Explicitly correlated quantum chemistry methods for high-spin open-shell molecules

Von der Fakultät Chemie der Universität Stuttgart
zur Erlangung der Würde eines Doktors der
Naturwissenschaften (Dr.rer.nat) genehmigte Abhandlung

Vorgelegt von
Gerald Knizia
aus Halle an der Saale

Hauptberichter:	Prof. Dr. Hans-Joachim Werner
Mitberichter:	Prof. Dr. Hermann Stoll

Tag der mündlichen Prüfung: 25. Feb. 2010

Institut für Theoretische Chemie der Universität Stuttgart
2010

DE Kurzfassung

An der Berechnung der Elektronenstruktur von Molekülen allein aus ersten Prinzipien wird weltweit intensiv geforscht. Ziel der Anstrengungen ist dabei, effiziente rechnerische Methoden zur Bestimmung von Eigenschaften von Molekülen zu entwickeln. Solche Methoden sind auf quantenmechanischen Vielteilchentheorien aufgebaut, die eine korrekte Beschreibung der komplexen korrelierten Bewegung von Elektronen ermöglichen. Trotz der immensen Entwicklung der Rechentechnik innerhalb der letzten Jahrzehnte sind aber hochgenaue Rechnungen nach wie vor auf Moleküle mit weniger als etwa zehn bis zwanzig Atomen beschränkt. In diesem Kontext ist eine der größten Herausforderungen, dass die intrinsische Genauigkeit der verwendeten Vielteilchenmethoden in praktischen Rechnungen bisher nur schwierig erreicht werden konnte. Dies ist eine Nebenwirkung der Expansion molekularer Wellenfunktionen mittels einer beschränkten Basis aus Einteilchenfunktionen ("Molekülorbitalen"): Sehr viele Basisfunktionen pro Atom werden benötigt, um die aus der divergenten Coulomb-Wechselwirkung resultierende Mikrostruktur der Wellenfunktion adäquat zu beschreiben. Die Berechnungen werden sehr schnell sehr teuer.

Explizit korrelierte Methoden stellen einen vielversprechenden Ansatz dar, um dieses Problem der Basissatzkonvergenz zu umgehen. Dabei handelt es sich im wesentlichen um Varianten der gängigen Vielteilchenmethoden (z.B. Störungsrechnung oder Coupled-Cluster-Methoden), bei denen jedoch der gewöhnliche Wellenfunktionsansatz um zusätzliche Terme erweitert wird, die interelektronische Abstände explizit enthalten. Formen dieser Methoden mit Potenzial für Routineanwendungen wurden erstmalig 1991 von Kutzelnigg und Klopper demonstriert. Innerhalb der letzten Jahre wurden explizit korrelierte Methoden dann so weit weiterentwickelt, dass sie jetzt ihre konventionellen Varianten im Genauigkeit-zu-Kosten-Verhältnis deutlich übertreffen.

Im Rahmen dieser Dissertation werden solche explizit korrelierten Methoden entwickelt. Es handelt sich dabei um offenschalige Varianten von Møller-Plesset Störungsrechnung und Coupled Cluster Singles&Doubles.

Desweiteren wird auf theoretische Aspekte eingegangen, wie z.B. die Interpretation von F12-Wellenfunktionen oder spezielle Erscheinungen bei der Behandlung offenschaliger Spezies. Auch technische Aspekte der praktischen Umsetzung der Theorie werden diskutiert, beispielsweise die zuverlässige iterative Lösung der auftretenden Gleichungen und die Vermeidung numerischer Probleme. Weiterhin werden Methoden entwickelt, um im Rahmen explizit korrelierter Rechnungen die zugrundeliegenden Hartree-Fock-Energien zu verbessern, und im Falle der Coupled-Cluster Methoden werden einfache Möglichkeiten untersucht, auch die Basissatzkonvergenz von perturbativen Dreifachanregungen zu beschleunigen.

Um die praktische Verwendung neuer Berechnungsmethoden zu ermöglichen sind umfangreiche Testrechnungen essenziell. Dies gilt insbesondere, wenn ein “black box”-Charakter der Methoden angestrebt wird, wie es hier der Fall ist. Infolge dessen beschäftigt sich ein Großteil der Arbeit mit der Durchführung und Auswertung von Rechnungen, um die Zuverlässigkeit der entwickelten Methoden sicherzustellen, die zu erwartende Genauigkeit zu ermitteln, und optimale Berechnungsparameter abzuleiten. Hierbei werden etwa verschiedene Atomorbital-Basissatzfamilien auf ihre Eignung für F12-Rechnungen untersucht und die Wirkung des freien Längenskalen-Parameters der Theorie getestet. Um einen möglichst weitgehenden Anwendungsbereich der entwickelten Methoden abzudecken, wird bei den Rechnungen besonderes Augenmerk auf schwierige Fallbeispiele gelegt, wie zum Beispiel die Berechnung von Atomisierungsenergien, von Elektronenaffinitäten und Ionisierungsenergien, sowie von Potenzialkurven zweiatomiger Moleküle und von atomaren Dipol-Polarisierbarkeiten.

Die hier entwickelten Quantenchemie-Methoden können bei Rechnungen mit hoher Genauigkeit zu Rechenzeitersparnissen von mehr als zwei Größenordnungen führen, da aufgrund ihrer schnellen Basissatzkonvergenz sehr viel kleinere Basissätze ausreichend sind, um eine gewünschte Genauigkeit zu erreichen. Dadurch leistet die Arbeit einen Beitrag dazu, die Reichweite an Molekülen, für die hochgenaue Rechnungen möglich sind, deutlich zu erweitern.

EN

Abstract

The calculation of the electronic structure of molecules solely from first principles is a subject of intensive research programs worldwide. The goal of these efforts is to develop efficient computational methods for determining molecular properties. Such methods are based on quantum mechanical many-body methods, which allow for a correct description of the complex correlated motion of electrons. However, despite the immense advances in the computational capabilities during the last decades, highly accurate calculations are still limited to molecules containing no more than ten to twenty atoms. In this context, one of the most severe problems is that the intrinsic accuracy of the employed many-body methods is generally hard to reach in practical calculations. This is a side effect of the expansion of molecular wave functions in terms of a finite basis of one-particle functions ("molecular orbitals"): A large number of basis functions per atom is required for accurately describing the microstructure of the wave function which is caused by the divergent Coulomb interaction. Calculations become very expensive very quickly.

Explicitly correlated methods offer a promising approach to circumvent this problem of basis set convergence. Basically, such methods are variants of the common many-body methods of quantum mechanics (e.g., perturbation theory or Coupled Cluster methods) in which the common wave function ansatz is extended by additional terms which contain the inter-electronic distance explicitly. Methods of this form with a potential for routine applications were first proposed by Kutzelnigg and Klopper in 1991. During the last years the explicitly correlated methods were developed to such a degree that they now can clearly surpass the accuracy-to-cost ratio of their conventional counterparts.

Such explicitly correlated methods were developed as a part of this dissertation. Concretely, open-shell variants of Møller-Plesset perturbation theory and of Coupled Cluster Singles&Doubles were produced. Apart from that, also theoretical aspects are discussed, like the interpretation of F12 wave functions or special issues arising when handling open-shell species. Also technical aspects are considered, like the robust solution

of the occurring equation systems and the avoidance of numerical problems. Furthermore, methods are developed to correct for Hartree-Fock basis set deviations in the context of explicitly correlated calculations, and simple approximations are investigated which accelerate the basis set convergence of perturbative triple excitations in Coupled Cluster calculations.

Extensive benchmark calculations are essential for facilitating the practical application of new calculation methods. This applies particularly if the new methods are supposed to have a "black box" character, like it is the case here. As a consequence, a great part of this work is concerned with performing and evaluating calculations, in order to establish the accuracy of the developed methods and in order to derive optimal default calculation parameters. In this regard, different families of atomic orbital basis sets are investigated for their compatibility with F12 calculations, and the effect of the free length scale parameter of the theory is quantified. In pursuance of a wide range of application scenarios, the new methods are tested on a great variety of different calculation tasks. A special emphasis is placed on difficult problems, like the calculation of atomization energies, of electron affinities and ionization potentials, and of potential energy curves of diatomic molecules and of atomic dipole polarizabilities.

The quantum chemical methods developed here can reduce the computational cost of highly accurate calculations by more than two order of magnitude since, due to their fast basis set convergence, less basis functions are required to achieve a desired accuracy. Therefore this work significantly extends the range of molecules for which such highly accurate calculations are possible.

Contents

Abbreviations used in this work

	Quantum chemistry methods
F12	Explicitly correlated
RHF	Spin-restricted Hartree-Fock (open- or closed-shell)
ROHF	Open-shell RHF
RSPT	Rayleigh-Schrödinger perturbation theory [cf. Sec. 2.10]
CI	Configuration Interaction
CCSD	Coupled Cluster Singles & Doubles
UCCSD	Unrestricted open-shell CCSD (here: on RHF reference)
RCCSD	Partially spin-restricted open-shell CCSD
MPn	Møller-Plesset perturbation theory, nth order.
UMPn	Unrestricted MPn (based on UHF reference)
RMPn	MPn-analog based on RHF reference [cf. Sec. 4.1]
	Basis sets & Integrals
MO	"Molecular orbital" (conventional one-particle orbitals)
AO	"Atomic orbital" (refers to basis MOs are expanded in)
CBS	Complete basis set limit (estimate) [cf. Sec. 2.3]
BSSE	Basis set superposition error
DF	Density fitting (for integral evaluations. cf. Sec. 2.4)
JKFIT	Auxiliary DF basis set for Fock matrix computations
MP2FIT	Auxiliary DF basis set for K^{ij}_{ab} computations
RI	Resolution of the identity (in F12 methods)
CA	"Complementary auxiliary" (orthogonal to MO space)
CABS	CA basis set (for approximate RIs in F12 methods)
	Derivations & Approximations
lhs	left hand side (of an equation)
rhs	right hand side (of an equation)
$a := b$	a is defined as b
$a \mathrel{+}= b$	a is changed to b plus a's previous value.
$a \stackrel{!}{=} b$	a is supposed to equal b (e.g., after iterations. Sec. 1.5)
GBC	"Generalized Brillouin condition" (assuming $f^i_x = 0$)
EBC	"Extended Brillouin condition" (assuming $f^a_x = 0$)
RMS	Root mean square (deviation). $\mathrm{rms}(\hat{x}) = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}$

1 Introduction

1.1 Basis set convergence in quantum chemistry methods

Most electronic structure methods use a finite set of one-electron basis functions (called "atomic orbitals") to represent one- and N-electron wave functions. Unfortunately, the basis set convergence of molecular properties is usually slow; several dozen basis functions per atom may be required to obtain accurate energies even for first-row atoms. This is a direct consequence of the singularity of the Coulomb interaction at vanishing inter-electronic distance r_{12}.

Currently this problem is addressed either by a combination of systematic hierarchies of basis sets and basis set extrapolation (see Sec. 2.3), or by using prescribed composite quantum chemistry methods, which sometimes employ empirical parameters (e.g., Ref. [1]). Both approaches have drawbacks. Basis set extrapolation can only be applied when basis sets of a certain minimum size can be employed (e.g., AVQZ, which more often than not is already too large) because otherwise the extrapolation premises are violated and the extrapolation fails. On the other hand, composite methods are adjusted to some specified problem region, usually thermochemical databases, and have unpredictable accuracy outside of their fitting domain. In any case, large basis set calculations with accurate quantum chemistry methods, like coupled cluster methods, are often desirable, but in practice not possible due their computational cost.

Explicitly correlated methods (called F12 methods in the following) address this problem by representing electronic wave functions in the electron-coincidence regions in a non-conventional manner. Compared to conventional quantum chemistry treatments, speedups of two to three orders of magnitude can be obtained because less basis functions are required to reach the same accuracy. This increase in efficiency allows accurate quantum chemistry treatments of much larger systems than previously possible, thus allowing investigations of more interesting chemical systems. This work is concerned with the development of practical F12 methods. In particular, we aim for a black-box character of the developed methods, such that they are as easy to use as their conventional counterparts—only more accurate.

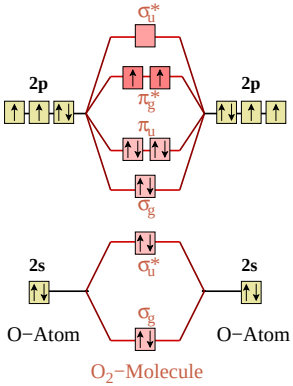

Figure 1.1: The O_2 molecule: Not treatable without open-shell methods

1.2 High-spin open-shell molecules

Most stable molecules occurring in nature have an even number of electrons, and all their electrons pair up perfectly. Under this condition the electronic structure can be qualitatively described by a single Slater determinant formed from doubly occupied molecular orbitals. Such molecules are accessible to closed-shell electronic structure methods.

However, most non-noble-gas single atoms and many interesting stable or metastable molecules do not fall into this category. Radicals have open shells and are usually very reactive, which makes them an important target for theoretical investigations. Most molecules of this class can be described with high-spin open-shell methods. The high-spin state plays this distinguished role because in many scenarios it is the most stable open-shell state; this preference for high-spin states is, for example, expressed by the well-known rules of Friedrich Hund. While high-spin open-shell methods can be constructed with similar recipes as closed-shell methods (due to still being *single-reference* methods), they tend to be susceptible to lots of subtle theoretical and practical problems. This is exemplified by the fact that even today, more than 30 years after the closed-shell CCSD equations have been put into their near-final form, there is still no *implemented* fully spin-adapted high-spin open-shell CCSD program which handles closed electron shells as efficiently as a closed-shell program.

This work is primarily concerned with the open-shell aspects of explicitly correlated methods. In particular, it establishes the open-shell second-order perturbation method RMP2-F12 and the open-shell coupled cluster methods UCCSD-F12x (x=a,b). But due to the parallel development of closed- and open-shell methods, considerable mixing with the closed-shell development occurred.

1.3 Literature review

Already in 1929 Hylleraas performed an accurate electronic structure calculation of the Helium atom [2], which featured a model wave function with terms depending explicitly on the inter-electronic distance r_{12}. This real-space ansatz for the wave function accounted for the Coulomb singularity efficiently, so only few variational parameters were required for a highly accurate description of the system. This technique allowed him to obtain numerical results which were far ahead of their time, but extensions to multi-electron systems proved difficult: To this day, methods based on real-space analytical wave function expressions remain an active research topic, but utility for general molecules of chemical interest could never be established. Such methods are mainly used to calculate properties of few-electron systems (1 to 6) up to sub-cm^{-1} accuracy, thus serving as benchmark reference for more approximate many-body methods, or to study various physical effects where such high accuracy is required (e.g., originating from quantum electrodynamics). Recent advances include accurate wave functions for S and P states of the Beryllium atom [3, 4], highly accurate binding energies of H_2 and He_2 [5], compact wave functions for two-electron systems [6], and a very general, systematic method of obtaining analytic wave function expansions for atoms and molecules [7].

While such methods are very useful for investigating various physical effects, and even for probing quantum mechanics itself, they are unlikely to become competitive with more traditional methods of quantum chemistry for systems with more than ten electrons: The formalism of second quantization cannot be used unless a formal one-particle basis set expansion is introduced, and without second quantization, anti-symmetrization issues of the wave function will quickly become intractable.

Therefore techniques for improving on the defects of traditional quantum chemistry techniques—foremost slow basis set convergence—seem preferable. At this moment, there are two promising, complementary approaches for doing this. One approach is to split the Coulomb interaction into its singular short-range part and its non-singular long-range part, and apply traditional wave function based correlation calculations only on the long-range part. Since this eliminates the Coulomb singularity, the basis set convergence of the wave function part becomes fast. The short-range part is then treated by density functional theory, which does not have basis set scaling defects due to treating electron correlation in an implicit, integrated manner [8–17]. The main problem with this approach is that the accuracy of the density functional part cannot be improved on in a systematic manner. The other way is to follow in the footsteps of Kutzelnigg and Klopper, who developed a practical way of including r_{12}-dependent terms (i.e., terms depending explicitly on the inter-electronic distance) into conventional quantum chemistry methods [18–20]. This is the approach followed in this work.

Kutzelnigg, Klopper, Noga, Gdanitz, and others demonstrated that the basis set

limit of correlated methods can be closely approached with R12 methods within the formalism of the "Standard Approximation" [21–28]. Unfortunately, some of the simplifications made within these first generation R12 methods are only valid if *very* large conventional orbital basis sets are used. Consequently, also these methods were basically restricted to non-routine benchmark calculations, although already many more electrons could be treated than with the fully analytical wave function expansions discussed before. Starting in 2002, the issues which caused these restrictions began to be addressed. Separate auxiliary basis sets for RI approximations, introduced by Klopper and Samson [29], allowed using much smaller orbital basis sets than before. Complementary auxiliary basis sets (CABS), introduced by Valeev [30], increased the efficiency and accuracy of RI approximations. Density fitting (DF) approximations for the numerous new integrals arising in F12 theory were introduced by Ten-no and Manby [31, 32] and greatly increased the numerical efficiency of F12 methods, particularly for "simple methods" like MP2-F12. Ten-no [33] and Manby et al. [34] also realized that non-linear correlation factors, especially the Slater geminal, are much more efficient in describing electron correlation than the linear correlation factors used before. Further important improvements were F12 wave function extensions adjusted to electronic cusp conditions [35], systematic investigations of approximations [36] and ansaetze [37] for MP2-F12, simplified CC2-R12 and CCSD-F12 approximations [38–40], full CCSD-F12 implementations using a CABS [41–43], and F12 treatments for triple excitations [44] and linear response excitation energies [45, 46]. Even local approximations were developed [36, 47, 48], and these can be done in such a way that the F12 corrections greatly reduce their inherent domain approximation error [49–51].

When this work was started, only few explicitly correlated methods existed which allowed treating open-shell systems. The only practical demonstrations of open-shell methods we were aware of were the MRCI-R12 calculations by Gdanitz et al., [23–26], and UMP2-R12 as well as UHF-CCSD-R12 atomization energies and electron affinities by Noga et al. [52, 53]. These calculations were limited, however, to atoms and very small molecules, due to the use of the "Standard Approximation". Later also CASPT2-F12 calculations by Ten-no [54] and Tew et al. [55] were reported, as well as UMP2-R12 calculations for the uracil anion by Bachorz et al. [56]. Very recently, open-shell F12 coupled cluster programs not employing the "Standard Approximation" were demonstrated also by other groups [41, 43, 57, 58]. However, two of them are based on automated implementations [41, 43] and are rather slow (compared with a manually optimized program like ours), one [57] appears to require manual transport of calculation intermediates from one program to another and to produce results significantly worse than our F12x approximations, while the last one used in Ref. [58] is probably fast, but was not yet officially presented and is not yet available for public use.

1.4 Publications and programs arising from this work

This work summarizes, details and extends results which were published in:

- G. Knizia, H.-J. Werner: Explicitly correlated RMP2 for high-spin open-shell reference states, J. Chem. Phys. **128**, 154103 (2008)
- T.B. Adler, G. Knizia, H.-J. Werner: A simple and efficient CCSD(T)-F12 approximation, J. Chem. Phys. **127**, 221106 (2007)
- G. Knizia, T.B. Adler, H.-J. Werner: Simplified CCSD(T)-F12 methods: Theory and benchmarks, J. Chem. Phys. **130**, 054104 (2009)

While the RMP2-F12 article was published only after the first CCSD(T)-F12 article, both the RMP2-F12 and UCCSD(T)-F12a programs were already working at this earlier time. The main reasons for the delay of the open-shell publications were synchronization issues between the open- and closed-shell programs (i.e., making sure that the open-shell program reproduces the results of the closed-shell program when applied to closed-shell cases, which required large-scale modifications in both programs), and also the involved benchmark calculations. The second CCSD(T)-F12 article was delayed because the proposed F12b approximations, unlike F12a, worked only badly in the open-shell case unless the open-shell wave function was modified; and it took me a long period of time to figure out what exactly was going wrong.

For technical reasons the RMP2-F12 program is nearly completely separate from the closed-shell MP2-F12 program; this excludes the evaluation of transformed integrals, for which existing routines of the latter were used. The open-shell CCSD-F12 implementation was achieved by extending the UCCSD/RCCSD program of Knowles, Hampel, and Werner [59, 60] with the attached (T)-triples program of Deegan and Knowles [61].

Some aspects also important for the closed-shell method were first realized in the open-shell program. This includes the perturbative CABS-singles correction to Hartree-Fock energies and some numerical stability improvements. Also open-shell calculations were first used to collect evidence that simply *adding* RMP2-F12 energy corrections to conventional UCCSD energies already results in greatly improved basis set convergence. This fueled the presumption that including only a few essential coupling terms could result in a well working CCSD-F12 method, which turned out to be true. The concrete terms to be included were suggested by Hans-Joachim Werner based on formulas for full closed-shell CCSD-F12 derived by Thomas B. Adler. After their inclusion in the closed-shell CCSD program, they were implemented into the open-shell program with a few weeks delay.

The RMP2-F12 and open-shell CCSD-F12 programs are available for public use in Molpro starting at version 2008.1. We spent large amounts of time on trying to make them as robust and accessible as possible.

1.5 Notation

Orbitals. The indices r,s refer to spin-orbitals inside the molecular orbital (MO) basis, i,j,...,p to occupied spin-orbitals, and α,β,γ,δ to a formally complete set of virtual spin-orbitals. The complete virtual space is divided into the set of virtual orbitals a,b,c,d in the MO basis and the remaining complementary auxiliary (CA) space (indices x, y), so that $\{\alpha\} = \{a\} \cup \{x\}$. In certain terms of the working equations the CA orbitals occur explicitly, and then they are approximated by the union of the atomic orbital (AO) basis and an auxiliary RI basis set. The occupied orbitals indices m, n, o, p are used for F12 generator orbitals [Sec. 3.7]. We assume non-relativistic equations throughout, thus complex numbers are avoided for orbitals and wave functions, and real number arithmetic is applied. All orbitals are orthonormal.

Basis functions. Basis function indices are μ, ν, ρ, σ for the AO basis, η, χ for the RI basis, and A, B for the DF basis (not to be confused with spin labels). Basis functions are not orthonormal.

Tensors and contractions. For non-orthonormal basis sets, lower tensor indices are covariant and upper indices are contravariant. Basis functions and raw integrals are understood to represent purely covariant quantities (e.g., $\left\langle \mu\nu \,\middle|\, r^{-1} \,\middle|\, \sigma\rho \right\rangle$ has four lower labels). For orthonormal basis sets, upper or lower placement of indices has no special significance. Summation over dummy indices occurring twice will be implied in all expressions ("Einstein summation convention").

Spin and spin-orbitals. Spin-labels are called A (alpha or up) or B (beta or down). Combined spin-space coordinates are denoted $\boldsymbol{x} = (\boldsymbol{r}, s)$. We assume that all spin-orbitals are formed as $\varphi(\boldsymbol{x}) = \varphi(\boldsymbol{r}) \cdot \Sigma(s)$, where φ is a spatial orbital and $\Sigma \in \{A, B\}$ is a spin function.

Operators. Unless required for clarity, operator hats will not be used. For any one-particle operator a, the notation a_n means that the operator is applied to particle n and acts as identity operator on the other particles. a_{12} means $a_1 + a_2$. Typically one-particle operators are written in lowercase (e.g., Fock operator f), and two-particle operators with capitals (e.g., correlation factor F_{12}).

Matrix elements. By default, the Dirac notation is employed:

$$\langle rs \,|\, W_{12} \,|\, tu \rangle := \iint \varphi_r(\boldsymbol{x}_1)\varphi_s(\boldsymbol{x}_2) W_{12}(\boldsymbol{x}_1, \boldsymbol{x}_2)\varphi_t(\boldsymbol{x}_1)\varphi_u(\boldsymbol{x}_2)\, \mathrm{d}^4 x_2\, \mathrm{d}^4 x_1.$$

Mulliken notation may be used for Coulomb integrals: $(rs|tu) = \left\langle rt \,\middle|\, r^{-1} \,\middle|\, su \right\rangle$.

Operators in equations. The notation "$a \mathrel{+}= b$" means that b is added to a. The notation $a \stackrel{!}{=} b$ expresses that a is demanded to be equal to b. For example, in an iterative process, $r_a^i \stackrel{!}{=} 0$ means that the goal of the iterations is bring to r_a^i down to zero.

2 Basic Techniques

This section briefly reviews general techniques commonly employed in quantum chemistry which also have central roles in this work. Apart from giving an overview, it is also supposed to fix the notation and to indicate the specific form of the techniques used. A familiarity with closed-shell quantum chemical methods is assumed. This section does not contain original research of me.

2.1 One-electron and N-electron wave functions

An important concept in quantum mechanics is that fundamental particles are indistinguishable. As a result of this, the total N-particle wave function is either *symmetric* (for Bosons) or *antisymmetric* (for Fermions, e.g. electrons) with respect to the interchange of any two particle coordinates:

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_N) = \pm\Psi(\boldsymbol{x}_2, \boldsymbol{x}_1, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_N).$$

In this work we are concerned with electronic wave functions exclusively, which are anti-symmetric.

A N-particle wave function fulfilling the total anti-symmetry condition can be built from one-particle functions ("spin-orbitals") by inserting them into a *determinant*:

$$\begin{aligned}\Phi(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_N) &= \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \varphi_1(\boldsymbol{x}_1) & \ldots & \varphi_1(\boldsymbol{x}_N) \\ \vdots & \ddots & \vdots \\ \varphi_N(\boldsymbol{x}_1) & \ldots & \varphi_N(\boldsymbol{x}_N) \end{pmatrix} \\ &= \frac{1}{\sqrt{N!}} \sum_{\pi \in S_N} \operatorname{sgn}(\pi)\, \varphi_1(\boldsymbol{x}_{\pi(1)})\, \varphi_2(\boldsymbol{x}_{\pi(2)}) \ldots \varphi_N(\boldsymbol{x}_{\pi(N)}), \end{aligned} \tag{2.1}$$

(S_N: permutation group of order N (has $N!$ elements), $\operatorname{sgn}(\pi)$: sign of the permutation π). Wave functions of this form are called *Slater determinants*.

The following statements hold:

- If the one-particle basis is complete, the set of determinants buildable from it forms a basis for the anti-symmetric subspace of N-particle wave functions. That means *any* anti-symmetric wave function can be decomposed into determinants built from a single one-particle orbital set.

- A unitary transformation among the orbitals does not change the determinant built from them.
- If two N-particle determinants are built from the same set of orthonormal spin-orbitals, they are orthogonal (have zero N-particle overlap), unless they consist of exactly the same orbitals. A determinant built from orthonormal spin-orbitals is normalized: $\langle\Phi\,|\,\Phi\rangle = 1$.
- If the set of spin-orbitals used to form the determinant is linearly dependent (e.g., if two are equal), the determinant vanishes. This effect is commonly described with "an orbital cannot be occupied with more than one electron" and called *Pauli principle*. It is a direct consequence of the anti-symmetry of the wave function.

For single-reference correlation methods, which is what we are concerned with here, the calculation starts with determining a *reference determinant* for describing the coarse electronic structure of a molecule. This is done by executing a Hartree-Fock, or, rarely, a Kohn-Sham calculation for the molecule, which determines the orbitals of this function. The orbitals forming the reference determinant are called *occupied orbitals* and labeled with $i,j,\ldots,p$. The number of occupied spin-orbitals is equal to the number of electrons in the molecule. Most kinds of correlated electron motion cannot be represented by a wave function consisting of only a single determinant. The reference determinant is therefore mainly used as starting point for a more involved *correlation calculation*, in which trial wave functions with more degrees of freedom are generated from the reference function. The difference between the system energy obtained in a correlated calculation and its parent reference function is called *correlation energy*.

The space of one-electron functions which is orthogonal to all occupied orbitals is called *virtual space*, and the functions spanning it are called *virtual orbitals* and labeled with $a,b,\ldots,f$. For a given model space, the number of virtual orbitals usually exceeds the number occupied orbitals by far. Determinants which are built from the reference by replacing occupied orbitals with virtual orbitals are called *excited determinants*. They are used in trial wave functions in correlation calculations.

For non-relativistic calculations (or other calculations with small spin-orbit coupling effects) one usually *restricts* the set of spin-orbitals to the set which can be formed by multiplying alpha (A) or beta (B) spin functions to a single set of spatial orbitals:

$$\forall\Sigma \in \{\hat{A},\hat{B}\}: \quad \varphi_{r\Sigma}(\boldsymbol{x}) = \varphi_r(\boldsymbol{r})\Sigma(s), \tag{2.2}$$

where φ_r denotes the spatial orbitals, and the spin functions $\Sigma(s)$ are defined by $\hat{A}(A) = 1$, $\hat{A}(B) = 0$, $\hat{B}(A) = 0$, $\hat{B}(B) = 1$ (in the following, spin functions and spin labels will not be distinguished anymore). For closed-shell and high-spin open-shell systems, spin-adapted N-electron wave functions can then be realized by single

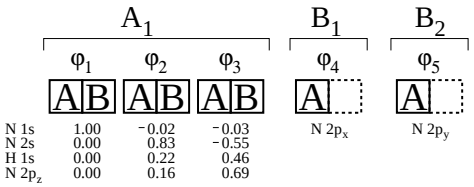

Figure 2.1: A high-spin reference determinant with two open shells: Ground state of NH ($^3\Sigma^-$) = $\det[\varphi_1\hat{A}, \varphi_1\hat{B}, \varphi_2\hat{A}, \varphi_2\hat{B}, \varphi_3\hat{A}, \varphi_3\hat{B}, \varphi_4\hat{A}, \varphi_5\hat{A}]$. The symmetry labels refer to irreducible representations of the C_{2v} subgroup of $C_{\infty}v$. The spatial orbitals are characterized by their composition from atomic orbitals in a minimal basis. $\varphi_1 \ldots \varphi_3$ are the closed-shell orbitals, and φ_4, φ_5 are active orbitals.

determinants: Any determinant built by assembling pairs of A- and B-spin spin-orbitals formed from a single spatial orbital (closed shells) and single A-spin spin-orbitals (active orbitals) is a spin eigenfunction and called a *high-spin determinant*. This is illustrated in Fig. 2.1 for the ground state of the NH ($^3\Sigma^-$) molecule. The "high-spin" refers to the fact that the quantum number s is equal to 1/2 times the number of active orbitals (the determinant is an eigenfunction of operator S^2 for eigenvalue $s(s+1)$). For a given s, there are $2s+1$ orthogonal high-spin wave functions, but only two of them (if $s > 0$) are determinants, the rest being more complicated. Since the non-relativistic Hamiltonian does not depend on spin, all are degenerate and lead to the same energy values, such that it is sensible to pick the simplest of them as reference state for calculations.

In derivations, dealing with determinants directly is painful and error-prone. This can be avoided by using the formalism of second quantization (see Sec. 2.5), which is simply a reformulation of "first quantization" which handles the permutation symmetry constraints of the wave function implicitly.

2.2 Orbital- and auxiliary basis sets

In mainstream quantum chemistry, the spatial part of one-electron molecular wave functions is expressed as a linear combination within a finite set of basis functions (usually local and atom-centered). The form of these basis functions is determined by the need to evaluate molecular integrals over them (for example, kinetic energy matrix elements and electron repulsion integrals), which restricts their form to simple analytical functions. Nowadays basis functions in ab initio methods are nearly exclusively expanded into Gauss type orbitals (GTOs). A spherical primitive GTO with center $\boldsymbol{R}$, angular momentum l and exponent ζ has the functional form

$$\varphi_{\boldsymbol{R}l\zeta}(\boldsymbol{r}) = N(l,\zeta)\, S_{lm}(\boldsymbol{r}-\boldsymbol{R}) \exp(-\zeta\, \|\boldsymbol{r}-\boldsymbol{R}\|^2), \tag{2.3}$$

where $N(l,\zeta)$ is some normalization factor and $S_m^l(\boldsymbol{r})$ is the mth real solid harmonic of angular momentum l [62], which is a polynomial in r_x, r_y and r_z of total degree l. This analytic form allows comparatively simple formulas for various types of integrals which can be evaluated efficiently using recurrence relations [62]. Because primitive GTOs poorly represent occupied atomic orbitals (in particular, they cannot contain any radial nodes), typically multiple ones with different exponents are set into a fixed linear combination (contracted GTO). These are then the actual basis functions used:

$$\varphi_{\boldsymbol{R}ln}(\boldsymbol{r}) = \sum\nolimits_i c_{ni}\varphi_{\boldsymbol{R}l\zeta_i}(\boldsymbol{r}), \tag{2.4}$$

where n indexes contracted functions, and c_{ni} is the contraction coefficient for primitive exponent ζ_i.

Basis functions are usually placed on atoms, but bond-centered functions are also seen occasionally (especially for potential energy surface calculations, see Ref. [63] and articles citing [63]). The exponents and contraction coefficients of the basis functions are taken from a basis set library, which contains pre-optimized basis sets for various kinds of applications.

Some examples of systematic basis set series are listed in Tab. 2.1. These systematic basis sets come in different sizes (denoted by "zeta" cardinalities), from basis sets allowing only qualitative correlation calculations (double zeta), to accurate variants (quadruple zeta and larger). For coupled cluster calculations, a rule of thumb is that going from one basis set cardinality to the next increases the calculation time by one order of magnitude.

Additionally to the orbital basis sets, density fitting (DF) and "resolution of the identity" (RI) techniques require auxiliary basis sets which are optimized to represent quantities other than orbitals (e.g., products of occupied orbitals for DF sets designed for exchange matrix calculations). These additional basis sets have the same functional form as the orbital basis sets, but different exponents, and they usually are largely uncontracted. Like the orbital basis sets, they are taken from basis set libraries.

2.3 Basis set extrapolation

While the basis set convergence of electron correlation calculations is very slow, it is usually quite smooth. The correlation energy approximately follows the law [64,65]

$$E_n \approx E_{\mathrm{CBS}} + a/n^3 \tag{2.5}$$

where a is some system-specific constant, n is the AO basis set cardinality, E_n is the correlation energy for this basis set cardinality, and E_{CBS} is the complete basis set limit. Helgaker and coworkers [64–66] derived this formula from partial

Table 2.1: Popular systematic series of basis sets

cc-pVnZ, (n=D,T,Q,5,6): Dunning-style "correlation consistent polarized valence n-tuple zeta" basis sets. These are basis sets optimized for systematic convergence of *valence correlation energies*. The cc-pV(n+d)Z variants should always be used for second-row elements.
aug-cc-pVnZ: As above, but *augmented* by additional diffuse functions such that excited states, anions, and weak interactions can be described more accurately.
cc-pwCVnZ: weighted core+valence correlation basis sets. Used for all-electron calculations describing valence, core-core and core-valence interactions. These also exist in augmented variants.
def2-SVP, def2-TZVPP, def2-QZVPP: "split valence polarized" (double-ζ), triple- and quadruple zeta valence polarized. These are the Turbomole basis sets, used for both density functional and correlated calculations. They are built with a different optimization strategy and may be able to describe some weak interactions or core-valence correlation effects where cc-pVnZ-type basis sets already require specialized sets. They are available for each element from H to Rn (except Lanthanides), including density fitting basis sets adjusted for them.

wave expansions for single atoms. It holds under the condition that the basis set is saturated for all angular momenta $l \leq l_{\max}$, but the formula also covers the dominant basis set scaling for molecules and when using realistic, unsaturated basis sets (see, e.g. [67]).[1]

This relation is exploited by performing two correlation energy calculations for the same molecule, one with basis set of cardinality n and one with $n-1$, and then using the extrapolated energy,

$$E_{\text{CBS[n-1,n]}} = \frac{n^3 E_n - (n-1)^3 E_{n-1}}{n^3 - (n-1)^3}, \tag{2.6}$$

as actual calculation result. Because the relation is not exact, including more than two calculations in the fitting does not improve the extrapolated energies. Because it assumes basis set saturation for low angular momenta, it generally does not work well unless the $n-1$ basis set has at least triple-zeta quality.

[1] It is noteworthy that first-generation R12 methods were essential in establishing this extrapolation formula. The formula itself was a major breakthrough in quantitative quantum chemistry.

In our work, basis set extrapolation of conventional correlation energy calculations was essential in establishing the basis set limits we require for benchmarking our F12 results.

Hartree-Fock energies do not need to be extrapolated, because they converge quickly with increasing basis set cardinality, and because extrapolation works poorly for relative energies anyway [68]. AVQZ or AV5Z basis sets are usually sufficient to obtain reasonably converged Hartree-Fock results.

2.4 Density fitting and integral evaluation

In quantum chemistry calculations typically large amounts of time are spent with

- Evaluating molecular integrals over basis functions and
- Transforming the resulting integrals into some orthonormal basis.

This particularly applies to F12 methods, which require many different classes of two-electron integrals, involving novel integral kernels and various different orbital spaces. Density fitting (DF)[2] is an approximation technique which can greatly accelerate such tasks. If applied properly, no noteworthy losses in accuracy occur.

The form of DF applied here is robust density fitting, which in the context of F12 theory was pioneered by Manby [31, 32]. The central idea of DF is to approximate four-index two-electron integrals like

$$\langle ij \,|\, W_{12} \,|\, ab\rangle = \iint \varphi_i(\boldsymbol{r})\varphi_j(\boldsymbol{r}')W_{12}(\left\|\boldsymbol{r}-\boldsymbol{r}'\right\|)\varphi_a(\boldsymbol{r})\varphi_b(\boldsymbol{r}')\,\mathrm{d}^3r\,\mathrm{d}^3r' \tag{2.7}$$

by sums of products of two- and three-index two-electron integrals. This is done by fitting the orbital product electron density distributions $|ia) \equiv \varphi_i(\boldsymbol{r})\varphi_a(\boldsymbol{r})$ with an expansion in a single basis set $\{|A)\}$:

$$|ia) \approx |\widetilde{ia}) = D^A_{ia}|A). \tag{2.8}$$

Using this expansion, an integral $\langle ij \,|\, W_{12} \,|\, ab\rangle = (ia|W_{12}|jb)$ can be approximated with the robust fitting formula [32]

$$(ia|W_{12}|jb) \approx D^A_{ia}(A|W_{12}|jb) + (ia|W_{12}|B)D^B_{jb} - D^A_{ia}(A|W_{12}|B)D^B_{jb}. \tag{2.9}$$

As shown in Ref. [32], for any reasonable fitting expansion D^A_{ia}, the errors of the approximated integral $(ia|W_{12}|jb)$ are only quadratic in the incompleteness of the

[2] DF is also called "resolution of the identity" (RI) by many groups. We use "RI" with different meaning in the context of F12 methods.

fitting basis $\{|A)\}$. In practice, the fitting coefficients D^A_{ia} are determined by

$$D^A_{ia}(A|r_{12}^{-1}|B) \stackrel{!}{=} (ia|r_{12}^{-1}|B), \tag{2.10}$$

which corresponds to minimizing $(ia - \widetilde{ia}|ia - \widetilde{ia})$ and thus the difference of the electric fields emitted by $|ia)$ and $|\widetilde{ia})$ [32, 69]. With this metric, integrals over the Coulomb operator have the simple robust form $(ia|jb) \approx D^A_{ia}(A|jb)$, but integrals over other kernels require the full formula (2.9).

The programs developed in the context of this work use the recursive intermediate reusing DF-integral program of Ref. [36] to evaluate transformed four-index integrals. This program manages the numerous DF intermediates and final integral assembly automatically. The molecular integrals themselves are calculated by integral routines of May and Manby, as described in Ref. [32]. I would like to state explicitly that this work would not have been possible without these integral drivers.

2.5 Second quantization and Wick theorems

For practical derivations of quantum chemical methods, it is *very helpful* (although not strictly necessary) to formulate physical operators and wave functions in terms of *second quantization*. Working with determinants directly becomes cumbersome very quickly. The second-quantized operator algebra offers a simple way around that by absorbing all fermionic antisymmetry properties implicitly. Thus it is impossible to violate the Pauli principle as long as one works in second quantization exclusively.

In this work a many-body formalism based on "Quantum Chemistry in Fock Space" is applied [70–73]. This particular formalism has the appealing practical properties of avoiding both continuous basis sets and cumbersome notation, and avoiding time-dependent equation derivations. It is centered around particle number conserving orbital substitution operators instead of using creation- and destruction operators directly[3]. This leads to great simplifications, in particular for spin-free [70] and multi-reference theories [73].

A substitution operator transforms a N-particle determinant into another N-particle determinant by replacing any number of orbitals it is composed of. This is done by first destroying a set of old orbitals, and then creating new orbitals in reverse order:

$$a^r_s := a^\dagger_r a_s, \qquad a^{rs}_{tu} := a^\dagger_r a^\dagger_s a_u a_t, \qquad a^{rst}_{uvw} := a^\dagger_r a^\dagger_s a^\dagger_t a_w a_v a_u, \qquad \ldots, \tag{2.11}$$

($a^\dagger_r$: r-creation operator, a_r: r-destruction operator). Determinants not containing the old orbitals are annihilated by the destructors. Determinants already containing

[3] Ref. [70] uses the term "excitation operator" instead of "substitution operator". We reserve the first term for operators which create more virtual orbitals than they destroy occupied orbitals, and "de-excitation operator" in the reverse sense.

the new orbitals are annihilated by the creators. The operators are defined on the full N-particle Fock space by linearly extending their definition on the determinant basis functions.

In terms of substitution operators, physical one- and two-particle operators (which may not differentiate between identical particles) can be compactly expressed as:

$$\sum_{i=1}^{N} h(i) = \sum_{rs} \langle r \,|\, h \,|\, s \rangle \, a_s^r = h_r^s a_s^r \tag{2.12}$$

$$\sum_{i,j>i}^{N} W(i,j) = \frac{1}{2} \sum_{rstu} \langle rs \,|\, W \,|\, tu \rangle \, a_{tu}^{rs} = \tfrac{1}{2} W_{rs}^{tu} a_{tu}^{rs}, \tag{2.13}$$

where $r \ldots u$ index a complete orthonormal basis set.

Equations (2.12) and (2.13) state operator identities within the N-particle Fock space. Note that the rhs of (2.12) and (2.13) do not depend on the particle number N. This is the natural way of using operators in second quantization. In this sense, the electronic Hamilton operator becomes

$$H = \sum_{rs} \langle r \,|\, h \,|\, s \rangle \, a_s^r + \frac{1}{2} \sum_{rstu} \left\langle rs \,\middle|\, r_{12}^{-1} \,\middle|\, tu \right\rangle a_{tu}^{rs} =: h_r^s a_s^r + \tfrac{1}{2} W_{rs}^{tu} a_{tu}^{rs}, \tag{2.14}$$

where $h = t + v$ is the *core Hamiltonian* consisting of the electron kinetic energy operator t $(= -\frac{1}{2}\Delta)$ and the electrostatic potential emitted by the atom cores v_{nuc} $(= \sum_I \frac{Z_I}{\|\boldsymbol{R}_I - \boldsymbol{r}\|})$. W_{rs}^{tu} denotes matrix elements of the Coulomb interaction.

