement and, if appropriate, a link to the final published version hosted on Science Direct.

B.2 Chapter 5: SSpG: A Strongly Orthogonal Geminal Method with Relaxed Strong Orthogonality

Permission is granted to all authors to include published work in a dissertation or thesis; as seen on: http://publishing.aip.org/authors/copyright-reuse accessed on 4/21/15.

Q: May I include my AIP article in my thesis or dissertation?

AIP permits authors to include their published articles in a thesis or dissertation. It is understood that the thesis or dissertation may be published in print and/or electronic form and offered for sale, as well as included in a university's repository. Formal permission from AIP is not needed. If the university requires written permission, however, we are happy to supply it.