Substitution operators have the following properties:

- $\left(a_{s_1 \ldots s_n}^{r_1 \ldots r_n}\right)^\dagger = a_{r_1 \ldots r_n}^{s_1 \ldots s_n}$
- Invariance to permutations applied simultaneously to upper and lower indices. Spin-orbital substitutions additionally are anti-symmetric with respect to permutations within the upper indices alone or within the lower indices alone.
- Applicability of "Wick theorems": A product of substitution operators can be decomposed into a sum of normal ordered substitutions by summing up all possible contractions [73]:

$$\begin{aligned}
a_{tu}^{rs} a_w^v &= \left(a_{tuw}^{rsv} + \delta_t^v a_{uw}^{sr} + \delta_u^v a_{tw}^{rs}\right) \\
a_s^r a_{vw}^{tu} &= \left(a_{svw}^{rtu} + \delta_s^t a_{vw}^{ru} + \delta_s^u a_{vw}^{tr}\right) \\
a_{tu}^{rs} a_{xy}^{vw} &= \left(a_{tuxy}^{rsvw} + \delta_t^v a_{uxy}^{srw} + \delta_u^v a_{txy}^{rsw} + \delta_t^w a_{uxy}^{svr} + \delta_u^w a_{txy}^{rvs} + \delta_t^v \delta_u^w a_{xy}^{rs} + \delta_u^v \delta_t^w a_{xy}^{sr}\right).
\end{aligned}$$

  That means summing up all possible combinations of contracting indices on the upper right with indices on the lower left and putting the terms which are

left over above each other like shown in the examples. As long as pairing is preserved (i.e., the terms left over by a contraction are put above each other), no sign rules have to be observed.

The matrix elements of substitution operators are called *reduced density matrices*:

$$\gamma^r_s = \left\langle \Phi \,\middle|\, a^r_s \,\middle|\, \Phi \right\rangle \qquad \gamma^{rs}_{tu} = \left\langle \Phi \,\middle|\, a^{rs}_{tu} \,\middle|\, \Phi \right\rangle \qquad \gamma^{rst}_{uvw} = \left\langle \Phi \,\middle|\, a^{rst}_{uvw} \,\middle|\, \Phi \right\rangle \qquad \ldots . \tag{2.15}$$

If $|\Phi\rangle$ is a single Slater determinant and the spin-orbitals $\{|r\rangle\}$ are subdivided into the occupied $\{|i\rangle\}$ and the virtual $\{|a\rangle\}$ orbitals, the γ matrix elements have simple explicit forms: They vanish (i) if any index corresponds to a virtual orbital or (ii) if any lower index occurs more than once or (iii) if the set of lower indices differs from the set of upper indices. Otherwise they are given by the sign of the permutation aligning the upper indices to the lower indices:

$$\gamma^i_j = \delta^i_j \qquad \gamma^{ij}_{kl} = \delta^i_k \delta^j_l - \delta^i_l \delta^j_k \qquad \ldots . \tag{2.16}$$

The substitution operators introduced in eq. (2.11) were *normal ordered* with respect to the *genuine vacuum*, which means that the vacuum expectation value of any such an operator vanishes. By linearly combining the operators amongst themselves, one can also achieve normal ordering with respect to the *reference function* [72, 73] (also called *hole/particle vacuum*). To accomplish this, the operators are deliberately combined in such a way that their reference expectation values vanish and a generalized Wick theorem holds. In the case of a single Slater determinant $|\Phi\rangle$, the first transformed operators are

$$e^r_s := a^r_s - \gamma^r_s \tag{2.17}$$

$$e^{rs}_{tu} := a^{rs}_{tu} - \left(\gamma^r_t e^s_u + \gamma^s_u e^r_t - \gamma^r_u e^s_t - \gamma^s_t e^r_u + \gamma^{pq}_{rs}\right). \tag{2.18}$$

The general recursive construction can be found in Ref. [73]. It is noteworthy that the notion of creation and annihilation operators for "hole states" is *not* interchanged, in contrast to the practice in other branches of many-body physics. Only a new form of normal order is introduced [72]. Transformed to these operators, the electronic Hamiltonian (2.14) takes the form

$$H = E_{\text{ref}} + f^s_r e^r_s + \tfrac{1}{2} W^{tu}_{rs} e^{rs}_{tu}, \tag{2.19}$$

where $f^s_r = h^s_r + W^{si}_{ri} - W^{is}_{ri}$ denotes the matrix elements of the Fock operator and $E_{\text{ref}} = \langle \Phi \,|\, H \,|\, \Phi \rangle$ is the reference energy.

Again, a generalized Wick theorem can be formulated which allows the decomposition of any e operator product into sums of single normal-ordered e operators.

This Wick theorems now looks more complicated [73]. For example:

$$e^p_q e^r_s = e^{pr}_{qs} + \delta^r_q(1-n_q)e^p_s - \delta^p_s n_p e^r_q + \delta^r_q \delta^p_s n_p(1-n_q), \tag{2.20}$$

where n_p is p's occupation number in the reference determinant (i.e., 0 or 1).

- Additionally to the *particle contractions* from upper right to lower-left, also *hole contractions* from upper left to lower right and mixed contractions occur.
- Particle contractions involving occupied orbitals vanish, as do hole-contractions involving virtual orbitals. This is expressed by the $(1-n_q)$ and n_p factors above.
- There is a factor for (-1) for each hole contraction and for each closed loop.
- Apart from that, the Wick reduction procedure works like for the a operators.

Since the reference expectation value of any single e-operator vanishes (with respect to the reference it is normal-ordered to), the reference-expectation value of any *non-normal ordered* e-operator product is given solely by the *full contractions* in the Wick theorem (i.e., the scalar terms *not* containing operators). This also holds for products of more than two e-operators.

Due to the "full contraction" property, the Wick theorem of e-operators offers a convenient way of translating formal working equations in terms of operators into a series of tensor contractions. Therefore these rules make it possible to avoid the cumbersome direct handling of determinants, as well as replacing the Condon-Slater rules, which are less general. The process can be further broken down by introducing a diagrammatic representation of equation terms and of the Wick theorem [72], but this is not necessary.

2.6 Spin-free substitution operators

Spin-free substitution operators play a central role in closed-shell and restricted open-shell theories. They are built from their spin-orbital counterparts by spin summation:

$$E^{r_1 \ldots r_n}_{s_1 \ldots s_n} := \sum_{\sigma_1 \ldots \sigma_n \in \{A,B\}} e^{r_1 \sigma_1 \ldots r_n \sigma_n}_{s_1 \sigma_1 \ldots s_n \sigma_n}, \tag{2.21}$$

for example

$$E^r_s = e^{rA}_{sA} + e^{rB}_{sB}, \tag{2.22}$$

$$E^{rs}_{tu} = e^{rA\,sA}_{tA\,uA} + e^{rA\,sB}_{tA\,uB} + e^{rB\,sA}_{tB\,uA} + e^{rB\,sB}_{tB\,uB}. \tag{2.23}$$

In total they correspond to *spatial* orbital replacements.[4]

Spin free substitution operators commute with both the total spin operator S^2 and the spin projection operator S_z [74]. It is easily shown that all operators which do not depend on spin (e.g., the non-relativistic Hamiltonian) can be written in terms of such operators. In the chemical physics literature, the one-particle substitution operators E^r_s are often called "generators of the unitary group".[5]

A set of Wick theorems holds also for spin-free operators directly, so an algebraic reduction to spin-orbital operators is not necessary when deriving working equations for fully spin-free theories (no matter if open- or closed shell) [70, 73]. This route will not be followed in this work, however.

2.7 Unrestricted open-shell Hartree-Fock

The UHF wave function $|\Phi\rangle$ is a single N-particle Slater determinant which is characterized by stationarity of its energy expectation value with respect to variations of the occupied spin-orbitals.

Using the methods of Sec. 2.5, the equations resulting from that condition can easily be derived. We can parameterize a perturbed wave function $|\Phi'\rangle$ as follows:

$$\left|\Phi'\right\rangle = \exp\left(\tfrac{1}{2}\alpha_{ia}(e^i_a - e^a_i)\right)|\Phi\rangle\,, \tag{2.24}$$

where α_{ia} are the *rotation angles*. Since $Y = (e^i_a - e^a_i)$ is an anti-Hermitian operator, $\exp(Y)$ is unitary, and thus $\langle\Phi'\,|\,\Phi'\rangle = \langle\Phi\,|\,\Phi\rangle = 1$. Demanding stationarity of $E_0 = \langle\Phi\,|\,H\,|\,\Phi\rangle$ with respect to changes in the orbitals implies

$$\forall i,a: \quad \left.\frac{\partial}{\partial\alpha_{ia}}\right|_{\alpha_{ia}=0} \left\langle\Phi'\,\middle|\,H\,\middle|\,\Phi'\right\rangle \stackrel{!}{=} 0 \tag{2.25}$$

$$\Leftrightarrow \quad \left.\frac{\partial}{\partial\alpha_{ia}}\right|_{\alpha_{ia}=0} \left\langle\Phi\,\middle|\,\left(1+\tfrac{1}{2}\alpha_{ia}(e^i_a - e^a_i)\right)^\dagger H\left(1+\tfrac{1}{2}\alpha_{ia}(e^i_a - e^a_i)\right)\,\middle|\,\Phi\right\rangle = 0 \tag{2.26}$$

$$\Leftrightarrow \quad \left\langle\Phi\,\middle|\,He^a_i\,\middle|\,\Phi\right\rangle = 0. \tag{2.27}$$

Since (2.24) expresses an unitary rotation, no side conditions had to be enforced in the minimization procedure. Since we consider a *first* derivative ∂_α at $\alpha = 0$, only linear orders in α need to be considered. Thus the first-order expansion of

[4] The symbol E is also commonly used to denote non-normal ordered operators (e.g., $E^r_s = a^{rA}_{sA} + a^{rB}_{sB}$ instead of $E^r_s = e^{rA}_{sA} + e^{rB}_{sB}$). But as long as the orbital spaces of the upper and lower labels do not overlap, the a-operators and e-operators are equivalent. This is usually the case, for example, in excitation operators used to parameterize wave functions.

[5] This naming is somewhat unfortunate, because the term omits stating which unitary group is referred to; and the E^r_s themselves are not actually generators of any unitary group (unlike their anti-Hermitian counterparts $\frac{1}{2}(E^r_s - E^s_r)$, which generate unitary rotations of spatial orbitals).

the exponential (2.24) was inserted in eq. (2.26) and subsequently linearized before arriving at eq. (2.27).

Eq. (2.27) can be evaluated by inserting the Hamiltonian (2.19) and using the Wick theorem (2.20):

$$\left\langle \Phi \,\middle|\, H e_i^a \,\middle|\, \Phi \right\rangle = \left\langle \Phi \,\middle|\, (E_0 + f_r^s e_s^r + \tfrac{1}{2} W_{rs}^{tu} e_{tu}^{rs}) e_i^a \,\middle|\, \Phi \right\rangle \tag{2.28}$$

$$= \left\langle \Phi \,\middle|\, (f_r^s e_s^r) e_i^a \,\middle|\, \Phi \right\rangle \tag{2.29}$$

$$= f_i^a \stackrel{!}{=} 0 \tag{2.30}$$

On the rhs of (2.28), neither $E_0 e_i^a$ nor $e_{tu}^{rs} e_i^a$ can lead to any full contractions in the Wick theorem, therefore the respective reference expectation values vanish and only (2.29) remains. Eq. (2.30) is called the *Brillouin* condition, and any set of orbitals satisfying it makes the UHF energy stationary (not necessarily minimal), and thus defines a Hartree-Fock wave function. In the "self-consistent field" (SCF) iterative Hartree-Fock method, one typically introduces *canonical spin-orbitals*, which diagonalize the Fock operator f:

$$f \,|r\rangle = \varepsilon_r \,|r\rangle \,. \tag{2.31}$$

This equation implies eq. (2.30), since f_i^a are non-diagonal matrix elements of f. Note that f is the Fock operator built from the reference function $|\Phi\rangle$ (we made that assumption in eq. (2.28)), and thus it depends explicitly on the orbitals it is supposed to diagonalize. Consequently, eq. (2.31) expresses a non-linear relation which needs to be solved iteratively[6].

2.8 Spin-restricted open-shell Hartree-Fock

In the high-spin open-shell case, we typically employ spin-restricted Hartree-Fock (RHF) reference wave functions. In this case the reference function is a high-spin determinant (as described in Sec. 2.1), built from a set of closed-shell orbitals, doubly-occupied with both alpha (A)- and beta (B)-spin, and singly occupied A-spin active orbitals. Such high-spin determinants are spin eigenfunctions.

The RHF wave function is characterized by stationarity of the energy expectation value with respect to variations in the occupied spatial orbitals (not spin orbitals). Compared to eq. (2.24), the variational ansatz thus changes to

$$\left|\Phi'\right\rangle = \exp\left(\tfrac{1}{2}\alpha_{ia}(E_a^i - E_i^a)\right) |\Phi\rangle \,, \tag{2.32}$$

[6] Though not necessarily with the SCF method. Second-order orbital optimization techniques [62,75] may be used to obtain HF solutions in the very rare cases in which standard SCF cannot be forced to converge.

where again $|\Phi\rangle$ is the current trial HF wave function (here: a high-spin determinant) and $|\Phi'\rangle$ is a perturbed wave function. i and a now refer to spatial orbitals instead of spin-orbitals, and E_i^a is a spin-free excitation operator as described in Sec. 2.6. The index i runs over closed-shell and active orbitals (occupied for A-spin), and a runs over active (unoccupied for B-spin) and virtual orbitals. The ansatz eq. (2.32) includes rotations from closed-shell to active orbitals, from closed-shell to virtual and from active to virtual orbitals, but no rotations leaving the determinant energy invariant (i.e., no closed–closed, no active–active, no virtual–virtual).

In the same way as before, we then obtain

$$\forall i, a: \quad \left.\frac{\partial}{\partial \alpha_{ia}}\right|_{\alpha_{ia}=0} \left\langle \Phi' \middle| H \middle| \Phi' \right\rangle \stackrel{!}{=} 0 \tag{2.33}$$

$$\Leftrightarrow \quad \left\langle \Phi \middle| H E_i^a \middle| \Phi \right\rangle = 0, \tag{2.34}$$

which can be simply unpacked into its spin-orbital components:

$$\left\langle \Phi \middle| H E_i^a \middle| \Phi \right\rangle = \left\langle \Phi \middle| H e_{iA}^{aA} \middle| \Phi \right\rangle + \left\langle \Phi \middle| H e_{iB}^{aB} \middle| \Phi \right\rangle \tag{2.35}$$

$$= \begin{cases} f_{iA}^{aA} + f_{iB}^{aB} & \text{if } i \in \text{closed and } a \in \text{virtual} \\ f_{iA}^{aA} & \text{if } i \in \text{active and } a \in \text{virtual} \\ f_{iB}^{aB} & \text{if } i \in \text{closed and } a \in \text{active} \end{cases} \stackrel{!}{=} 0. \tag{2.36}$$

Depending on the spatial orbital subspaces in which i and a lie in eq. (2.35), either the first or the second term may vanish. If i is an active orbital, then the spin-orbital iA is occupied but iB is not, so e_{iB}^{aB} annihilates the reference function $|\Phi\rangle$. Similarly, for an active a, e_{iA}^{aA} annihilates $\langle\Phi|$, because aA is already occupied. For non-vanishing combinations, the matrix elements are taken from eq. (2.30).

The equations eq. (2.36) are called *RHF Brillouin conditions* and any high-spin determinant fulfilling them is a restricted Hartree-Fock wave function. Again, the Fock operators refer to the ones built form the actual reference function.

Unlike in the closed-shell or unrestricted open-shell case, there are multiple different sensible ways to set up an iterative diagonalization based SCF method for these conditions. For example, one can duplicate the active space (such that it occurs separately in its roles as occupied and partially-unoccupied space), set up a Fock matrix like in Fig. 2.2, diagonalize, and use the occupied section to form new occupied closed-shell and active orbitals. One can also delete the parts of the matrix corresponding to active orbitals in their virtual role to arrive at a matrix with the AO dimension. The choice of the non-shaded regions is rather free in principle, and only determines the convergence properties of the iterative scheme, but not its result.

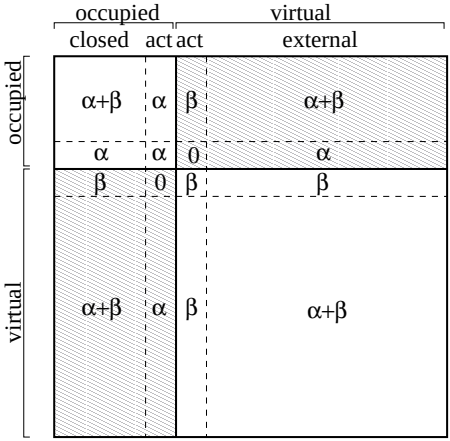

Figure 2.2: One possible way of assembling a single open-shell Fock matrix from A- and B-spin-Fock matrices (in MO basis). In this matrix, all singly occupied spatial orbitals ("active orbitals") are duplicated in the MO basis and occur as both occupied- and unoccupied orbitals separately. If the shaded regions vanish, the MO set fulfills the RHF conditions (2.36). An equivalent matrix can be obtained by removing the non-shaded parts of the active orbital rows and columns.

2.9 The problem of spin contamination

The non-relativistic electronic Hamiltonian (2.14) commutes with the spin operators S^2 and S_z. Consequently, the exact eigensystem of H can be chosen in such a way that H, S^2 and S_z are diagonalized at the same time, so that each eigenvector of H has clear s and m_s quantum numbers. Unfortunately, this property does not automatically propagate to approximate wave functions. While any determinant built exclusively from A- and B-spin-orbitals is an eigenvector of S_z, in general it is not an eigenvector of S^2. For example, UHF reference functions, which employ different spatial parts for orbitals of A- and B-spin, may describe mixtures of states with multiple different s (e.g., singlet and triplet states) at the same time. This violation of spin symmetry is called *spin contamination*.

Spin contamination can be a serious problem for correlation calculations based on UHF reference functions. These correlation calculations "try" to compensate for the wrong physics due to broken spin symmetry, and deteriorate the energetics in the process. The literature documents several accounts of UHF-based calculations failing due to large spin contamination (see [76–80] and references therein).

On the other hand, the high-spin determinants employed in ROHF are eigenvectors of the spin operators S^2 and S_z. As a result, correlation treatments improving on them are not as affected by spin contamination as they would be when using UHF

reference functions—even when the correlation treatment itself is unrestricted. For example, it is well known that the unrestricted RHF-UCCSD(T) and the partially spin-restricted RHF-RCCSD(T) methods show little difference in calculation results (e.g., Refs. [59, 80]).

2.10 Rayleigh-Schrödinger Perturbation Theory

Rayleigh-Schrödinger perturbation theory (RSPT) is the most widely applied size-consistent perturbation method in quantum physics. It is described in detail in Ref. [62] and many other textbooks. Some details are repeated here in order to point out mathematical subtleties. RSPT is based on partitioning the full Hamiltonian of the system in question into an unperturbed Hermitian part H_0 and the Hermitian perturbation V:

$$H = H_0 + \lambda V. \tag{2.37}$$

Later we substitute $\lambda = 1$, which is the case we are interested in. But first the eigenenergies and eigenvectors are expanded in orders of λ around $\lambda = 0$:

$$E_\lambda := \sum_{i=0}^{\infty} \lambda^i E_i \tag{2.38}$$

$$|\Psi_\lambda\rangle := \sum_{i=0}^{\infty} \lambda^i |\Psi_i\rangle . \tag{2.39}$$

We employ "intermediate normalization" (i.e., $\langle \Psi_\lambda | \Psi_0 \rangle := 1$), and assume a non-degenerate system, which allows asserting $\forall i > j : |\Psi_i\rangle \perp |\Psi_j\rangle$. The equations for $|\Psi_n\rangle$ and E_n are obtained by inserting these $|\Psi_\lambda\rangle$ and E_λ into the time-independent Schrödinger equation

$$H |\Psi_\lambda\rangle = E_\lambda |\Psi_\lambda\rangle \tag{2.40}$$

$$\sum_i \lambda^i (H_0 + \lambda V) |\Psi_i\rangle = \sum_{i,j} \lambda^{i+j} E_j |\Psi_i\rangle \tag{2.41}$$

and collecting orders of λ:

$$H |\Psi_\lambda\rangle = \lambda^0 H_0 |\Psi_0\rangle + \lambda^1 (H_0 |\Psi_1\rangle + V |\Psi_0\rangle) + \lambda^2 (H_0 |\Psi_2\rangle + V |\Psi_1\rangle) + \ldots \tag{2.42}$$

$$E_\lambda |\Psi_\lambda\rangle = \lambda^0 E_0 |\Psi_0\rangle + \lambda^1 (E_1 |\Psi_0\rangle + E_0 |\Psi_1\rangle) + \lambda^2 (E_2 |\Psi_0\rangle + E_1 |\Psi_1\rangle + E_0 |\Psi_2\rangle) + \ldots \tag{2.43}$$

Thus, the first equations are:

$$\lambda^0: \qquad (H_0 - E_0)\,|\Psi_0\rangle = 0 \tag{2.44}$$

$$\lambda^1: \qquad (H_0 - E_0)\,|\Psi_1\rangle + (V - E_1)\,|\Psi_0\rangle = 0 \tag{2.45}$$

$$\lambda^2: \qquad (H_0 - E_0)\,|\Psi_2\rangle + (V - E_1)\,|\Psi_1\rangle + (0 - E_2)\,|\Psi_0\rangle = 0. \tag{2.46}$$

Eq. (2.44) tells us that $|\Psi_0\rangle$ *must* be an eigenvector of H_0 for this expansion to be valid. By projecting and re-arranging the remaining equations we obtain

$$\forall \Xi \perp |\Psi_0\rangle: \quad \langle \Xi\,|\,H_0 - E_0\,|\,\Psi_1\rangle + \langle \Xi\,|\,H\,|\,\Psi_0\rangle = 0 \tag{2.47}$$

as equation for the first-order wave function $|\Psi_1\rangle$ (amplitude equation), and

$$E_0 + E_1 = \langle \Psi_0\,|\,H_0\,|\,\Psi_0\rangle + \langle \Psi_0\,|\,V\,|\,\Psi_0\rangle = \langle \Psi_0\,|\,H\,|\,\Psi_0\rangle = E_{\text{ref}}, \tag{2.48}$$

$$E_2 = \langle \Psi_0\,|\,H\,|\,\Psi_1\rangle \tag{2.49}$$

as equations for $E_0 + E_1$ and for the second order energy E_2. By inserting $|\Xi\rangle = |\Psi_1\rangle$ in eq. (2.47) and adding the result to E_2, we obtain the alternative expression:

$$\begin{aligned} E_2 &= \langle \Psi_1\,|\,H_0 - E_0\,|\,\Psi_1\rangle + \langle \Psi_1\,|\,H\,|\,\Psi_0\rangle + \langle \Psi_0\,|\,H\,|\,\Psi_1\rangle \\ &= \langle \Psi_1\,|\,H_0 - E_0\,|\,\Psi_1\rangle + 2\,\text{Re}\,\langle \Psi_1\,|\,H\,|\,\Psi_0\rangle\,, \end{aligned} \tag{2.50}$$

which is denoted as *Hylleraas functional*. Since we deal with real wave functions exclusively, the Re operator in eq. (2.50) can be omitted. Eqs. (2.49) and (2.50) are identical if eq. (2.47) is fulfilled exactly. A difference occurs if this is not the case. If a small perturbation $\delta\,|\Xi\rangle$ (where $|\Xi\rangle \perp |\Psi_0\rangle$ is the error direction) is added to an exact solution $|\Psi_1\rangle$, the direct E_2 expression eq. (2.49) depends linearly on the error δ, while the Hylleraas functional depends on the error only quadratically:

$$\begin{aligned} &\left.\frac{\partial}{\partial\delta}\right|_{\delta=0} \langle \Psi_1 + \delta\Xi\,|\,H_0 - E_0\,|\,\Psi_1 + \delta\Xi\rangle + 2\,\text{Re}\,\langle \Psi_1 + \delta\Xi\,|\,H\,|\,\Psi_0\rangle \\ &\quad = \langle \Xi\,|\,H_0 - E_0\,|\,\Psi_1\rangle + \langle \Psi_1\,|\,H_0 - E_0\,|\,\Xi\rangle + 2\,\text{Re}\,\langle \Xi\,|\,H\,|\,\Psi_0\rangle = 0. \end{aligned} \tag{2.51}$$

Eq. (2.47) was used to find eq. (2.51) to vanish.

In standard Møller-Plesset perturbation theory, one would now use the Hartree-Fock reference function $|\Phi\rangle$ as $|\Psi_0\rangle$, the Fock operator f (which gives a *mean-field* one-particle approximation of the true Hamiltonian H) as H_0, and treat the *fluctuation potential* $V = H - f = W + E_{\text{ref}}$ as the perturbation. Because V then contains the singular Coulomb interaction, which is not a small "perturbation" at all, the Møller-Plesset series shows poor or erratic convergence behavior more often than not [81, 82], and going beyond second-order in MPn is generally rather futile [62].

Nevertheless, MP2 is a popular method due to its favorable accuracy/computation time tradeoff, and for many molecular systems it is the best treatment possible within reasonable amounts of computer time. In particular, MP2 is the simplest first-principles method capable of describing dispersive weak molecular interactions to some extent, which are missing completely in mean field theories like Kohn-Sham Density Functional Theory or Hartree-Fock.

2.11 Coupled-cluster expansions

Nowadays the most popular methods for accurate treatments of dynamic correlation in molecules are based on coupled-cluster expansions of the wave function. A cluster expansion is characterized by building an approximate wave function from the ansatz

$$|\Psi\rangle = \exp(T)\,|\Phi\rangle\,, \tag{2.52}$$

where $|\Phi\rangle$ is the reference determinant, and the cluster-operator T excites occupied electrons into the virtual space. T is expressed as

$$T = T_1 + T_2 + T_3 + \ldots\,, \tag{2.53}$$

where the individual cluster operators T_i represent i-particle excitations:

$$T_1 = \sum_{ai} t_a^i e_i^a \qquad T_2 = \frac{1}{2}\sum_{abij} T_{ab}^{ij} e_{ij}^{ab} \qquad T_3 = \frac{1}{6}\sum_{abcijk} T_{abc}^{ijk} e_{ijk}^{abc} \qquad \ldots \tag{2.54}$$

The variables t_a^i and T_{ab}^{ij} are called singles- and doubles cluster amplitudes, respectively[7]. In the CCSD ("coupled cluster singles&doubles") wave function, the operator is truncated after T_2, such that $T = T_1 + T_2$. This is by far the most common choice. Higher excitations (triples, quadruples) are treated explicitly only in benchmark calculations aiming at sub-kJ/mol accuracy [83, 84]. However, a perturbative treatment of triple excitations based on converged CCSD wave functions is usually required to obtain accurate energies. The so-far best performing methods of this kind are CCSD(T) [85], which has become the de-facto standard, and CCSD-T [61] which is somewhat more complete theoretically, and offers comparable accuracy in practice. A concise account of the (T)-correction working equations is given in Ref. [86].

The exponential parameterization of the wave function in (2.52) allows for size

[7] The normalization of the cluster amplitudes used here is not the one usually employed for spin-orbitals. For example, prefactor 1/4 is usually used instead of 1/2 for the T_2 operator. This was done in order to simplify equations at RMP2 level, where the current choice eliminates many factors of 2 and 1/2.

consistency and size extensivity, which are important desirable asymptotic wave function properties [62, 87]. In short, size extensivity is the property of describing extended regular systems of interacting, identical units correctly, while size consistency means that super-systems of arbitrary non-interacting sub-systems can be dissociated correctly.

In the standard CCSD method, the equations for the energy and for the cluster amplitudes in (2.52) are determined by left-projecting the Schrödinger equation

$$H \exp(T) |\Phi\rangle = E \exp(T) |\Phi\rangle \qquad \Longleftrightarrow \qquad (H - E) \exp(T) |\Phi\rangle = 0 \tag{2.55}$$

onto the configuration subspace spanned by the reference determinant, and the singly and doubly excited configurations, respectively. Thus we get

$$E = \left\langle \Phi \,\middle|\, H \exp(T) \,\middle|\, \Phi \right\rangle \tag{2.56}$$

$$r_a^i = \left\langle \Phi \,\middle|\, e_i^a (H - E) \exp(T) \,\middle|\, \Phi \right\rangle \stackrel{!}{=} 0 \tag{2.57}$$

$$R_{ab}^{ij} = \tfrac{1}{2} \left\langle \Phi \,\middle|\, e_{ij}^{ab} (H - E) \exp(T) \,\middle|\, \Phi \right\rangle \stackrel{!}{=} 0. \tag{2.58}$$

The corresponding matrix elements can be transformed into formal tensor contraction sequences using the Wick theorems sketched in Sec. 2.5 (if done manually, this would be a rather lengthy exercise). In these equations, the expansion of the exponentials terminates, because the Hamiltonian (2.19) can de-excite at most two virtual orbitals, and higher excitations produced by the exponential do not couple to the singles- and doubles space eq. (2.55) is projected on. This form of the residual equations is equivalent to other forms commonly found in the literature which employ a similarity transformed Hamiltonian.

The coupled-cluster equations are solved iteratively. In each iteration, the residual tensors r_a^i and R_{ab}^{ij} are calculated using the current set of cluster amplitudes t_a^i and T_{ab}^{ij}. The residuals determine to which extent the current cluster amplitudes *violate* the CCSD equations. Then the residuals are used to update the amplitudes in a semi-perturbative manner in order to reduce the residual size [88].

2.12 Programming with spin-orbitals

Most formulas in this work involve tensors which are defined with respect to spin-orbital indices. This refers, for example, to the cluster amplitudes of eq. (2.54), and it will later refer to various kinds of calculation intermediates arising in practical computations. When programming open-shell correlation methods based on spin-orbitals, it is necessary to explicitly represent such tensors numerically. A naïve implementation, in which the spatial orbital basis is simply doubled (each orbital occurring once with A- and once with B-spin) would be rather wasteful. This is the

case because we deal with non-relativistic computations exclusively, in which the electronic Hamilton operator is spin-free. As pointed out in Sec. 2.9, this spin-freeness should also be present for approximate wave operators, at least ideally.

In principle this can be achieved by formulating the entire theory in terms of the spin-free excitation operators mentioned in Sec. 2.6. Since the RHF reference function is a clean spin eigenfunction, this would also imply clean approximate wave functions. Unfortunately, due to various technical difficulties of varying degree of severeness, this turns out to be very hard in practice. To my best knowledge, up to this day an efficient spin-free CCSD program could not be produced for any general class of open-shell reference states (see Ref. [89] for a discussion of the problems). Not even the only-approximately spin adapted theory of Ref. [59], which features near-optimal computational scaling, was actually implemented due to its complexity. For this reason we resort to the simpler spin-orbital based formulation.

While it is quite complicated to ensure that an approximate wave function $|\Psi\rangle$ is a proper S^2 eigenvector, making it an eigenvector of the spin projection operator S_z is easy. For this it is sufficient to only include spin-orbital excitation operators $e^{abc\cdots}_{ijk\ldots}$ in the wave function ansatz (e.g., eq. (2.54)) which create as many A-spin spin-orbitals as they destroy (and therefore, also create as many B-spin orbitals as they destroy). This leaves us with non-redundant single excitations of the form

$$t^i_a e^a_i = t^{iA}_{aA} e^{aA}_{iA} + t^{iB}_{aB} e^{aB}_{iB}, \tag{2.59}$$

(lhs indices: spin-orbitals, rhs: spatial orbitals multiplied by indicated spin function), and no $A \mapsto B$ or $B \mapsto A$ excitations need to be considered. For the double excitations, we must also consider the anti-commutation relations of orbital creation and destruction operators:

$$e^{ab}_{ij} = e^{ba}_{ji} = -e^{ba}_{ij} = -e^{ab}_{ji}. \tag{2.60}$$

As a result, of the 16 different conceivable spin label combinations, only double excitations of the form

$$\begin{aligned} T^{ij}_{ab} e^{ab}_{ij} &= T^{ij,AA}_{ab} e^{aA\,bA}_{iA\,jA} + T^{ij,BB}_{ab} e^{aB\,bB}_{iB\,jB} + \\ &\qquad T^{ij,AB}_{ab} \left(e^{aA\,bB}_{iA\,jB} + e^{aB\,bA}_{iB\,jA} - e^{aB\,bA}_{iA\,jB} - e^{aA\,bB}_{iB\,jA} \right) \\ &= T^{ij,AA}_{ab} e^{aA\,bA}_{iA\,jA} + T^{ij,BB}_{ab} e^{aB\,bB}_{iB\,jB} + 4\, T^{ij,AB}_{ab} e^{aA\,bB}_{iA\,jB} \end{aligned} \tag{2.61}$$

are non-redundant. Thus, the spin-orbital t^i_a are represented by two independent amplitude sets (A,B), and the spin-orbital T^{ij}_{ab} are represented by three sets (AA,BB,AB). In the doubles case, the AA and BB amplitudes have different properties compared to the AB ones, though. For AA and BB it is sufficient to store T^{ij}_{ab} for $i > j$

and $a > b$ due to eq. (2.60). For AB no comparable symmetry exists—all $ijab$ are required. Consequently, a spin-orbital formulation has about three times as many doubles amplitudes as a spin-free formulation (in which one set of T_{ab}^{ij} is required for $i \geq j$, all ab, since spin-free double excitations have only one symmetry: $E_{ij}^{ab} = E_{ji}^{ba}$).

Working equations in terms of spin-orbitals then need to be split into their individual spin components for a practical application. Depending on the situation, this may or may not involve subtle sources of errors. To illustrate, let us investigate the spin-orbital energy expression

$$\langle \Phi \,|\, H T_2 \,|\, \Phi \rangle = T_{ab}^{ij} K_{ab}^{ij} = \sum_{ijab} T_{ab}^{ij} \left\langle ij \,\middle|\, r^{-1} \,\middle|\, ab \right\rangle . \tag{2.62}$$

On the rhs, the integral vanishes whenever spin(i) $\neq$ spin(a) or spin(j) $\neq$ spin(b), since r^{-1} acts as identity on the spin space. Consequently, only AA, BB, AB→AB and BA→BA T-amplitudes lead to contributions. Accordingly, we get

$$T_{ab}^{ij} K_{ab}^{ij} = \sum_{ijab} (T_{ab}^{ij,AA} + T_{ab}^{ij,BB} + 2\, T_{ab}^{ij,AB}) K_{ab}^{ij} \tag{2.63}$$

as unpacked expression. However, in practice we often absorb another factor of 2 into the $T_{ab}^{ij,AB}$ amplitudes (e.g., in eq. (4.9) later by not multiplying K_{ab}^{ij} by 1/2), such that the factor 2 for AB in eq. (2.63) can be omitted. On the other side, for AA and BB the summations are carried out explicitly only for $i > j$, such that the partial results must be multiplied by 2 to account for $i < j$.

A different picture arises when computing the wave function squared norm

$$\|T_2\|^2 = \left\langle \Phi \,\middle|\, T_2^\dagger T_2 \,\middle|\, \Phi \right\rangle = \left\langle \Phi \,\middle|\, (\tfrac{1}{2} T_{ab}^{ij} e_{ab}^{ij})(\tfrac{1}{2} T_{cd}^{kl} e_{kl}^{cd}) \,\middle|\, \Phi \right\rangle = T_{ab}^{ij} T_{ab}^{ij}. \tag{2.64}$$

In this expression, also AB→BA and BA→AB excitations have contributions, because unlike the integral K_{ab}^{ij}, the corresponding doubles amplitudes do not vanish. This leaves us with

$$\|T_2\|^2 = \sum_{ijab} \left(T_{ab}^{ij,AA\,2} + T_{ab}^{ij,BB\,2} + 4\, T_{ab}^{ij,AB\,2} \right). \tag{2.65}$$

Also in this expression no factor arises for the AB contribution if the mentioned factor of 2 is absorbed into $T_{ab}^{ij,AB}$.

Integral adjustments may or may not be necessary. For example, in the spin-orbital formalism, the one-particle part of the Hamiltonian, $f = f_r^s e_s^r$, where

$$f_r^s = h_r^s + W_{ri}^{si} - W_{ri}^{is}, \tag{2.66}$$

becomes spin-dependent. If this expression in terms of spin-orbitals is translated into

spatial orbitals, we obtain two different spatial components of the single operator f, one for $\langle rA\,|\,f\,|\,rA\rangle$ and one for $\langle rB\,|\,f\,|\,rB\rangle$:

$$f^{s_A}_{r_A} = h^{s_A}_{r_A} + \left(W^{s_A\,i_A}_{r_A\,i_A} + W^{s_A\,i_B}_{r_A\,i_B}\right) - W^{i_A\,s_A}_{r_A\,i_A}$$
$$f^{s_B}_{r_B} = h^{s_B}_{r_B} + \left(W^{s_B\,i_A}_{r_B\,i_A} + W^{s_B\,i_B}_{r_B\,i_B}\right) - W^{i_B\,s_B}_{r_B\,i_B}. \tag{2.67}$$

These differ in the last term, the "exchange" part, because $W^{is}_{ri} = \left\langle is \,\middle|\, r^{-1} \,\middle|\, ri \right\rangle$ vanishes unless spin(i) = spin(r) and spin(s) = spin(i). Although the spatial parts of the orbitals i_A and i_B are equal, the range of orbitals summed over is different, because for A-spin the active orbitals are occupied, while for B-spin they are not. Again, the actual (spatial) matrix elements of the raw integrals h and W do not depend on the spin labels, as long as the upper spin labels are equal to the lower ones (e.g., $h^{sA}_{rA} = h^{sB}_{rB} = h^{s}_{r}$). So apart from the labels indicating the i summation range, all spin labels can be omitted:

$$f^{s_A}_{r_A} = h^{s}_{r} + \left(\sum_{i\in\text{clos}} 2\cdot W^{si}_{ri} + \sum_{i\in\text{acti}} W^{si}_{ri}\right) - \sum_{i\in\text{clos}} W^{is}_{ri} - \sum_{i\in\text{acti}} W^{is}_{ri} \tag{2.68}$$

$$f^{s_B}_{r_B} = h^{s}_{r} + \left(\sum_{i\in\text{clos}} 2\cdot W^{si}_{ri} + \sum_{i\in\text{acti}} W^{si}_{ri}\right) - \sum_{i\in\text{clos}} W^{is}_{ri}. \tag{2.69}$$

The rhs refers only to spatial orbitals. The spin-orbital operator f can be reassembled from these expressions as

$$f = f^{s_A}_{r_A} e^{rA}_{sA} + f^{s_B}_{r_B} e^{rB}_{sB}. \tag{2.70}$$

In summary, as long as a single set of spatial orbitals is employed, the different spin quantities differ only by either including or excluding summations over active occupied or unoccupied orbitals within the contractions they are formed of. The exact expressions in terms of spatial orbitals can always be determined by tracing intermediate quantities back to either the spin-free integrals used to calculate them, or to the amplitudes of the wave function ansatz. In the following, these spatial expressions will generally *not* be given, because they are very redundant and require a baroque notation, which would confuse the essential aspects of the theory. Nonetheless, some of the less clear aspects of the spin-orbital to spatial transformation in the RMP2-F12 case are described in Sec. 4.1.6.

3 F12 Wave Functions

3.1 The Coulomb interaction singularity

The pathological short-range behavior of the electronic Coulomb interaction lies at the heart of the basis set convergence problems of conventional quantum chemistry methods: The Coulomb potential,

$$W(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|}, \tag{3.1}$$

is singular at $r_{12} := \|\boldsymbol{r}_1 - \boldsymbol{r}_2\| \to 0$. This potential occurs in the electron Hamiltonian

$$H = \sum_i^N v_{\text{nuc}}(\boldsymbol{r}_i) - \frac{1}{2}\sum_i^N \Delta_{\boldsymbol{r}_i} + \sum_{i,j>i}^N W(\boldsymbol{r}_i, \boldsymbol{r}_j), \tag{3.2}$$

where N is the number of electrons, $\boldsymbol{r}_i$ are the electron coordinates, v_{nuc} is the electrostatic potential emitted by the atom cores, and the kinetic energy operator $-\frac{1}{2}\Delta_i$ is the Laplacian acting on coordinate $\boldsymbol{r}_i$.

When any two electron coordinates $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$ approach coalescence, the repulsive interaction potential W in $H\Psi$ goes to infinity. On the other hand, the Schrödinger equation

$$(H - E)\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = 0 \tag{3.3}$$

implies that $H\Psi$ is finite if $E\Psi$ is finite, since in the latter E is a scalar. Therefore, if $\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \neq 0$, the Coulomb singularity in H must be canceled by another singularity. As shown by Kato [90], the singularity cancellation occurs through the kinetic energy terms, which in turn implies derivative discontinuities in the wave function Ψ.

The balance of these singularities is not well described by determinant expansions in terms of single-particle wave functions [cf. Sec. 3.3]. As a result, correlated wave function theories, which attempt to explicitly model the correlated movement of electrons induced by the interplay of their interactions, suffer from basis set convergence problems. A detailed analysis of these issues in terms of partial wave expansions is given in Refs. [18] and [91].

3.2 Form of the wave function in F12 methods

We first discuss the closed-shell case. The open-shell aspects are addressed in Sec. 3.7. In the F12 methods pioneered by Kutzelnigg and Klopper, a conventional expansion of the wave function in terms of determinants of single orbitals is augmented by additional explicitly correlated configurations. As explained later, these explicitly correlated terms improve the wave function's description of electrons coming near to each other. Technically, this wave function extension is accomplished by replacing the conventional double excitation operator T_2 by

$$T_2 = \frac{1}{2} T_{ab}^{ij} E_{ij}^{ab} + \frac{1}{2} \mathcal{T}_{\alpha\beta}^{ij} E_{ij}^{\alpha\beta}, \tag{3.4}$$

where $E_{ij}^{\alpha\beta}$ are the spin-free excitation operators [see Sec. 2.6]. The $T_{ab}^{ij} E_{ij}^{ab}$ part of Eq. (3.4), which describes excitations into the standard virtual orbitals a,b, is the same as used in conventional wave function methods. The additional amplitudes $\mathcal{T}_{\alpha\beta}^{ij}$ describe double-excitations into a formally complete virtual orbital space (indexed by α,β); they are given implicitly by

$$\mathcal{T}_{\alpha\beta}^{ij} = T_{mn}^{ij} \mathcal{F}_{\alpha\beta}^{mn}, \tag{3.5}$$

$$\mathcal{F}_{\alpha\beta}^{mn} = \big\langle mn \big| F_{12} Q_{12} \big| \alpha\beta \big\rangle, \tag{3.6}$$

$$Q_{12} = (1 - o_1)(1 - o_2)(1 - v_1 v_2). \tag{3.7}$$

In these expressions $o = |i\rangle\langle i|$ and $v = |a\rangle\langle a|$ are the projectors onto the occupied and virtual orbital spaces, respectively, $F_{12} \equiv F(r_{12})$ is a suitable short-range correlation factor, and T_{mn}^{ij} are the actual amplitudes used in the F12-treatment. The occurring projector Q_{12} assures orthogonality of the F12-configurations

$$|\Phi_{ij}^{mn}\rangle = \mathcal{F}_{\alpha\beta}^{mn} E_{ij}^{\alpha\beta} |\Phi\rangle \tag{3.8}$$

to all conventional configurations. Thus, $\mathcal{F}_{rs}^{mn} = 0$ and therefore $\mathcal{T}_{rs}^{ij} = 0$. This particular form of Q_{12} maximally decouples the F12 excitations from the conventional excitations; it was introduced as *modified ansatz 2* in Refs. [30,92] and called *ansatz 3* in Ref. [36]. Combining the conventional and F12 excitations, in total the doubles excitation amplitudes in the complete virtual basis are given explicitly by

$$T_{\alpha\beta}^{ij} = \delta_\alpha^a \delta_\beta^b T_{ab}^{ij} + \mathcal{T}_{\alpha\beta}^{ij} = \delta_\alpha^a \delta_\beta^b T_{ab}^{ij} + T_{mn}^{ij} \mathcal{F}_{\alpha\beta}^{mn}. \tag{3.9}$$

In derivations the F12 configurations can thus be treated as *externally contracted* doubly excited configurations regarding a complete virtual space.

Compared to the conventional trial CI expansion in terms of Slater determinants, the F12 terms add additional doubles configurations which can be interpreted as

modifying the reference $|\Phi\rangle$ by replacing occupied orbital products $\varphi_i(\boldsymbol{r}_1)\varphi_j(\boldsymbol{r}_2)$ by short-ranged pair correlation functions

$$|u_{ij}(\boldsymbol{r}_1,\boldsymbol{r}_2)\rangle = T^{ij}_{mn} Q_{12} F_{12}(r_{12}) |\varphi_m(\boldsymbol{r}_1)\varphi_n(\boldsymbol{r}_2)\rangle. \tag{3.10}$$

The most important terms are for $mn = ij$ and $mn = ji$. The function $|u_{ij}\rangle$ represents a negative short-range *hole* in the orbital product $|ij\rangle$. If added to the reference function, it *directly* suppresses the probability of finding the two $|ij\rangle$ electrons in a spatial configuration in which they are close to each other (i.e., near the repulsion singularity). In contrast to orbital products, the short-range correlation factor $Q_{12}F_{12}|\varphi_m\varphi_n\rangle$ can describe the wave function cusp for $r_{12} \to 0$ [cf. Sec. 3.3] correctly. Both aspects counter major deficiencies of conventional wave function expansions.

For the correlation factor F_{12}, we use an exponential function fitted to a set of Gaussian geminals in practice:

$$F(r_{12}) = -\frac{1}{\gamma}\exp(-\gamma r_{12}) \approx \sum_i c_i \exp(-\alpha_i r_{12}^2). \tag{3.11}$$

The exponent γ is a free length scale parameter. In the current work we use six Gaussians, and the coefficients are fitted as described in Ref. [36]. This choice and other non-linear factors have been investigated previously [33,34,93], and the simple Slater function was found to work very well. This ansatz resembles analytic solutions of the hydrogen problem and parts of the helium problem, as already discussed by Hylleraas [2]. It is thus expected to cover the coarse form of the electron correlation hole decently.

This extension of the conventional wave function ansatz by F12 terms leads to dramatic improvements in the basis set convergence of correlation energies; this will be shown later. Consequently, one could get the impression that conventional excitations are not really necessary. Notwithstanding, F12 configurations are not better than conventional configurations in every respect—they are just bad in a very different way, so that combinations of both cancel each others deficiencies. In particular, the size and form of the correlation hole described by the F_{12} correlation factor is constant across the entire molecule, for all orbital pairs ij. In contrast, a realistic correlation hole in a molecule is position dependent, sometimes highly anisotropic [94], and most importantly, varies widely in spatial extent depending on the electron density [see Refs. [95,96], where the correlation hole of the homogeneous electron gas is parameterized. Its extent scales approximately with the Seitz radius r_s, which is a function of the electron density]. None of those aspects is retained in a description by F12 terms alone. The F12 terms simply average a coarse form of F_{12} over the entire molecule, and thus rely on approximate error cancellation. As a result, for a good total description of the electronic structure it is essential

that F12 configurations are combined with conventional configurations. This was amply demonstrated in Ref. [97], by the disastrous performance of no-conventional F12 methods for relative energies.

On the other hand, if conventional excitations are present, no such problems arise. Due to the projector Q_{12} in (3.10), only those parts of the F12 configurations are retained which are not expressible by products of virtual orbitals. Thus, even if the F12 configurations are approximated inaccurately, the coarse form of the wave function is still treated accurately by the actual quantum chemical method in question, because it is well described by the conventional contributions in the wave function expansion.

3.3 Cusp conditions and fixed amplitudes

Let us investigate the effect of electron coalescence on a purely spatial wave function $\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ which is an eigenvector of the molecular Hamilton operator eq. (3.2). Concrete cusp conditions for symmetric and antisymmetric spatial wave functions were first derived by Pack and Byers Brown [98]; they are stated next. Assume that the two coalescenting coordinates are $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, while all other $\boldsymbol{r}_i$ are well separated. If we introduce the coordinate transformation

$$\boldsymbol{s}_{12} = (\boldsymbol{r}_1 + \boldsymbol{r}_2)/2 \qquad \boldsymbol{r}_{12} = \boldsymbol{r}_1 - \boldsymbol{r}_2, \tag{3.12}$$

they derive [98] the following asymptotic short range expansions for symmetric and anti-symmetric wave functions, respectively[1]:

$$\Psi_+(\boldsymbol{r}_{12}, \boldsymbol{s}_{12}, \boldsymbol{r}_3, \ldots, r_N) = \left(1 + \tfrac{1}{2} r_{12}\right) \Psi_+(0, \boldsymbol{s}_{12}, \boldsymbol{r}_3, \ldots, r_N) + O({r_{12}}^2), \tag{3.13}$$

$$\Psi_-(\boldsymbol{r}_{12}, \boldsymbol{s}_{12}, \boldsymbol{r}_3, \ldots, r_N) = \left(1 + \tfrac{1}{4} r_{12}\right) \left(\boldsymbol{r}_{12} \cdot \nabla_{\boldsymbol{r}_{12}} \Psi_-(0, \boldsymbol{s}_{12}, \boldsymbol{r}_3, \ldots, r_N)\right) + O({r_{12}}^3). \tag{3.14}$$

On the other hand, let us consider a symmetric or antisymmetric two-particle wave function built from a product of two smooth one-particle functions:

$$\sqrt{2}\, \Phi_\pm(\boldsymbol{r}_1, \boldsymbol{r}_2) = \varphi_i(\boldsymbol{r}_1)\varphi_j(\boldsymbol{r}_2) \pm \varphi_i(\boldsymbol{r}_2)\varphi_j(\boldsymbol{r}_1). \tag{3.15}$$

In the transformed coordinates this becomes

$$\begin{aligned}\sqrt{2}\, \Phi_\pm(\boldsymbol{r}_{12}, \boldsymbol{s}_{12}) = {} & \varphi_i(\boldsymbol{s}_{12} + \tfrac{1}{2}\boldsymbol{r}_{12})\varphi_j(\boldsymbol{s}_{12} - \tfrac{1}{2}\boldsymbol{r}_{12}) \pm \\ & \varphi_i(\boldsymbol{s}_{12} - \tfrac{1}{2}\boldsymbol{r}_{12})\varphi_j(\boldsymbol{s}_{12} + \tfrac{1}{2}\boldsymbol{r}_{12}),\end{aligned} \tag{3.16}$$

[1] Symmetric/anti-symmetric relative to $\boldsymbol{r}_{12}$, the other coordinates do not matter for the derivation.

and explicit calculation of the $\boldsymbol{r}_{12}$ derivatives at $r_{12} = 0$ shows that

$$\left.\frac{\partial}{\partial r_u}\right|_{r=0} \Phi_+(\boldsymbol{r}_{12}, \boldsymbol{s}_{12}) = 0 \tag{3.17}$$

$$\left.\frac{\partial}{\partial r_u}\frac{\partial}{\partial r_v}\right|_{r=0} \Phi_-(\boldsymbol{r}_{12}, \boldsymbol{s}_{12}) = 0, \tag{3.18}$$

where u and v index Cartesian components of $\boldsymbol{r}_{12}$. That means that symmetrized product functions have no linear r_{12} component and antisymmetrized ones have no quadratic component, and thus the asymptotic short range expansions of the products are[2]

$$\Phi_+(\boldsymbol{r}_{12}, \boldsymbol{s}_{12}) = \Phi_+(0, \boldsymbol{s}_{12}) + O({r_{12}}^2), \tag{3.19}$$

$$\Phi_-(\boldsymbol{r}_{12}, \boldsymbol{s}_{12}) = \boldsymbol{r}_{12} \cdot \nabla_{\boldsymbol{r}_{12}} \Phi_-(0, \boldsymbol{s}_{12}) + O({r_{12}}^3). \tag{3.20}$$

The linear (+) and quadratic cusp (−) terms in (3.13) and (3.14) are missing—products of one-electron functions violate the cusp conditions; they cannot represent the short-range asymptotic structure dictated by the Hamiltonian. On the other hand, (3.19) and (3.20) also demonstrate that the functions

$$(1 + \tfrac{1}{2} r_{12}) \Phi_+(\boldsymbol{r}_{12}, \boldsymbol{s}_{12}) \tag{3.21}$$

$$(1 + \tfrac{1}{4} r_{12}) \Phi_-(\boldsymbol{r}_{12}, \boldsymbol{s}_{12}) \tag{3.22}$$

fulfill the respective cusp conditions. That means we can simply multiply $(1 + \frac{1}{2} r_{12})$ or $(1 + \frac{1}{4} r_{12})$ on spatially symmetric or anti-symmetric product functions, respectively, to satisfy the cusp conditions. [Or instead of r_{12} itself, multiply any other function F_{12} with r_{12} short-range behavior]. This discussion still applies for wave functions built as (anti-) symmetrized products of multiple spatial one-electron functions; it is especially relevant for the pair functions of doubly excited configurations if they are demanded to fulfill the cusp conditions of exact wave functions. Problems only arise if space- and spin degrees of freedom cannot be easily separated.

Anyway, this short-range asymptotic behavior of the exact wave function can be used to great advantage in F12 theory. As pointed out by Ten-no [35], in order to approximately fulfill these asymptotic conditions for a closed-shell wave function $|\Phi\rangle + T_{mn}^{ij} \left|\Phi_{ij}^{mn}\right\rangle$ [cf. Eq. (3.8)], one can fix the amplitudes T_{mn}^{ij} to

$$T_{ii}^{ii} = t_+ \qquad T_{ij}^{ij} = \frac{1}{2}(t_+ + t_-) \qquad T_{ji}^{ij} = \frac{1}{2}(t_+ - t_-), \tag{3.23}$$

where $i \neq j$, $t_+ = 1/2$, $t_- = 1/4$, and the remaining amplitudes are zero. Let us intro-

[2] Anti-symmetric functions vanish for $r_{12} = 0$. Thus, $\boldsymbol{r}_{12} \cdot \nabla_{\boldsymbol{r}_{12}}$ is the first-order Taylor expansion—there is no constant term.

duce $\mathcal{S}_p$ as the purely spatial projector onto the symmetric ($p = 1$) or antisymmetric ($p = -1$) spatial subspaces of the pair functions:

$$\mathcal{S}_p := \frac{1}{2}\left(|rs\rangle\langle rs| + p\,|sr\rangle\langle rs|\right) \tag{3.24}$$

(i.e., the operator acts as *identity* in the spin space). The pair functions $\left|u_{ij}\right\rangle$ of eq. (3.10) then take the explicit form

$$\left|u_{ij}\right\rangle = Q_{12}F_{12}\,|mn\rangle\, T^{ij}_{mn} = Q_{12}\sum_{p=\pm 1} t_p F_{12}\mathcal{S}_p\,|ij\rangle\,. \tag{3.25}$$

So in this case, the cusps are simply added by artificially multiplying F_{12} factors on occupied RHF orbital products in the sense of eq. (3.21) and eq. (3.22)[3]. As a result, the short-range electron behavior not expressible by the virtual-orbital products of conventional excitations (Q_{12}) is fixed in consistence with the cusp conditions.

Of course this does not mean that these additions are exact. Additionally to the possibility that the RHF function might not be a good coarse description of the electronic structure in question, the cusp conditions only determine the wave function behavior at $r_{12} \to 0$. But the F_{12} functions we add have sizable spatial extent (with $\gamma \approx 1/a_{\text{bohr}}$), which is not fixed by the conditions, and which is more important for the properties of the molecule because it covers a larger volume element. Accordingly, flexible wave function expansions with T^{ij}_{mn} not fixed at all, or with T^{ij}_{mn} restricted to diagonal excitations (allowing only mn with $mn = ij$ or $mn = ji$) may or may not lead to better practical performance, even if the cusp conditions are violated.

3.4 Ansatz restrictions for T^{ij}_{mn}

As already indicated, the ansatz for explicitly correlated configurations in eq. (3.5) is not always used in full, and some elements of T^{ij}_{mn} may be assumed to vanish or to be fixed. The following ansaetze are in common use for ground state calculations:

- *The diagonal ansatz:* In this case only T^{ij}_{mn} are allowed where the spatial parts of either $|ij\rangle$ and $|mn\rangle$ or of $|ij\rangle$ and $|nm\rangle$, coincide (i, j, m, n are spin-orbitals). This is the original ansatz of Kutzelnigg and Klopper [19, 20]. This ansatz is the simplest to interpret, because it directly corresponds to cutting short-range correlation holes $\left|u_{ij}\right\rangle_{\pm} \equiv t_{ij,\pm}Q_{12}F_{12}\,|ij\rangle_{\pm}$ into orbital products $|ij\rangle_{\pm}$, where $\pm$ indicates projections onto the symmetric/anti-symmetric pair function space in the sense of eq. (3.24).

  A problem with this ansatz is that is not unitarily invariant: If the reference determinant $|\Phi\rangle$ is expressed by a different but equivalent orbital set, properties

[3] The 1-term is implicit because these functions are added to the reference function.

derived from the correlated wave function change. Additionally, the ansatz may violate size-consistency unless it is applied with localized occupied orbitals.

On the other hand, the amplitudes $t_{ij,\pm}$ are variable. This makes this ansatz rather insensitive to the choice of the geminal exponent γ. Also, the equation systems determining the amplitudes are well-behaved typically, and only the parts R_{ij}^{ij} and R_{ji}^{ij} of the F12 residuals need to be evaluated.

- *The general orbital invariant ansatz:* No restrictions are imposed on the T_{mn}^{ij}. The ansatz was introduced by Klopper [99] to remove the lack of unitary invariance of the diagonal ansatz. It is not easily interpreted. And despite giving the wave function more freedom, it was found to actually *worsen* the F12-treatment in certain cases. This can happen due to a geminal basis set superposition error (BSSE) [36, 100], which cannot easily be corrected for. Additionally, the equation system determining the T_{mn}^{ij} sometimes becomes singular due to redundancies in the F12 configuration manifold [cf. Sec. 4.1.10]. While the singularities can be projected out easily, they make calculation results somewhat ambiguous and might lead to discontinuity problems when potential energy surfaces are evaluated.

- *The fixed amplitude ansatz*: This is a scalar ansatz introduced by Ten-no [35]. As discussed in Sec. 3.3, the T_{mn}^{ij} are fixed in such a way that they are consistent with the wave function cusp conditions determined from the RHF reference function—they are therefore completely independent of the actual correlation method being employed.

  Like the general ansatz, this ansatz is unitarily invariant. Like the diagonal ansatz (of which it can be considered a variant), it is free of geminal BSSE. Unlike both, it is sensitive to the choice of the geminal exponent γ. This sensitivity makes it necessary to account for the γ-dependence in some way. This is done by evaluating the F12 part of the correlation energy by an exact or approximate Lagrange functional (e.g., the Hylleraas functional eq. (2.51) in RSPT2), which adjusts for the non-solved F12 amplitude equations in the intermediate r_{12}-range. Because the F12 residuals R_{ij}^{ij}/R_{ji}^{ij} form a part of these functionals, the residuals still need to be evaluated at least in an approximate way, even though the T_{mn}^{ij} amplitudes are not varied.

Despite the apparently reduced complexity of the diagonal and fixed ansaetze compared with the general one, in many cases the actual numerical work required to evaluate any of them does not differ fundamentally: All of them require evaluating the same transformed integrals, and this is usually the most time consuming step in a MP2-F12 calculation. However, the general ansatz contains some $O(N^6)$ terms which the diagonal ones do not have. These terms can become time limiting for

large MP2-F12 calculations. For our approximate CCSD(T)-F12 methods, the -F12 part is cheap compared to the conventional CCSD(T) part anyway, regardless of the ansatz.

Nevertheless, more rigorous CCSD-F12 treatments can profit from the fixed amplitudes because several coupling terms between the F12 and conventional parts of the wave function remain constant during the iterations. Additionally, the evaluation of the F12 residual R_{mn}^{ij}, which may be expensive if evaluated rigorously, is then required only a single time for the F12 Lagrange functional.

3.5 Strong orthogonality projectors

Various matrix elements involving the strong orthogonality projector [eq. (3.7)] arise when deriving working equations of F12 methods. In principle it is always possible to calculate such matrix elements exactly, but doing so nearly always results in an intractable computational burden. For example, the Q_{12} projector in $\langle F_{12} Q_{12} F_{12} \rangle$ contains a part

$$\langle mn \,|\, F_{12} o_1 F_{12} \,|\, op \rangle = \langle mn| \, F_{12} \, |i\varrho\rangle \, \langle i\varrho| \, F_{12} \, |op\rangle = \ldots , \tag{3.26}$$

where ϱ runs over a hypothetical complete orthonormal basis set. The rhs can be resolved exactly into a three-electron integral by inserting the real-space completeness relation $|\varrho\rangle\langle\varrho| \equiv \int_{r_\varrho} \left|r_\varrho\right\rangle \left\langle r_\varrho\right| \mathrm{d}^3 r_\varrho$ (and $\langle r_\varrho | r'_\varrho \rangle = \delta(r_\varrho - r'_\varrho)$, the Dirac delta distribution):

$$\begin{aligned}
\ldots &= \int_{r_\varrho} \Bigg[\iint m(\boldsymbol{r}_1) n(\boldsymbol{r}_2) F(\boldsymbol{r}_1, \boldsymbol{r}_2) i(\boldsymbol{r}_1) \delta(\boldsymbol{r}_2 - \boldsymbol{r}_\varrho) \, \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \cdot \\
&\qquad\qquad \iint i(\boldsymbol{r}_3) \delta(\boldsymbol{r}_4 - \boldsymbol{r}_\varrho) F(\boldsymbol{r}_3, \boldsymbol{r}_4) o(\boldsymbol{r}_3) p(\boldsymbol{r}_4) \, \mathrm{d}^3 r_3 \mathrm{d}^3 r_4 \Bigg] \mathrm{d}^3 r_\varrho \\
&= \iiint m(\boldsymbol{r}_1) n(\boldsymbol{r}_\varrho) i(\boldsymbol{r}_3) F(\boldsymbol{r}_1, \boldsymbol{r}_\varrho) F(\boldsymbol{r}_3, \boldsymbol{r}_\varrho) i(\boldsymbol{r}_1) p(\boldsymbol{r}_\varrho) o(\boldsymbol{r}_3) \, \mathrm{d}^3 r_1 \mathrm{d}^3 r_\varrho \mathrm{d}^3 r_3 \\
&= \langle mni \,|\, F_{12} F_{32} \,|\, ipo \rangle
\end{aligned} \tag{3.27}$$

(Here the orbitals $\varphi_x(\boldsymbol{r}_1)$ are denoted as $x(\boldsymbol{r}_1)$). Such integrals can indeed be evaluated in closed form, but there are simply too many of them. Methods employing such techniques can therefore never become computationally competitive with conventional methods. How such many-electron integrals can be avoided using RI approximation was one of the key discoveries by Kutzelnigg and Klopper [18–20], which paved the way to practical quantum chemical methods involving explicitly correlated wave functions.

In modern F12 treatments employing the CABS approach, the Q_{12} projector is always treated in one of three different ways. Ordered by preferability, these are [36]:

Fully analytically: Since Q_{12} is a projection operator, it is idempotent:

$$Q_{12}Q_{12} = Q_{12}. \tag{3.28}$$

A related property involves the formal double RI in the virtual space:

$$Q_{12} \underbrace{|\alpha\beta\rangle\langle\alpha\beta|}_{=(1-o_1)(1-o_2)} = Q_{12}. \tag{3.29}$$

These two relations often allow evaluating contractions between different $\mathcal{F}^{kl}_{\alpha\beta}$ [eq. (3.6)] by exactly resolving the summations over the formally complete virtual basis $\{|\alpha\rangle\}$.

Partly analytically : When a completely analytical treatment is impossible, it may still be possible to treat the dominant parts of Q_{12} exactly. This involves rewriting

$$\begin{aligned} Q_{12} &= (1-o_1)(1-o_2)(1-v_1v_2) \\ &= 1 - o_1(1-(o_2+v_2)) - (1-(o_1+v_1))o_2 - (o_1+v_1)(o_2+v_2) \\ &\approx 1 - |ix\rangle\langle ix| - |xi\rangle\langle xi| - |rs\rangle\langle rs|\,. \end{aligned} \tag{3.30}$$

In the second and third term, in which $1-(o+v)$ projects onto "everything not describable within the MO basis", the approximation

$$1 - |r\rangle\langle r| \approx |x\rangle\langle x| \tag{3.31}$$

is made by expanding these terms in an extended (but finite) complementary auxiliary basis set. All other projectors are treated exactly. As a result, in eq. (3.30) at most *one* numerical RI is performed per term, not *two* which would occur in a straight-forward RI expansion of Q_{12}. It should be noted that in eq. (3.30) i must run over *all* occupied orbitals, including frozen core orbitals if present.

Fully numerically: Sometimes a straight-forward numerical expansion of Q_{12} as double RI is either necessary or convenient. It is necessary if integrals resulting from the unit operator in eq. (3.30) cannot be resolved analytically; and it is convenient if the numerical resolution converges quickly and is cheaper to evaluate than the analytical one. In either case, we employ:

$$\begin{aligned} Q_{12} &= (1-o_1)(1-o_2)(1-v_1v_2) \\ &= v_1(1-(o_2+v_2)) + (1-(o_1+v_1))v_2 + (1-o_1-v_1)(1-o_2-v_2) \\ &\approx |ax\rangle\langle ax| + |xa\rangle\langle xa| + |xy\rangle\langle xy|\,. \end{aligned} \tag{3.32}$$

With this formula, the quadratic basis incompleteness of the combined MO and CABS basis is isolated in the last term, which can often be dropped to achieve computational savings [36, 101–103].

All of these approaches avoid the explicit calculation of 3- and 4-electron integrals completely, and replace them by sums of products of 2-electron integrals. On the downside, an additional numerical RI basis set needs to be introduced, and the concrete evaluation of terms often becomes cumbersome.

In early R12 methods the "standard approximation" was employed, which assumed that the AO basis is "large enough" for explicit RIs. In the modern language, this implies that the approximation $|x\rangle\langle x| \approx 0$ is applied throughout (i.e., that all terms involving $|x\rangle$ orbitals are neglected). While this is a rather coarse approximation, it can still be applied for numerical or computational gains in many less-important terms as we will demonstrate later.

This applies particularly to expressions which require the use of eq. (3.32). Employing this projector corresponds to evaluating the term in which it occurs completely conventionally, just in a bigger basis (MO+CABS). It is therefore a "last resort"-kind of approach—if it would be applied throughout in preference to eq. (3.29) and eq. (3.30), F12 methods could not work better than non-F12 methods at all. On the other hand, if a term *requires* the use of eq. (3.32) (i.e., if eq. (3.30) cannot possibly be applied instead), it also means that the term in question *is not* one which attempts to perform the problematic partial wave expansion of

$$1 = r_{12}^{-1} r_{12} = \sum\nolimits_l \langle (r_{12}^{-1})_l (r_{12})_l \rangle$$

(with r_{12}^{-1} from the Coulomb interaction and r_{12} from the wave function cusp), which Ref. [18, cf eq. 7.2] argues to lie at the root of the slow basis set convergence of conventional quantum chemistry methods. Therefore, terms *requiring* eq. (3.32) are expected to be represented much better by the regular MO basis expansion than the terms where this is not the case.

3.6 Numerical construction of the complementary auxiliary (CA) basis

All formulas of Sec. 3.5 become exact in the limit of a complete CA basis set $\{|x\rangle\}$. But as mentioned, in practice the terms involving $|x\rangle$ orbitals are explicitly represented numerically. One way of dealing with these terms (which was used in the RMP2-F12 program, but originally not in the LMP2-F12 program [36]), is to explicitly calculate a set of orthonormal $|x\rangle$ orbitals. These are generated by projecting the MO basis out of a given auxiliary RI basis set, thus assuring

$\forall x, r : \langle r \,|\, x \rangle = 0$. As a result, $|x\rangle$ orbitals have components in both the AO and the RI bases.

This projection can be done either directly in the basis of non-orthogonal RI basis functions, or in the basis of an auxiliary orthonormalized set of RI functions. In the first case, the projected RI functions

$$|\tilde{\chi}\rangle = (1 - |r\rangle\langle r|)\,|\chi\rangle = |\chi\rangle - |\mu\rangle\, C_r^\mu C_r^\nu \langle \nu \,|\, \chi\rangle \tag{3.33}$$

are formed directly from the basis functions, and subsequently orthonormalized. Here χ,η denote RI basis functions, μ,ν denote AO basis functions, and C_r^μ are the molecular orbital coefficients (and $C_r^\mu C_r^\nu = [\boldsymbol{S}_{\mathrm{AO}}^{-1}]^{\mu\nu}$, but it is numerically favorable not to do this contraction first). The orthonormal basis is formed by first explicitly constructing the overlap matrix of the projected functions,

$$\begin{aligned}
[\boldsymbol{S}]_{\tilde{\eta}\tilde{\chi}} = \langle \tilde{\eta} \,|\, \tilde{\chi}\rangle &= \left\langle \eta \,\middle|\, (1 - |r\rangle\langle r|)^2 \,\middle|\, \chi \right\rangle = \langle \eta \,|\, (1 - |r\rangle\langle r|)\,|\chi\rangle \\
&= \langle \eta \,|\, \chi\rangle - \langle \eta \,|\, r\rangle\langle r \,|\, \chi\rangle \\
&= \langle \eta \,|\, \chi\rangle - \langle \eta \,|\, \mu\rangle\, C_r^\mu C_r^\nu \langle \nu \,|\, \chi\rangle\,,
\end{aligned} \tag{3.34}$$

[where $\langle \eta \,|\, \chi\rangle$ and $\langle \eta \,|\, \mu\rangle$ are the RI/RI and RI/AO overlap matrices, respectively] then constructing $[\boldsymbol{S}^{-1/2}]_{\tilde{\eta}\tilde{\chi}}$ by spectral decomposition, and finally obtaining the orthonormal set of orbitals as

$$\begin{aligned}
\{|x\rangle\} &= \{[\boldsymbol{S}^{-1/2}]_{\tilde{\eta}\tilde{\chi}}\,|\tilde{\chi}\rangle\,, \eta = 1 \ldots N_{\mathrm{RI}}\}. \\
&= \{[\boldsymbol{S}^{-1/2}]_{\tilde{\eta}\tilde{\chi}}(|\chi\rangle - |\mu\rangle\, C_r^\mu C_r^\nu \langle \nu \,|\, \chi\rangle), \eta = 1 \ldots N_{\mathrm{RI}}\}.
\end{aligned} \tag{3.35}$$

Numerically, the set is represented by two matrices, `RiCabs` and `AoCabs`, with dimensions `nRi×nCabs` and `nAo×nCabs`, respectively.

This is the direct way of forming CABS vectors, and if all $|x\rangle$ functions are retained then it is also the favorable approach. But in practice, the projected overlap matrix $[\boldsymbol{S}]_{\tilde{\eta}\tilde{\chi}}$ generally contains very small eigenvalues ($\ll 10^{-8}$), such that serious numerical instabilities would be introduced if the corresponding eigenfunctions were not deleted. Thus, the full span of the RI basis can usually not be retained as auxiliary space. This is an expression of the near linear dependency which arises within the CA space due to two different sources: First, the overlap within the RI basis itself. For example, if multiple very diffuse functions are present, small eigenvalues appear in S because the norm of the difference between two adjacently centered diffuse functions is small compared the norm of the functions themselves (=one). But this does not necessarily mean that one of them can be discarded easily. The difference may be important (e.g., for polarization in outer molecule regions). On the other hand, we also get small overlap eigenvalues because the initial RI basis may have

a large overlap with the regular AO basis, and this overlap is projected out[4]. As a result, some of the initial (normalized) RI functions are diminished and acquire a small norm.

I believe that if we have to delete functions, then we should rather delete functions which are troublesome due to their redundancy with the AO space rather than due to diffuseness. But in the direct CABS construction approach both aspects cannot be separated: We have to delete functions corresponding to small $[\boldsymbol{S}]_{\bar{\eta}\bar{\chi}}$ eigenvalues, regardless of the reason which made it small. In fact, both aspects modulate each other, so a direct construction is biased towards deleting diffuse RI functions with moderate AO basis overlap.

An alternative is to first construct an orthonormal auxiliary basis (ABS) out of the RI basis, and then construct the CABS by projecting out MOs from the ABS instead of the RI basis directly. In this case, typically no functions need to be deleted in the ABS construction, and the AO/RI redundancy issue is isolated in the second part. While results are inconclusive, the second approach appeared at least as good, and sometimes superior to the direct construction in nearly all trials where both have been tested. Apart from using ABS vectors instead of direct RI functions in the CABS construction, the construction recipe is unchanged.

3.7 F12 generator orbitals in the open-shell case

In Sec. 3.2 we explained that the F12 configurations are formed by multiplying F_{12} correlation factors on occupied orbital pairs $|mn\rangle$, and Sec. 3.3 argued that the spatial symmetry of the resulting pair functions plays a crucial role for the electron cusp conditions. In the genuine open-shell case, these symmetry restrictions create some subtle issues; simply transcribing the closed-shell formulas into spin-orbitals results in a wave function ansatz that is incompatible with the cusp conditions.

This problem is isolated in the AB-spin $F_{12}|mn\rangle$ configurations, which are controlled by T^{ij}_{mn} amplitudes [cf. (3.4)]. In the closed-shell case and for AA and BB-spin, the spatially symmetric and anti-symmetric combinations $\frac{1}{2}(|mn\rangle \pm |nm\rangle)$ can readily be formed by linearly combining F12 configurations within the manifest ansatz. As in eq. (3.25), one can then multiply $t_p F_{12}$ factors to the results to fulfill the cusp conditions. On the other hand, for the AB-spin T^{ij}_{mn} amplitudes, clearly i and m must have A-spin and j and n must have B-spin [cf. Sec. 2.12]. But if m and n are restricted to be occupied orbitals of those spins, then there are less n orbitals than m orbitals, and consequently the (anti-)symmetrized mn combinations

[4] In principle, the AO/RI overlap issue can be avoided by using dedicated RI basis sets which do not have significant overlap with the AO basis in the first place. But such dedicated sets are not widely available at this moment, thus more general basis sets, like density-fitting auxiliary sets, are commonly used.

of Sec. 3.3 cannot be formed. In particular, for an active orbital i and a closed-shell orbital j, the direct diagonal excitation $t_{i_A j_B}^{i_A j_B}$ is alright, but the reversed diagonal excitation $t_{j_A i_B}^{i_A j_B}$ refers to the *unoccupied* spin-orbital i_B.

The problem can be solved by extending the ansatz for the F12 wave function. We include additional orbital pairs $|mn\rangle$ in which some of the m,n refer to unoccupied spin orbitals. The new F12-part of the double excitation operator is [cf. (3.4),(3.5)]

$$T_{\mathrm{F12}} = \frac{1}{2} T_{mn}^{ij} \mathcal{F}_{\alpha\beta}^{mn} e_{ij}^{\alpha\beta}, \tag{3.36}$$

where $e_{ij}^{\alpha\beta}$ is a spin-orbital excitation operator, i and j run over occupied spin-orbitals as usual, while for *all* spins m and n run over *both* closed-shell and active orbitals—including B-spin, where active spatial orbitals are unoccupied. The ansatz extension does not pose orthogonality problems, because the projectors Q_{12} still project out the overlap of F12 configurations with the conventional space. In principle one is free in the choice of the orbitals used for generating the F12 excitation manifold. The current choice means that F12 generator orbitals m, n, o, p run over all spin-orbitals which can be formed by multiplying A- or B- spin functions to any partially occupied spatial orbital, while i, j, k, l still run only over occupied spin-orbitals.

With this extended wave function ansatz, the fixed amplitude ansatz [Sec. 3.4] is not a problem anymore. The spin-orbital T_{mn}^{ij} of occupied spin-orbital pairs ij are fixed according to eq. (3.23), where $t_s = 0$ for AA and BB spin.

This ansatz extension is supposedly equivalent [Ten-no, private communication, 20.8.2009] to the one mentioned by Bokhan et al. in Ref. [104] in the context of spin-unrestricted orbitals, and it was already indicated by Ten-no in Ref. [35]. While, in retrospect, I agree that it was indicated in Ref. [35], I still find myself unable to follow the explanations of Ref. [104]. Addressing these matters in the context of spin degrees of freedom ("singlet" pairs, "triplet" conditions, "spin flipped orbitals" etc.), makes them much more confusing than necessary. According to my current understanding [Sec. 3.3], the cusp conditions actually depend *only* on *spatial symmetry* properties of the wave function, and are related to spin indirectly at best. The fact that clean singlet- and triplet pairs cannot be formed for AB pairs, is therefore *not relevant*.

3.8 Choice of Q_{12} projectors in the open-shell case

By default, we used spin-orbital projectors Q_{12}. That means in $Q_{12} = (1 - o_1)(1 - o_2)(1 - v_1 v_2)$ the projectors $o = |i\rangle\langle i|$ and $v = |a\rangle\langle a|$ run over the spin-orbitals occupied or unoccupied in the respective spin case (A,B).

Another possibility is to use projectors Q_{12} with identical spatial parts for all spins. Then for both A and B spin, in $o = |i\rangle\langle i|$ i runs over both closed- and active orbitals

and in $v = |a\rangle\langle a|$ a runs over pure external orbitals. As a result, the F12 wave function would be spin-free if fixed amplitudes were used, and the semi-internal contributions of all F12 pair functions would be removed (i.e., the overlap with virtual spin-orbital pairs involving active spatial orbitals is projected out).

Experiments indicated that the effect of these changes is only significant for double-zeta (aug-cc-pVDZ) basis sets. RMP2-F12 benchmarks showed better convergence when the semi-internal contributions were not removed. Thus, the spin-orbital spin-specific projectors are employed throughout. The spin-unspecific projector, in which $(1 - o)$ removes all overlap with any partially occupied spatial orbital, might be more useful in multi-reference theories. There it offers a simple way to eliminate overlap of the F12 wave function with the (very complicated) reference function.

3.9 Other possible wave functions

This section discusses variations of the F12 wave function ansatz described above which are relevant if other quantities than ground state energies are considered, or if theories with higher excitations than double excitations are used. These approaches were not followed in the current work, but are mentioned in order to display the limits of the current approach, and how they can be circumvented.

3.9.1 $F_{12}\,|ia\rangle$ configurations for excited states

It is possible to obtain fairly accurate vertical excitation energies in a non-state-specific manner. The typical way of doing this (e.g., Refs. [105], [62]) is based on linear response functions of a ground state wave function. Basically, one starts with the time dependent Schrödinger equation

$$[H + v(t)]\,|\Psi\rangle = i\hbar\partial_t\,|\Psi\rangle\,,$$

in which a time-dependent (adiabatically switched) electric field perturbation of form

$$v(t) = v_\omega \exp(-i\omega t + \alpha) \quad \text{with} \quad \alpha \to +0$$

is added to the regular time-independent Hamiltonian H. Then the equation is Fourier-transformed from the time domain to the energy/frequency domain, and the resonance frequencies are determined by finding the frequencies ω which make the system singular. In practice this amounts to solving a generalized eigenvalue equation based on the coupled cluster Jacobian [106].

When Fliegl et al. followed this standard recipe for the CC2(R12) method in Ref. [107], they found that the basis set convergence of excitation energies was actually not improved compared to conventional calculations. The problem was

identified to lie in the ansatz for the F12 wave function: A wave function of form (3.5) is strongly biased to stabilize the reference wave function (i.e., the ground state) compared to the excited states, because the correlation holes are only formed by applying F12 factors to *occupied* orbital products. A subsequent analysis by Neiss and coworkers [45] showed that the basis set convergence of excitations energies can be strongly improved by including F12 configurations which are generated by virtual orbitals. In particular, in eqs. (3.5) and (3.6), the range of the generator orbitals *mn* needs to be extended such that non-occupied orbitals are also included. For example, a F12 ansatz of the form

$$\widehat{T}_{\mathrm{F12}} = \frac{1}{2} T^{ij}_{rn} \mathcal{F}^{rn}_{\alpha\beta} E^{\alpha\beta}_{ij} \quad \text{or} \quad \widehat{T}_{\mathrm{F12}} = \frac{1}{2} T^{ij}_{rs} \mathcal{F}^{rs}_{\alpha\beta} E^{\alpha\beta}_{ij}, \tag{3.37}$$

where r, s run over both occupied and (some) virtual orbitals, should be sufficient to stabilize the singly excited states that are well approximated by some singly excited $E^a_i |\Phi\rangle$ determinant (left) or also doubly excited ones (right). Unfortunately, such extensions make the F12 methods much more expensive if all virtual orbitals are included. Neiss et al. [45] recommended the inclusion of a small number of pseudo-natural virtual orbitals with high occupation numbers in MP2 pseudo-density matrices.

Another reasonable choice would be to let both *m* and *n* run not only over occupied orbitals, but over the full valence space (i.e., over a set of orbitals with a linear span roughly equivalent to that of a minimal basis set, apart from core orbitals). While this approach certainly cannot describe Rydberg states, for valence excited states it might be sufficient. To my knowledge, this has not been investigated.

As shown by Köhn [46, 108], variants of the fixed amplitude ansatz (Sec. 3.3) can also be applied in response calculations. In this case [46, 108], the doubles wave function ansatz (eq. (3.4)) is augmented by double excitations of the form

$$T_{\mathrm{fix,ph}} \equiv \sum_{p=\pm 1} t_p t^j_c \langle \alpha\beta \,|\, Q_{12} F_{12} \,|\, \mathcal{S}_p ic \rangle E^{\alpha\beta}_{ij},$$

where t^j_c are the T_1 singles amplitudes and t_p and $\mathcal{S}$ are given by (3.23)ff. While this ansatz is only suitable for singly excited states, this is not such a big problem because for most common single-reference wave function methods (e.g., LR-CC2 or LR-CCSD/EOM-CCSD), the performance for doubly excited states is anyway rather unimpressive.

In passing I would like to mention that the extensions discussed in this section are non-issues for state specific calculations, in which the reference function $|\Phi\rangle$ actually describes the excited state. If such calculations are made, excited states can also be handled well within the "ground state" formalism discussed in the rest of this text. The main problems here are that (i) one calculation must be done for each

excited state, (ii) it may problematic to find out which excited states are relevant, or to obtain reference functions for them, and (iii) excited states often are not well described in a single-determinant picture—in particular, they may require low-spin reference functions. Still, there is a sizable number of cases where (ii) can be solved and (iii) does not apply, and then the single-reference open-shell methods work fine.

3.9.2 2+1 triply excited F12 configurations

In order to make single reference wave function methods really accurate, it is necessary to account for triply excited trial configurations in some way. For example, the popular CCSD(T) method is very accurate, while for many properties CCSD itself is not a huge improvement over MP2.

In absolute terms, the (T) correction is usually one order of magnitude smaller than the doubles correlation energy, and therefore its basis set error is also small. Nevertheless, after double substitutions and Hartree-Fock energies have been accounted for with F12 techniques, in practical calculations the basis set dependence of (T) often forms a major part of the residual basis set error [44, 103].

This also applies to higher excitations. Shiozaki et al. [109, 110] implemented a full CCSDTQ-F12 method, but included only F12 configurations of the doubly excited form (3.5). Their calculations revealed that the basis set convergence with respect to triply- and higher excited configurations is not improved by such wave functions.

As pointed out by Köhn [44], the main reason for this is not that simultaneous triple electron coalescence is not handled by the ansatz. Rather, the combination of conventional single excitations with simultaneous F12 double excitations need to be accounted for, such that F12 triple excitations of the general form

$$T^{ijk}_{mnc}\mathcal{F}^{mn}_{\alpha\beta}E^{\alpha\beta c}_{ijk}\,|\Phi\rangle \tag{3.38}$$

(and the related permuted quantities) need to be included, where T^{ijk}_{mnc} are the F12 triples amplitudes. Köhn showed [44] that it is possible to derive fixed amplitude ansaetze also for the case of triple excitations, and that such kinds of higher excitations significantly improve the basis set convergence of the (T) correction.

4 Method Developments

4.1 Second-Order Perturbation Theory: RMP2-F12

Second-order Rayleigh-Schrödinger perturbation theory (RSPT) is the simplest way of addressing the electron correlation problem if the Fock operator f is used as H_0. Such methods are generally called Møller-Plesset perturbation theory (MPn) or "many body perturbation theory" (MBPT). Both terms are synonymous.

The theory of open-shell MPn methods is transparent if a spin-orbital formalism and unrestricted Hartree-Fock (UHF) reference wave functions are employed (MPn is then called UMPn). In this case no substantial differences to the closed-shell case occur. On the other hand, it is well known that UMP2 based on UHF reference functions may suffer severely from spin contamination, and the results of UMP2 treatments often cannot be trusted (see, e.g., Refs. [76–79, 111]). This problem can be avoided by using the RMP2 method of Knowles et al. [76] and Lauderdale et al. [112, 113], which is based on spin-restricted Hartree-Fock (RHF) reference functions.[1]

Here I describe the explicitly correlated version of this RMP2 method which I developed during the course of my graduate studies. The theory is presented in an orbital invariant form, but the possibility of closed form solutions is pointed out when present. The orbital invariant form allows the use of localized orbitals and simplifies the introduction of local approximations [47, 48] (which here were not implemented or tested, though). Sections 4.1.1 and 4.1.2 present the theory and discuss issues arising because ROHF reference functions violate the spin-orbital Brillouin conditions. Sections 4.1.3 to 4.1.10 deal with various aspects of the concrete implementation.

The new DF-RMP2-F12 method is modeled after closed-shell DF-MP2-F12 methods developed earlier in the Stuttgart group [36]. For technical reasons, however, there are separate programs for open-shell and closed-shell systems.

[1] The RMP2 method itself is actually an *unrestricted* correlation treatment based on spin-orbitals. But as mentioned in Sec. 2.9, a spin-clean *reference* function is sufficient to remove spin contamination nearly completely from the correlated calculation as well. Nevertheless, in analogy to the RHF-UCCSD method it most closely corresponds to, "RHF-UMP2" might have been a better naming than "RMP2".

4.1.1 Recapitulation of RMP2

RMP2 is a second-order Rayleigh-Schrödinger perturbation method as discussed in Sec. 2.10. However, it differs from Møller-Plesset perturbation theory: If an open-shell RHF reference wave function $|\Phi\rangle$ is employed, the Fock operator f cannot be used as H_0 directly. The reason is that $|\Phi\rangle$ then is not an eigenfunction of the Fock operator, thus violating eq. (2.44). This is a consequence of the constrained variation of the spin-orbitals described in Sec. 2.8, which leaves the spin-orbital Brillouin conditions [eq. (2.30)] unfulfilled.

RMP2 solves this problem by using a projected Fock operator g as H_0 for RSPT:

$$g = ofo + (1-o)f(1-o), \tag{4.1}$$

where $o = |i\rangle\langle i|$ is the projector onto the occupied spin-orbitals. Consequently, the operator g can be written as [eq. (2.12)]

$$g = f^i_j e^j_i + f^\alpha_\beta e^\beta_\alpha \tag{4.2}$$

in second-quantization. If applied to the reference determinant $|\Phi\rangle$, only the part $f^i_j e^j_i$ gives non-vanishing contributions, and of this term only the summands with $i = j$. The reason for this is that in $e_i|\Phi\rangle$ all spin-orbitals j are already occupied except for $j = i$, and thus $e^j(e_i|\Phi\rangle)$, $i \neq j$ vanishes. Consequently, we have

$$g|\Phi\rangle = f^i_j e^j_i |\Phi\rangle = \sum\nolimits_i f^i_i e^i_i |\Phi\rangle = \left(\sum\nolimits_i f^i_i\right)|\Phi\rangle = E_0|\Phi\rangle\,. \tag{4.3}$$

The reference function is an eigenvector of g, with the sum of occupied spin-orbital energies as eigenvalue E_0 (note that $E_0 \neq E_\mathrm{ref}$; see Sec. 2.10).

The RMP2 amplitude equations is obtained from eq. (2.47) by inserting the ansatz

$$|\Psi\rangle = |\Phi\rangle + |\Psi_1\rangle = \left(1 + t^k_\gamma e^\gamma_k + \tfrac{1}{2} T^{kl}_{\gamma\delta} e^{\gamma\delta}_{kl}\right)|\Phi\rangle \tag{4.4}$$

for the first-order wave function (all indices spin-orbitals) and projecting with $\Xi \in \{e^\alpha_i|\Phi\rangle, \tfrac{1}{2}e^{\alpha\beta}_{ij}|\Phi\rangle\}$:

$$2R^{ij}_{\alpha\beta} := \left\langle \Phi^{ij}_{\alpha\beta} \,\middle|\, H \,\middle|\, \Phi\right\rangle + \left\langle \Phi^{ij}_{\alpha\beta} \,\middle|\, g - E_0 \,\middle|\, \Psi_1\right\rangle \tag{4.5}$$

$$r^i_\alpha := \left\langle \Phi^i_\alpha \,\middle|\, H \,\middle|\, \Phi\right\rangle + \left\langle \Phi^i_\alpha \,\middle|\, g - E_0 \,\middle|\, \Psi_1\right\rangle. \tag{4.6}$$

By inserting the normal ordered form of the Hamiltonian we then obtain

$$\begin{aligned} 2R^{ij}_{\alpha\beta} = &\left\langle \Phi \,\middle|\, e^{ij}_{\alpha\beta}(E_0 + f^r_s e^s_r + \tfrac{1}{2}K^{rs}_{tu}e^{tu}_{rs}) \,\middle|\, \Phi\right\rangle + \\ &\left\langle \Phi \,\middle|\, e^{ij}_{\alpha\beta}(g^r_s e^s_r - E_0)\left(t^k_\gamma e^\gamma_k + \tfrac{1}{2}T^{kl}_{\gamma\delta}e^{\gamma\delta}_{kl}\right)\right\rangle. \end{aligned} \tag{4.7}$$

Employing the Wick theorem eq. (2.20)ff and simplifying the result, we get

$$R^{ij}_{\alpha\beta} = \frac{1}{2}(1-\pi_{ab})(K^{ij}_{\alpha\beta} - t^i_\beta g^\alpha_j + g^\alpha_i t^j_\beta) - g^i_k T^{kj}_{\alpha\beta} - T^{ik}_{\alpha\beta} g^j_k + g^\gamma_\alpha T^{ij}_{\gamma\beta} + T^{ij}_{\alpha\gamma} g^\gamma_\beta \tag{4.8}$$

$$= \frac{1}{2}(K^{ij}_{\alpha\beta} - K^{ij}_{\beta\alpha}) - g^i_k T^{kj}_{\alpha\beta} - T^{ik}_{\alpha\beta} g^j_k + g^\gamma_\alpha T^{ij}_{\gamma\beta} + T^{ij}_{\alpha\gamma} g^\gamma_\beta. \tag{4.9}$$

In the last step $g^i_\alpha = 0$ [due to eq. (4.1)] was used; note that g and f as defined above are equal for all terms remaining in (4.9). The singles projection equation gives:

$$r^i_\alpha = \left\langle \Phi^i_\alpha \middle| H \middle| \Phi \right\rangle + \left\langle \Phi^i_\alpha \middle| g - E_0 \middle| \Psi_1 \right\rangle \tag{4.10}$$

$$= f^\alpha_i - g^i_k t^k_\alpha + g^\gamma_\alpha t^i_\gamma + 2 g^\gamma_k T^{ik}_{\alpha\gamma} \tag{4.11}$$

$$= f^\alpha_i - g^i_k t^k_\alpha + g^\gamma_\alpha t^i_\gamma. \tag{4.12}$$

The resulting equations (4.9) and (4.12) look like their orbital invariant closed-shell counterparts. They are equivalent to the formulas given by Lauderdale et al. [112, 113], who extracted them from the CCSDT equations by removing unused terms.

The second-order energy can be obtained by

$$E^{(2)} = \left\langle \Psi^{(1)} \middle| H \middle| \Phi \right\rangle = t^i_\alpha f^i_\alpha + T^{ij}_{\alpha\beta} K^{ij}_{\alpha\beta} \tag{4.13}$$

(direct RSPT2 expression, eq. (2.49)) or by evaluating the Hylleraas functional (eq. (2.50))

$$E^{(2)} = \left\langle \Psi^{(1)} \middle| H^{(0)} - E^{(0)} \middle| \Psi^{(1)} \right\rangle + 2 \left\langle \Psi^{(1)} \middle| H \middle| \Phi \right\rangle$$

$$= t^i_\alpha \left(f^i_\alpha + r^i_\alpha \right) + T^{ij}_{\alpha\beta} \left(K^{ij}_{\alpha\beta} + R^{ij}_{\alpha\beta} \right). \tag{4.14}$$

As mentioned in Sec. 2.10, the latter variational expression is more stable with respect to inaccuracies in $\left|\Psi^{(1)}\right\rangle$, but both expressions are equivalent if (2.47) is exactly fulfilled within the given model space (here: if $R^{ij}_{\alpha\beta}$ and r^i_α are exactly zero). It is noteworthy that in eq. (4.13) spin contaminations of the first-order wave function $\langle\Psi_1|$ are projected out; the ROHF reference function $|\Phi\rangle$ is spin-clean and the Hamilton operator is spin-free. Spin contaminations—if present—can therefore only enter the energy indirectly, by influencing the spin-clean part of $|\Psi_1\rangle$. This is a significant difference to UHF based perturbation theories, where spin-contaminated parts of $|\Psi_1\rangle$ have immediate energy contributions.

Both Knowles et al. [76] and Lauderdale et al. [112, 113] pointed out that the RMP2 equations can be solved in closed form by introducing semi-canonical orbitals fulfilling

$$g \left| r \right\rangle = \varepsilon_r \left| r \right\rangle \tag{4.15}$$

in the spin-orbital sense (i.e., introducing different spatial parts for A- and B-spin orbitals). The second-order energy equation then becomes

$$E_2 = -\sum_{ijab} \frac{\frac{1}{2}\left(K^{ij}_{\alpha\beta} - K^{ij}_{\beta\alpha}\right)K^{ij}_{\alpha\beta}}{\varepsilon_\alpha + \varepsilon_\alpha - \varepsilon_i - \varepsilon_j} - \sum_{ia} \frac{\left(f^i_\alpha\right)^2}{\varepsilon_\alpha - \varepsilon_i}, \tag{4.16}$$

which, apart from featuring single-excitations and employing spin-orbitals, looks nearly exactly like the textbook MP2 formulas. In this sense RMP2 may be the most natural extension of MP2 for spin-restricted reference functions. Nevertheless, switching to a semi-canonical formulation at this point in time would have disadvantages in our case. In MP2-F12 methods the evaluation and transformation of integrals requires a large share of the total program runtime; and having equal spatial orbitals for all spin cases—in particular, for the occupied orbitals—allows for sharing various calculation intermediates across the different spin cases.

4.1.2 Derivation of the RMP2-F12 equations

Eq. (4.9) is the RMP2 doubles equation for a formally complete virtual basis set. But in reality we only have a finite number of related wave function parameters, namely T^{ij}_{ab} and T^{ij}_{mn} (see Sec. 3.2), so we need to constrain this equation to an adequate model space. This is done by projecting it onto the subspace spanned by the independent doubly excited configurations which are actually contained in the first order wave function. For the conventional excitations (expanded by T^{ij}_{ab}), we can simply substitute $\alpha \mapsto a, \beta \mapsto b$ on the lhs and rhs of (4.9), because the conventional virtual space (a,b,…) is a subspace of the complete virtual space (α,β,…). We get

$$R^{ij}_{ab} = \frac{1}{2}\left(K^{ij}_{ab} - K^{ij}_{ba}\right) - g^i_k T^{kj}_{ab} - T^{ik}_{ab} g^j_k + g^\gamma_a T^{ij}_{\gamma b} + T^{ij}_{a\gamma} g^\gamma_b \tag{4.17}$$

$$= \frac{1}{2}\left(K^{ij}_{ab} - K^{ij}_{ba}\right) - g^i_k T^{kj}_{ab} - T^{ik}_{ab} g^j_k + g^c_a T^{ij}_{cb} + T^{ij}_{ac} g^c_b + g^x_a T^{ij}_{xb} + T^{ij}_{ax} g^x_b. \tag{4.18}$$

In eq. (4.17) we cannot directly substitute $\gamma \mapsto c$, because there are non-vanishing contributions from the non-finite-basis part of the double excitations; these are represented by the last two terms of eq. (4.18). By inserting eq. (3.9) for these T^{ij}_{ax}/T^{ij}_{xb}, we obtain:

$$g^x_a T^{ij}_{xb} + T^{ij}_{ax} g^x_b = \left(g^x_a \mathcal{F}^{mn}_{xb} + \mathcal{F}^{mn}_{ax} g^x_b\right) T^{ij}_{mn} = \ldots,$$

and, by inserting the $\mathcal{F}$ matrix elements [eq. (3.6)],

$$\begin{aligned} \ldots &= \left(g^x_a \langle xb \,|\, Q_{12} F_{12} \,|\, mn\rangle + g^x_b \langle ax \,|\, Q_{12} F_{12} \,|\, mn\rangle\right) T^{ij}_{mn} \\ &= \langle ab \,|\, (g_1 + g_2) Q_{12} F_{12} \,|\, mn\rangle \, T^{ij}_{mn} =: C^{mn}_{ab} T^{ij}_{mn} \end{aligned} \tag{4.19}$$

with

$$C_{ab}^{mn} = \langle ab \,|\, (g_1 + g_2) Q_{12} F_{12} \,|\, mn \rangle \,. \tag{4.20}$$

Eq. (3.32) allowed moving the Fock operators into the bracket in eq. (4.19). Thus, the amplitude equation for the conventional part of the RMP2-F12 wave function is given by the conventional RMP2 amplitude equation, augmented by a single F12/conventional coupling term $C_{ab}^{mn} T_{mn}^{ij}$:

$$R_{ab}^{ij} = \frac{1}{2}\big(K_{ab}^{ij} - K_{ba}^{ij}\big) - g_k^i T_{ab}^{kj} - T_{ab}^{ik} g_k^j + g_a^c T_{cb}^{ij} + T_{ac}^{ij} g_b^c + C_{ab}^{mn} T_{mn}^{ij}. \tag{4.21}$$

For the F12 doubles equation, we contract $R_{\alpha\beta}^{ij}$ of eq. (4.9) with $\mathcal{F}_{\alpha\beta}^{mn}$ of eq. (3.6), thus projecting the equation onto manifold of independent F12 configurations:

$$\begin{aligned} R_{mn}^{ij} &= \mathcal{F}_{\alpha\beta}^{mn} R_{\alpha\beta}^{ij} \\ &= \mathcal{F}_{\alpha\beta}^{mn} \left(\tfrac{1}{2}(K_{\alpha\beta}^{ij} - K_{\beta\alpha}^{ij}) - g_k^i T_{\alpha\beta}^{kj} - T_{\alpha\beta}^{ik} g_k^j + g_\alpha^\gamma T_{\gamma\beta}^{ij} + T_{\alpha\gamma}^{ij} g_\beta^\gamma \right). \end{aligned} \tag{4.22}$$

We investigate the individual terms. Inserting the definition of $\mathcal{F}_{\alpha\beta}^{mn}$, and using $Q_{12} \,|\alpha\beta\rangle\, \langle\alpha\beta| = Q_{12}$ [eq. (3.29)], we get:

$$\begin{aligned} \mathcal{F}_{\alpha\beta}^{mn} K_{\alpha\beta}^{ij} &= \langle mn \,|\, F_{12} Q_{12} \,|\, \alpha\beta \rangle \left\langle \alpha\beta \,\middle|\, r_{12}^{-1} \,\middle|\, ij \right\rangle \\ &= \left\langle mn \,\middle|\, F_{12} Q_{12} r_{12}^{-1} \,\middle|\, ij \right\rangle =: V_{mn}^{ij}. \end{aligned} \tag{4.23}$$

The terms with direct $\mathcal{F}_{\alpha\beta}$–$T_{\alpha\beta}$ contractions are resolved using eq. (3.9):

$$\mathcal{F}_{\alpha\beta}^{mn} T_{\alpha\beta}^{kj} = \mathcal{F}_{\alpha\beta}^{mn} \left(\delta_\alpha^a \delta_\beta^b T_{ab}^{kj} + \mathcal{F}_{\alpha\beta}^{op} T_{op}^{kj} \right) = \ldots,$$

which is simplified with $Q_{12}^2 = Q_{12}$ [eq. (3.28)]:

$$\begin{aligned} \ldots &= \underbrace{\langle mn \,|\, F_{12} Q_{12} \,|\, ab \rangle}_{Q_{12}|ab\rangle = 0} T_{ab}^{kj} + \langle mn \,|\, F_{12} Q_{12} \,|\, \alpha\beta \rangle \, \langle \alpha\beta \,|\, Q_{12} F_{12} \,|\, op \rangle \, T_{op}^{kj} \\ &= \langle mn \,|\, F_{12} Q_{12} F_{12} \,|\, op \rangle \, T_{op}^{kj} =: X_{mn,op} T_{op}^{kj}. \end{aligned} \tag{4.24}$$

For the first $\mathcal{F}$–g–T contractions we obtain

$$\begin{aligned} \mathcal{F}_{\alpha\beta}^{mn} g_\alpha^\gamma T_{\gamma\beta}^{ij} &= \mathcal{F}_{\alpha\beta}^{mn} g_\alpha^\gamma (\delta_\gamma^c \delta_\beta^b T_{cb}^{ij} + \mathcal{F}_{\gamma\beta}^{op} T_{op}^{ij}) \\ &= \mathcal{F}_{\alpha\beta}^{mn} g_\alpha^c T_{cb}^{ij} + \mathcal{F}_{\alpha\beta}^{mn} g_\alpha^\gamma \mathcal{F}_{\gamma\beta}^{op} T_{op}^{ij} = \ldots \end{aligned}$$

and, again using eq. (3.29),

$$\begin{aligned} \ldots &= \langle mn \,|\, F_{12} Q_{12} g_1 \,|\, cb \rangle \, T_{cb}^{ij} + \langle mn \,|\, F_{12} Q_{12} g_1 \,|\, \gamma\beta \rangle \, \langle \gamma\beta \,|\, Q_{12} F_{12} \,|\, op \rangle \, T_{op}^{ij} \\ &= \langle mn \,|\, F_{12} Q_{12} g_1 \,|\, cb \rangle \, T_{cb}^{ij} + \langle mn \,|\, F_{12} Q_{12} g_1 Q_{12} F_{12} \,|\, op \rangle \, T_{op}^{ij}. \end{aligned} \tag{4.25}$$

A similar matrix element, in which g acts on the second electron instead of the first electron, is generated by the $g^{\gamma}_{\beta} T^{ij}_{\alpha\gamma}$ term in eq. (4.22). As sum of both we obtain

$$\mathcal{F}^{mn}_{\alpha\beta} \left(g^{\gamma}_{\alpha} T^{ij}_{\gamma\beta} + T^{ij}_{\alpha\gamma} g^{\gamma}_{\beta} \right) = C^{mn}_{ab} T^{ij}_{ab} + B_{mn,op} T^{ij}_{op} \tag{4.26}$$

with C^{mn}_{ab} from eq. (4.20) and

$$B_{mn,op} := \langle mn \,|\, F_{12} Q_{12} (g_1 + g_2) Q_{12} F_{12} \,|\, op \rangle \,. \tag{4.27}$$

After collecting terms, the total F12 residual equation eq. (4.22) turns into:

$$R^{ij}_{mn} = \tfrac{1}{2}(V^{ij}_{mn} - V^{ij}_{nm}) - X_{mn,op} \left(g^{i}_{k} T^{kj}_{op} + T^{ik}_{op} g^{j}_{k} \right) + B_{mn,op} T^{ij}_{op} + C^{mn}_{ab} T^{ij}_{ab}. \tag{4.28}$$

The conventional singles equation (4.12) is unaffected by the F12 configurations if the Fock operator is projected according to eq. (4.1), because then it does not couple to doubles at all. On the other hand, it formally refers to the complete virtual orbital space (α), and simply restricting it to the conventional space would imply neglecting the inhomogeneous matrix elements f^{i}_{x}, and thus imply GBC [Sec. 4.1.5]. This issue is discussed in Sec. 4.2.

For convenience, the combined formulas for the RMP2-F12 equations and F12 matrices are compiled in Tab. 4.1 (page 67). The evaluation of the matrix elements is discussed in the next two sections.

4.1.3 Formal calculation of F12 matrix elements

There are multiple ways of evaluating the V, X, B and C matrix elements derived in Sec. 4.1.2. My program implements the rigorous ansatz "3C" [36]. This ansatz is characterized by not neglecting any terms, and applying analytical integral evaluation techniques where it is due. The equivalent MP2-R12/C method for closed-shell systems was defined in Ref. [114], but the derivation and the working equations presented here are much simpler. The derivations follow the closed-shell method of [36], but several details were changed due to the projection of the H_0-Hamiltonian.

We investigate the matrix elements one by one. For the F12 configuration overlap

$$X_{mn,op} = \langle mn \,|\, F_{12} Q_{12} F_{12} \,|\, op \rangle$$

we obtain closed formulas directly by inserting eq. (3.30) for Q_{12}. This yields

$$\begin{aligned} X_{mn,op} &= \left\langle mn \,\middle|\, F_{12} (1 - |ix\rangle\langle ix| - |xi\rangle\langle xi| - |rs\rangle\langle rs|) F_{12} \,\middle|\, op \right\rangle \\ &= \langle mn \,|\, F_{12} F_{12} \,|\, op \rangle - F^{mn}_{ix} F^{op}_{ix} - F^{mn}_{xi} F^{op}_{xi} - F^{mn}_{rs} F^{op}_{rs} \end{aligned} \tag{4.29}$$

where $F^{mn}_{pq} = \langle mn \,|\, F_{12} \,|\, uv \rangle$ (u, v: arbitrary orbitals). The occurring F_{12} and F^2_{12}

integrals can be evaluated analytically using the techniques mentioned in Sec. 2.4.

For the Coulomb matrix V_{mn}^{kl}, we get completely analogously

$$\begin{aligned} V_{mn}^{kl} &= \left\langle mn \,\middle|\, F_{12}Q_{12}r_{12}^{-1} \,\middle|\, kl \right\rangle \\ &= \left\langle mn \,\middle|\, F_{12}r_{12}^{-1} \,\middle|\, kl \right\rangle - F_{ix}^{mn}K_{ix}^{kl} - F_{xi}^{mn}K_{xi}^{kl} - F_{rs}^{mn}K_{rs}^{kl}, \end{aligned} \tag{4.30}$$

where $K_{uv}^{kl} = \left\langle kl \,\middle|\, r_{12}^{-1} \,\middle|\, uv \right\rangle$ are the matrix elements of the Coulomb interaction. A difference between the two is that V_{mn}^{kl} has both pair labels (kl) and F12 generator labels (mn), while $X_{mn,op}$ has only F12 generator labels. As pointed out in Sec. 3.7, under certain circumstances the generator space is not equal to the space of occupied orbitals. This also affects the diagonal ansatz of the F12 wave function: While V_{mn}^{kl} becomes restricted to the diagonal form, $X_{mn,op}$ is still dense when the 3C(D) or 3C(FIX) form of the wave function is employed.

As next term we address the C matrix elements, which represent mean field couplings between F12- and conventional configurations. Employing form eq. (3.32) of Q_{12} we get

$$\begin{aligned} C_{ab}^{mn} &= \langle ab \,|\, (g_1 + g_2)Q_{12}F_{12} \,|\, mn \rangle \\ &= \langle ab \,|\, (g_1 + g_2)(|cx\rangle\langle cx| + |xc\rangle\langle xc| + |xy\rangle\langle xy|)F_{12} \,|\, mn \rangle \\ &= \langle ab \,|\, g_1 \,|xc\rangle\langle xc|\, F_{12} \,|\, mn \rangle + \langle ab \,|\, g_2 \,|cx\rangle\langle cx|\, F_{12} \,|\, mn \rangle \,. \end{aligned} \tag{4.31}$$

Because g_1/g_2 act only on the first/second electron and act as identity element on the other one, we have $\langle ab \,|\, g_1 \,|\, xc \rangle = \delta_{bc} \langle a \,|\, g \,|\, x \rangle$, and thus

$$\begin{aligned} C_{ab}^{mn} &= \langle a \,|\, g \,|\, x \rangle \langle xb \,|\, F_{12} \,|\, mn \rangle + \langle b \,|\, g \,|\, x \rangle \langle ax \,|\, F_{12} \,|\, mn \rangle \\ &= g_{ax}F_{xb}^{mn} + F_{ax}^{mn}g_{xb}. \end{aligned} \tag{4.32}$$

Like in X, C^{mn}s upper labels (mn) denote F12 generator orbitals, not (ij) pairs.

The evaluation of

$$B_{mn,op} = \langle mn \,|\, F_{12}Q_{12}(g_1 + g_2)Q_{12}F_{12} \,|\, op \rangle \tag{4.33}$$

is more complicated. In order to calculate $B_{mn,op}$, we first notice that

$$Q_{12}g_{12}Q_{12} = (1 - P_{12})g_{12}Q_{12} = g_{12}Q_{12} - P_{12}g_{12}Q_{12}. \tag{4.34}$$

If we do not evaluate the operators Q_{12} and P_{12} exactly, the approximations we introduce might break Bs symmetry. In order to avoid that, we explicitly symmetrize B with respect to the bra- and ket labels using the symmetrizing operation

$$\mathcal{S}B_{mn,op} = \frac{1}{2}(B_{mn,op} + B_{op,mn}). \tag{4.35}$$

In this sense we have

$$B_{mn,op} = \left\langle mn \middle| F_{12}\left[\mathcal{S}\left(g_{12}Q_{12} - P_{12}g_{12}Q_{12}\right)\right]F_{12}\middle| op\right\rangle \tag{4.36}$$

$$= \mathcal{S}\left[A_{mn,op} - Z_{mn,op} - F^{mn}_{ab}C^{op}_{ab}\right] \tag{4.37}$$

$$A_{mn,op} = \left\langle mn \,\middle|\, F_{12}g_{12}Q_{12}F_{12} \,\middle|\, op\right\rangle \tag{4.38}$$

$$Z_{mn,op} = \left\langle mn \,\middle|\, F_{12}\left(o_1g_1 + o_2g_2\right)Q_{12}F_{12} \,\middle|\, op\right\rangle. \tag{4.39}$$

The terms $Z_{mn,op}$ and $F^{mn}_{ab}C^{op}_{ab}$ arise from inserting

$$\begin{aligned} P_{12} &= 1 - Q_{12} = 1 - (1 - o_1)(1 - o_2)(1 - v_1v_2) \\ &= o_1 + o_2 - o_1o_2 + v_1v_2 \end{aligned} \tag{4.40}$$

(where v is the projector onto the virtual spin-orbitals) into eq. (4.36):

$$F_{12}P_{12}g_{12}Q_{12}F_{12} = F_{12}(o_1 + o_2 - o_1o_2 + v_1v_2)(g_1 + g_2)Q_{12}F_{12} \tag{4.41}$$

$$= F_{12}(o_1g_1 + o_2g_2 + v_1v_2g_{12})Q_{12}F_{12}. \tag{4.42}$$

When going from (4.41) to (4.42) various terms vanish because Q_{12} projects out any overlap with occupied spaces (e.g., $o_2g_1Q_{12} = g_1o_2Q_{12} = 0$).

Due to the projected form of g [cf. eq. (4.1)], we have og=ogo and thus $Z_{mn,op} = 0$ since Q_{12} projects out orbital pairs containing occupied orbitals [cf. eq. (3.7)]. This is a difference to the method originally denoted as MP2-F12/3C in [36], where Z needed to be neglected or evaluated and included explicitly. In the current method this is not necessary because the effect of Z is handled implicitly as perturbation by using the projected Fock operator g as H_0 for RSPT. The resulting total formulas were shown to be equal to that of [36]'s MP2-F12/3C(+Z) method for the closed-shell case by H.-J. Werner. Therefore the projection technique is now also employed in the closed shell program.

For calculating the A-term [eq. (4.38)] we split g into $g = h - n$ with

$$h = t + v_{\text{nuc}} + j \tag{4.43}$$

$$n = h - g = k - (2ofo - of - fo) =: k - \Delta g. \tag{4.44}$$

The Hartree operator h contains the kinetic energy operator t and the parts of g which commute with F_{12} (here Coulomb field j and nuclear potential v_{nuc}), while n contains the entire rest (here $\Delta g = f - g$ and exchange k). The A-terms are then evaluated in the following way:

$$\begin{aligned} A_{mn,op} &= \left\langle mn \,\middle|\, F_{12}(h_{12} - n_{12})Q_{12}F_{12} \,\middle|\, op\right\rangle \\ &= \left\langle mn \,\middle|\, \left[F_{12}h_{12} - h_{12}P_{12} - n_{12}Q_{12}\right]F_{12} \,\middle|\, op\right\rangle = \ldots \end{aligned} \tag{4.45}$$

If we now interchange F_{12} and h_{12} in the first term, only one additional term arises because all parts of h except t commute with F_{12} [eq. (4.43)]:

$$\begin{aligned}\ldots = &\left\langle mn \,\middle|\, h_{12}F_{12}F_{12} \,\middle|\, op\right\rangle - \left\langle mn \,\middle|\, [t_{12}, F_{12}]\, F_{12} \,\middle|\, op\right\rangle \\ &- \left\langle mn \,\middle|\, F_{12}h_{12}P_{12}F_{12} \,\middle|\, op\right\rangle - \left\langle mn \,\middle|\, F_{12}n_{12}Q_{12}F_{12} \,\middle|\, op\right\rangle. \end{aligned} \tag{4.46}$$

This commutator trick avoids using a double RI for the term involving the kinetic energy operator, which would be slowly convergent in partial wave expansions [19, 20]. For the remaining terms in expression eq. (4.46), we expand $\langle mn| h_{12}$ (first term) in the combined AO and RI basis, and apply a straight RI [eq. (3.30)] for P_{12}. We then obtain

$$\begin{aligned}A_{mn,op} = &F^2_{\bar{m}n,op} + F^2_{m\bar{n},op} + U^F_{mn,op} - \bar{F}^{mn}_{rs}F^{op}_{rs} \\ &- \bar{F}^{mn}_{ix}F^{op}_{ix} - \bar{F}^{mn}_{xi}F^{op}_{xi} - \widetilde{Y}_{mn,op}\end{aligned} \tag{4.47}$$

where

$$U^F_{mn,op} = \langle mn\,|\,[F_{12}, t_{12}]F_{12}\,|\,op\rangle \tag{4.48}$$

$$F^2_{mn,op} = \langle mn\,|\,F_{12}F_{12}\,|\,op\rangle \tag{4.49}$$

are integrals that can be evaluated analytically (for U^F: after symmetrizing),

$$\bar{F}^{mn}_{uv} = F^{mn}_{\bar{u}v} + F^{mn}_{u\bar{v}} \tag{4.50}$$

(u, v: arbitrary orbitals) represents F_{12} acting on h_{12}-transformed mn orbitals, which are expanded in the combined AO and RI basis in practice:

$$|\bar{k}\rangle = h\,|k\rangle\,, \tag{4.51}$$

and, finally, a residual exchange term

$$\widetilde{Y}_{mn,op} = \left\langle mn \,\middle|\, F_{12}n_{12}Q_{12}F_{12} \,\middle|\, op\right\rangle. \tag{4.52}$$

This term $\widetilde{Y}_{mn,op}$, which contains the exchange contribution and Δg [eq. (4.44)f], can be approximated by a straight double-RI using eq. (3.32) for the projector:

$$\widetilde{Y}_{mn,op} = \widetilde{F}^{mn}_{ax}F^{op}_{ax} + \widetilde{F}^{mn}_{xa}F^{op}_{xa} + \widetilde{F}^{mn}_{xy}F^{op}_{xy} \tag{4.53}$$

where $\widetilde{F}$ refers to F_{12} acting on n_{12}-transformed orbitals:

$$|\tilde{k}\rangle = n\,|k\rangle \tag{4.54}$$

$$\widetilde{F}^{mn}_{pq} = F^{mn}_{\tilde{p}q} + F^{mn}_{p\tilde{q}}. \tag{4.55}$$

Again, the combined AO and RI basis is used to expand $|\tilde{k}\rangle$. In the atomic case this double-RI is known to converge very fast for the exchange term. Despite containing kinetic energy contributions, also the additional Δg-term does not pose a problem because of its specific form (4.44): All contributions except of are projected out by Q_{12}, and for the remaining expression the formal double-RI is in fact only a single RI.

In passing I would like to note that the $\widetilde{Y}_{mn,op}$ term defined in this work slightly differs from the $\widetilde{Y}_{mn,op}$ term of Ref. [36]. This is a result from the Fock-operator projection employed here. If used in the closed shell case, the projected 3C method suggested here reproduces the results of the unprojected 3C(+Z) method reported in Ref. [36]. This happens because the (extremely small) effect of the $Z_{mn,op}$-term (which is not exactly zero without the projection (4.1)), is handled implicitly as part of the perturbation. All other matrices with equal names have equal meaning in the closed-shell case, since the minimally-projected form (4.1) of g then only affects GBC-terms.

One special problem of F12 methods—not only of open-shell methods—is that Fock operators which cannot be split according to eq. (4.44) need special handling. This applies when core Hamiltonian operators different from $t + v_{\mathrm{nuc}}$ are employed, for example in the following situations: (i) When external fields are added to h for finite field calculations (e.g., finite dipole fields to calculate dipole moments or dipole polarizabilities via the Hellman-Feynman theorem), (ii) When the effect of core electrons is absorbed in semi-local effective core potentials (ECPs), or (iii) When scalar-relativistic calculations are attempted which use a different kinetic energy operator (e.g., Douglas-Kroll schemes). This is not a serious problem as long as the core Hamiltonian modifications represent local potentials (e.g., multipole fields); they then can be simply added to the h part in eq. (4.44) and the theory resumes unchanged. If the modifications do not commute with F_{12}, then one is left with two options: (i) Try to resolve the modified core Hamiltonian patch in the combined AO/RI basis by adding it to the n part of eq. (4.44), and leave the theory unchanged. This appears to work sufficiently well for ECPs, for example. (ii) Handle the new term like the kinetic energy operator t, which means that additional commutator integrals $[F_{12}, \Delta h_{12}]F_{12}$ need to be programmed [cf. eq. (4.48)]. This might be necessary for scalar relativistic kinetic energy operators. Unlike their conventional counterparts, the F12 methods are not agnostic about the concrete version of the Hamiltonian applied.

4.1.4 Practical evaluation of F12 matrix elements

Sec. 4.1.3 featured a number of matrix elements (V,X,A,C,Y) which have to be evaluated from basic electronic integrals with non-trivial processes. How this was done in practice is described in an overview in this section. The premise is that

we have the orbital matrix $(\boldsymbol{C}) \equiv C_r^\mu$ and the two matrices $(\boldsymbol{R}) \equiv R_x^\mu$ and $(\bar{\boldsymbol{R}}) \equiv R_x^\chi$ representing `AoCabs` and `RiCabs` of the CABS [cf. Sec. 3.6]:

$$|r\rangle = C_r^\mu |\mu\rangle \qquad\qquad |x\rangle = R_x^\mu |\mu\rangle + R_x^\chi |\chi\rangle . \tag{4.56}$$

(Here and in the following, μ, ν denote AO-basis indices and χ, η denote RI basis indices.) Further, we need an integral driver capable of calculating two-index one-electron integrals (Fock, Exchange) and four-index two-electron integrals (all other F12 integrals) with respect to the AO and RI basis sets. We now need to devise a scheme of calculating the necessary integral transformations (e.g., for f_{xr}, (4.50), or (4.55)), the integrals based on that, and finally for assembling the target quantities V,X,A,C,Y.

A. Transforming from AO/RI to MO/CABS. In order to calculate a set of operators f_{rs}, f_{xs} and f_{xy}, which are specified with respect to the orthonormal MO/CABS basis sets, the series of $f_{\mu\nu}$, $f_{\chi\nu}$ and $f_{\chi\eta}$ with respect to the direct basis functions are required. It is assumed that the latter ones are calculated from integrals directly, and they thus have only covariant indices. We can then apply the following transformations into the orthonormal bases:

$$f_{rs} = C_r^\mu f_{\mu\nu} C_s^\nu \tag{4.57}$$

$$f_{xs} = (R_x^\mu f_{\mu\nu} + R_x^\chi f_{\chi\nu}) C_s^\nu , \tag{4.58}$$

and, for the CABS/CABS quantity:

$$f_{xy} = (R_x^\mu f_{\mu\nu} + R_x^\chi f_{\chi\nu}) R_y^\nu + (R_x^\mu f_{\mu\eta} + R_x^\chi f_{\chi\eta}) R_y^\eta . \tag{4.59}$$

The formulas would look simpler if we could merge the AO and RI basis sets, because the CABS vectors are expanded with respect to both. However, the actual programs did not support this (fully); also, the basis merge would remove some possibilities of CPU time saving (because some AO quantities are required for both CABS and MO separately), and some redundancies of symmetric quantities then could not be exploited.

These transformations need to be done for the A-spin and B-spin Fock operators and the Hartree operator [eq. (4.43)]. The projected Fock operators g [eq. (4.1)] can be obtained from the full Fock operators by setting all matrix elements g_{ix} and g_{ia} to zero in the orthogonal basis. The sum of exchange operators and Δg [eq. (4.44)f] can then be obtained as difference between g and the Hartree operator h.

B. Orbital transformations for h and n. The quantities $\bar{F}^{mn}$ [(4.50)] and $\widetilde{F}^{mn}$ [(4.55)] can be assembled directly from AO/RI quantities, provided the transformed orbitals from eqs. (4.51) and (4.54) are represented in a suitable manner. For this, we need matrices for h and n with one index MO or CA (designating the target domain)

and the other index contravariant AO or contravariant RI (which will be contracted to the covariant AO/RI indices of the to-be-calculated integrals). We can obtain such matrices from the completely orthonormal MO/CA sets from (*A*.) by just re-expanding one of the indices. For example

$$n_x{}^{\mu} = n_{xr} C_r^{\mu} + n_{xy} R_y^{\mu} \qquad\qquad n_x{}^{\chi} = n_{xy} R_y^{\chi} \tag{4.60}$$

will raise indices μ (left)/χ(right). Here the orbital and CA vector matrices [same as in eq. (4.56)] are used to implicitly represent the pseudo-inverse overlap matrix over the combined AO+RI basis sets. Note that the contractions in this expression run over orthonormal orbitals, for which no distinction between covariance and contravariance is necessary.

C. Calculation of integrals in the orthonormal basis. With the matrix representations of orbital vectors, CA vectors and transformed orbital vectors available, all the integrals occurring in Sec. 4.1.3 can now be formed. This is done by combining sets of AO/RI basis quantities just like in (*A*.), but this time in larger quantities and for more diverse target domains. For example, once we have evaluated a set of matrices $F_{\mu\nu}^{mn}$, $F_{\chi\nu}^{mn}$, and $F_{\chi\eta}^{mn}$ using density fitting techniques [cf. Sec. 2.4], from these all of the following sets can be assembled: F_{ax}^{mn} [for *C*, eq. (4.32)], F_{rs}^{mn}, F_{xi}^{mn} [for *X*,*V*,*A*,$C_{ab}^{mn} F_{ab}^{op}$ in *B*, eq. (4.37)]; $F_{\bar{r}s}^{mn}$, $F_{\bar{x}i}^{mn}$, $F_{x\bar{i}}^{mn}$ [for *A*, eq. (4.47)]; $F_{\bar{x}a}^{mn}$, $F_{x\tilde{a}}^{mn}$, F_{xy}^{mn}, $F_{\bar{x}\bar{y}}^{mn}$, [for $\widetilde{Y}$, eq. (4.53). *n*-transformed quantities are required for spin A and B separately]. For example, $F_{\bar{x}\bar{y}}^{mn}$ is assembled as

$$F_{\bar{x}\bar{y}}^{mn} = (n_x{}^{\mu} F_{\mu\nu}^{mn} + n_x{}^{\chi} F_{\chi\nu}^{mn}) R_y^{\nu} + (n_x{}^{\mu} F_{\mu\eta}^{mn} + n_x{}^{\chi} F_{\chi\eta}^{mn}) R_y^{\eta}. \tag{4.61}$$

D. Assembly of F12 matrices. Once the integrals in the orthonormal basis are available, the F12 matrices can be calculated. The non-trivial operations in here are contractions of F_{uv}^{mn} matrices (*uv*: *rs*, *xi*, *xa*, *xy*) to other matrices W_{uv}^{op} of compatible shape (e.g., K_{uv}^{op} for V_{mn}^{op}, F_{uv}^{op} for $X_{mn,op}$, $\bar{F}_{uv}^{mn}$ for *A*, $\widetilde{F}_{uv}^{mn}$ for *Y*):

$$\text{Target}_{mn,op} \mathrel{+}= F_{uv}^{mn} W_{uv}^{op}. \tag{4.62}$$

As the set of rhs matrices can become fairly large (larger than the available memory), it is necessary to calculate these incrementally using a paging algorithm. In this algorithm, the program leaves room for a small number ($\approx$ 16) of W^{op} matrices in memory, and uses the rest of the memory to load a maximal batch of F^{mn} matrices. With these loaded matrices, all F12 matrix increments depending on them are calculated by cycling all of the corresponding W^{op} matrices (K, F, $\widetilde{F}$,...) through the space left; and evaluating eq. (4.62) via a matrix multiplication over the *uv* space (if all *mn*,*op* are required) or matrix dot products over the *uv* space (if only $mn = op$ or $mn = po$ are required). While this is not a minimum-IO algorithm (in

principle, eq. (4.62) can be evaluated without any IO overhead if the *uv* labels instead of the *mn* labels are chosen as slow indices), it keeps the IO cost reasonable because typically at most a few passes over the W^{op} matrix sets are required. Compared to the evaluation of the transformed integrals themselves, the F12 matrix computation is usually cheap, especially if a diagonal wave function ansatz is used. In the latter case, paging is also not necessary, because of all F12 matrices except $X_{mn,op}$ only the diagonal elements are required (e.g., only B^{ij}_{mn} with $mn = ij$ or $mn = ji$).

4.1.5 Approximations to F12 matrix elements

Until now the theory was presented without any approximations beyond DF and RI. However, some terms from the full theory have minor impact on calculation results and can generally be omitted or approximated without much loss of accuracy.

The prime candidate for such a treatment is the $\widetilde{Y}_{mn,op}$ term of eq. (4.53). For larger molecules, its evaluation requires a large share of the total time for F12 integrals, but its effect on calculation results is small. As in the closed-shell theory [36], it is possible to drop the last term of (4.53) without significant loss of accuracy (approximation "3C(HY2)"). As $\Delta g_{xy} = 0$ this affects only the approximation of the exchange contributions. A more severe approximation is to neglect the matrix $\widetilde{Y}_{mn,op}$ entirely (approximation "3C(HY1)"); but as noted in Sec. 3.5, the terms represented by $\widetilde{Y}$ are well representable within the MO basis itself.[2] For these "hybrid approximations" the storage for the transformed integrals scales only linearly with the size of the RI-basis. The computational effort for some transformation steps still scales quadratically, but significant savings are possible, in particular for approximation HY1.

Unrelated to HYn, two other important simplifications are often employed: Assuming the applicability of the "Generalized Brillouin Condition" (GBC) or of the "Extended Brillouin Condition" (EBC). GBC states that the Fock operator does not connect occupied orbitals to CA orbitals:

$$f_{ix} = 0. \tag{4.63}$$

This relation actually holds for single atoms and highly saturated AO basis sets, but under conditions of realistic calculations it generally does not apply strictly. Nevertheless, the effect of assuming GBC is very small for most terms, and in most cases neglecting inconvenient GBC terms does not lead to noticeable relative energy differences [36]. A notable exception to this rule are the MP2-level singles

[2] This approximation is similar (but not identical) to Klopper's hybrid approximation [101]. In the latter work Ansatz 2 was used, which differs from the Ansatz 3 we use by omission of the factor $1 - v_1 v_2$ in the projector Q_{12} [cf. eq. (3.7)]. As discussed previously [36], the two ansätze lead to different results if hybrid approximations are introduced.

Table 4.1: Equations for RMP2-F12 residuals and intermediates. The equations are compiled from sections 4.1.2 to 4.1.4 and are given in terms of spin-orbitals.

Residual- and energy equations

$$
\begin{aligned}
r_\alpha^i &= f_\alpha^i - g_k^i t_\alpha^k + g_\alpha^\gamma t_\gamma^i \qquad (\alpha,\beta,\gamma \text{ virtual indices run over MO+CA}) \\
R_{ab}^{ij} &= \frac{1}{2}\left(K_{ab}^{ij} - K_{ba}^{ij}\right) - g_k^i T_{ab}^{kj} - T_{ab}^{ik} g_k^j + g_a^c T_{cb}^{ij} + T_{ac}^{ij} g_b^c + C_{ab}^{mn} T_{mn}^{ij} \\
R_{mn}^{ij} &= \frac{1}{2}\left(V_{mn}^{ij} - V_{nm}^{ij}\right) - X_{mn,op}\left(g_k^i T_{op}^{kj} + T_{op}^{ik} g_k^j\right) + B_{mn,op} T_{op}^{ij} + C_{ab}^{mn} T_{ab}^{ij} \\
E_{\mathrm{corr}} &= t_a^i \left(f_a^i + r_b^i\right) + T_{ab}^{ij}\left(K_{ab}^{ij} + R_{ab}^{ij}\right) + T_{mn}^{ij}\left(V_{mn}^{ij} + R_{mn}^{ij}\right)
\end{aligned}
$$

Definition and evaluation of F12 Matrices

$$
\begin{aligned}
C_{ab}^{mn} &= \langle ab \,|\, (g_1 + g_2) Q_{12} F_{12} \,|\, mn \rangle = g_{ax} F_{xb}^{mn} + F_{ax}^{mn} g_{xb} \\
X_{mn,op} &= \langle mn \,|\, F_{12} Q_{12} F_{12} \,|\, op \rangle \\
&= \langle mn \,|\, F_{12} F_{12} \,|\, op \rangle - F_{ix}^{mn} F_{ix}^{op} - F_{xi}^{mn} F_{xi}^{op} - F_{rs}^{mn} F_{rs}^{op} \\
V_{mn}^{kl} &= \left\langle mn \,\middle|\, F_{12} Q_{12} r_{12}^{-1} \,\middle|\, kl \right\rangle \\
&= \left\langle mn \,\middle|\, F_{12} r_{12}^{-1} \,\middle|\, kl \right\rangle - F_{ix}^{mn} K_{ix}^{kl} - F_{xi}^{mn} K_{xi}^{kl} - F_{rs}^{mn} K_{rs}^{kl} \\
B_{mn,op} &= \langle mn \,|\, F_{12} Q_{12} g_{12} Q_{12} F_{12} \,|\, op \rangle \\
&= \frac{1}{2}\left[A_{mn,op} + A_{op,mn} - F_{ab}^{mn} C_{ab}^{op} - F_{ab}^{op} C_{ab}^{mn}\right]
\end{aligned}
$$

Intermediates

$$
\begin{aligned}
h &= t + v_{\mathrm{nuc}} + j \qquad |\bar{k}\rangle = h\,|k\rangle \qquad |\tilde{k}\rangle = (g - h)\,|k\rangle \\
A_{mn,op} &= F_{\bar{m}n,op}^2 + F_{m\bar{n},op}^2 - \bar{F}_{ix}^{mn} F_{ix}^{op} - \bar{F}_{xi}^{mn} F_{xi}^{op} \\
&\qquad - \bar{F}_{rs}^{mn} F_{rs}^{op} + U_{mn,op}^F - \widetilde{Y}_{mn,op} \\
\widetilde{Y}_{mn,op} &= \widetilde{F}_{ax}^{mn} F_{ax}^{op} + \widetilde{F}_{xa}^{mn} F_{xa}^{op} + \widetilde{F}_{xy}^{mn} F_{xy}^{op} \\
U_{mn,op}^F &= \langle mn \,|\, [F_{12}, t_{12}] F_{12} \,|\, op \rangle \\
F_{mn,op}^2 &= \langle mn \,|\, F_{12} F_{12} \,|\, op \rangle \\
\bar{F}_{uv}^{mn} &= F_{\bar{u}v}^{mn} + F_{u\bar{v}}^{mn} \\
\widetilde{F}_{pq}^{mn} &= F_{\tilde{p}q}^{mn} + F_{p\tilde{q}}^{mn}
\end{aligned}
$$

terms, where not assuming GBC leads to a very important Hartree-Fock relaxation [cf. Sec. 4.2]. In this work, GBC is not applied for RMP2-F12, because it is not necessary. At the CCSD level, however, many terms are dropped; some of them would vanish under GBC.

The EBC states that the Fock operator also does not connect virtual orbitals to CA orbitals (i.e., that the MO space is closed under application of the Fock operator):

$$f_{ax} = 0. \tag{4.64}$$

As a consequence, the matrix elements C_{ab}^{mn} [cf. (4.20)] vanish and the F12 doubles decouple from the conventional doubles [cf. eqs. (4.21), (4.22)]. This condition is generally not fulfilled—not even approximately—and assuming it leads to sizable changes in total energies and, to a smaller extent, also in relative energies. The changes are not always for the worse, however. Methods employing the EBC approximation are denoted with a star in their method designations (e.g., "3*C" or "3*C(FIX)" denote MP2-F12 methods employing EBC). As under EBC $C_{ab}^{mn} = 0$, also the term $F_{ab}^{mn} C_{ab}^{op}$ is removed from eq. (4.37) in our programs.

4.1.6 Transformation to spatial orbitals

The previous equations were written in terms of spin orbitals. Since we use a RHF reference wave function, they can easily be adapted to restricted spatial orbitals. In our current situation, the most favorable way to do this is to keep the *single* orbital set of the reference function (i.e., not following the way of the original RMP2 paper of obtaining separate A- and B-spin orbitals) and obtaining the spin orbitals as spatial RHF orbitals multiplied by either alpha (A) or beta (B) spin functions. That means that closed-shell orbitals occur with identical spatial part as occupied A- and B-orbitals, active orbitals occur as occupied A- and virtual B-orbitals, and external orbitals occur as virtual A- and B-orbitals. As pointed out Sec. 2.12, we then need T-amplitudes in AA-, BB-, and AB-spin varieties. For our F12 theory, this applies for T_{ab}^{ij} and T_{mn}^{ij} amplitudes, as well as for the corresponding residuals and F12 matrices referenced in the residuals. Some comments on the transformation to spatial orbital expressions:

- For all F12 tensors of Sec. 4.1.3 which depend on Q_{12}, in practice the spin-orbital matrix set is represented by three spatial orbital matrix sets for AA, BB and AB spin. For example

$$C_{ab}^{mn} = g_{ax} F_{xb}^{mn} + F_{ax}^{mn} g_{xb} \tag{4.65}$$

of eq. (4.32) turns into

$$C_{ab}^{mn,\sigma\varsigma} = g_{a^\sigma x}^{\sigma} F_{xb^\varsigma}^{mn} + F_{a^\sigma x}^{mn} g_{xb^\varsigma}^{\varsigma}, \tag{4.66}$$

where $\sigma, \varsigma \in \{A, B\}$, g^σ denotes the σ-spin projected Fock matrix, and a^σ denotes orbitals unoccupied for spin σ [i.e., only external orbitals for A-spin, active and external orbitals for B-spin].

Also the matrix elements X and V must be provided in three different versions, although they actually do not depend on Fock matrices. This results from the summation ranges in Q_{12} and P_{12} being spin-dependent: Depending on spin, the sums over occupied or virtual spatial orbitals in the projectors (3.30) and (3.32) either *include* or *exclude* the active spatial orbitals.

- A splitting also occurs for the tensors $F_{\bar{k}l,mn}^{2,\sigma\varsigma}$, $F_{k\bar{l},mn}^{2,\sigma\varsigma}$, and $\widetilde{Y}_{kl,mn}^{\sigma\varsigma}$ in (4.47), which involve n-transformed integrals (n [eq. (4.44)] contains the spin specific exchange contribution). However, while n does depend on spin and occurs in two versions as n^A and n^B, h is actually spin free. Like $F_{\mu\nu}^{kl}$ and $K_{\mu\nu}^{kl}$, the same $\bar{F}_{\mu\nu}^{kl}$ can be used in all spin configurations.

- The doubles residual equations eq. (4.21) and eq. (4.28) both contain an anti-symmetrized Coulomb interaction as inhomogeneous term (e.g., $\frac{1}{2}(K_{ab}^{ij} - K_{ba}^{ij})$ for the conventional part). The anti-symmetrization is written with respect to spin-orbitals, however. Consequently, in both cases the second term vanishes for AB spin because the Coulomb interaction acts as identity in the spin space and $\langle AB \,|\, BA \rangle = 0$.

- Like the AB-amplitudes, AB-integral quantities have less symmetry than their closed-shell counterparts.

- In order to calculate $\widetilde{Y}_{kl,mn}^{\sigma\varsigma}$, the transformed sets $\widetilde{F}_{xa}^{kl}$ and $\widetilde{F}_{ax}^{kl}$ are required. Unlike for the other integrals, for the AB-spin case one cannot obtain one from the other and both have to be explicitly calculated (or, equivalently, only $\widetilde{F}_{xa}^{kl}$ is calculated, but for AA, BB, AB and BA spin).

Although most F12 matrices and some integral matrices have to be calculated in three to four different versions, due the use of common spatial internal orbitals, the same half-transformed integrals can be used for all of them. Compared to a theory based on unrestricted orbitals, this leads to significant savings in the (expensive) integral evaluation step. Furthermore, in RMP2 (but not in higher order theories), the residual equations of AA, BB, and AB are completely decoupled and can be solved independently.

4.1.7 Iterative solution of F12 amplitude equations

Eq. (4.28) gives the conditions which the RMP2-F12 T_{mn}^{ij} amplitudes need to fulfill: For a fixed $T_{ab}^{ij}C_{ab}^{mn}$, and F12 matrix elements V,B,X, we need to calculate the set of T_{mn}^{ij} amplitudes that results in

$$R_{mn}^{ij} = \widetilde{V}_{mn}^{ij} - X_{mn,op}\left(g_k^i T_{op}^{kj} + T_{op}^{ik} g_k^j\right) + B_{mn,op}T_{op}^{ij} \stackrel{!}{=} 0, \tag{4.67}$$

where (using spin-orbital anti-symmetry for C)

$$\widetilde{V}_{mn}^{ij} := \frac{1}{2}(1 - \pi_{[m\leftrightarrow n]})\left(V_{mn}^{ij} + T_{ab}^{ij}C_{ab}^{mn}\right). \tag{4.68}$$

If g was a diagonal matrix (i.e., $g_j^i = \varepsilon_i \delta_j^i$), eq. (4.67) would reduce to

$$R_{mn}^{ij} = \widetilde{V}_{mn}^{ij} + \left(B_{mn,op} - (\varepsilon_i + \varepsilon_j)X_{mn,op}\right) T_{op}^{ij} \stackrel{!}{=} 0. \tag{4.69}$$

Numerical problems notwithstanding, this can be solved in closed form with

$$T_{mn}^{ij} = \left[\left(\boldsymbol{B} - (\varepsilon_i + \varepsilon_j)\boldsymbol{X}\right)^{-1}\right]_{mn,op} \widetilde{V}_{op}^{ij} \tag{4.70}$$

if $\boldsymbol{B}$ and $\boldsymbol{X}$ are regarded as matrices with the combined (mn) labels as row index and the combined (op) labels as column index, and $()^{-1}$ indicating matrix inversion in this space. This approach suffers from two problems:

- In the basis of RHF orbitals, g_j^i is diagonal dominant, but not diagonal. Thus, eq. (4.70) is not exact in this case. In principle, g_j^i can be made diagonal by switching to a semi-canonical basis with different spatial orbitals for A- and B-spin electrons, but this would create other problems. When using localized occupied orbitals, g_j^i is never diagonal.
- A straight-forward exact evaluation of eq. (4.70) would require one $mn \cdot op$ matrix inversion/LU decomposition for each orbital pair ij; if o is the number of correlated occupied orbitals, this would lead to $O(o^8)$ scaling in CPU time. This scaling is worse than that of conventional MP2 (scaling as $O(N^5)$ where N is a measure of the molecule size) and worse than that of the F12 matrix element computation ($O(N^6)$ at most). Therefore this step can become a serious performance problem for molecules with many occupied orbitals.

Nevertheless, this formula illustrates a way of constructing an efficient preconditioner for an iterative solution of the T_{mn}^{ij} equations. Instead of solving the equations in closed fashion, we iterate the following process:

- Calculate R_{mn}^{ij} for current T_{mn}^{ij}

- Calculate energy, check if R^{ij}_{mn} and energy change are small enough. If yes, exit
- Update T^{ij}_{mn} by:

$$\Delta T^{ij}_{mn} := -\left[\left(\boldsymbol{B} - (\varepsilon_i + \varepsilon_j)\boldsymbol{X}\right)^{-1}\right]_{mn,op} R^{ij}_{op} \tag{4.71}$$

$$T^{ij}_{mn} \leftarrow T^{ij}_{mn} + \Delta T^{ij}_{mn}. \tag{4.72}$$

  Note that this process would converge in one step if g was diagonal, for any choice of initial T^{ij}_{mn}.

- Apply iterative subspace convergence accelerator (DIIS [115, 116] or KAIN [117]) on ΔT^{ij}_{mn} and T^{ij}_{mn}

By iterating the process, the issue of g^i_j being non-diagonal is solved. The scaling issue is addressed in one of two ways: (i) by not calculating $\left(\boldsymbol{B} - (\varepsilon_i + \varepsilon_j)\boldsymbol{X}\right)^{-1}$ exactly, but rather approximating it or (ii) by calculating the action of this operator on R^{ij}_{mn} exactly, without calculating the matrix inverse. First, option (i):

Numerical experiments indicated that the commutator norm $\|[\boldsymbol{X}, \boldsymbol{B}]\|$ is usually smaller than $\|\boldsymbol{X}\| \cdot \|\boldsymbol{B}\|$ by one to two orders of magnitude. As a result, $\boldsymbol{X}$ and $\boldsymbol{B}$ can be represented approximately by a *single* common eigensystem $\boldsymbol{Q} = \{\boldsymbol{q}\}$:

$$\boldsymbol{X} \approx \sum_q x_q \boldsymbol{q}\boldsymbol{q}^T \qquad\qquad \boldsymbol{B} \approx \sum_q b_q \boldsymbol{q}\boldsymbol{q}^T \tag{4.73}$$

with "eigen"values x_q and b_q. This allows calculating the approximate inverse as

$$\left(\boldsymbol{B} - (\varepsilon_i + \varepsilon_j)\boldsymbol{X}\right)^{-1} \approx \sum_q \boldsymbol{q}\, \frac{1}{b_q - (\varepsilon_i + \varepsilon_j)x_q}\, \boldsymbol{q}^T. \tag{4.74}$$

Using this formula, the update in eq. (4.71) can be calculated in a (fast) $O(o^6)$ process, by first multiplying $\boldsymbol{R}$ with the matrix of eigenvectors $\boldsymbol{Q}$, then scaling the resulting rows with the scalar inverse factors of eq. (4.74), then multiplying with $\boldsymbol{Q}$ again to transform back.

The eigensystem of $\boldsymbol{X} + \boldsymbol{B}$ works well for approximating both $\boldsymbol{B}$ and $\boldsymbol{X}$ via (4.73) at the same time. The reason why $\boldsymbol{B}$ and $\boldsymbol{X}$ are numerically related is indicated by their definitions:

$$\begin{aligned} B_{mn,op} &= \langle mn \,|\, F_{12} Q_{12} (g_1 + g_2) Q_{12} F_{12} \,|\, op \rangle\,, \\ X_{mn,op} &= \langle mn \,|\, F_{12} Q_{12}\, 1\, Q_{12} F_{12} \,|\, op \rangle\,. \end{aligned} \tag{4.75}$$

They both connect F12 generator orbital pairs in the same way, only with different kernels. It is reasonable to assume that the orbital pairs with strong overlap (X) also

have a strong mean field interaction (B). In fact, in some semi-empirical quantum chemistry methods (e.g., Extended Hückel), orbital interaction is approximated using orbital overlap.

For the second option, which evaluates eq. (4.71) exactly, we transform:

$$\begin{aligned}\left(\boldsymbol{B} - (\varepsilon_i + \varepsilon_j)\boldsymbol{X}\right)^{-1} &= \left[\boldsymbol{X}^{\frac{1}{2}}\boldsymbol{X}^{-\frac{1}{2}}\left(\boldsymbol{B} - (\varepsilon_i + \varepsilon_j)\boldsymbol{X}\right)\boldsymbol{X}^{-\frac{1}{2}}\boldsymbol{X}^{\frac{1}{2}}\right]^{-1} \\ &= \boldsymbol{X}^{-\frac{1}{2}}\left[\boldsymbol{X}^{-\frac{1}{2}}\boldsymbol{B}\boldsymbol{X}^{-\frac{1}{2}} - (\varepsilon_i + \varepsilon_j)\boldsymbol{1}\right]^{-1}\boldsymbol{X}^{-\frac{1}{2}}. \end{aligned} \tag{4.76}$$

This re-expresses the preconditioner in the basis of orthonormal geminal pair functions, where it can then be subjected to a spectral decomposition. Let $\boldsymbol{Q}' = \{\boldsymbol{q}'\}$ denote the eigensystem of $\boldsymbol{X}^{-\frac{1}{2}}\boldsymbol{B}\boldsymbol{X}^{-\frac{1}{2}}$, with eigenvectors b'_q. Then eq. (4.76) can be written as

$$\left(\boldsymbol{B} - (\varepsilon_i + \varepsilon_j)\boldsymbol{X}\right)^{-1} = \boldsymbol{X}^{-\frac{1}{2}}\sum_{q'}\boldsymbol{q}'\,\frac{1}{b'_q - (\varepsilon_i + \varepsilon_j)}\,\boldsymbol{q}'^T\boldsymbol{X}^{-\frac{1}{2}}, \tag{4.77}$$

which can be evaluated in a $O(o^6)$ step just as eq. (4.74). In fact, both variants can be handled with common code if $\boldsymbol{Q} = \boldsymbol{X}^{-\frac{1}{2}}\boldsymbol{Q}'$ is introduced instead of the $\boldsymbol{Q}$ defined above (the x_q are then all equal to 1). A possible problem with this approach is that $\boldsymbol{X}^{-\frac{1}{2}}$ needs to be calculated. As $\boldsymbol{X}$ is the overlap matrix of the geminal pair functions, it is supposed to be symmetric and positive semi-definite. Despite this, artifacts introduced (probably) by the RI or DF approximations led to some very rare situations in which negative $\boldsymbol{X}$ eigenvalues occurred; simply deleting said eigenvalues there had noticeable impact on the total energy.

The main advantage of these spectral preconditioning schemes is that they work well also when simpler preconditioning schemes like $\Delta\boldsymbol{T} \approx -\boldsymbol{B}^{-1}\boldsymbol{R}$ break down. In particular, they work well for core-valence correlation calculations, in which a wide range of occupied orbital energies $\varepsilon_i + \varepsilon_j$ can occur.

4.1.8 Iterative solution of conventional amplitude equations

The conventional amplitude equations can be solved in the standard form usually employed in coupled cluster programs (e.g., Ref. [88]). In each iteration, the doubles residual R^{ij}_{ab} is calculated according to eq. (4.21), then DIIS is applied to residual and amplitudes, then a perturbative amplitude update is calculated as

$$\Delta T^{ij}_{ab} := -\frac{R^{ij}_{ab}}{\varepsilon_a + \varepsilon_b - \varepsilon_i - e_j}, \tag{4.78}$$

$$T^{ij}_{ab} \leftarrow T^{ij}_{ab} + \Delta T^{ij}_{ab}, \tag{4.79}$$

where $\varepsilon_r = f_r^r$ are the diagonal elements of the Fock matrix. Again, this iterative scheme would converge in one iteration (for fixed T_{mn}^{ij}) if f was diagonal. Usually four to eight iterations are necessary in the open-shell case, because then f is only diagonal dominant.

Five aspects are noteworthy:

- The doubles equations are decoupled from the singles equations due to the projected form of the H_0-Hamiltonian [cf. eq. (4.9)f]. Therefore the singles equations are solved independently of the doubles, in closed form.
- In RMP2, the equations for the individual spin parts (AA,BB,AB) are also decoupled. Therefore the AA, BB and AB equations can be solved independently.
- However, unless EBC is assumed or a different projected H_0 Hamiltonian is used, the F12 equations are coupled to the conventional equations over the $C_{ab}^{mn} T_{mn}^{ij}$ term (see eq. (4.21)). Since solving the F12 equations is much cheaper than solving the conventional equations, it is advisable to iterate the F12 equations up to complete convergence in each single iteration of the conventional equations. It is noteworthy that the spectral decomposition in eq. (4.76) is required only once, not once per macro-iteration.
- For calculations with many occupied orbitals, the evaluation of the coupling term $C_{ab}^{mn} t_{mn}^{ij}$ in the R_{ab}^{ij} residual becomes the rate-limiting step, because it scales as $O(v^2 o^4)$. This can be avoided by using fixed or diagonal F12 amplitudes.
- As realized by H.-J. Werner, the most expensive step in evaluating the MP2-F12 R_{ab}^{ij} residual (the matrix multiplication by the external Fock matrices $\boldsymbol{f}\boldsymbol{T}^{ij}$ and $\boldsymbol{T}^{ij}\boldsymbol{f}$) can be avoided by using semi-canonical *virtual* orbitals. For this, it is not necessary to introduce semi-canonical *occupied* orbitals, too.

By default the RMP2-F12 program first computes a conventional RMP2 wave function without coupling to F12 amplitudes (by employing EBC), and then enables the coupling and calculates a fully coupled wave function. This is only done in order to obtain conventional RMP2 energies. Technically this is by no means necessary, as the fully coupled calculation converges just as well as the uncoupled one.

4.1.9 Non-Iterative coupled energy evaluation for fixed F12 amplitudes

If fixed T_{mn}^{ij} amplitudes are used, the coupled RMP2-F12 energy can also be evaluated non-iteratively in closed form. Again, this was suggested by H.-J. Werner for the closed-shell case. Starting from the conventional doubles- and energy equations

(4.21) and (4.13),

$$R_{ab}^{ij} = \frac{1}{2}(K_{ab}^{ij} - K_{ba}^{ij}) - g_k^i T_{ab}^{kj} - T_{ab}^{ik} g_k^j + g_c^a T_{cb}^{ij} + g_c^b T_{ac}^{ij} + t_{mn}^{ij} C_{ab}^{mn} \tag{4.80}$$

$$E_{(2),\mathrm{conv}} = T_{ab}^{ij} K_{ab}^{ij}, \tag{4.81}$$

we can transform to the semi-canonical orbital basis diagonalizing the Fock operator in a spin-orbital sense

$$f\,|r\rangle = \varepsilon_r\,|r\rangle\,, \tag{4.82}$$

(this implies different spatial orbital parts for different spins), and get

$$R_{ab}^{ij} = \widetilde{K}_{ab}^{ij} + (\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j) T_{ab}^{ij} \stackrel{!}{=} 0 \tag{4.83}$$

$$\widetilde{K}_{ab}^{ij} = \frac{1}{2}(1 - \pi_{[m\leftrightarrow n]}) \left(K_{ab}^{ij} + t_{mn}^{ij} C_{ab}^{mn} \right) \tag{4.84}$$

(cf. eq. (4.68)f) and thus

$$E_{(2),\mathrm{conv}} = \sum_{ijab} \frac{-K_{ab}^{ij} \widetilde{K}_{ab}^{ij}}{(\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j)}. \tag{4.85}$$

However, eq. (4.84) is deceiving: The ij of t_{mn}^{ij} occurring here are relative to the semi-canonical spin-orbital basis; but in this basis, the formulas (3.23) for the fixed amplitudes do not apply. The fixed t_{mn}^{ij} must be evaluated in terms of equal spatial orbitals for each spin, and then the ij labels must be transformed into the semi-canonical basis. As a result, the t_{mn}^{ij} are not diagonal anymore, unlike in the closed-shell case. Also the ab labels of C_{ab}^{mn} must be explicitly transformed.

With the T_{ab}^{ij} thus obtained by eq. (4.83), the tensor $T_{ab}^{ij} C_{ab}^{mn}$ can be formed and transformed back into the spatial orbital basis [for a $O(N^5)$ evaluation of the term, the ij labels of T_{ab}^{ij} must be transformed to RHF orbitals first]. From there it can be used to evaluate the Hylleraas functional of the fixed amplitude F12 equations as usual, in the RHF orbital basis.

Performing this process requires roughly the same amount of memory as an iterative solution of the RMP2-F12 equations, and it has exactly the same scaling in terms of computer time. Nevertheless, there may be combinations of basis set size and molecule size for which the iterative solution requires a significant fraction of the total computer time, and there this transformation would be advantageous. This region would be limited, however, because for small systems evaluating the raw integrals is dominant (a very slow $O(N^3)$ process) and calculating the transformed 4-index integrals is always expensive (a not quite as slow $O(N^5)$ process, but still slower

than the $O(N^5)$ iterative solution). Due to the limited scope, this transformation has not been implemented in order to reduce the code complexity.

4.1.10 Normalization of F12 configurations and B-Matrix singularities

Singularities and negative eigenvalues in the B-matrix [eq. (4.27)] have been reported by several groups. Concretely, the equation system determining the full T_{mn}^{ij} amplitudes may become ill-conditioned or singular. This effect mainly appears when larger systems are being investigated (see Ref. [118]).

The main reason for these singularities is that for extended systems the explicitly correlated configurations (3.8) become linearly dependent. This is easily seen when considering the configurations in a localized orbital basis. Assume that in

$$\mathcal{F}_{\alpha\beta}^{mn} = \left\langle \alpha\beta \,\middle|\, Q_{12}F_{12} \,\middle|\, mn \right\rangle,$$
$$Q_{12} = (1 - o_1)(1 - o_2)(1 - v_1v_2)$$

the orbitals m and n are well-localized and far apart. Then inevitably in $\mathcal{F}_{\alpha\beta}^{mn}$ only such F_{12} can have non-vanishing impact for which r_{12} is large, since in

$$\langle \alpha\beta \,|\, F_{12} \,|\, mn \rangle = \iint \varphi_\alpha^*(\boldsymbol{r}_1)\varphi_\beta^*(\boldsymbol{r}_2)F_{12}(\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|)\varphi_m(\boldsymbol{r}_1)\varphi_n(\boldsymbol{r}_2)\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2$$

$\boldsymbol{r}_1$ is essentially restricted to the domain of the orbital m and $\boldsymbol{r}_2$ to that of n. Since F_{12} falls off sharply with r_{12} (for real Slater functions with $\exp(-\beta r_{12})$ and for linear combinations of Gaussians as $\exp(-\alpha_i r_{12}^2)$), the resulting explicitly correlated configurations become almost zero in norm.

This problem could be approached by orthonormalization of the explicitly correlated configurations with respect to $\langle nm \,|\, F_{12}F_{12} \,|\, op \rangle$ (note that X is their overlap matrix). But ultimately, this would have limited value since the configurations with large r_{12} (i) don't help in the treatment of the correlation problem anyway, (ii) numerical problems are still to be expected in cases with near linear dependencies, and (iii) the explicitly correlated configurations are *legitimately* diminished by Q_{12}, since their effect can be achieved with the conventional wave function.

Due to their limited impact, we simply project out the offending configurations in the iterative solution process. This is done by removing eigenvectors of the residual R_{mn}^{ij}'s (mn) space which overlap with the eigenspace of the preconditioning matrix $\boldsymbol{B} + (\varepsilon_i + \varepsilon_j)\boldsymbol{X}$ with eigenvalues smaller than numerical zero (default value 10^{-10}) in absolute value. If using a variant of the spectral preconditioner described in Sec. 4.1.7, this can be done trivially in the denominator scaling step (by setting rows corresponding to singular denominators to zero). This is not an exact treatment, but it appears to work well in practice. However, one must make sure that the initial

T^{ij}_{mn} do not overlap with the singular space, because this singular space cannot be updated in the iterations (i.e., one must start the iterations with $T^{ij}_{mn} = 0$, not with fixed amplitudes).

Inadequate orbital or auxiliary RI basis sets may also lead to singular F12 equations in small molecules. This can happen, for example, if trying to calculate core-core or core-valence-correlation with valence-only basis sets. If corresponding warnings occur, better basis sets should be employed. Nevertheless, in order to force convergence also with unsuitable basis sets, the option `bdiag=1` is implemented, which directly removes all eigenvectors with negative eigenvalues from the $\boldsymbol{B}$ matrix. While $\boldsymbol{B}$ (unlike $\boldsymbol{X}$, at least for the general ansatz) is formally not strictly required to be positive semi-definite, one would certainly expect that the $|ax\rangle$, $|xa\rangle$ and $|xy\rangle$ virtual orbital pairs represented by the F12 configurations have no lower $f_1 + f_2$ energies than at least some of the MO basis virtual pairs $|ab\rangle$. Thus, negative eigenvalues of $\boldsymbol{B}$ are very likely to represent calculation artifacts (for valence correlation they usually vanish when the RI basis set is extended with diffuse functions), so deleting them is not really illegitimate.

For fixed F12 amplitude methods we do not employ any projections and simply take all F12 matrices as they are calculated.

4.2 Corrections to Hartree-Fock energies

The RMP2 method contains energy contributions from single excitations. These do not couple to the doubles equations and depend only on the A- and B-Fock matrices [eq. (4.12)]. In RMP2-F12 we can thus calculate the single excitations not only within the normal MO basis, but within the combined MO+CA basis, where the CABS orbitals are treated like normal virtual orbitals. This extension is cheap because we have a full set of Fock matrices to represent this extended MO+CA space anyway (the MO/CA and CA/CA Fock matrices are required to evaluate A^{kl}_{mn} in eq. (4.47)). Concretely, the solution of the RMP2 singles equation in eq. (4.12) can be written down in closed form if it is transformed into a basis in which g is diagonal (that means diagonalizing (f_{ij}) and $(f_{\alpha\beta})$ independently):

$$t^i_\alpha = \frac{f^i_\alpha}{g_{ii} - g_{\alpha\alpha}} \qquad\qquad E^{(2)}_{\text{singles}} = \sum_{i\alpha} \frac{(f^i_\alpha)^2}{g_{ii} - g_{\alpha\alpha}}, \tag{4.86}$$

where α runs over all virtual orbitals (the unoccupied MOs and the CABS).

This basis extension has two consequences: (i) The basis set convergence of the singles correlation contribution is accelerated, and (ii) The Hartree-Fock orbitals are perturbatively relaxed by the CABS. The second effect is by far dominant in magnitude, and it is very helpful in practical calculations, since it greatly diminishes

the deviation of the Hartree-Fock energy from the basis set limit. This is significant because, while with conventional electron correlation calculations the accuracy of the Hartree-Fock treatment is usually not a serious issue, this changes when using explicitly correlated methods. If not corrected for, the basis set deviation due to Hartree-Fock is often larger than the remaining MP2-F12 correlation error.[3]

We can isolate the part of the singles energy correction that corresponds to a spin-free relaxation of the RHF function [as opposed to RMP2 singles correlation]. By following similar steps as in Sec. 2.8, we now solve

$$0 = \left\langle \bar{\Phi}^i_\alpha \,\middle|\, H \,\middle|\, \Phi \right\rangle + \sum_{k\gamma} t^k_\gamma \left\langle \bar{\Phi}^i_\alpha \,\middle|\, g \,\middle|\, \bar{\Phi}^k_\gamma \right\rangle \tag{4.87}$$

as perturbation equation, where $\left|\bar{\Phi}^i_\alpha\right\rangle = E^\alpha_i |\Phi\rangle$ and $E^\alpha_i = e^{\alpha A}_{iA} + e^{\alpha B}_{iB}$ are the spin-free excitation operators. The corresponding numerical equations can be obtained from the spin-orbital equation (4.12) by using the spin-orbital amplitudes

$$\begin{aligned} t^{iA}_{\alpha A} &= t^{iB}_{\alpha B} = t^i_\alpha \\ t^{uA}_{\alpha A} &= t^u_\alpha \text{ and } t^{uB}_{\alpha B} = 0 \\ t^{iB}_{uB} &= t^i_u \text{ and } t^{iA}_{uA} = 0, \end{aligned} \tag{4.88}$$

and demanding

$$r^{iA}_{\alpha A} + r^{iB}_{\alpha B} = r^{uA}_{\alpha A} = r^{iB}_{uB} = 0, \tag{4.89}$$

where i denotes closed, u active, and α external spatial orbitals. The values of $t^{uA}_{\alpha A}$ and $t^{uB}_{\alpha B}$ only *appear* to be different, because $e^{\alpha B}_{uB} |\Phi\rangle = 0$ and thus the B-configuration and corresponding residual intrinsically vanish. Due to the RHF Brillouin conditions [eq. (2.36)],

$$f^{iA}_{aA} + f^{iB}_{aB} = f^{uA}_{aA} = f^{iB}_{uB} = 0, \tag{4.90}$$

no closed-to-active relaxation occurs for optimized RHF functions and furthermore, the total energy correction vanishes if α and γ are constrained to the MO basis.

4.3 Coupled Cluster Singles and Doubles: CCSD-F12

As already mentioned in Sec. 2.10, a MP2 method is not an end in itself. In practical calculations, the accuracy of MP2 methods is not usually limited by the basis sets that are employed, but rather by the intrinsic (in-)accuracy of the perturbation approach itself.

[3] In one case found in our laboratory the basis set limit deviation of the Hartree-Fock treatment at AVTZ level was ten times as large as that of the correlation deviation with DF-MP2-F12.

This changes in coupled-cluster treatments. The CCSD(T) method offers an excellent intrinsic accuracy for a wide variety of applications in quantum chemistry. As long as a single-determinant reference function is adequate, CCSD(T) obtains energy differences which reliably differ by no more than a few kJ/mol from a near-exact quantum many-body calculations [83, 84]—within a given basis set. However, the method is very expensive. Therefore it cannot be applied to many systems of interest with sufficiently large basis sets, and the practical accuracy suffers considerably from that.

Unfortunately, the CCSD equations, in particular in their open-shell form, are much more complex than MP2-type equations. While exact CCSD-F12 treatments have appeared in the literature very recently [27, 41–43], these are much more expensive than the already expensive conventional CCSD methods, and the methods using modern F12 techniques [41–43] could only be realized with automatic implementation tools. Approximate CCSD-F12 methods can fare much better in performance. For example, Fliegl et al. [38, 39] introduced the simplified CC2(R12) and CCSD(R12) methods, which were extended to CCSD(F12) by Tew et al. [40, 119]; CCSD(F12) was described as being only three times as slow as (their) conventional CCSD. Additionally, this approach is also much simpler to implement than a full CCSD-F12. However, still many terms arise additionally to conventional CCSD, because various contributions with small impact (e.g., contributions of single excitations) are handled in an exact manner.

Encouraged by the good performance of crude F12 treatments like MP2-F12 in its varieties 3*C(FIX,HY1) [Sec. 5] or 3*A(DX) [Ref. [36]], we [H.-J. Werner, T.B. Adler and I] searched for a way to include F12 terms in the CCSD equations in such a form that the performance does not deteriorate at all compared to the conventional method with the same basis set. Further, we attempted to minimize the number of additional terms such that a future implementation of analytic energy gradients (with respect to the atom coordinates) and a full local implementation are less complex: Both extensions would add considerable practical value to CCSD-F12 methods, and in both extensions every single unnecessary additional term could lead to serious implementation problems. The results of our developments of simplified CCSD methods are the CCSD-F12x (x=a,b) methods described below.

The sections 4.3.1 to 4.3.3 are concerned with general aspects of CCSD-F12 theory, while sections 4.3.4 and 4.3.5 define our F12x methods and discuss practical implementation aspects. The theory description is based on closed-shell formulas. This is done because in the CCSD theory no new open-shell specific aspects arise which have not already been discussed in Sections 3 and 4.1; presenting the less transparent open-shell formulas would thus not aid in the discussion of the F12x approximations. Consequently, in this section all indices refer to spatial orbitals, not spin orbitals.

4.3.1 The CEPA(0)-Doubles equations

The CEPA(0)-Doubles terms form the core of the CCSD equations. They are obtained formally from the CCSD equations by removing all terms that are non-linear in the amplitudes and all terms that involve singles amplitudes. Because we look only for a crude treatment of F12 excitations, concentrating on the CEPA(0)-part is sufficient: We only need to identify additional terms to include in the already existing CCSD programs. Concretely, in the closed-shell case the following equations are obtained from Ref. [88] (or Tab. 4.2):

$$\boldsymbol{R}^{ij} = \boldsymbol{R}^{ij}_{\mathrm{MP2}} + \boldsymbol{K}(\boldsymbol{T}^{ij}) + \boldsymbol{K}^{ij}_{kl}\boldsymbol{T}^{kl} + \boldsymbol{G}^{ij} + \boldsymbol{G}^{ji\dagger} \tag{4.91}$$

$$[\boldsymbol{K}(\boldsymbol{T}^{ij})]_{\alpha\beta} = \left\langle \alpha\beta \middle| r^{-1} \middle| \gamma\delta \right\rangle T^{ij}_{\gamma\delta} \tag{4.92}$$

$$\boldsymbol{R}^{ij}_{\mathrm{MP2}} = \boldsymbol{K}^{ij} + \boldsymbol{f}\boldsymbol{T}^{ij} + \boldsymbol{T}^{ij}\boldsymbol{f} - \boldsymbol{T}^{ik}f_{kj} - f_{ik}\boldsymbol{T}^{kj} \tag{4.93}$$

$$\boldsymbol{G}^{ij} = \widetilde{\boldsymbol{T}}^{ik}\boldsymbol{K}^{kj} - \boldsymbol{T}^{ik}\boldsymbol{J}^{kj} - \boldsymbol{J}^{kj}\boldsymbol{T}^{ik} \tag{4.94}$$

$$\widetilde{\boldsymbol{T}}^{ik} = 2\boldsymbol{T}^{ik} - \boldsymbol{T}^{ik\dagger}. \tag{4.95}$$

Here matrices have the dimension of the virtual space, which is assumed to be complete (indices α, β, ...), and the integral matrices are given by $K^{ij}_{\alpha\beta} = \left\langle \alpha\beta \middle| r^{-1} \middle| ij \right\rangle$ and $J^{ij}_{\alpha\beta} = \left\langle \alpha i \middle| r^{-1} \middle| \beta j \right\rangle$ as usual.

Let us look at the conventional part of the residual equation, which is formed by restricting (4.91) to α and β lying in the regular virtual space (a,b), and inserting the doubles amplitudes eq. (3.9):

$$T^{ij}_{\alpha\beta} = \delta^a_\alpha\delta^b_\beta T^{ij}_{ab} + \underbrace{\mathcal{T}^{ij}_{\alpha\beta}}_{T^{ij}_{mn}\mathcal{F}^{mn}_{\alpha\beta}}. \tag{4.96}$$

The form of the terms arising from $\boldsymbol{R}^{ij}_{\mathrm{MP2}}$ was already discussed in Sec. 4.1.2, and is not repeated here. Basically, we only get one contribution

$$R^{ij}_{ab} \mathrel{+}= T^{ij}_{mn}C^{mn}_{ab}, \tag{4.97}$$

which needs to be treated explicitly additionally to the conventional MP2 residual terms [which are already accounted for by the conventional CCSD program]. The term $\boldsymbol{K}(\boldsymbol{T}^{ij})$ leads to additional contributions:

$$\begin{aligned} R^{ij}_{ab} \mathrel{+}= [K(\boldsymbol{T}^{ij})]_{ab} &= \left\langle ab \,\middle|\, r^{-1} \,\middle|\, \gamma\delta \right\rangle \left(\delta^c_\gamma\delta^d_\delta T^{ij}_{cd} + T^{ij}_{mn}\mathcal{F}^{mn}_{\gamma\delta}\right) \\ &= \left\langle ab \,\middle|\, r^{-1} \,\middle|\, cd \right\rangle T^{ij}_{cd} + V^{mn}_{ab}T^{ij}_{mn}. \end{aligned} \tag{4.98}$$

The first term is part of the regular CCSD program, while the second one

$$V_{ab}^{mn} T_{mn}^{ij} = \left\langle ab \left| r^{-1} Q_{12} F_{12} \right| mn \right\rangle T_{mn}^{ij} \tag{4.99}$$

is new. Its evaluation will be discussed in Sec. 4.3.5. The term $\boldsymbol{K}_{kl}^{ij} \boldsymbol{T}^{kl}$ does not lead to additional contributions to the conventional CCSD residual because $\mathcal{T}_{ab}^{ij} = 0$. The terms arising from $\boldsymbol{G}^{ij}$ lead to additional contributions which require the use of eq. (3.32) for a resolution:

$$\begin{aligned} G_{ab}^{ij} &= \widetilde{T}_{a\gamma}^{ik} K_{\gamma b}^{kj} - T_{a\gamma}^{ik} J_{\gamma b}^{kj} - J_{a\gamma}^{kj} T_{\gamma b}^{ik} \\ &= \widetilde{T}_{ac}^{ik} K_{cb}^{kj} - T_{ac}^{ik} J_{cb}^{kj} - J_{ac}^{kj} T_{cb}^{ik} + \left(\widetilde{\mathcal{T}}_{ax}^{ik} K_{xb}^{kj} - \mathcal{T}_{ax}^{ik} J_{xb}^{kj} - J_{ax}^{mn} \mathcal{T}_{xb}^{kj} \right). \end{aligned} \tag{4.100}$$

Since the $\mathcal{T}$s occur in their explicit basis representation, according to eq. (3.9) they are given simply by

$$\mathcal{T}_{ax}^{ij} = t_{mn}^{ij} \mathcal{F}_{ax}^{mn} = t_{mn}^{ij} F_{ax}^{mn} \tag{4.101}$$

—the Q_{12} projector has no effect on $|xa\rangle$ and $|xb\rangle$. Due to the structure of the G-terms, no exact or semi-exact resolutions of the identity could be used: There is no way to avoid eq. (3.32).

Altogether, at the linearized doubles-only level only the following F12 contributions arise additionally to the conventional part of the CCSD equation:

$$R_{ab}^{ij} \mathrel{+}= t_{mn}^{ij} C_{ab}^{mn} + V_{ab}^{mn} T_{mn}^{ij} + \Delta G_{ab}^{ij} + \Delta G_{ba}^{ji} \tag{4.102}$$

$$\Delta G_{ab}^{ij} = \widetilde{\mathcal{T}}_{ax}^{ik} K_{xb}^{kj} - \mathcal{T}_{ax}^{ik} J_{xb}^{kj} - J_{ax}^{mn} \mathcal{T}_{xb}^{kj}. \tag{4.103}$$

We now consider the amplitude equations for the explicitly correlated configurations. For this, we form

$$R_{mn}^{ij} = \mathcal{F}_{\alpha\beta}^{mn} R_{\alpha\beta}^{ij} \tag{4.104}$$

with $R_{\alpha\beta}^{ij}$ from Eq. (4.91) and $T_{\alpha\beta}^{ij} = \delta_\alpha^a \delta_\beta^b T_{ab}^{ij} + \mathcal{T}_{\alpha\beta}^{ij}$ from Eq. (3.4). The MP2 terms have been discussed in Sec. 4.1.2. Due to the terms $\boldsymbol{K}(\boldsymbol{T}^{ij})$ and $\boldsymbol{K}_{kl}^{ij} \boldsymbol{T}^{kl}$ we obtain:

$$R_{mn}^{ij} = \ldots + H_{mn}^{kl} T_{kl}^{ij} + V_{cd}^{mn} T_{cd}^{ij} + K_{kl}^{ij} T_{op}^{kl} X_{mn}^{op} + \ldots \tag{4.105}$$

$$\begin{aligned} H_{mn}^{kl} &= \mathcal{F}_{\alpha\beta}^{kl} \left\langle \alpha\beta \left| r_{12}^{-1} \right| \gamma\delta \right\rangle \mathcal{F}_{\gamma\delta}^{mn} \\ &= \left\langle kl \left| F_{12} Q_{12} r_{12}^{-1} Q_{12} F_{12} \right| mn \right\rangle \end{aligned} \tag{4.106}$$

$$V_{cd}^{ij} = \left\langle ij \left| F_{12} Q_{12} r_{12}^{-1} \right| cd \right\rangle. \tag{4.107}$$

The F12 matrices H_{mn}^{kl} and V_{cd}^{ij} will be discussed later in Sec. 4.3.5. Again, the terms stemming from $\mathcal{F}_{\alpha\beta}^{mn} G_{\alpha\beta}^{ij}$ can only be evaluated using eq. (3.32). For example,

$\boldsymbol{T}^{ik}\boldsymbol{J}^{kj}$ of (4.94) turns into

$$\mathcal{F}^{mn}_{\alpha\beta}T^{ik}_{\alpha\gamma}J^{kj}_{\gamma\beta} = \left\langle mn\middle|F_{12}Q_{12}\middle|ax\right\rangle T^{ik}_{ac}J^{kj}_{cx} + \left\langle mn\middle|F_{12}Q_{12}\middle|\alpha\beta\right\rangle T^{ik}_{op}\left\langle op\middle|F_{12}Q_{12}\middle|\alpha\gamma\right\rangle J^{kj}_{\gamma\beta}. \tag{4.108}$$

The second summand of this expression is problematic: Since none of the tensors share more than one external index, no exact RI can be used and *all* external indices need to be resolved simultaneously with approximate RIs in the combined MO+CA basis. Furthermore, the approximate $G^{ij}_{\alpha\beta} = \ldots - T^{ik}_{\alpha\gamma}J^{kj}_{\gamma\beta}$ must be explicitly formed as first contraction because $\mathcal{F}$ shares only one index with either T or J. This also holds for the other contributions to G. As a result, in the G-terms nothing is saved compared with actually doing the CEPA calculation with the combined MO+CA bases *as orbital basis*: Formally, eq. (4.94) must be evaluated exactly as written, where α, β, γ are approximated by the combined regular virtual and CA basis sets, and in this basis an approximate $T^{ij}_{\alpha\beta}$ is explicitly built from T^{ij}_{ab} and $F^{mn}_{\alpha\beta}t^{ij}_{mn}$. The MO/MO parts $G^{ij}_{\alpha\beta}$ would enter the conventional residual (eq. (4.100)) while the other parts enter the F12 residual after contraction with F:

$$\begin{aligned} R^{ij}_{mn} &= \ldots + (G^{ij}_{\alpha\beta} + G^{ji}_{\beta\alpha})\mathcal{F}^{mn}_{\alpha\beta} + \ldots \\ &= \ldots + (G^{ij}_{xa} + G^{ji}_{ax})F^{mn}_{xa} + (G^{ij}_{by} + G^{ji}_{by})F^{mn}_{by} + (G^{ij}_{xy} + G^{ji}_{yx})F^{mn}_{xy} + \ldots \end{aligned} \tag{4.109}$$

Thus, even for the linearized model, a full evaluation of R^{ij}_{mn} would imply a cost of $O(o^3(v+c)^3)$, where o denotes the number of correlated occupied orbitals, v the number of virtual orbitals and c the number of CA orbitals. In a typical application, where $c \approx 2v \ldots 3v$, the computational cost of this term would increase by one to two orders of magnitude. Since even in a conventional CCSD program calculating the $\boldsymbol{G}^{ij}$ terms is one of the most expensive steps, a full evaluation of this equation in each iteration is out of question.

In total, the CEPA(0)D R^{ij}_{mn} is given by

$$R^{ij,\mathrm{CEPA}}_{mn} = R^{ij,\mathrm{MP2}}_{mn} + H^{kl}_{mn}T^{ij}_{kl} + V^{mn}_{cd}T^{ij}_{cd} + K^{ij}_{kl}T^{kl}_{op}X^{op}_{mn} + (G^{ij}_{\alpha\beta} + G^{ji}_{\beta\alpha})\mathcal{F}^{mn}_{\alpha\beta}, \tag{4.110}$$

where $|\alpha\rangle\langle\alpha|$ and $|\beta\rangle\langle\beta|$ are approximated by $|a\rangle\langle a| + |x\rangle\langle x|$ within the finite MO and CA basis sets as indicated.

The direct expression for the correlation energy is given by

$$\begin{aligned} E_{\mathrm{corr}} &= \left\langle \Phi \middle| H\exp(T_2) \middle| \Phi \right\rangle - E_0 = \langle \Phi | HT_2 | \Phi \rangle \\ &= \left(2T^{ij}_{ab} - T^{ij}_{ba}\right)K^{ij}_{ab} + \left(2T^{ij}_{mn} - T^{ij}_{nm}\right)V^{ij}_{mn} \end{aligned} \tag{4.111}$$

just like in the MP2 case.

4.3.2 Additional considerations for CCSD

Further considerations are necessary when including the terms discussed in Sec. 4.3.1 in a CCSD program. First, due to the straight exponential parameterization of the CCSD wave function, the *actual* amplitudes of double excitations are not only given by the cluster amplitudes T_{ab}^{ij} (which describe only the so-called "connected" doubles), but also by quadratic single excitations:

$$\begin{aligned}\exp(T_1 + T_2) &= T_1 + T_2 + \frac{1}{2}T_1^2 + \dots \\ &= t_a^i E_i^a + \frac{1}{2}\big(T_{ab}^{ij} + t_a^i t_b^j\big)E_{ij}^{ab} + \dots \end{aligned} \tag{4.112}$$

As a result, the T_{ab}^{ij} terms in CEPA(0) in most cases must be replaced by

$$\bar{T}_{ab}^{ij} = T_{ab}^{ij} + t_a^i t_b^j \tag{4.113}$$

in CCSD.[4] For example, the conventional closed-shell CCSD correlation energy is given by

$$E_{\text{corr}} = \big(2\bar{T}_{ab}^{ij} - \bar{T}_{ba}^{ij}\big)K_{ab}^{ij}, \tag{4.114}$$

rather than the same CEPA(0)-equation with Ts instead of $\bar{T}$s. However, if the replacement eq. (4.113) is done indiscriminately in all equations, then some additional disconnected terms may arise: For example $T_{ab}^{ij}f_a^i f_b^j$ is connected, while $t_a^i t_b^j f_a^i f_b^j = (t_a^i f_a^i)(t_b^j f_b^j)$ is disconnected. While such terms arise in some CCSD formulation (e.g., when evaluating eq. (2.56) instead of a similarity transformed formulation), they are generally undesirable because they are cancelled by other terms in the residual equations at convergence (if evaluated exactly), and break the size consistency of the method if not handled properly. Thus, it is generally advisable to not include such terms in the first place. These considerations also apply to F12 corrections: When going from CEPA(0) to CCSD, all occurrences of regular T_{ab}^{ij} amplitudes are replaced by $\bar{T}_{ab}^{ij}$, provided that the additional $t_a^i t_b^j$ terms would not lead to disconnected contributions in the conventional theory.

A further formal obstacle lies in the fact that we used the projected Fock operator (4.1) in the derivation of RMP2-F12 matrix elements. In CCSD, we are not free in the choice of the H_0-Hamiltonian: The residual equations contain f, not the projected g. It can be shown that the F12 intermediates X, B and V [Sec. 4.1.3] are, in fact, not affected by this modification because $Q_{12}g_{12}Q_{12} = Q_{12}f_{12}Q_{12}$. However, if $g = f$,

[4] In principle, $\bar{T}_{ab}^{ij}$ can be avoided completely if $\exp(T_1+T_2-\frac{1}{2}T_1^2)\,|\Phi\rangle$ is used as CCSD wave function ansatz instead of the usual $\exp(T_1 + T_2)\,|\Phi\rangle$. Both approaches are exactly equivalent because the wave functions have the same flexibility. However, this ansatz is itself disconnected, and would lead to size consistency problems if approximations were introduced into the equations.

additional coupling terms between the conventional singles and F12 doubles occur even at the MP2 level, because in the last step of the derivations of (4.9) and (4.12) $g^i_\alpha = 0$ does not hold exactly anymore. The additional coupling matrix elements arising from that have the form

$$r^i_a \mathrel{+}= f^\gamma_k \mathcal{T}^{ik}_{a\gamma} = f^x_k F^{mn}_{ax} t^{ik}_{mn}, \tag{4.115}$$

where only the F12 part of T_2 was considered because the conventional part is already included in the conventional CCSD equations. We neglect these terms because they are *very small*: They vanish under GBC (then $f^x_k = 0$) and they affect only single excitations, which themselves only have minor impact on the correlation energy. This has also been confirmed experimentally by explicitly programming the terms in the UCCSD program.

4.3.3 Wave function ansatz and energy functional

For CCSD-F12x, we exclusively employ the fixed amplitude ansatz. That means that we do not attempt to actually solve the F12 amplitude equation (4.110), but rather determine the F12 amplitudes from the short-range asymptotic behavior of the wave function [cf. Sec. 3.3]. Because the amplitudes are fixed to reproduce semi-exact wave function properties, they are expected to be quite good, even if the F12 equations are not fully solved.

As already discussed, this fixed amplitude ansatz has various advantages to its alternatives [cf. Sec. 3.4], but on the downside it introduces a comparatively large dependence of the energy expression eq. (4.111) on the geminal exponent γ (eq. (3.11)). As a result, it is essential that the energy is not evaluated with the direct energy expression eq. (4.111), but rather with an approximate Lagrange functional which accounts for the non-solved F12 equations.

For this functional, we start from the ansatz

$$L(\{t_i\}, \{\lambda_i\}) = E(\{t_i\}) + \lambda_i R_i(\{t_j\}), \tag{4.116}$$

where E is the direct energy functional (which depends only on the amplitudes and integrals), $\{t_i\}$ denotes the set of all amplitudes of the model (t^i_a, T^{ij}_{ab}, T^{ij}_{mn}), R_i denotes the corresponding residuals (r^i_a, R^{ij}_{ab}, R^{ij}_{mn}), and $\{\lambda_i\}$ is the set of corresponding Lagrange multipliers. In principle, the amplitudes t_i and multipliers λ_i should be determined by demanding

$$\frac{\partial}{\partial t_i} L = 0 \qquad\qquad \frac{\partial}{\partial \lambda_i} L = 0. \tag{4.117}$$

It is seen that the ∂_{λ_i}-conditions just reproduces the regular amplitude equations (i.e.,

$\forall i : R_i = 0$), and if the amplitude equations are exactly solved (i.e., all R_i vanish), then $L = E$. However, since $\partial_{t_i} L = 0$, the Lagrange functional only depends on *second order* on the amplitudes t_i, and hence also their errors if the corresponding residual equations are *not* solved completely.

In the case of CEPA(0), the equations for the λ_i-multipliers can be shown to be the bi-orthogonal equivalents of the equations for the t_i-amplitudes. For example, in the closed-shell case explicit calculation shows:

$$\begin{aligned} \lambda_a^i &= 2t_a^i \\ \lambda_{ab}^{ij} &= 2T_{ab}^{ij} - T_{ba}^{ij} \\ \lambda_{mn}^{ij} &= 2T_{mn}^{ij} - T_{nm}^{ij}. \end{aligned} \tag{4.118}$$

This simple relationship has a deeper root in the fact that like RMP2, also CEPA(0) can be derived as a second order perturbation theory [120], and therefore its Lagrangian is exactly given by the CEPA(0)-Hylleraas functional eq. (2.50) just like in the RMP2 case (with the CEPA(0)-residuals instead of RMP2 residuals). In the case of full CCSD, the λ_i-multipliers do not have simple relationships to the t_i-amplitudes—rather they must be determined for fixed t_i amplitudes in an explicit calculation [62] which is approximately as expensive as CCSD itself. However, as already mentioned, the CCSD equations are dominated by the CEPA(0) terms, so the λ_i given by eq. (4.118) still give a good approximation to the exact λ_i. Therefore this replacement is a common approximation, usually employed whenever an exact Lagrange functional is not strictly required.

Taking these considerations into account, our basic F12 energy functional (4.116) becomes

$$E_{2\mathrm{nd}} \approx E_{\mathrm{direct}} + (2t_a^i) r_a^i + (2T_{ab}^{ij} - T_{ba}^{ij}) R_{ab}^{ij} + (2T_{mn}^{ij} - T_{nm}^{ij}) R_{mn}^{ij}, \tag{4.119}$$

where E_{direct} is the direct linear energy expression depending only on the amplitudes (e.g., (4.111) for CEPA(0)D). In principle, the residuals r_a^i and R_{ab}^{ij} correspond to the full residuals given in eq. (4.3.1). *If* during the iterative solution of the CCSD equations some terms of r_a^i and R_{ab}^{ij} are neglected, then the effect of these neglected terms can be estimated by a-posteriorly re-introducing them in the r_a^i and R_{ab}^{ij} of the $E_{2\mathrm{nd}}$ functional. For example, one could solve the CCSD equations *without any F12 contributions*, and then obtain the effect due to the F12 amplitude coupling via eq. (4.119) with the r_a^i and R_{ab}^{ij} including F12 terms. This is the way Valeev went in Ref. [121], although using a different formal perturbation theory as starting point. If no additional a-posteriori effects on conventional amplitudes are considered, then $r_a^i = R_{ab}^{ij} = 0$ and only $(2T_{mn}^{ij} - T_{nm}^{ij}) R_{mn}^{ij}$ leads to a nonzero energy contribution.

Table 4.2: Factorized closed-shell CCSD equations. Based on Ref. [88], but here no disconnected terms occur, terms generated by $K(D)$ are split into $K(\bar{T})$ and singles, and some intermediates are re-aligned. Based on derivations from H.-J. Werner.

Modified amplitudes

$$\bar{T}^{ij}_{ab} = T^{ij}_{ab} + t^i_a t^j_b \qquad \widehat{T}^{ij}_{ab} = \tfrac{1}{2} T^{ij}_{ab} + t^i_a t^j_b$$

Integral quantities (u, v: arbitrary orbitals)

$$K^{ij}_{uv} = \left\langle uv \left| r^{-1} \right| ij \right\rangle \qquad J^{ij}_{uv} = \left\langle ui \left| r^{-1} \right| vj \right\rangle \qquad L^{ij}_{uv} = 2K^{ij}_{uv} - K^{ij}_{vu}$$

3- and 4-external integral–amplitude contractions

$$K(\bar{T}^{ij})_{uv} = \left\langle uv \left| r^{-1} \right| ab \right\rangle \bar{T}^{ij}_{ab} \qquad K(E^{ij})_{ab} = \left\langle ab \left| r^{-1} \right| ic \right\rangle t^j_c$$
$$J(E^{ij})_{ab} = \left\langle ai \left| r^{-1} \right| bc \right\rangle t^j_c$$

Intermediates

$$\begin{aligned}
A_{ab} &= L^{kl}_{ac} T^{lk}_{cb} \\
\alpha^{ij}_{kl} &= K^{ij}_{kl} + t^i_c K^{kl}_{cj} + t^j_c K^{lk}_{ci} + \bar{T}^{ij}_{ab} K^{kl}_{ab} \\
\beta_{ki} &= f^i_k + t^l_c L^{lk}_{ci} + f^c_k t^i_c + \bar{T}^{il}_{ab} L^{kl}_{ab} \\
v^k_a &= f^k_a + L^{kl}_{ac} t^l_c \\
X_{ab} &= f_{ab} - A_{ab} - v^k_a t^k_b + \left[2J(E^{kj})_{ab} - K(E^{kj})_{ab}\right] \delta^j_k \\
w^{ij}_{ak} &= K^{ij}_{ak} + K^{ik}_{ac} t^j_c + J^{kj}_{ac} t^i_c + K(\bar{T}^{ij})_{ak} \\
G^{ij}_{ab} &= T^{ij}_{ac} X_{bc} - \beta_{ki} T^{kj}_{ab} - w^{ij}_{ak} t^k_b + K(E^{ij})_{ab} \\
&\quad + [2T^{ki}_{ca} - T^{ki}_{ac}] Y^{kj}_{cb} - T^{ik}_{cb} Z^{kj}_{ca} - \tfrac{1}{2} T^{ik}_{ca} Z^{kj}_{cb} \\
Y^{kj}_{ab} &= K^{kj}_{ab} - \tfrac{1}{2} J^{kj}_{ab} + \tfrac{1}{2} L^{kl}_{ac} \left[2\widehat{T}^{lj}_{cb} - \widehat{T}^{lj}_{bc}\right] \\
&\quad + K(E^{kj})_{ab} - \tfrac{1}{2} J(E^{kj})_{ab} - \tfrac{1}{2} L^{kl}_{aj} t^l_b + (f^k_a - v^k_a) t^j_b \\
Z^{kj}_{ab} &= J^{kj}_{ab} + J(E^{kj})_{ab} - K^{kl}_{ca} \widehat{T}^{lj}_{bc} - K^{lk}_{aj} t^l_b
\end{aligned}$$

Residuals

$$\begin{aligned}
r^i_a &= f^i_a + (f_{ac} - A_{ac}) t^i_c - \beta_{ki} t^k_a - T^{lk}_{ac} L^{kl}_{ci} + [2T^{ik}_{ac} - T^{ik}_{ca}] v^k_c \\
&\quad + \left[2K(\bar{T}^{ik})_{am} - K(\bar{T}^{ki})_{am}\right] \delta^m_k + [2K^{ik}_{ac} - J^{ik}_{ac}] t^k_c \\
R^{ij}_{ab} &= K^{ij}_{ab} + K(\bar{T}^{ij})_{ab} + \alpha^{ij}_{kl} \bar{T}^{kl}_{ab} + G^{ij}_{ab} + G^{ji}_{ba}
\end{aligned}$$

4.3.4 Retained terms in F12x (x=a,b)

In order to obtain a practical method we now need to decide which terms of the exact CCSD-F12 equations we intend to handle exactly, which ones perturbatively using eq. (4.119), and which ones to neglect completely. As commented on in Sec. 4.3.1, for the F12 part we exclude all non-CEPA(0)D terms from the considerations; compared to the CEPA(0) ones, they are unlikely to be significant. For reference, a variant of the full closed-shell CCSD equations is displayed in Tab. 4.2.

In the F12x methods we employ the fixed amplitude approximation by default. That means that the F12 amplitudes T^{ij}_{mn} are fixed according to their short-range asymptotic behavior as described sections Sec. 3.3 and Sec. 3.4; consequently no F12 equations need to be solved, and the only free wave function amplitudes are t^i_a and T^{ij}_{ab}, just like in conventional CCSD. These t^i_a and T^{ij}_{ab} amplitudes, however, are affected by the F12 terms indirectly. We allow these amplitudes to adjust to the presence of the F12 configurations by including all couplings between F12 and conventional amplitudes self-consistently (i.e., in the iterations). Consequently, in our approximations the energy functional eq. (4.119) is used only for the F12 term, as for the final CCSD amplitudes $r^i_a = 0$ and $R^{ij}_{ab} = 0$. The R^{ij}_{mn} contribution, however, does not vanish, because we use fixed F12 amplitudes.

Before describing the retained terms, let us first note that if *no* coupling between F12 and conventional amplitudes are introduced, and the approximate CCSD-F12 energy is calculated by simply *adding* the MP2-F12 energy correction[5] to the CCSD energy, already very accurate results are obtained [102]. This is somewhat surprising, because in composite quantum chemistry methods it is recognized that the basis set dependence of MP2 and CCSD is slightly different [122,123]. Nevertheless, for small to medium AO basis set sizes, in all investigated cases this correction consistently lead to good results, nearly on par with actual CCSD-F12 approximations [102, suppl. mat.]. One can speculate that the good performance is a result of the asymptotically correct F12 cusp conditions—which, just like the direct F12 energy expressions [eq. (4.111)] themselves—are equal for CCSD and MP2. Consequently, the energy functional is *only* required to account for the γ-dependence of the direct energy expression (4.111). This interpretation is supported by the fact that the difference between the energy functional and the direct energy expression practically vanishes for optimized γ exponents [102].

This simple approach can be made self-consistent by including all double excitation terms arising at second-level perturbation theory (i.e., terms involving the Fock operator f) in the CCSD iterations. That means that in eq. (4.102) we include $T^{ij}_{mn}C^{mn}_{ab}$ and in eq. (4.110) we include $R^{ij,\mathrm{MP2}}_{mn}$, which is evaluated exactly like described in Sec. 4.1, but in both expressions, the MP2 T^{ij}_{ab} amplitudes are replaced by CCSD

[5] The correction is defined as $E_{\mathrm{MP2\text{-}F12}} - E_{\mathrm{MP2}}$, where MP2-F12 is the full MP2-F12 method including F12/conv coupling. MP2-F12 means MP2-F12/3C(FIX) here.

Table 4.3: Illustration of the performance of MP2-level corrections to CCSD(T) energies. The basis set limits are (1.249 ± 2) eV and (2167 ± 1) cm^{-1}, respectively. Calculation details are described in Sec. 5.4.1. $\gamma = 1.0, 1.2, 1.4, 1.6$ are used.

	AVDZ	AVTZ	AVQZ	AV5Z
Electron affinity of C atom/eV				
Pure CCSD(T)	1.143	1.218	1.236	1.242
...+ΔMP2-F12[1]	1.230	1.245	1.251	1.251
...+ΔMP2-F12(EBC)[2]	1.233	1.246	1.250	1.250
...+ΔMP2-F12(SC)[3]	1.229	1.245	1.250	1.250
Vibrational frequency ω_e of CO molecule/cm^{-1}				
Pure CCSD(T)	2104	2144	2160	2164
...+ΔMP2-F12[1]	2158	2162	2166	
...+ΔMP2-F12(EBC)[2]	2160	2162	2166	
...+ΔMP2-F12(SC)[3]	2159	2162	2166	

(1) Adds E[MP2-F12/3C(FIX)]-E[MP2] to CCSD(T) energy
(2) Adds E[MP2-F12/3C(FIX,EBC)]-E[MP2] to CCSD(T) energy
(3) Includes C^{ij} coupling self-consistently in CCSD iterations

T^{ij}_{ab} or $\bar{T}^{ij}_{ab} = T^{ij}_{ab} + t^i_a t^j_b$ as described in Sec. 4.3.2. If only these approximations are made (i.e., if *no* further terms beyond MP2-F12 are considered), the resulting method already works remarkably well. Tab. 4.3 gives an indication of the performance of such MP2-F12-level methods.

Considering that (i) the F12 amplitudes are asymptotically correct (ii) the (dominant) direct F12 energy expression eq. (4.111) is exact and can easily be evaluated exactly and (iii) the second order energy functional is only required to account non-optimized γ-values, and this corrections already works quite well if considering *only* mean-field terms, we figured that we might get away with rather drastic approximations. Thus, the only other major term we include in the F12a approximation is $V^{mn}_{ab} T^{ij}_{mn}$ given in eq. (4.98). This term was chosen for two reasons: First, at CEPA(0)D-level it is the only term in R^{ij}_{ab} that can be calculated without the insertion of eq. (3.32)-style resolutions of the identity; such resolutions are expensive to evaluate, and not as accurate as the other options outlined in Sec. 3.5. It is therefore also the only term in the R^{ij}_{ab} residual that would survive the application of the so called "Standard Approximation" used in first-generation R12 methods [19]. Secondly, the *conventional* term $\left\langle ab \left| r^{-1} \right| cd \right\rangle \bar{T}^{ij}_{cd}$, to which it is a F12 correction, is the only term depending on integrals with four virtual orbitals. Since, unlike the occupied space, the virtual space shrinks when the AO basis set is truncated, such a basis set truncation affects the interaction expressed by this term in the strongest possible

manner. In fact, Ref. [18] argues that the slow basis set convergence of conventional methods results almost exclusively from trying to express quantities like

$$\int_{\boldsymbol{r}_1}\int_{\boldsymbol{r}_2} \varphi_a(\boldsymbol{r}_1)\varphi_b(\boldsymbol{r}_1) r_{12}^{-1} u_{ij}(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\mathrm{d}^3 r_1\,\mathrm{d}^3 r_2 \tag{4.120}$$

by their atom-centered partial wave expansions, which is precisely what the term $\left\langle ab \,\middle|\, r^{-1} \,\middle|\, cd \right\rangle \bar{T}_{cd}^{ij}$ is doing. This is problematic because the exact pair functions $u_{ij}(\boldsymbol{r}_1,\boldsymbol{r}_2)$ have cusps with leading terms $\propto r_{12}$ [see Sec. 3.3], which are supposed to be canceled by r_{12}^{-1} of the Coulomb interaction—but this cancellation cannot be well described when using products of one-particle functions like $T_{ab}^{ij}|ab\rangle$, because they can contain only even orders of r_{12} [Sec. 3.3]. The $V_{ab}^{mn}T_{mn}^{ij}$ term is therefore clearly the dominant F12 contribution to the R_{ab}^{ij} residual. Ref. [18] also concludes that decent partial wave convergence rates are present for all terms in which such a cancellation is not attempted. Thus, the approximation of neglecting nearly all other terms is less severe than it looks like.

At the CCSD level (i.e., when single excitations and non-linear amplitudes are introduced), a few other terms of the general form $\left\langle rs \,\middle|\, r^{-1} \,\middle|\, cd \right\rangle \bar{T}_{cd}^{ij}$ arise [Ref. [88], Tab. 4.2]. Of these, there are two instances depending on 3-external integrals which couple singles- and doubles residuals:

$$R_{ab}^{ij} \mathrel{-}= K(\bar{T}^{ij})_{ak} t_b^k + t_a^k K(\bar{T}^{ij})_{kb} \tag{4.121}$$

$$r_a^i \mathrel{+}= 2K(\bar{T}^{ik})_{ak} - K(\bar{T}^{ki})_{ak} \tag{4.122}$$

For these terms, the F12 corrections $-\left(V_{ak}^{mn}t_b^k + t_a^k V_{kb}^{mn}\right) T_{mn}^{ij}$ and $V_{ak}^{mn}(2T_{mn}^{ik} - T_{mn}^{ki})$, respectively, are included in F12a. The first term is non-linear and does not occur at CEPA(0)-level. It is included mainly because in the actual CCSD program only a single set of $\boldsymbol{K}(\bar{T}^{ij})$ matrices is calculated, and all terms using these matrices are affected. The other term is linear and should be included. Two other non-linear terms in the full CCSD equations are of the general form $\left\langle kl \,\middle|\, r^{-1} \,\middle|\, cd \right\rangle \bar{T}_{cd}^{ij}$; they affect the intermediates $\alpha_{ij,kl}$ and β_{ik}. The F12 corrections for these are neglected.

In total, the complete CCSD-F12a method is given by adding the terms

$$R_{ab}^{ij} \mathrel{+}= \left(V_{ab}^{mn} + C_{ab}^{mn} - V_{ak}^{mn}t_b^k - t_a^k V_{kb}^{mn}\right) T_{mn}^{ij} \tag{4.123}$$

$$r_a^i \mathrel{+}= V_{ak}^{mn}\left(2T_{mn}^{ik} - T_{mn}^{ki}\right) \tag{4.124}$$

to the conventional CCSD residuals R_{ab}^{ij}/r_a^i, and estimating the γ-dependence of the direct energy expression via the double excitation part of the MP2-F12 Hylleraas functional:

$$E_{\mathrm{F12a}} = \left(2\bar{T}_{ab}^{ij} - \bar{T}_{ba}^{ij}\right)K_{ab}^{ij} + \left(2T_{mn}^{ij} - T_{nm}^{ij}\right)\left(V_{mn}^{ij} + R_{mn}^{ij,\mathrm{F12a}}\right) \tag{4.125}$$

$$R_{mn}^{ij,\mathrm{F12a}} = V_{mn}^{ij} - X_{mn,op}\left(g_k^i T_{op}^{kj} + T_{op}^{ik} g_k^j\right) + B_{mn,op} T_{op}^{ij} + C_{ab}^{mn} T_{ab}^{ij}. \tag{4.126}$$

The resulting CCSD-F12a method overestimates correlation energies, but as long as small to medium-sized basis sets are employed, it works excellently for relative energies. However, after the publication of Ref. [102] we realized that its basis set *convergence* for large sets is slow. While results at the AVDZ and AVTZ levels are so good that there is actually *no reason* to try to improve on them, results obtained with AVQZ or AV5Z basis sets do not necessarily lie closer to the basis set limit than results obtained with AVTZ sets. While, in theory, the method must converge to the CBS limit at some point in time—simply because all F12 corrections go to zero—the situation is certainly unsatisfactory.

This overestimation of correlation energies is corrected in the F12b method by including additional terms in the F12 Lagrange functional. Namely, in addition to the terms of eq. (4.126), also the F12 analog of the term $V_{ab}^{mn} T_{mn}^{ij}$ [eq. (4.98)] and its singles-analog are considered in the R_{mn}^{ij} residual. This term stems from the same part of the complete-basis CEPA(0) equations which also gave rise to the $V_{ab}^{mn} T_{mn}^{ij}$ part of F12a's R_{ab}^{ij}. Thus, F12b has one additional residual contribution of

$$R_{mn}^{ij,\mathrm{F12b}} = R_{mn}^{ij,\mathrm{F12a}} + V_{ab}^{mn} \bar{T}_{ab}^{ij} + V_{ib}^{mn} t_b^j + t_a^i V_{aj}^{mn}, \tag{4.127}$$

with V_{rs}^{mn} as defined in (4.107). Apart from that, the same formulas as in F12a are used.

4.3.5 Evaluation of additional F12 matrix elements

Apart from the F12 matrices encountered already at MP2-level [Sec. 4.1.3], also some additional or extended F12 matrices arise in CCSD-F12. Our main interest lies in the tensor

$$\begin{aligned} V_{ab}^{mn} &= \left\langle ab \,\middle|\, r^{-1} Q_{12} F_{12} \,\middle|\, mn \right\rangle \\ &= \left\langle ab \,\middle|\, r^{-1} (1 - |ix\rangle\langle ix| - |xi\rangle\langle xi| - |rs\rangle\langle rs|) F_{12} \,\middle|\, mn \right\rangle \\ &= \left\langle ab \,\middle|\, r^{-1} F_{12} \,\middle|\, mn \right\rangle - K_{ab}^{rs} F_{rs}^{mn} - K_{ab}^{ix} F_{ix}^{mn} - K_{ab}^{xi} F_{xi}^{mn}, \end{aligned} \tag{4.128}$$

which occurs in eqs. (4.99) and (4.107) and which in CCSD-F12 mediates the dominant F12 contribution beyond MP2-F12; it is therefore explicitly handled in the F12x methods. Formally, this tensor is a simple extension of the V_{kl}^{mn} encountered already in eq. (4.30) at MP2 level. However, the lower labels of V_{ab}^{mn} now refer to virtual dimensions instead of the much smaller occupied dimensions, so evaluating it efficiently is more complicated. We discuss eq. (4.128) term-by-term.

The first summand in eq. (4.128) can be evaluated as an analytic integral just like in the MP2 case and does not pose a problem.

The second summand in eq. (4.128), $K^{rs}_{ab}F^{mn}_{rs}$, resembles a similar term of the conventional CCSD equations (eq. (4.92) for CEPA or $K(\bar{T})$ in Tab. 4.2) which has the form

$$K(\boldsymbol{D}^{ij})_{rs} = \left\langle rs \,\middle|\, r^{-1} \,\middle|\, tu \right\rangle D^{ij}_{tu}. \tag{4.129}$$

Due to the massive size of the K^{tu}_{rs} integrals (easily several terabytes), processing this amplitude–4-external-integral contraction is a complex task with high computational cost; typically evaluating eq. (4.92) requires at least half of the time of a CCSD iteration. As a consequence, for this contraction refined special techniques have been developed. In Molpro, the basic contraction (4.129) is performed by transforming D^{ij}_{tu} into the contravariant AO basis, then contracting the resulting $D^{ij,\mu\nu}$ with the raw 4-index AO-basis integrals, and finally transforming the contraction result back into MO basis [124]:

$$K(\boldsymbol{D}^{ij})_{rs} = C^{\rho}_{r}C^{\sigma}_{s}\left[\left\langle \rho\sigma \,\middle|\, r^{-1} \,\middle|\, \mu\nu \right\rangle \left(C^{\mu}_{t}C^{\nu}_{u}D^{ij}_{tu}\right)\right]. \tag{4.130}$$

This way K^{tu}_{rs} never needs to be computed explicitly, and furthermore, the sparsity of the AO basis integrals can be used for computational savings [125].

This technique of calculating the external exchange contribution to the CCSD equations also allows us to treat our additional F12a terms completely transparently. We just replace the default external exchange operators

$$K(\bar{T}^{ij})_{rs} = \left\langle rs \,\middle|\, r^{-1} \,\middle|\, ab \right\rangle \bar{T}^{ij}_{ab} \tag{4.131}$$

by modified versions which have our F12 terms [eqs. (4.98), (4.128)] included:

$$K(\bar{T}^{ij})_{rs} = C^{\rho}_{r}C^{\sigma}_{s}\left[\hat{W}^{mn}_{\rho\sigma}T^{ij}_{mn} + \left\langle \rho\sigma \,\middle|\, r^{-1} \,\middle|\, \mu\nu \right\rangle \left(C^{\mu}_{a}C^{\nu}_{b}\bar{T}^{ij}_{ab} - C^{\mu}_{r}C^{\nu}_{s}F^{mn}_{rs}T^{ij}_{mn}\right)\right] \tag{4.132}$$

where

$$\hat{W}^{mn}_{\rho\sigma} = \left\langle \rho\sigma \,\middle|\, r^{-1}F_{12} \,\middle|\, mn \right\rangle - \left\langle \rho\sigma \,\middle|\, r^{-1}\left[\,|ix\rangle\langle ix| + |xi\rangle\langle xi|\,\right]F_{12} \,\middle|\, mn \right\rangle. \tag{4.133}$$

This way the additional F12a terms due to V^{mn}_{ab} propagate to all occurrences of $K(\bar{T})$ in the CCSD equations [cf. Tab. 4.2] automatically, without any further programming. Since these terms are treated together with the conventional external exchange operators (which must be calculated anyway), their inclusion incurs only insignificant computational cost.

The right hand summand of eq. (4.133) represents the $|ix\rangle\langle ix| + |xi\rangle\langle xi|$ contributions of V^{mn}_{ab} in eq. (4.128). While not quite as fearsome as the external exchange operators themselves, also they pose non-trivial computational problems. Nevertheless, also these $|ix\rangle\langle ix|$ contributions can be calculated without storing large intermediates.

In a density fitting scheme, we have

$$\langle \mu\nu | r_{12}^{-1} | ix \rangle \langle ix | F_{12} | mn \rangle \approx \left[(\nu x|A) D_A^{\mu i} \right] F_{ix}^{mn}, \tag{4.134}$$

where A is a fitting basis index, and $D_A^{\mu i} = J_{AB}^{-1}(\mu i|B)$ is the DF coefficient giving $|\mu i) \approx D_A^{\mu i}|A)$. Since Coulomb integrals are concerned, the fitting formula is robust already in this form [cf. Sec. 2.4, Ref. [32]]. Eq. (4.134) can be calculated for each ν basis function shell independently: For each ν-shell, the integrals $(\nu x|A)$ are computed for all A and x, and contracted first with $D_A^{\mu i}$ and then F_{ix}^{mn} (which are pre-calculated and kept in memory). Once all ν-shells are done, the final result can be assembled and added to eq. (4.133). However, this evaluation requires significant computer time since two virtual indices and one CA index are involved. Numerical experiments have shown that these terms have minor impact on relative energies. Therefore we neglect them by default. Thus, unless otherwise noted, the projector in V_{ab}^{mn} is approximated as $1 - |rs\rangle\langle rs|$ and in eq. (4.133) only the first term on the rhs is kept.

The other new F12 matrix arising at CEPA(0) level is the

$$H_{mn}^{kl} = \left\langle kl \,\middle|\, F_{12} Q_{12} r_{12}^{-1} Q_{12} F_{12} \,\middle|\, mn \right\rangle \tag{4.135}$$

encountered in eq. (4.106) and eq. (4.110). Contrary to the V_{ab}^{ij} term discussed above, the corresponding quantity is not required for F12x calculations. Nevertheless, H_{mn}^{kl} can be calculated with similar techniques: By inserting eq. (3.30) for the right hand Q_{12} operator in eq. (4.135), we obtain

$$\begin{aligned} H_{mn}^{kl} &= \left\langle kl \,\middle|\, F_{12} Q_{12} r_{12}^{-1} F_{12} \,\middle|\, mn \right\rangle - \left\langle kl \,\middle|\, F_{12} Q_{12} r_{12}^{-1} P_{12} F_{12} \,\middle|\, mn \right\rangle \\ &= \widetilde{V}_{mn}^{kl} - V_{rs}^{kl} F_{rs}^{mn} - V_{xi}^{kl} F_{xi}^{mn} - V_{ix}^{kl} F_{ix}^{mn} . \end{aligned} \tag{4.136}$$

Here, for example, $\widetilde{V}_{mn}^{kl}$ is basically a V_{mn}^{kl} [eq. (4.30)] with different analytic integrals:

$$\begin{aligned} \widetilde{V}_{mn}^{kl} &= \left\langle mn \,\middle|\, F_{12} Q_{12} r_{12}^{-1} F_{12} \,\middle|\, kl \right\rangle \\ &= \left\langle mn \,\middle|\, [F_{12}]^2 r_{12}^{-1} \,\middle|\, kl \right\rangle - F_{ix}^{mn} W_{ix}^{kl} - F_{xi}^{mn} W_{xi}^{kl} - F_{rs}^{mn} W_{rs}^{kl} \end{aligned} \tag{4.137}$$

$$W_{uv}^{mn} = \left\langle uv \,\middle|\, F_{12} r_{12}^{-1} \,\middle|\, mn \right\rangle \tag{4.138}$$

while

$$\begin{aligned} V_{xi}^{kl} &= \left\langle xi \,\middle|\, r^{-1} Q_{12} F_{12} \,\middle|\, kl \right\rangle = \left\langle xi \,\middle|\, r^{-1} F_{12} \,\middle|\, kl \right\rangle - \left\langle xi \,\middle|\, r^{-1} P_{12} F_{12} \,\middle|\, kl \right\rangle \\ &= W_{xi}^{kl} - K_{rs}^{xi} F_{rs}^{kl} - K_{yj}^{xi} F_{yj}^{kl} - K_{jy}^{xi} F_{jy}^{kl} \end{aligned} \tag{4.139}$$

is a more complicated version of the tensor (4.128) from before. Assembling H_{mn}^{kl} in

an efficient and numerically stable manner is not trivial, but can be done. However, we did not program its evaluation because, as noted, H_{mn}^{kl} is neglected in the F12x methods.

4.3.6 A possible next step

While the F12a and F12b methods work in practice, and greatly improve on conventional methods, they do have defects. These most likely result from the less than optimal theoretical consistency of some of the approximations employed. It is therefore desirable to develop a CCSD-F12 approximation—at least for single-point energy evaluations—which is more complex, but is also expected to cover all significant F12 terms in some way. In my opinion, such a method should be based on the following principles:

- As in F12a/F12b, all terms arising from second order perturbation theory should be included exactly. These are the dominant contributions.
- As in F12a/F12b, the fixed-amplitude ansatz should be used by default.
- Unlike F12a/F12b, couplings between F12 amplitudes and conventional amplitudes should not be included self-consistently, but rather perturbatively based on the approximate Lagrange functional eq. (4.119). That means that a purely conventional CCSD calculation is followed by a single step, in which F12-including residuals R_{ab}^{ij}, r_a^i and R_{mn}^{ij} are evaluated and used for eq. (4.119).

  This proposition in based on the following reasons, in order of importance:

  – As pointed out by Köhn [44], conventional perturbative (T) corrections calculated from self-consistent F12-T_{ab}^{ij}/t_a^i amplitudes differ more from the CBS limit than the same (T) corrections calculated from purely conventional T_{ab}^{ij}/t_a^i amplitudes. This is important because the residual CCSD(T) basis set error is dominated by (T)-contributions after F12 corrections for CCSD and RHF are applied.
  – When the residuals involving F12 terms are calculated only once, and not once in each iteration, more expensive terms can be included without making the whole method much slower.
  – Nearly all of my calculations, with *very* few exceptions, showed that MP2-F12 methods employing the EBC approximation have a better basis set convergence than the formally more correct MP2-F12 methods without it. This applies to small basis sets in particular.

    My impression is that the "F12 correlation field", expressed by the presence of F12 configurations in the wave function, is actually not such a great description of the exact field [see comments on defects of F12

configurations in Sec. 3.2]. It is likely much better suited for correcting global properties (like energies) than for describing the actual microscopic structure of the wave function, particularly if fixed amplitudes are used. Thus, forcing the conventional amplitudes to adjust to the F12 configurations might not be optimal.

- When using fixed amplitudes, we do not obtain full self-consistency anyway, because the F12 amplitudes are not adjusted. Thus, a quasi-perturbative approach then cannot be avoided.

- For the F12 correction step, all terms arising from CEPA(0)D should be included in both R_{ab}^{ij} and R_{mn}^{ij}, but only to the linear order in the size of the CA basis. For example, for $G_{\alpha\beta}^{ij}$ (eq. (4.94)) this means:
  - G_{xy}^{ij} is neglected completely.
  - G_{xa}^{ij} and G_{ax}^{ij}, which enter the R_{mn}^{ij} residual according to eq. (4.109), are calculated. But all contributions involving contractions over CA orbitals are neglected. This leads to

$$G_{xb}^{ij} \approx \widetilde{T}_{xc}^{ik} K_{cb}^{kj} - T_{xc}^{ik} J_{cb}^{kj} - J_{xc}^{ik} T_{cb}^{kj} \tag{4.140}$$

$$G_{ax}^{ij} \approx \widetilde{T}_{ac}^{ik} K_{cx}^{kj} - T_{ac}^{ik} J_{cx}^{kj} - J_{ac}^{ik} T_{cx}^{kj}. \tag{4.141}$$

  - G_{ab}^{ij} in eq. (4.100) is evaluated as written.

  I expect this combination of approximations to be well-balanced and, at least in the closed-shell case, to closely approximate full CCSD-F12 with linearized T_{mn}^{ij} amplitudes. The latter was shown to be very accurate by Klopper et al. [126]. Due to the linear CA restriction, a good performance in practical matters could still be obtained. Additionally, unnecessary expensive terms could be identified and removed *after* assessing their impact in realistic situations.

Since this method involves lots of terms, a manual implementation would still be rather laborious, despite the approximations employed. This applies particularly as various aspects important for an efficient implementation (e.g., support for molecular symmetry and inline formation of large intermediates) are not easily handled in our current F12 programs. However, these problems can be circumvented. I intent to program this method with the help of a general tensor framework I designed for quantum chemical applications, which allows a high-level description of the algorithms. At the point of writing, the framework is already sufficiently functional that I managed to implement closed-shell CCSD algorithm based on it [using the formulas in Tab. 4.2 and Molpro's native routines for integral transformation and contractions of amplitudes with 4-external and 3-external integrals]. The resulting CCSD program is even competitive with Molpro's native implementation, which is

widely regarded to be one of the fastest in existence. The main issue keeping me from implementing the method right now is that I did not yet program an interface to F12 basis sets and F12 integral shell driver routines. While this is not a major obstacle, it is also not trivial (especially as symmetry-adapted basis functions need to be emulated), and requires some development effort which, at this moment in time, is better spent elsewhere.

4.4 Coarse treatment of perturbative (T)-triples

CCSD itself is not a very accurate method. Most applications demand an inclusion of perturbative (T)-triples correlation energy corrections. But while the (T)-triples contribution is very important to obtain accurate energies, it is also fairly small—typically around one order of magnitude smaller than the CCSD correlation energy itself. For this reason, sufficiently accurate (T) corrections can usually be computed from smaller basis sets than required for obtaining similarly accurate doubles correlation energies or Hartree-Fock energies [62].

Following this reasoning, by default we compute (T)-triples contributions just like in the conventional CCSD(T) method also in CCSD(T)-F12 calculations. Since in our F12 methods the conventional T_1/T_2-amplitudes adjust to the presence of the F12 configurations, the triples energy is still slightly affected—usually it becomes somewhat smaller because the conventional T amplitudes are suppressed. (Meanwhile it turned out that relative (T) correction energies suffer from that [44], but not too badly).

Despite its smallness, the basis set incompleteness error of the (T) correction now often become the main component of the residual basis set error of CCSD(T)-F12 total energy differences. While F12 corrections are also conceivable for (T) corrections [44], the method of calculating them is highly non-trivial, and it is not yet clear if it is possible to evaluate them without increasing the computational cost of the (T) correction. For the time being, we therefore settle for more pragmatic correction approaches. We investigated the performance of two very simple basis set corrections, obtained by scaling different energy contributions.

In the first form, denoted (T*), the (T) triples are scaled by a basis set incompleteness factor estimated from MP2-F12 energies:

$$E[(\mathrm{T}^*)] = E[(\mathrm{T})] \cdot \frac{E_{\mathrm{corr}}^{\mathrm{MP2\text{-}F12}}}{E_{\mathrm{corr}}^{\mathrm{MP2}}}. \tag{4.142}$$

This assumes that the basis set incompleteness affects the perturbative triples energy to a similar degree as it affects the MP2 correlation energy. We chose to base this factor on the MP2-F12/MP2 ratio instead of the CCSD-F12/CCSD ratio mainly because our MP2-F12 treatments are formally exact, while our CCSD-F12 corrections

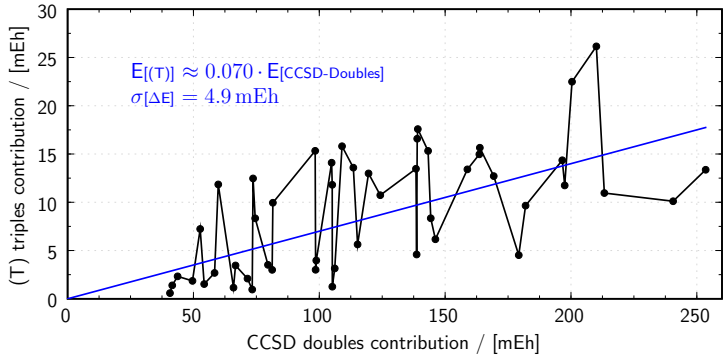

Figure 4.1: Fit of CBS[56] (T) correlation contributions to pair contributions for atomization energies in the AE benchmark set.

are not as complete formally. In section 5 it will be demonstrated that on average this scaling significantly improves relative (T) energy contributions. Unfortunately, there are exceptions (sometimes this factor over-estimates the errors, sometimes it underestimates it), and also it does not always lead to better total energy differences. At this moment it is not clear if the defects are a result of our approximate CCSD-F12 treatments or the (T*) correction itself.

This correction may slightly break size consistence (but not size extensivity), if the ratios $E_{\mathrm{corr}}^{\mathrm{MP2\text{-}F12}}/E_{\mathrm{corr}}^{\mathrm{MP2}}$ are different for different dissociation products of a supersystem. However, this effect is minor and even for the case of atomization energies, which one would expect to be the most susceptible to these kinds of errors, the correction greatly improves the triples contributions.

In the second form of triples corrections, denoted (T$^+$), the F12 basis set correction to the doubles energy is scaled by a fixed factor of 1.070. This value was obtained by least-squares fitting $\alpha \cdot E_{\mathrm{Doubles}}$ to $E_{(\mathrm{T})}$ at conventional CBS[56] level for the atomization energies of the G2 benchmark set (see Sec. 5.1.1). For these 49 molecules, the RMS contribution of the (T)-triples was 11.0 mEh. The (T) energy could be approximated with a RMS deviation of 4.9 mEh by scaling the doubles energy contributions by $\alpha = 0.070$ as shown in Fig. 4.1. Since the fitting was performed at CBS[56], possibly differing basis set convergence rates of (T) and doubles contributions were circumvented and the results should be independent of the actually used basis sets. The (T$^+$)-approach does not suffer from size-consistency defects.

For the scaling approach, I also examined if either fitting total energies instead of relative energies, or using different factors for singlet- and triplet contributions offers advantages. Fitting total energies did not work well for relative energy calculations, although the total CCSD/total (T) energy plot corresponding to Fig. 4.1 looked much

more systematic. Different singlet/triplet factors lead only to a minor reduction in the RMS deviation in the training set. Furthermore, they also yielded imbalanced relative energies for energy differences in practical calculations, particularly if open-shell molecules were involved, so this approach was not followed further.

Like all kinds of F12 corrections, both the (T*) and the (T^+) corrections will eventually converge to the correct limit with increasing basis set size, since F12 corrections go to zero if the orbital basis set becomes complete.

4.5 Technical Aspects

4.5.1 Mixing of exact and density-fitted integrals

Our CCSD(T)-F12 methods are implemented on top of the previous conventional CCSD programs of Molpro. The density fitting approximation does not offer advantages to conventional, non-local CCSD, so in these programs DF approximations are not applied. These methods are also based on conventional RHF usually, where DF approximations do also not occur.

So how should we then handle conventional integral derived quantities, like $f_{\mu\nu}$ Fock matrices or $K^{ij}_{\mu\nu}$ exchange operators? Of these we can choose to either use the exact conventional versions, made by the conventional programs, or to assemble new quantities with density fitting. Unfortunately, the obvious choice of using exact conventional quantities whenever they are present, and never calculating DF quantities of them, fails dreadfully in some non-obvious way: Whenever raw AO basis and RI basis integrals are used to construct CABS quantities (e.g., in equations (4.56) and (4.58)), it is absolutely essential that all the raw integrals are calculated in a compatible manner. For example, in eq. (4.58)

$$f_{xs} = (R^{\mu}_{x} f_{\mu\nu} + R^{\chi}_{x} f_{\chi\nu}) C^{\nu}_{s},$$

either both $f_{\mu\nu}$ and $f_{\chi\nu}$ must be calculated with the DF approximation, or both must be calculated without it. The reason for this is that the CABS transformation $\boldsymbol{R}$ is responsible for projecting AO contributions out of RI functions, and if these contributions are not calculated in exactly the same manner, then the rather small differences between density fitted and exact operators are amplified by many orders of magnitude. Such kinds of problems will often go unseen, but in other cases may lead to badly conditioned F12 singles equations, and to obscure numerical instability problems in the F12 equations.

The solution to this problem is to re-calculate the occurring quantities with DF approximations, and use exact integral quantities only in orthogonal basis sets. For example, a new density fitted $f_{\mu\nu}$ must be assembled to calculate f_{xs} with eq. (4.58), but afterwards the exact $f_{\mu\nu}$ can be used to form f_{rs}, because there the CABS

transformation is not present. Quantities in the orthogonal MO and CA function spaces can be mixed freely. For example, in Tab. 4.1, V^{kl}_{mn} can be calculated from a density fitted K^{kl}_{ix} and an exact K^{kl}_{rs}, but the density fitted $K^{kl}_{\mu\nu}$ must be used to assemble K^{kl}_{ix}.

4.5.2 Freezing of auxiliary basis sets

Sec. 3.6 described how to build a set of orthonormal CA vectors. As explained, it is often necessary to delete some of the generated CA or ABS functions due to a high degree of near-linear dependency, because otherwise the numerical stability of subsequent calculations would suffer. While this is not a serious problem for single point calculations, it can become one when evaluating local potential energy surfaces (for example, in calculations of molecular vibrational frequencies). This deletion may introduce small, unphysical steps in the surface which can produce unpleasant side effects, like increased basis set deviations, bad convergence in geometry optimizations, symmetry breaking, and so on. These steps occur when in a surface calculation different auxiliary functions are deleted at adjacent molecular geometries; an effect which is caused by functions either slipping above or below the numerical CABS thresholds, or by being ordered in a different manner.

This problem can be avoided by *freezing* the auxiliary basis set at some reference geometry. For the distorted geometries, the auxiliary sets are then built in such a way that they resemble the frozen set as far as possible. Concretely, that means that if auxiliary vectors must be deleted, we attempt to delete exactly these ones which correspond to the vectors deleted at the reference point.

In order to accomplish this, first the deleted vectors $\boldsymbol{D}$[6] must be stored at the reference geometry. This can be, for example, the neutral geometry when evaluating numerical energy derivatives (in Molpro this is indicated to the program automatically for the commands `optg`, `freq`, and `surf`). At the distorted geometry, the set of new orthonormal basis vectors $\boldsymbol{V}$ is then first built as usual, but with CABS thresholds set to zero, such that no functions are removed. Then the vectors of $\boldsymbol{V}$ which resemble the deleted reference vectors most closely are deleted. For this, first the deleted reference vectors $\boldsymbol{D}$ are loaded from disk and orthonormalized with respect to the new basis functions:

$$\boldsymbol{S}' = \boldsymbol{D}^T \boldsymbol{S} \boldsymbol{D} \tag{4.143}$$

$$\boldsymbol{D}' = \boldsymbol{D} \boldsymbol{S}'^{-\frac{1}{2}}, \tag{4.144}$$

where $\boldsymbol{D}'$ are the components of the orthonormalized deleted vectors, and $\boldsymbol{S}$ is the

[6] Rows: components of AO and/or RI basis functions. For CABS vectors, both AO and RI components are required. Each column stands for one vector.

basis function overlap matrix at the new geometry (which, for example, would be assembled from the AO/AO, AO/RI and RI/RI overlap matrices when considering CABS vectors). Then for each of the new vectors $\boldsymbol{v}$ of $\boldsymbol{V}$, the squared norm of its projection $\sum_d |d'\rangle \langle d'||v\rangle$ onto the linear span of the orthonormalized deleted vectors is determined:

$$\begin{aligned} p(v) &= \sum_{d'} \left\langle v \,\middle|\, d' \right\rangle \left\langle d' \,\middle|\, v \right\rangle \qquad \text{(no sum over } v\text{)} \\ &= \left(\boldsymbol{v}^T \boldsymbol{S} \boldsymbol{D'}\right) \left(\boldsymbol{D'}^T \boldsymbol{S} \boldsymbol{v}\right). \end{aligned} \tag{4.145}$$

The new vectors $\boldsymbol{V}$ are then ordered by their overlap with the deleted space $p(v)$, and the n_{del} vectors with the highest $p(v)$ values are discarded, where n_{del} is the number of vectors deleted at the reference geometry (=number of columns in $\boldsymbol{D}$, $\boldsymbol{D'}$, and $\boldsymbol{S'}$).

The described process is not completely safe when degenerate eigenvectors occur, because the eigenvectors might be mixed differently at the the reference geometry and at the distorted geometry. This problem can be bypassed to a high degree by making sure that if one nearly-degenerate eigenvector is deleted at the reference geometry, then all the other ones are deleted, too. In Molpro this is accomplished by slightly modifying the CA construction scheme in Sec. 3.6: When a vector is deleted due to its overlap eigenvalue, all vectors with overlap eigenvalues differing by less than 10% from it are also deleted, even if it they do not exceed the CABS thresholds. This process is repeated until there is more than 10% difference between the last deleted vector and the first non-deleted vector.

This freezing process must be done independently for each basis set that is to be frozen. For example, when the CABS is built from an ABS like described at the end of Sec. 3.6, it must be done for both the ABS and the CABS (and both times relative to the raw basis functions). Molpro can also use different CA basis sets for some different F12 terms, and then the process is again done independently for the ABS and CA sets of each.

4.6 Summary & Outlook: F12 Methods

F12 quantum chemistry methods — General

- Kutzelnigg & Klopper have shown that wave functions expressed in orbital expansions converge fast in general shape, but slow for the $r_{12} \to 0$ regions. As remedy they proposed to augment the conventional doubly excited configurations by additional explicitly correlated configurations [18–20].

- This extensions leads to tremendous enhancements in AO basis set convergence, but building practical F12 methods is quite sophisticated: There are many new types integrals (some depending on three electrons, making RI approximation necessary); some badly conditioned equations leading to numerical instabilities; there are lots of possible approximations, of which the impact must be carefully tested, etc. Constructing robust explicitly correlated methods becomes challenging even at MP2 level.

- While originally linear correlation factors were used ("R12 methods"), nonlinear short-range correlation factors F_{12} yield drastically improved accuracy compared with $F_{12} = r_{12}$ [33, 34].

F12 quantum chemistry methods — Techniques

- F12 methods can be understood in terms of standard second quantization. The (exact) full space of doubly excited configurations is approximated with two subsets:

$$T^{kl}_{\gamma\delta} e^{\gamma\delta}_{kl} \approx T^{kl}_{cd} e^{cd}_{kl} + T^{kl}_{mn} \mathcal{F}^{mn}_{\gamma\delta} e^{\gamma\delta}_{kl} \tag{4.146}$$

  (γ, δ: virtual orbitals regarding a fictional complete basis). The first term contains conventional double excitations, while the second term represents F12 pair functions. Here

$$\mathcal{F}^{mn}_{\alpha\beta} := \langle \alpha\beta \,|\, Q_{12} F_{12} \,|\, mn \rangle \,,$$
$$Q_{12} = (1 - o_1)(1 - o_2)(1 - v_1 v_2)$$

  are the projected correlation factor and a projector. As pair correlation factor $F_{12} = \sum c_i \exp(-\gamma_i r_{12}^2)$, fitted to a Slater function, is used in our work.

- F12 configurations can be viewed as "externally contracted" doubly excited configurations. They are short ranged and well suited to describe the correlation cusp of the wave function for $r_{12} \to 0$ and thus *dynamical* correlations.

- The projector Q_{12} ensures orthogonality to the occupied and conventional doubly excited space ($\mathcal{F}^{mn}_{rs} = 0$).

- In principle, the equations of any F12 method can be be constructed in a clear-cut way. First, the conventional residual equations are recast into a form referring to a fictional complete virtual basis (CVB) set. Then, in these equations, the F12 ansatz for the doubles excitation operator is inserted and the equations are resolved and simplified with RI and other approximations.

About RMP2-F12

- Based on the outlined F12 techniques, generally known from closed-shell theory (e.g., Ref. [36]), I derived RMP2-F12 equations and constructed a DF-RMP2-F12 program.
- The program can evaluate the theory virtually exactly, the only necessary approximations being DF- and RI approximations (the first of which is very accurate and easy to control, and the second cannot be avoided). Unlike in most early R12/F12 methods, no terms are neglected a-priori, not the least because this turned out to be unnecessary.
- The theory was derived in a systematic and transparent way. Aspects specific to open-shell theories were clarified. Derivations were based on second quantization, such that connections to higher many-body methods could be seen.

About full CCSD-F12

- Full CCSD-F12 iterations require at least ten times the cost of conventional CCSD.
- Full CCSD-F12 would likely face serious numerical stability problems if applied to larger molecules. A really efficient implementation is most challenging.
- When presenting their generated full CCSD-F12 method, even Shiozaki et al. acknowledged that simplified methods are likely better suited for practical calculations than full CCSD-F12 [42].

Our simplified method: CCSD-F12x (x=a/b)

- Only the important couplings between conventional and F12 configurations are kept: Mean-field couplings (MP2-like) and couplings over integrals with three or four external indices. These are the terms most severely affected by finite basis set truncations.
- The F12 configurations are fixed to their short-range asymptotic cusp behavior as suggested by Ten-no [35]. The error thus introduced is estimated by a truncated F12-Lagrangian.

- Consequently, only few terms additional to CCSD and MP2-F12 arise. The methods are very simple.
  - ⇒ There is nearly no additional cost of CCSD-F12x compared to conventional CCSD.
  - ⇒ The numerical stability is decent.
  - ⇒ A timely extension to local DF-LCCSD possible. For the closed shell method, the simulated method is already implemented and shows promising results.
- Despite its simplicity, the accuracy of CCSD(T)-F12x still is very good: Deviation of E_{corr} from CBS is not worse than in full CCSD-F12 [103]. Relative energies are about as good as with the more rigorous CCSD(F12) method [38–40] from the Klopper group [102, 103].

Corrections to Hartree-Fock energies

- If left uncorrected, the Hartree-Fock basis set deviation from CBS is a severe problem in F12 calculations. Often it is larger than the remaining correlation error, especially for AVTZ and smaller basis sets.
- In the MP2-F12 calculations, a large auxiliary basis set $|x\rangle$ is employed, and Fock matrices f_{rs}, f_{xr} and f_{xy} are required.
- These Fock matrices can be reused for a perturbative correction to Hartree-Fock energies, which in the closed-shell case would look like

$$E_s = -\sum_{i,\alpha} \frac{2(f^i_\alpha)^2}{\varepsilon_\alpha - \varepsilon_i} \quad \text{where} \quad |\alpha\rangle\langle\alpha| = |a\rangle\langle a| + |x\rangle\langle x|$$

- This correction typically reduces RHF errors by one order of magnitude (Refs. [102, 103, 127], vide infra). Afterwards they are not problematic anymore.

Outlook

- A development of local variants of the open-shell coupled cluster methods would be very helpful for applications to larger molecules. For the closed-shell programs, local approximations for R12 and F12 methods have already been demonstrated [36, 47, 48], and recently it was shown that these local F12 variants can be done in such a way that the inherent domain approximation error of local methods is greatly diminished [49–51]. These articles still refer to simulated variants of the closed-shell LCCSD-F12 methods, but a real implementation is underway, and a timely extension to open-shell cases is expected once Molpro's DF-LUCCSD(T) program is in a stable form.

- It would be helpful if, in a single calculation, multiple different F12 correlation factors could be employed instead of a single Slater geminal. This could aid the basis set convergence for calculations of core-valence correlation energies and for other tasks where relevant correlation effects occur on multiple length scales.

  In the absence of DF approximations, this problem could perhaps be addressed directly at the integral level: For example, when calculating F12 integral shells like $\langle \mu\nu \mid F_{12} r^{-1} \mid \rho\sigma \rangle$, one could simply load different Slater geminal expansions based on the current shells of μ and ν basis functions. This scheme would even allow retaining the fixed amplitude approximation without major issues. In our DF methods, however, integrals are never calculated in this form and some other way of achieving a similar effect should be sought.

- One shortcoming of the presented CCSD(T)-F12x methods is that no direct F12 contributions to (T)-triples of CCSD(T) are included. In some cases (see below) this causes the major part of the residual basis set deviation. This problem can be approached with the methods proposed by Köhn [44] as outlined in Sec. 3.9.2.

- Analytic energy gradient methods for F12 methods are being developed in our group at the present time. These will be most helpful for geometry optimizations and potential energy surface constructions.

5 Benchmark Calculations

For novel computational methods it is essential that extensive benchmark calculations are performed; especially if the methods are intended to have "black box" character, like in our case. There are several reasons for this:

- The range of applicability must be carefully investigated, particularly if approximations are made of which the impact is not completely transparent. For the present methods this applies to numerical stability thresholds in the CABS construction, the fixed amplitude ansatz, hybrid approximations, and the F12x (x=a,b) treatments.
- The expected accuracy under various conditions must be established. In particular, good accuracy/computation time tradeoffs with regard to the used orbital- and auxiliary basis sets should be determined. Good basis set/geminal exponent combinations must be found.
- Trials for errors, numerical instabilities, and non-convergence issues must be performed on a large set of different molecules and different application types in order to find and repair occurring problems.

Consequently, the larger timeshare of my thesis project was spent performing, evaluating and responding to such calculations. In fact, the Python script I developed for benchmarking contains about as many lines of code as the RMP2-F12 program itself.

The benchmark section is organized in the following way: Sec. 5.1 introduces benchmark sets of atomic and molecular systems. These are used for determining the general effectiveness of the theory in Sec. 5.2 and for optimizing combinations of calculation parameters in Sec. 5.3. Sec. 5.4 then probes the usefulness of the theory in more complicated kinds of calculations.

5.1 Investigated systems

5.1.1 Atomization energies (AE), electron affinities (EA) and ionization potentials (IP)

We investigated atomization energies (AEs), electron affinities (EAs) and ionization potentials (IPs) from the original G2 test set by Curtiss et al. [1]. These systems

Table 5.1: Systems originating from the G2 benchmark set. EAs and IPs are adiabatic. RMP2/AVTZ geometries were used.

Electron Affinities: NO, CH_3, NH, O_2, CH_2, P, NH_2, PH, SiH_2, PO, CH, PH_2, SiH, C, SiH_3, Si, O, S_2, OH, S, SH, Cl_2, F, Cl, CN
Ionization Potentials: Al, Si, B, S_2, PH_3, $PH_2 \rightarrow PH_2^+(^1A_1)$, PH, S, NH_3, SH, $SH_2 \rightarrow SH_2^+$ ($^2B_1$), P, C_2H_4, P_2, SiH_4, C, CS, Cl_2, C_2H_2, O_2, ClF, $SH_2 \rightarrow SH_2^+$ ($^2A_1$), HCl, H_2O, CH_4, Cl, OH, O, CO, N, $N_2 \rightarrow N_2^+$ ($^2\Sigma_g$), HF, $N_2 \rightarrow N_2^+$ ($^2\Pi_u$), F
Atomization Energies: F_2, Cl_2, ClO, ClF, Si_2, NH, CH, HCl, OH, S_2, P_2, SiH_2 ($^3B_1$), SO, O_2, HF, SiH_2 ($^1A_1$), PH_2, NO, HOCl, CH_2 ($^1A_1$), CS, SH_2, NH_2, CN, CH_2 ($^3B_1$), SiO, SiH_3, PH_3, H_2O, N_2, CO, H_2O_2, SO_2, HCO, NH_3, CH_3, SiH_4, HCN, H_2CO, CH_3Cl, C_2H_2, CO_2, CH_4, N_2H_4, CH_3SH, CH_3OH, Si_2H_6, C_2H_4, C_2H_6

represent a broad range of different atoms and bonding situations. Additionally, they allow for accurate reference calculations because all of the contained molecules are very small. We excluded molecules containing alkali or alkali-earth atoms because JKFIT auxiliary basis sets are not yet available for them. Apart from that, all G2 systems were used. The systems retained are listed in Tab. 5.1.

For all molecules RMP2/AVTZ optimum geometries with tight thresholds were calculated. The IPs and EAs are adiabatic (i.e., the geometries for the neutral and ionic species are different).

AEs, IPs and EAs were benchmarked as representatives of certain worst case scenarios:

- An atomization energy is calculated as the difference between the energy of the molecule and the energies of its individual atoms. Since the electronic structures are very different, fortuitous error cancellations are unlikely. Furthermore, there is a huge BSSE effect, and whether its counter-poise correction is at all applicable for AEs is disputed.

- Ionization potentials and electron affinities are calculated as differences between species with a different number of electrons. For the additional/missing electron, there is nothing that could cancel an error in its description, and even highly systematic errors related to the interaction with the remaining electrons would inevitably show up in benchmark results.

For reference calculations, basis set limits were estimated with purely conventional methods (no F12, no DF), using l^3 extrapolation based on the AVnZ basis sets according to Sec. 2.3 [AVnZ is defined in Sec. 5.2]. The CBS limit estimate obtained from AV(n-1)Z and AVnZ calculations is denoted as CBS[(n-1)n] (e.g., CBS[45] for

AVQZ/AV5Z extrapolated energies). For the AE/IP/EA sets, CBS[56] reference values could be obtained for all systems.

5.1.2 Reaction energies of closed-shell (REc) and open-shell (REo) reactions

Reaction energies were investigated in order to evaluate the practical performance of the methods. Many reactions are taken from a benchmark set developed previously [128]. Further reactions, which can be formed with additional molecules from the AE/IP/EA set described above, were added by Thomas B. Adler.

All molecular geometries were optimized at the MP2/RMP2 level with AVTZ basis sets. The geometries and reference values can be found in the supplementary material of Ref. [103].

In total, there are 54 reactions involving exclusively closed-shell molecules, and 50 further reactions involving also open-shell molecules or atoms. The reference RHF, CCSD, and CCSD(T) values for the closed- and open-shell reactions are listed in Tables 5.2 and 5.3, respectively. We note that on the average the open-shell reactions studied here have larger reaction energies than the closed-shell reactions. In fact, there are only fourteen open-shell reactions with less than 200 kJ/mol reaction energy, while thirty-nine closed-shell reactions fall into this category. Therefore, larger mean absolute errors can be anticipated in the open-shell case. In most calculations in the following, the sets REc and REo are combined into a single set RE.

Reference calculations were done as for the AE/IP/EA set. Unfortunately, for some of the systems in the reaction set AV6Z calculations would have been too expensive. For these systems CBS[45] reference values are used; they are marked in Tabs. 5.2 and 5.3. These reactions without AV6Z/CBS[56] reference values were excluded from some benchmarks (either due to insufficient accuracy of CBS[45] or to save computer time). If done, this is indicated in affected statistics.

Table 5.2: Closed-shell reactions and CCSD(T) reference values [kJ/mol]

	Reaction	Total	RHF	ΔCCSD	Δ(T)
1	$CH_3COOH + NH_3 \rightarrow CH_3CONH2 + H_2O$	-0.13[a]	2.73	-3.05	0.19
2	$HCOOH + NH_3 \rightarrow HCONH_2 + H_2O$	-4.47[a]	-2.90	-1.90	0.32
3	$CH_3ONO \rightarrow CH_3NO_2$	-9.08[a]	3.35	-9.21	-3.22
4	$H_2SO_3 \rightarrow SO_2 + H_2O$	-9.12[a]	11.38	-12.56	-7.95
5	$HCONH_2 + CH_3OH \rightarrow HCOOCH_3 + NH_3$	-14.32[a]	-6.79	-5.22	-2.31
6	$H_2C{=}C{=}O + HCHO \rightarrow C_2H_4O + CO$	-16.97[a]	-24.51	2.98	4.56
7	$HCOOH + CH_3OH \rightarrow HCOOCH_3 + H_2O$	-18.79[a]	-9.69	-7.11	-1.99
8	$CO + H_2 \rightarrow HCHO$	-21.64	1.06	-23.43	0.73
9	$CO + H_2O \rightarrow CO_2 + H_2$	-26.88	0.24	-16.96	-10.16
10	$CH_3OH + HCl \rightarrow CH_3Cl + H_2O$	-33.61	-25.11	-6.66	-1.83
11	$H_2O + CO \rightarrow HCOOH$	-37.74	-7.59	-26.09	-4.05

12	$CO + NH_3 \rightarrow HCONH_2$	-41.92[a]	-10.49	-27.71	-3.73
13	$CH_3OH + H_2S \rightarrow CH_3SH + H_2O$	-45.59	-32.17	-11.29	-2.13
14	$CS_2 + 2\ H_2O \rightarrow CO_2 + 2\ H_2S$	-47.36	-122.28	57.28	17.65
15	$Furane + NH_3 \rightarrow Pyrrole + H_2O$	-54.43[a]	-46.09	-8.29	-0.05
16	$CO + CH_3OH \rightarrow HCOOCH_3$	-56.24[a]	-17.28	-32.92	-6.03
17	$C_2H_6 + H_2 \rightarrow 2\ CH_4$	-75.69	-88.77	10.64	2.44
18	$Furane + H_2S \rightarrow Thiophene + H_2O$	-80.87[a]	-49.65	-24.44	-6.77
19	$CH_3CN + H_2O \rightarrow CH_3CONH_2$	-81.10[a]	-80.31	-0.20	-0.59
20	$HNCO + H_2O \rightarrow CO_2 + NH_3$	-85.93	-96.84	9.60	1.30
21	$HNCO + NH_3 \rightarrow NH_2CONH_2$	-86.28[a]	-88.07	-3.88	5.67
22	$HCN + H_2O \rightarrow HCONH_2$	-92.02[a]	-98.95	6.85	0.08
23	$CH_4 + Cl_2 \rightarrow CH_3Cl + HCl$	-98.17	-110.01	8.81	3.03
24	$COCl_2 + 2\ NH_3 \rightarrow NH_2CONH_2 + 2\ HCl$	-102.72[a]	-147.80	33.97	11.11
25	$Cl_2 + F_2 \rightarrow 2\ ClF$	-113.25	-142.44	22.96	6.22
26	$CO + Cl_2 \rightarrow COCl_2$	-114.79	-56.53	-48.17	-10.09
27	$CO_2 + 3\ H_2 \rightarrow CH_3OH + H_2O$	-117.85	-118.69	-14.66	15.49
28	$C_2H_4 + Cl_2 \rightarrow C_2H_3Cl + HCl$	-117.85[a]	-117.09	-1.24	0.48
29	$HCHO + H_2 \rightarrow CH_3OH$	-123.09	-119.50	-8.19	4.61
30	$C_2H_2 + HCl \rightarrow C_2H_3Cl$	-132.91[a]	-120.03	-11.00	-1.89
31	$HCN + NH_3 + CO + H_2O \rightarrow 2\ HCONH_2$	-133.94[a]	-109.43	-20.85	-3.65
32	$CO + 2\ H_2 \rightarrow CH_3OH$	-144.73	-118.45	-31.62	5.34
33	$H_2C{=}C{=}O + H_2O \rightarrow CH_3COOH$	-156.66[a]	-152.44	-5.74	1.52
34	$H_2C{=}C{=}O + NH_3 \rightarrow CH_3CONH_2$	-156.79[a]	-149.71	-8.79	1.71
35	$C_2H_2 + H_2O \rightarrow CH_3CHO$	-161.15[a]	-169.99	7.75	1.09
36	$C_2H_4 + H_2 \rightarrow C_2H_6$	-165.48	-163.86	-5.92	4.30
37	$SO_3 + CO \rightarrow SO_2 + CO_2$	-180.95	-159.44	-15.12	-6.39
38	$C_2H_2 + HCN \rightarrow C_2H_3CN$	-185.16[a]	-179.87	-4.07	-1.22
39	$H_2 + Cl_2 \rightarrow 2\ HCl$	-190.59	-213.35	15.98	6.78
40	$C_2H_2 + H_2 \rightarrow C_2H_4$	-206.34	-216.39	5.68	4.37
41	$SO_2 + H_2O_2 \rightarrow SO_3 + H_2O$	-210.97	-231.70	16.38	4.34
42	$C_2H_4 + H_2O_2 \rightarrow C_2H_4O + H_2O$	-216.13[a]	-223.12	2.77	4.22
43	$CO + 3\ H_2 \rightarrow CH_4 + H_2O$	-270.76	-246.90	-31.12	7.26
44	$HCN + 3\ H_2 \rightarrow CH_4 + NH_3$	-321.02	-335.39	3.30	11.06
45	$H_2O_2 + H_2 \rightarrow 2\ H_2O$	-365.04	-391.37	18.22	8.11
46	$CH_3CHO + H_2O_2 \rightarrow CH_3COOH + H_2O$	-379.35[a]	-398.48	15.40	3.74
47	$CO + H_2O_2 \rightarrow CO_2 + H_2O$	-391.92	-391.13	1.26	-2.05
48	$2\ NH_3 + 3\ Cl_2 \rightarrow N_2 + 6\ HCl$	-407.43	-482.76	63.68	11.65
49	$3\ N_2H_4 \rightarrow 4\ NH_3 + N_2$	-439.08	-470.67	32.16	-0.57
50	$H_2 + F_2 \rightarrow 2\ HF$	-564.97	-610.74	32.89	12.88
51	$CH_3NH_2 + 3\ H_2O_2 \rightarrow CH_3NO_2 + 4\ H_2O$	-601.08[a]	-603.36	9.61	-7.33
52	$NH_3 + 4\ H_2O_2 \rightarrow HNO_3 + 5\ H_2O$	-746.89[a]	-759.28	17.75	-5.35
53	$CH_4 + 4\ H_2O_2 \rightarrow CO_2 + 6\ H_2O$	-1216.29	-1318.35	87.04	15.02
54	$2\ NH_3 + 3\ F_2 \rightarrow N_2 + 6\ HF$	-1530.57	-1674.93	114.42	29.93

(*a*): Row was calculated from CCSD(T)/CBS[45] values instead of CBS[56]

Table 5.3: Open-shell reactions and RHF-UCCSD(T) reference values [kJ/mol]

	Reaction	Total	RHF	ΔCCSD	Δ(T)
1	$HCl + H \rightarrow Cl + H_2$	-9.21	-12.52	-2.80	6.10

2	$H_2O + F_2 \rightarrow 2\ HF + O$	-49.82	-292.86	215.37	27.66
3	$CH_4 + OH \rightarrow CH_3 + H_2O$	-55.27	-9.73	-41.94	-3.60
4	$OH + H_2 \rightarrow H_2O + H$	-67.44	-27.53	-32.16	-7.75
5	$CO + OH \rightarrow CO_2 + H$	-94.32	-27.29	-49.12	-17.91
6	$S + 2\ HCl \rightarrow H_2S + Cl_2$	-118.93	5.02	-107.94	-16.02
7	$CH_3 + Cl_2 \rightarrow CH_3Cl + Cl$	-119.55	-140.33	15.80	4.97
8	$2\ NO + O_2 \rightarrow 2\ NO_2$	-124.25	55.87	-137.57	-42.54
9	$N + O_2 \rightarrow NO + O$	-140.59	-146.48	2.58	3.31
10	$4\ HCl + O_2 \rightarrow 2\ H_2O + 2\ Cl_2$	-154.42	-129.15	-25.67	0.40
11	$NO + H_2O_2 \rightarrow NO_2 + H_2O$	-159.37	-85.52	-53.71	-20.14
12	$2\ NO \rightarrow N_2 + O_2$	-173.60	-118.61	-50.82	-4.17
13	$2\ H_2O_2 \rightarrow 2\ H_2O + O_2$	-194.49	-226.90	30.16	2.25
14	$Cl_2 + H \rightarrow HCl + Cl$	-199.80	-225.87	13.18	12.89
15	$2\ SO_2 + O_2 \rightarrow 2\ SO_3$	-227.46	-236.50	2.60	6.44
16	$Cl + OH \rightarrow HOCl$	-246.11	-72.33	-152.64	-21.14
17	$4\ NO_2 + O_2 + 2\ H_2O \rightarrow 4\ HNO_3$	-252.28[a]	-218.35	-46.76	12.83
18	$H_2S + F_2 \rightarrow S + 2\ HF$	-255.45	-402.41	124.85	22.11
19	$2\ NH_2 \rightarrow N_2H_4$	-304.96	-162.50	-128.20	-14.27
20	$NO + N \rightarrow O + N_2$	-314.18	-265.09	-48.24	-0.86
21	$O + 2\ HCl \rightarrow H_2O + Cl_2$	-324.56	-104.53	-198.45	-21.57
22	$2\ SiH_3 \rightarrow Si_2H_6$	-332.57	-246.87	-78.86	-6.83
23	$CH_3 + SH \rightarrow CH_3SH$	-333.38	-223.96	-96.44	-12.98
24	$SO_2 + O \rightarrow SO_3$	-361.08	-158.21	-184.32	-18.55
25	$CS + O \rightarrow CO + S$	-366.89	-320.32	-52.72	6.15
26	$CH_3 + Cl \rightarrow CH_3Cl$	-368.62	-252.85	-101.75	-14.02
27	$CH_3OH + O \rightarrow HCHO + H_2O$	-392.06	-198.38	-174.29	-19.39
28	$CH_3 + OH \rightarrow CH_3OH$	-411.67	-267.79	-130.04	-13.84
29	$NH + H \rightarrow NH_2$	-415.32	-285.49	-124.28	-5.54
30	$Si + 2\ H_2 \rightarrow SiH_4$	-442.60	-385.34	-53.98	-3.29
31	$2\ C_2H_4 + O_2 \rightarrow 2\ CH_3CHO$	-446.20[a]	-463.08	9.53	7.36
32	$CS + S \rightarrow CS_2$	-450.31	-296.57	-128.41	-25.33
33	$NH_2 + H \rightarrow NH_3$	-482.10	-361.36	-114.96	-5.78
34	$2\ H_2 + O_2 \rightarrow 2\ H_2O$	-535.60	-555.84	6.28	13.97
35	$CO_2 + C \rightarrow 2\ CO$	-539.89	-438.83	-91.91	-9.16
36	$CO + O \rightarrow CO_2$	-542.03	-317.64	-199.44	-24.94
37	$C + H_2O \rightarrow CO + H_2$	-566.77	-438.59	-108.87	-19.31
38	$N_2H_4 + O_2 \rightarrow N_2 + 2\ H_2O$	-572.40	-607.88	27.50	7.98
39	$2\ NH \rightarrow N_2 + H_2$	-714.31	-434.59	-248.37	-31.34
40	$C + S_2 \rightarrow CS_2$	-732.75	-527.19	-172.71	-32.84
41	$2\ CO + 2\ NO \rightarrow N_2 + 2\ CO_2$	-762.95	-673.97	-78.46	-10.52
42	$CH_4 + 2\ O_2 \rightarrow CO_2 + 2\ H_2O$	-827.31	-864.55	26.72	10.51
43	$4\ NH_3 + 5\ O_2 \rightarrow 4\ NO + 6\ H_2O$	-930.91	-1115.75	151.99	32.84
44	$2\ NH_3 + 2\ NO + O \rightarrow 2\ N_2 + 3\ H_2O$	-1060.00	-835.05	-211.26	-13.69
45	$C + O_2 \rightarrow CO_2$	-1129.25	-994.19	-119.56	-15.50
46	$CS_2 + 3\ O_2 \rightarrow CO_2 + 2\ SO_2$	-1139.32	-1097.09	-42.07	-0.16
47	$CH_4 + 4\ NO \rightarrow 2\ N_2 + CO_2 + 2\ H_2O$	-1174.51	-1101.76	-74.92	2.17
48	$CH_4 + NH_3 + 3\ O \rightarrow HCN + 3\ H_2O$	-1224.42	-618.26	-550.74	-55.42
49	$2\ C + H_2 \rightarrow C_2H_2$	-1227.56	-901.96	-290.39	-35.20
50	$4\ NH_3 + 3\ O_2 \rightarrow 2\ N_2 + 6\ H_2O$	-1278.10	-1352.96	50.35	24.50

(*a*): Row was calculated from CCSD(T)/CBS[45] values instead of CBS[56]

5.2 Basic performance of the theory

This section is intended to demonstrate the general fitness of the developed methods. All of the following choices are default parameters in Molpro 2009.2. The calculations here thus represent black-box calculations that can be done by simply replacing commands like `uccsd(t)` by `uccsd(t)-f12` in the input file.

Calculations are carried out with the aug-cc-pVnZ basis sets [129] as orbital basis for first-row elements, and with the aug-cc-pV(n+d)Z basis sets [130] for second-row elements. These sets are denoted as "AVnZ" in the remainder of this text.

In DF-MP2-F12 calculations we require three auxiliary basis sets in addition to the orbital basis set: The RI basis set used to expand the CABS vectors [Sec. 3.6], a DF basis set for calculating Fock operators (JKFIT) and a DF basis set for other integral quantities (MP2FIT). Within this section the cc-pVnZ/JKFIT basis sets of Weigend [131] were used for both RI and DF/JKFIT. For the double-ζ variants no dedicated JKFIT sets are available, so in this case the VTZ/JKFIT sets were used, but the functions with the highest angular momentum were omitted (this is the default setting in Molpro 2009.2). For DF/MP2FIT, the MP2FIT sets of Weigend [132] were used for AVnZ (n=D,T,Q) orbital bases, and for AV5Z, the MP2FIT sets of Hättig [133] were used in that role.

For the actual CABS approximations, we employ orthonormal CABS orbitals which are formed in a two-step procedure (`ortho_cabs=1`) as described in Sec. 3.6. The Molpro 2009.2 default thresholds `thr_cabs=1e-8` and `thr_cabs_rel=1e-9` are used unless indicated otherwise.

As F12 correlation factor, we use a Slater geminal with exponent γ, which is fitted to six Gaussians as in eq. (3.11). The Gaussian geminal expansion coefficients are fitted as described in Ref. [36]. Unless otherwise noted, γ is set to 1.0 (in inverse atomic length units, $\equiv 1.0\, a_{\mathrm{Bohr}}^{-1}$).

5.2.1 RHF: Perturbative corrections to Hartree-Fock

We first investigate the basis set convergence of relative Hartree-Fock energies and, especially, its improvement due to methods described in Sec. 4.2. There are two basic ways of applying the correction:

CABS-singles: We add the energy difference of *"RMP2-Singles into the combined MO and CA basis"* minus *"RMP2-singles into the MO basis alone"* to the RHF energy.

CABS-relax: The energy correction from the spin-adapted singles perturbation theory is added to the RHF energy.

As described in Sec. 4.2, the first correction contains a part of the RMP2 singles correlation energy, while the second does not. Both approaches are equivalent in the

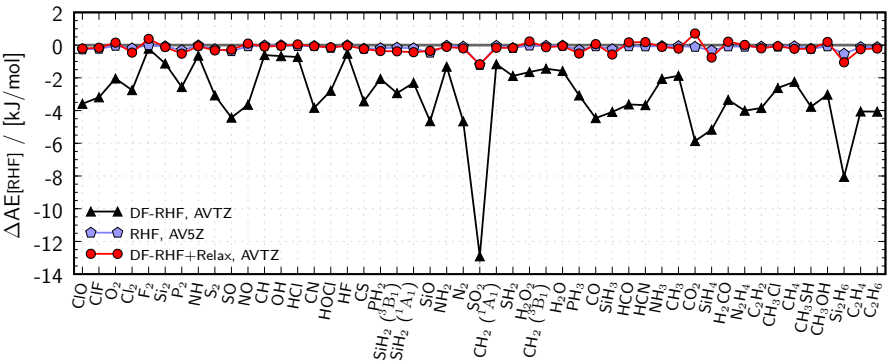

Figure 5.1: Deviation of Hartree-Fock atomization energies from RHF/AV6Z. The red curve is obtained by adding the CABS-relaxation correction to the black curve. The VTZ/JKFIT basis set is used to form the CA basis for the perturbative correction.

closed-shell case. The cost of calculating these corrections on their own is similar to that of a single RHF iteration in the combined MO+CA basis. However, since the occurring Fock matrices are required for the RMP2-F12 anyway, only the additional cost for block-diagonalizing the Fock matrix and/or performing a cheap iterative solution process remains.

Each of the two corrections can either be calculated for occupied valence orbitals only, or for both core- and valence orbitals; this is decided by the choice of the summation range of i in eq. (4.86). Including core excitations leads to better agreement with total basis set limit Hartree-Fock energies, but apart from few exceptions the relative energies produced by both variants are virtually indistinguishable. However, the corrections for core orbitals can have large numerical values. They therefore have a rather large dependence on the concrete method used to construct the CA basis and they thereby can introduce calculation artifacts. For this reason we now apply the correction for valence orbitals exclusively, contrary to what we still suggested in Refs. [127] and [102].

The performance of the perturbative correction for relative energies is illustrated in the case of atomization energies in Fig. 5.1. Tab. 5.4 further summarizes the performance of the different Hartree-Fock treatments for the other benchmark sets introduced in Sec. 5.1. Obviously both the CABS-singles and CABS-relax approaches greatly improve the DF-RHF energies. The only exception here are the AVQZ-level treatments of electron affinities. The AVQZ description of electron affinities is not significantly improved by the corrections because the RI basis set chosen here (non-augmented VnZ/JKFIT) does not add sufficient basis set flexibility for describing diffuse orbitals. Ref. [127] demonstrated that using more diffuse RI basis sets avoids this problem, and that also for the other quantities sometimes better Hartree-Fock

Table 5.4: RMS deviation of different Hartree-Fock treatments compared to near basis limit RHF values (AV5Z for reactions marked in Tab. 5.2–3, AV6Z for all others). The VnZ/JKFIT basis sets are used to form the CA basis for the perturbative corrections.

Method	RE[a]	AE	IP	EA
RHF/AVDZ	21.01	24.39	65.1	76.2
RHF/AVTZ	2.91	3.47	11.1	11.8
RHF/AVQZ	0.89	0.82	2.3	3.6
RHF/AV5Z[b]	0.30	0.25	0.6	1.0
DF-RHF+Singles/AVDZ	3.10	3.44	12.6	13.3
DF-RHF+Singles/AVTZ	0.59	0.41	3.1	2.9
DF-RHF+Singles/AVQZ[b]	0.21	0.14	1.1	3.0
DF-RHF+Relax/AVDZ	2.99	2.46	10.5	12.2
DF-RHF+Relax/AVTZ	0.58	0.36	2.9	3.0
DF-RHF+Relax/AVQZ[b]	0.21	0.13	1.0	3.0

(a) REs and AEs in kJ/mol, IPs and EAs in meV

(b) Reactions without AV6Z reference values were excluded

corrections are obtained with more diffuse RI basis sets. Nevertheless, the corrections generally work very well. Although with AVTZ basis sets conventional RHF/AV5Z accuracy is not obtained reliably, in any case the basis set error of the DF-RHF treatment becomes smaller again than that of the F12 correlation treatment (vide infra), which is the main goal.

The potential for correcting RHF energies of the CABS-singles and CABS-relax methods is very similar. The results of the spin-free CABS-relax approach look minimally better on average. We therefore employ the spin-free approach by default in all subsequent calculations.

5.2.2 DF-RMP2-F12: Choice of ansatz and approximation

In this section we want to establish the basic performance of various variations of the RMP2-F12 method. All F12 calculations are performed with the AVTZ orbital basis set, geminal exponent $\gamma = 1.0$, VTZ/JKFIT as basis for both RI- and DFJK, and AVTZ/MP2FIT as DF-basis. We compare density fitted RMP2(-F12) based on density fitted RHF to purely conventional (i.e., no DF, no F12) RHF + RMP2 calculations. By doing this we also assess the accuracy of the employed DF approximation, which greatly speeds up integral evaluation.

In order to obtain a general picture of the capabilities of F12 treatments, the figures 5.2 and 5.3 compare the finite basis set errors of conventional and explicitly correlated

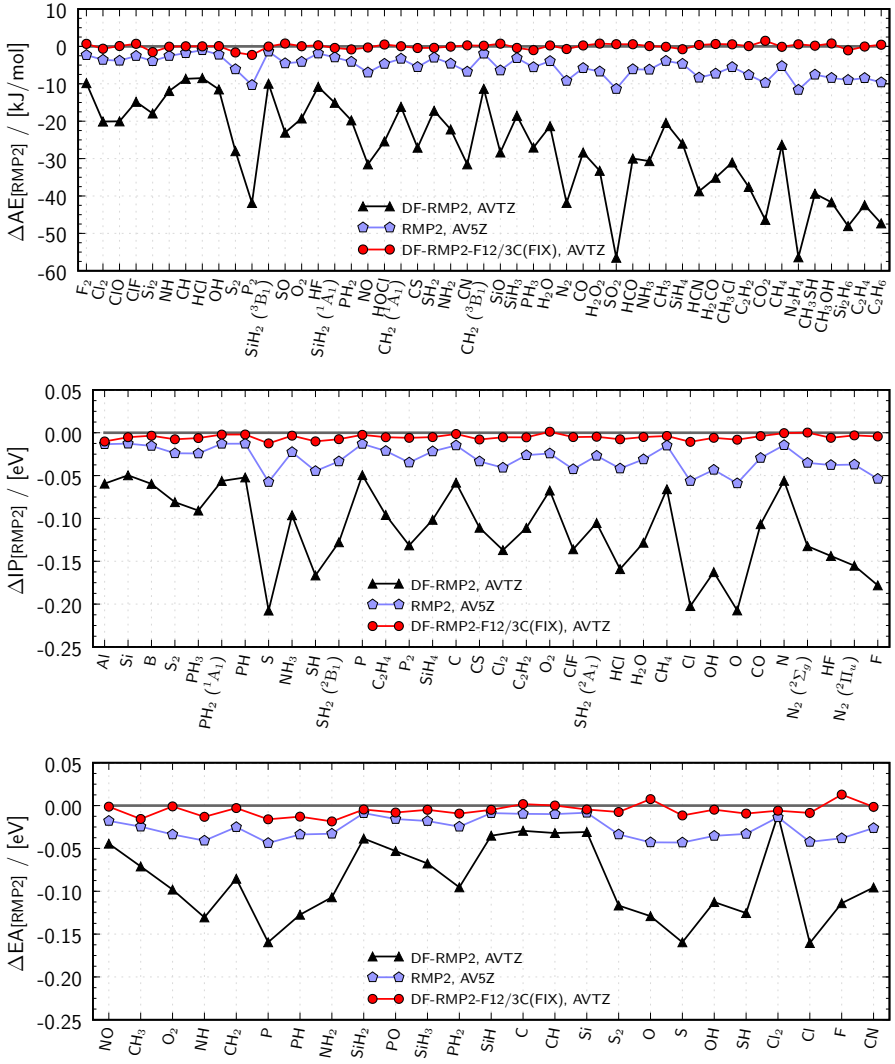

Figure 5.2: Deviation of total RMP2 atomization energies (top), ionization potentials (middle), and electron affinities (bottom) from RMP2/CBS. The red curves are obtained by adding F12/3C(FIX) and RHF corrections to the black curves. $\gamma = 1.0$ and default auxiliary basis sets were employed.

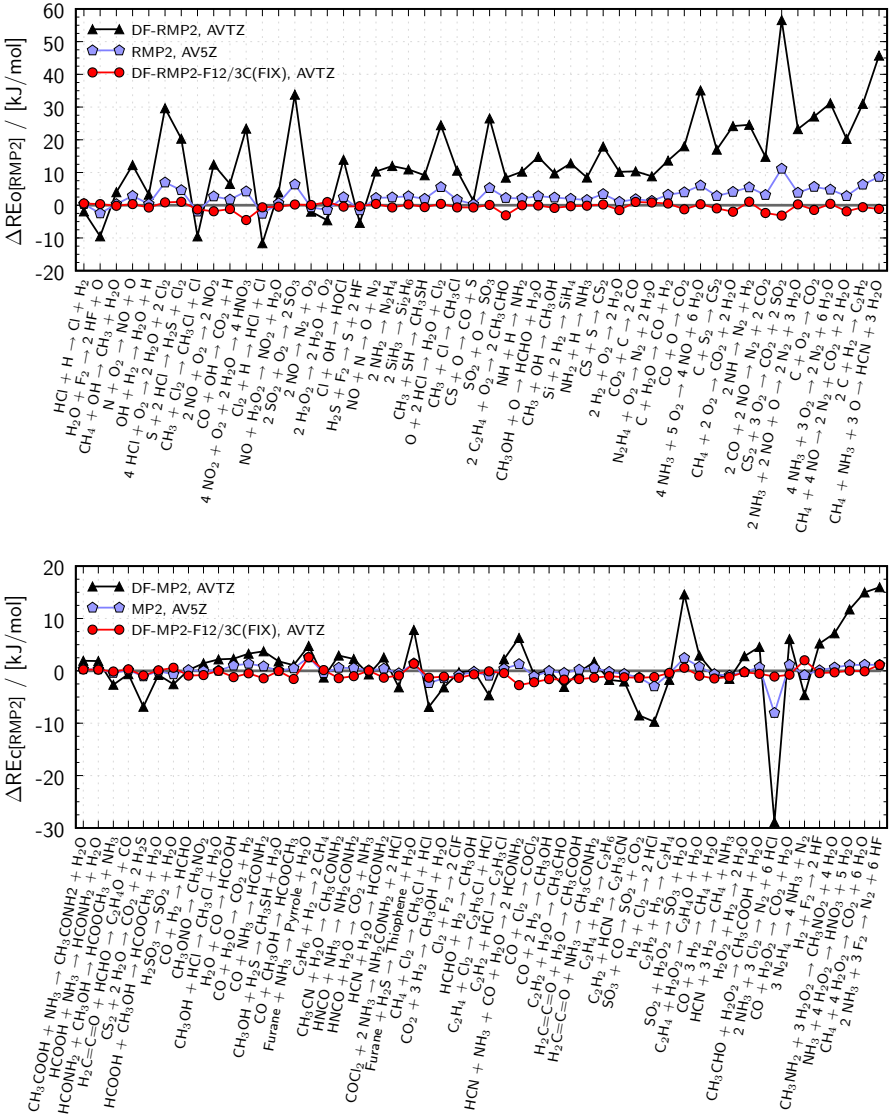

Figure 5.3: Deviation of total RMP2 open-shell reaction energies (top) and closed-shell reaction energies (bottom), from RMP2/CBS. The red curves are obtained by adding F12/3C(FIX) and RHF corrections to the black curves. $\gamma = 1.0$ and default auxiliary basis sets were employed.

RMP2 calculations directly. These figures display the deviation $E_{\text{method}} - E_{\text{CBS}}$, where E_{CBS} refers to the quantity in question (AEs, IPs, ...) calculated from extrapolations based on conventional treatments with very large basis sets (CBS[56] for all systems except the reactions marked in Tabs. 5.2 and 5.3; the marked reactions have CBS[45] reference values). Total energy differences are shown (i.e., Hartree-Fock and the perturbative Hartree-Fock correction are included), and the systems are ordered by reference value. These curves make it obvious that F12 treatments can massively diminish the basis set error of correlation methods. By adding F12 and RHF corrections to conventional DF-RMP2/AVTZ treatments, the basis set error can be eliminated almost entirely. It is also noteworthy that the density fitting approximations employed in the RHF and RMP2-F12 treatments do not spoil the accuracy of DF-RMP2-F12 at all, despite the fitting basis sets not being developed for F12 methods. For comparison, also curves of conventional RMP2/AV5Z are included. AV5Z calculations are already very large, and they are commonly used as reference point for very accurate correlation treatments. By using F12 methods, this high accuracy can clearly be surpassed with much smaller orbital basis sets—the AVTZ F12 treatments obviously outperform conventional ones with AV5Z basis sets.

It is also worth mentioning that the accurate F12 total energy differences shown in Fig. 5.2 could not possibly be obtained without the Hartree-Fock corrections investigated in Sec. 5.2.1. A direct comparison of Fig. 5.1 and Fig. 5.2 (upper) shows that the basis set errors of uncorrected Hartree-Fock are often much larger than the errors of the F12 correlation energy (mind the different scales in the plots).

In order to further quantify the performance of conventional methods, Tab. 5.5 displays the convergence behavior of conventional RMP2 treatments (of both direct energies and energies extrapolated according to Sec. 2.3). For the F12 methods, Tab. 5.6 displays basis set deviations of F12 treatments using different variants of the wave function ansaetze discussed in Sec. 3.4 and of the approximations discussed in Sec. 4.1.5. A direct comparison of the values of both tables shows that the accuracy of DF-RMP2-F12/AVTZ lies somewhere between CBS[34] and CBS[45]—regardless of the variant used.

Let us now discuss Tab. 5.6. In order to give a coarse indication of the relative computational costs of the different method variations for a semi-realistic calculation, the last column of Tab. 5.6 gives the timings for a DF-RMP2-F12 treatment of the aniline$^+$ cation[1]. While this molecule is not exactly large, it is certainly one one would not routinely attempt a basis set limit calculation for (getting AV6Z reference values for the C_2H_6 and Si_2H_6 molecules of the AE benchmark set was already quite straining).

In terms of accuracy the obtained results in Tab. 5.6 look puzzling. The most

[1] With Molpro 2009.2 PL 176 [compiled with g95 0.91, ACML 3.6.0] on a single core of an Opteron 2.5 GHz CPU. No molecular symmetry was used. DF-RHF takes 9.8 min, DF-RMP2 2.3 min.

Table 5.5: RMS deviation of conventional RMP2 (using the indicated orbital basis sets and basis set extrapolations) from the best CBS estimates.

Method	RE[a]	AE	IP	EA
RMP2/AVDZ	39.69	87.77	266.9	223.3
RMP2/AVTZ	14.27	29.58	122.6	104.6
RMP2/AVQZ	5.80	12.11	61.2	52.5
RMP2/AV5Z[b]	3.30	6.10	33.2	29.2
RMP2/AV6Z[b]	1.81	3.43	19.4	17.0
RMP2/CBS[23]	10.29	14.22	53.4	51.6
RMP2/CBS[34]	1.99	1.72	13.0	12.0
RMP2/CBS[45][b]	0.76	0.51	3.3	4.4

(a) REs and AEs in kJ/mol, IPs and EAs in meV

(b) Reactions without CBS[56] reference excluded

Table 5.6: RMS deviation of DF-RMP2-F12/AVTZ using the indicated ansaetze and approximations from the best conventional RMP2 CBS estimates. The last column gives the time in minutes required for a single-point calculation of the aniline$^+$ cation with the given method (on Opteron 2.5 GHz, with C_1 symmetry).

Method	RE[a]	AE	IP	EA	T(aniline$^+$)
Coupled RMP2-F12					
3C(FULL)	1.44	0.67	8.0	9.4	96.0 min
3C(D)	1.00	0.65	7.8	9.2	75.3 min
3C(FIX)	1.19	0.69	6.0	9.1	74.7 min
3C(FIX,HY2)	1.21	0.70	6.3	9.2	60.3 min
3C(FIX,HY1)	1.22	0.77	7.3	9.5	43.6 min
RMP2-F12 with EBC approximation					
3*C(FULL)	1.21	1.14	5.3	8.5	68.9 min
3*C(D)	0.95	1.25	4.7	8.2	67.3 min
3*C(FIX)	1.14	1.03	3.9	8.3	64.7 min
3*C(FIX,HY2)	1.15	1.04	3.9	8.4	51.5 min
3*C(FIX,HY1)	1.19	1.19	4.3	8.6	34.1 min

(a) REs and AEs in kJ/mol, IPs and EAs in meV

striking feature is probably the indifference of the accuracy with respect to the employed wave function ansatz (FIX,D,FULL)—the effect of choosing either one over the other is now rather small. This was still quite different when we employed only occupied spin orbitals as F12 generator orbitals in Ref. [127] instead of the extended generators described in Sec. 3.7. This generator extension has the interesting effect of making also the ansaetze (D) and (FULL) perform just like (FIX) in Ref. [127] (which's result change only slightly)[2]. The somewhat decreased accuracy of the (FULL) and (D) ansaetze for IPs may be related to a small imbalance in the variational freedom of F12 wave functions when calculating energy differences between closely related species with different numbers of open-shell orbitals. For example, if we split a closed-shell orbital into two open-shell orbitals, the number of F12 generator orbitals increases now, and with it the variational freedom of the F12 wave function. This effect appears to be of quite manageable proportions, however, as the results stay very good and are also much more balanced and predictable now.

The hybrid approximations HY1 and HY2 both perform well. The effect of the HY2 approximation on calculation accuracy is entirely negligible, and also HY1 does not deteriorate the accuracy significantly. The computational savings achieved by both approximations are not huge but significant, and they further increase with increasing RI basis size and molecule size (due to the $O(N^5)$ integral transformation steps becoming more important then). The effect of the EBC approximation is less transparent. According to these calculations, using the (rather coarse) EBC approximation *improves* the method accuracy significantly in several cases. In the presented table the only exceptions are atomization energies, which in the coupled case (no EBC) profit from a stochastic error cancellation between RHF corrections slightly biased for overcorrection, and F12 correlation corrections slightly biased for underestimating CBS limits (this cancellation is probably also responsible for AEs being more accurate than reaction energies here, although AEs are supposed to be harder to calculate). The reason for these improvements due to EBC are fuzzy at best. They could either be real effects, caused by problems of F12 configurations mentioned already in Sec. 3.2, or they could be artifacts of the employed basis sets—after all, they were not optimized for F12 applications. It is noteworthy that the effect of EBC looks quite different for the AVDZ and AVQZ basis sets, or when using the VnZ-F12 orbital basis sets of Peterson et al. [134] and corresponding optimized RI basis sets of Yousaf et al. [135]. At this moment no final conclusion can be made regarding this topic, and for now we continue using the theoretically more justified MP2-F12 methods without EBC approximation.

[2]Note that the EA results here cannot be compared directly to Ref. [127] because the latter used augmented JKFIT basis sets and added another set of diffuse s and p functions to the AVTZ orbital sets.

5.2.3 UCCSD(T)-F12: Basic performance

For realistic applications, calculations at the basis set limit are more interesting with coupled cluster methods than with MP2 methods. So let us now investigate the simplified UCCSD-F12x (x=a,b) treatments described in Sec. 4.3. Such calculations employ a DF-RMP2-F12 calculation as first step, and consequently they require the same kinds of basis sets and calculation parameters.

For a general overview, the figures 5.4 and 5.5 compare basis set errors of our explicitly correlated UCCSD(T)-F12a method (applied with a triple zeta basis set) to conventional calculations. Just like in the RMP2 case, the F12 corrections greatly reduce the basis set deviations—AVTZ calculations now offer results which are supposedly good enough for any kind of routine calculation. Contrary to the RMP2 case, the UCCSD(T) method has an excellent intrinsic accuracy; thus, reaching the basis set limit implies a great improvement in overall accuracy.

The curves in Figs. 5.5 and 5.4 do not look quite as flat as in the RMP2-F12 case (Figs. 5.3 and 5.2). This impression is also supported by Tab. 5.7, which compares conventional UCCSD treatments (without triples) to UCCSD-F12x (x=a,b)[3]—the basis set deviations of CCSD-F12 are larger than the ones of RMP2-F12. Nevertheless, CCSD-F12 treatments are still a huge improvement compared to their conventional counterparts. Even when basis set extrapolation is employed for conventional UCCSD, the F12 treatments with AVnZ basis sets still offer an accuracy comparable to extrapolated conventional methods employing AVnZ and AV(n+1)Z basis sets (even in default settings, and for purely energetic properties, for which extrapolation schemes work best). And these results are obtained with a computational cost only slightly exceeding that of only the smaller UCCSD/AVnZ calculation on its own.

Tab. 5.7 also demonstrates that F12a methods are more accurate than F12b for small and medium sized basis sets (AVDZ, AVTZ), while at the same time having basis set convergence problems for larger basis sets. F12b does not show these convergence problems, but works less well than F12a for genuine open-shell problems (IPs, EAs, reactions with open-shell species).

5.2.4 UCCSD(T)-F12: Corrections for (T)-triples

Tab. 5.7 only displayed the convergence of RHF-UCCSD; the perturbative (T) triples contribution of CCSD(T) was not applied. As discussed in Sec. 4.4, this contribution is not treated explicitly by our F12 methods, but due to its small magnitude, we might expect to get accurate total energy differences anyway. However, a closer look reveals

[3]The numbers in Tab. 5.7 slightly differ from the ones published in Ref. [103] due to using spin-unrestricted RHF-UCCSD instead of RHF-RCCSD as open-shell coupled cluster method.

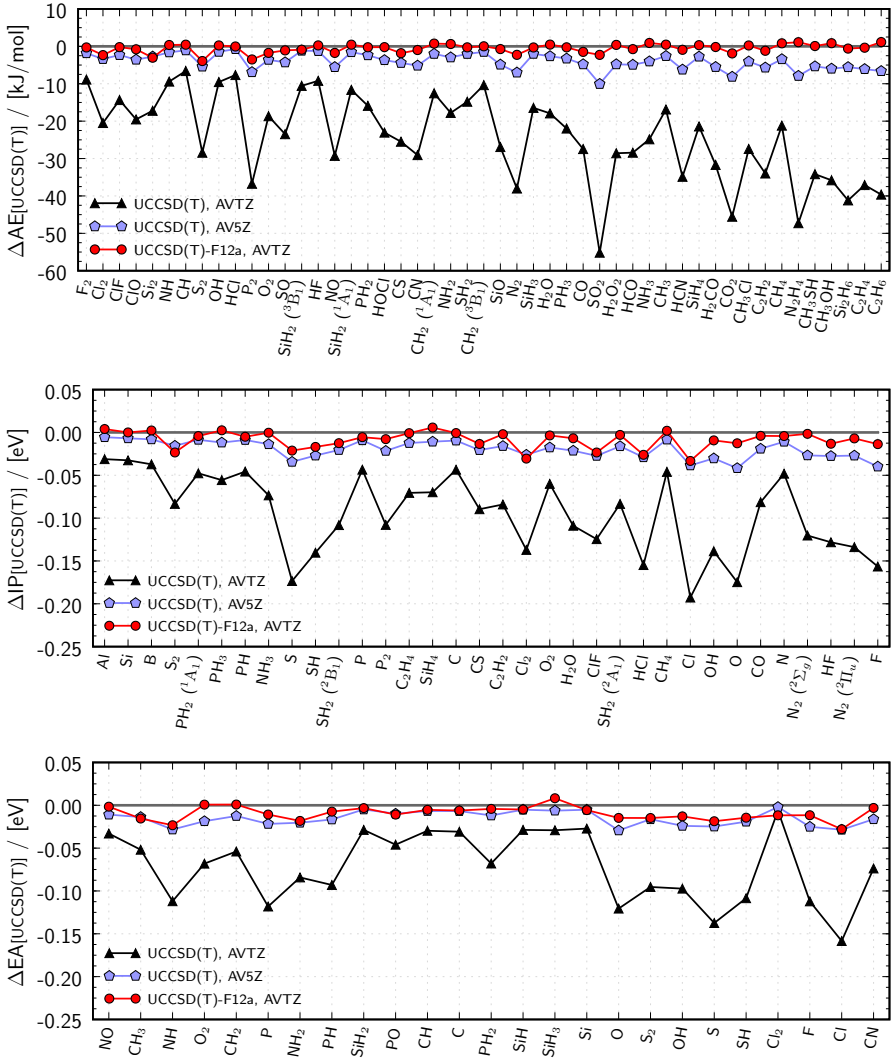

Figure 5.4: Deviation of total UCCSD(T) atomization energies (top), ionization potentials (middle), and electron affinities (bottom) from UCCSD(T)/CBS. $\gamma = 1.0$ and default auxiliary basis sets were employed for F12 treatments.

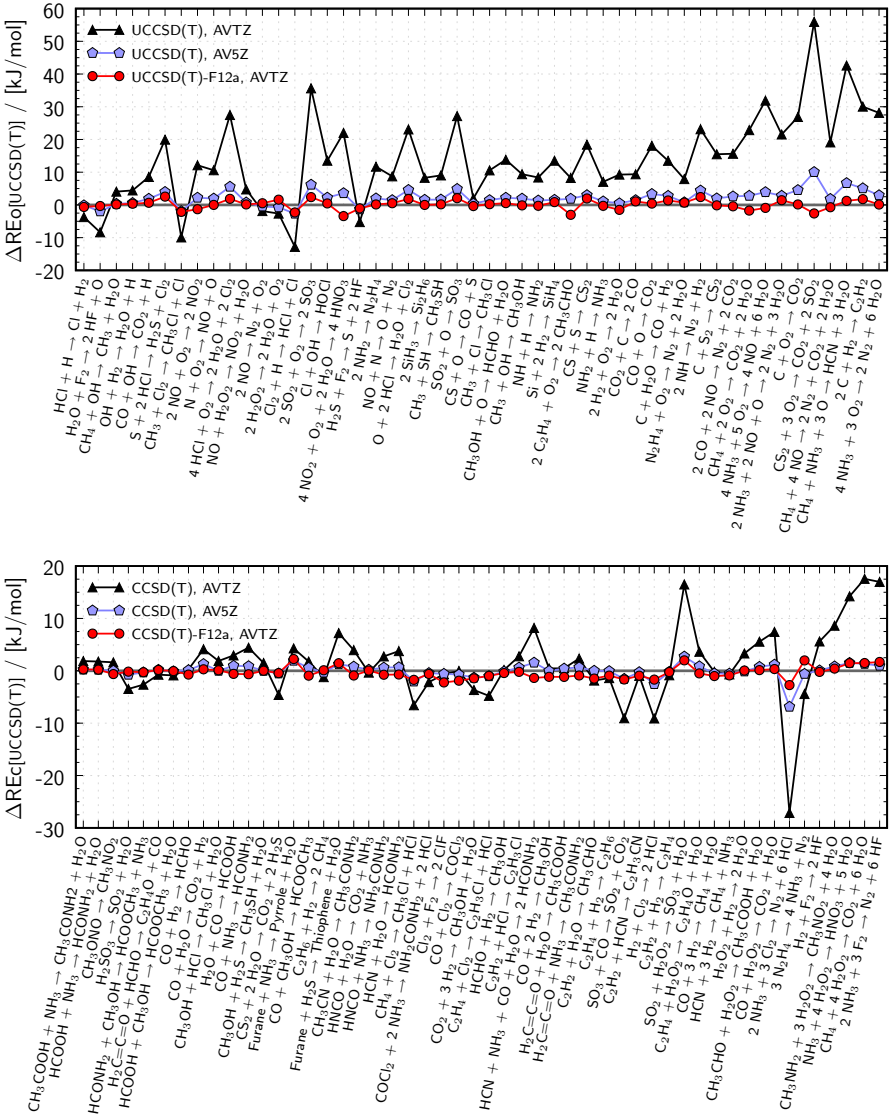

Figure 5.5: Deviation of total CCSD(T) open-shell reaction energies (top) and closed-shell reaction energies (bottom), from UCCSD(T)/CBS. $\gamma = 1.0$ and default auxiliary basis sets were employed for F12 treatments.

Table 5.7: RMS deviation of conventional RHF-UCCSD (using the indicated orbital basis sets and basis set extrapolations) and UCCSD-F12x treatments from the best CBS estimates. $\gamma = 1.0$ and default auxiliary basis sets are used for F12 methods.

Method	RE[a]	AE	IP	EA
UCCSD/AVDZ	37.98	79.84	244.3	175.5
UCCSD/AVTZ	13.74	24.86	97.9	72.0
UCCSD/AVQZ	4.97	8.84	40.1	28.1
UCCSD/AV5Z	2.32	4.17	20.5	14.7
UCCSD/AV6Z[b]	1.39	2.29	12.0	8.5
UCCSD/CBS[23]	9.99	11.18	31.6	30.6
UCCSD/CBS[34]	1.94	1.69	9.0	12.5
UCCSD/CBS[45][b]	0.56	0.39	1.5	1.8
UCCSD-F12a/AVDZ	4.04	7.50	49.3	37.6
UCCSD-F12a/AVTZ	1.29	1.96	9.2	7.6
UCCSD-F12a/AVQZ[b]	1.05	2.17	9.0	11.0
UCCSD-F12b/AVDZ	3.73	10.08	67.9	56.6
UCCSD-F12b/AVTZ	1.45	2.08	21.5	17.7
UCCSD-F12b/AVQZ[b]	0.65	0.71	5.2	5.3

(a) REs and AEs in kJ/mol, IPs and EAs in meV

(b) Reactions without CBS[56] reference were excluded

that in certain cases the largest part of the residual basis set error is caused by the finite basis (T) contribution.

The basis set dependence of both total CCSD(T) energies and of isolated (T) contributions is shown in Tab. 5.8. The following trends can be seen:

- Despite the small magnitude of the (T) correction, its basis set deviation becomes relevant when employing F12 methods. For example, the (T) contributions of conventional UCCSD(T)/AVTZ have RMS basis set deviations of 1.13 kJ/mol, 2.07 kJ/mol, 8.7 meV and 12.3 meV for REs, AEs, IPs and EAs, respectively, which *on their own* are as large as the total energy CBS deviations of DF-RMP2-F12/AVTZ or UCCSD-F12a/AVTZ [cf. Tabs. 5.6, 5.7]. This is a substantial difference to purely conventional methods, where the (T) basis set deviations indeed are approximately one order of magnitude smaller than the total energy deviations.

- At the same basis level, the CBS deviation of the direct unscaled (T) triples contribution slightly increases when these triples are calculated from CCSD-F12 wave functions instead of purely conventional CCSD wave functions (e.g.,

for AEs/AVTZ, when the (T) correction is calculated from conventional CCSD, it has 2.07 kJ/mol RMS deviation, while it deviates by 2.52 kJ/mol when based on CCSD-F12 wave functions). This effect is small, but not entirely negligible. It originates from the coupling of the conventional t_a^i/T_{ab}^{ij} amplitudes to the F12 configurations.

- The deviations of the isolated (T) contributions can be substantially reduced by applying the (T*) correction. (T^+) also reduces the deviations somewhat, but the effect is much smaller.

- These improvements in the (T) corrections do not translate one-to-one to improvements in total energies. While for F12b the (T*) correction typically leads to more accurate total energy differences, for F12a it tends to make things worse. In fact, UCCSD(T)-F12a (without triples corrections) generally shows a substantially *smaller* basis set deviation than the UCCSD-F12a it is based on [compare Tabs. 5.7 and 5.8]. This is clearly a result of an error cancellation between F12a's overestimated CCSD correlation energies and underestimated uncorrected (T) triples. However, this effect appears to be very consistent.

- The (T^+) correction appears to work decently for UCCSD(T)-F12b total energy differences at AVDZ and AVTZ level. However, since its actual reduction in the deviation of isolated (T) triples is rather limited [rhs columns in Tab. 5.8], applying it with larger basis sets might be a bad idea. Also its improvements in AVDZ and AVTZ treatments might be fortuitous, since it clearly works best for the benchmark set on which it was fitted (AEs).

- In previous studies [103] we showed that the CCSD(T)-F12b basis set error for large basis sets (≥AVQZ) is dominated by the deviation from the triples, and that it can be substantially reduced by the (T*) correction. While also present in Tab. 5.8, here the effect is less pronounced. This mainly results from choosing $\gamma = 1.0$ as geminal exponent instead of a larger number which would be more suitable for the doubles part.

The main effect of the (T*) correction is made obvious in Fig. 5.6: The scaled (T*) triples still scatter quite substantially, but the systematic bias is diminished; the scaled triples now scatter around the CBS limit.

While the previous table 5.7 certainly does not evidence that (T*) corrections should be applied in general, in special cases they turned out to be essential for obtaining very accurate relative energies [136, 137]. In Ref. [137] Marchetti and Werner found it necessary to apply a variation of the (T*) correction in a study of weakly interacting molecular dimers. Contrary to the (T*) correction discussed here, for which the scaling factors are calculated independently for each of the atoms

Table 5.8: RMS deviations of UCCSD(T) total energies and of isolated (T) triples contributions from CBS estimates if different types of (T) treatments are employed. (T) denotes standard triples based on the respective CCSD wave function. The (T*) and (T$^+$) rows refer to the F12 scaling corrections discussed in Sec. 4.4.

	RHF + CCSD(T)				(T) only			
Method/Basis	RE[a]	AE	IP	EA	RE	AE	IP	EA
Conventional UCCSD(T)								
(T)/AVDZ	39.75	86.62	270.6	208.2	3.64	7.84	31.1	41.5
(T)/AVTZ	14.10	26.73	105.3	83.0	1.13	2.07	8.7	12.3
(T)/AVQZ	5.15	9.62	43.2	33.3	0.50	0.86	3.6	5.7
(T)/AV5Z	2.40	4.53	22.1	17.5	0.25	0.40	1.7	3.0
(T)/AV6Z[b]	1.45	2.50	12.9	10.1	0.17	0.23	1.0	1.7
(T)/CBS[23]	9.94	10.95	31.0	30.5	0.58	0.59	2.2	2.1
(T)/CBS[34]	2.03	1.76	8.8	12.4	0.15	0.13	0.6	1.5
(T)/CBS[45][b]	0.59	0.45	1.7	2.1	0.05	0.09	0.2	0.4
Explicitly correlated UCCSD(T)-F12b								
(T)/AVDZ	3.68	14.28	94.8	93.7	4.31	8.91	33.6	44.2
(T)/AVTZ	1.79	3.91	30.0	31.0	1.35	2.52	10.6	14.6
(T)/AVQZ[b]	0.93	0.98	8.9	11.4	0.69	1.09	4.6	7.1
(T*)/AVDZ	3.81	11.13	87.5	79.9	2.74	3.74	24.4	28.5
(T*)/AVTZ	1.49	2.07	25.1	22.6	0.69	0.37	5.0	6.0
(T*)/AVQZ[b]	0.82	0.74	6.3	7.5	0.35	0.14	1.7	2.8
(T$^+$)/AVDZ	3.10	10.52	77.7	82.4	3.70	5.74	21.2	33.7
(T$^+$)/AVTZ	1.35	2.37	22.7	25.5	1.25	1.27	5.6	9.7
(T$^+$)/AVQZ[b]	0.94	0.67	5.7	9.2	0.64	0.55	2.3	4.8
Explicitly correlated UCCSD(T)-F12a								
(T)/AVDZ	3.50	8.99	73.6	66.8	4.31	8.91	33.6	44.2
(T)/AVTZ	1.25	1.30	13.1	12.4	1.35	2.52	10.6	14.6
(T)/AVQZ[b]	1.00	1.35	4.9	6.6	0.69	1.09	4.6	7.1
(T*)/AVDZ	4.23	7.04	67.3	55.4	2.74	3.74	24.4	28.5
(T*)/AVTZ	1.43	1.84	10.8	7.2	0.69	0.37	5.0	6.0
(T*)/AVQZ[b]	1.23	2.24	7.9	9.5	0.35	0.14	1.7	2.8
(T$^+$)/AVDZ	3.59	5.59	56.0	54.9	3.71	5.57	20.3	31.8
(T$^+$)/AVTZ	1.38	1.76	7.9	7.7	1.27	1.21	5.4	8.5
(T$^+$)/AVQZ[b]	1.34	2.11	9.2	9.2	0.64	0.52	2.3	4.0

(a) REs and AEs in kJ/mol, IPs and EAs in meV

(b) Reactions without CBS[56] reference values were excluded

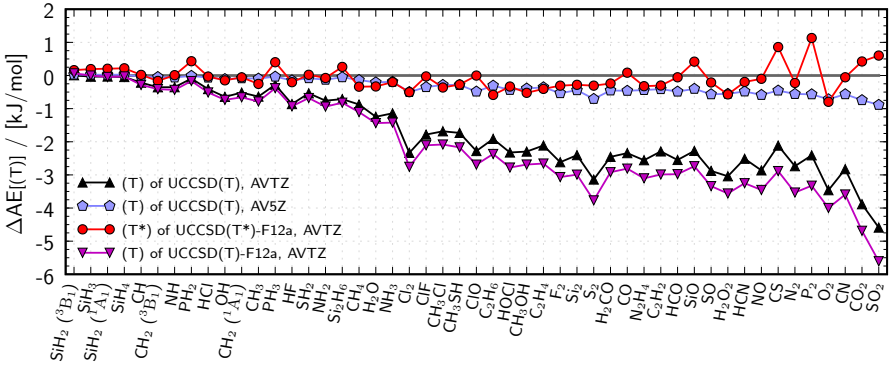

Figure 5.6: Deviation of the (T) correction from (T)/CBS[56] for atomization energies. (T*) refers to the scaled triples contribution discussed in Sec. 4.4. The non-scaled (T) corrections of UCCSD(T) and UCCSD(T)-F12a slightly differ because the F12 terms affect the conventional T_1 and T_2 amplitudes from which (T) is calculated.

or molecules involved in a relative energy balance, they used a common scaling factor obtained from the (T*) correction of the dimer and applied this factor to the individual monomers, too. This way of applying the scaling correction avoids the size consistency problems which the method used here has, and it is much less likely to produce calculation artifacts due to very different MP2-F12 scaling factors of different reaction reactants or products. On the downside, it is not obvious how chemical reactions with multiple different reactants *and* products should be treated, and it is less simple to automate. Nevertheless, using a common scaling factor is likely the better choice in general.

5.3 Basis set and parameter studies

Within this section additional basis set combinations and geminal exponents γ are investigated in order to establish a rough estimate of the accuracy and cost of F12 calculations under various conditions. This is done both to determine good calculation parameters and to check how robust the F12 methods are with regard to changes of the parameters.

5.3.1 Additional basis set combinations

Additionally to the AVnZ basis sets introduced in Sec. 5.2, we do calculations with the VnZ-F12 orbital basis sets of Peterson et al. [134], which are designed for F12 calculations, and with the Turbomole basis sets def2-SVP [138], def2-TZVPP [138], and def2-QZVPP [139].

In total, the Turbomole sets will be referred to as "def2-nZVPP". For individual basis sets, the def2- prefix will usually be omitted. The def2-nZVPP sets are used in their standard form, without diffuse orbital augmentation, and thus they have nearly exactly the same size as the cc-pVnZ (n=D,T,Q) basis sets for first row atoms, and as cc-pV(n+d)Z for second-row atoms. That means they are much smaller than the AVnZ and VnZ-F12 sets of the same formal cardinality. There are two reasons for including those sets: (i) In a recent study on atomization energies, Bischoff et al. [140] found the def2-TZVPP and def2-QZVPP basis sets to work well for F12 calculations, and (ii) Unlike the VnZ-F12 and AVnZ sets, both the def2- orbital basis sets themselves and the corresponding JKFIT and MP2FIT auxiliary sets are available for nearly every element across the periodic table at the current point in time[4]. If the def2-nZVPP sets would turn out to work well for F12 calculations, they would therefore considerably extend the applicability of F12 methods until suitable specialized orbital- and auxiliary sets are developed. Concerning this matter, the study of Bischoff et al. [140] leaves open some questions, because both the AVnZ sets and the nZVPP sets were only applied with $\gamma = 1.4$ and large AVnZ/MP2FIT sets for the RI basis, and both decisions are not optimal for a fair comparison of the sets.

The following choices are made for the auxiliary basis sets:

- By default, AVnZ sets are applied as before, with JKFIT sets as RI basis. Specialized basis sets for the RI approach, denoted AVnZ/OPTRI [141] are available for first- and second-row elements, but were only used when indicated.

- The VnZ-F12 basis sets were applied with the same MP2FIT sets as the AVnZ basis sets. For JKFIT, singly augmented versions of the same JKFIT sets as used for AVnZ were employed (this is Molpro's default choice). The VnZ-F12/OPTRI [135] of Yousaf et al. were used for RI.

- For the def2-nZVPP basis sets, the newer universal JKFIT basis sets of Weigend [142] were employed instead of the JKFIT sets of Ref. [131] (which are used for the other basis sets). Like in the AVnZ case, the functions with the highest angular momentum were omitted for the double-ζ set def2-SVP. For fitting integrals, the MP2FIT sets as specified by Hellweg et al. [143] were employed, which are based on Refs. [133, 144] for the elements considered here.

  For RI purposes, the universal JKFIT sets did not work too well in benchmark calculations. Particularly, they were ill-suited for the CABS approach due to

[4]Concretely: For all elements from H to Rn excluding Lanthanides (except Lanthan). For Rb to Rn they are used in conjunction with Stuttgart-Cologne energy-adjusted relativistic small-core pseudopotentials (ECPs).

Table 5.9: Relative sizes of basis set combinations introduced in Sec. 5.3.1: Number of basis functions in the molecule phenylphosphine (C_6H_7P) and timings of DF-MP2-F12 and conventional CCSD(T). The timings were done with Molpro 2009.2 [g95 0.91, ACML 3.6.0] running on a single core of an Opteron 2.6 GHz processor.

	Basis set size			Calculation time	
Basis	AO	RI	JK/CFIT[a]	MP2-F12	CCSD(T)
SVP	137	283	613/ 567	3min	24min
(RTZVP)	181	292	746/ 789	6min	1h 13min
AVDZ	233	693	693/ 685	14min	3h 54min
VDZ-F12	282	616	868/ 685	14min	7h 48min
TZVPP	326	500	746/ 789	15min	13h 20min
(RQZVP)	370	534	746/1405	30min	22h 16min
AVTZ	492	805	805/1105	45min	67h 56min
VTZ-F12	506	812	1092/1105	48min	76h 27min
QZVPP	622	863	746/1405	1h 16min	≈170h[b]
VQZ-F12	856	952	1568/1762	2h 35min	≈600h[b]
AVQZ	891	1141	1141/1762	2h 57min	≈700h[b]

(a) Number of basis functions for JKFIT/MP2FIT, respectively.

(b) Extrapolated from smaller basis sets by fit to $T = a \cdot n_{\text{virt}}^4 + b \cdot n_{\text{virt}}^3$.

large redundancy with the AO basis. For this reason, dedicated RI basis sets for the def2-nZVPP orbital bases were generated using a simple automated procedure described in Sec. 5.5. The resulting sets are usually much smaller than the corresponding universal JKFIT sets, and they lead to somewhat more accurate F12 treatments here.

For testing purposes, also two *reduced* orbital sets were created: One by mixing s and p functions from the QZVP set with higher angular momentum functions from the smaller TZVPP set, and one by mixing s and p functions from the TZVP set with higher angular momentum functions from SVP. For a lack of more imaginative naming, these are called RQZVP and RTZVP in the following. The rationale behind this combination will be discussed later.

In order to give an indication of the relative sizes of the introduced basis set combinations and the calculation cost associated with them, Tab. 5.9 reports the number of basis functions for an example molecule and gives calculation timings. The timings (done in C_1 symmetry) of conventional closed-shell CCSD(T) and MP2-F12 are given; the cost of a CCSD(T)-F12 calculation is approximately the sum of both (plus one half CCSD iteration for the F12b energy correction). For a given

Table 5.10: RMS deviations of DF-RHF and DF-RHF+(singles correction) from AV6Z reference values. Reactions without AV6Z reference values were excluded.

	DF-RHF				DF-RHF+CABS relax			
Basis	RE[a]	AE	IP	EA	RE	AE	IP	EA
AVDZ	24.91	24.92	65.9	78.5	3.31	2.46	10.5	12.2
AVTZ	3.60	3.73	11.4	13.8	0.60	0.36	2.9	3.0
AVQZ	1.23	0.98	2.5	5.8	0.21	0.13	1.0	3.0
VDZ-F12	14.68	12.79	34.2	53.1	2.58	2.13	6.2	9.3
VTZ-F12	2.16	2.15	4.1	5.7	0.60	0.48	1.6	2.7
VQZ-F12	0.40	0.40	0.7	1.8	0.15	0.14	0.3	1.5
SVP	43.06	30.54	133.7	796.4	18.01	15.87	60.7	300.0
TZVPP	5.00	4.70	12.3	246.4	1.56	1.10	4.2	98.7
QZVPP	1.15	1.52	3.3	98.9	0.34	0.42	1.0	38.4
(RTZVP)	42.13	36.10	81.5	247.1	15.39	12.04	24.2	100.9
(RQZVP)	3.68	4.23	6.7	100.5	1.10	1.01	1.9	39.5

(a) REs and AEs in kJ/mol, IPs and EAs in meV

molecule, the cost of both CCSD and the (T) correction depends formally as $O(v^4)$ on the number of virtual orbitals (there are also expensive $O(v^3)$ terms). As $n_{\text{virt}} \approx n_{\text{AO}}$, the computer time dependence on the number of basis functions is expected to show a very similar quartic scaling. Thus, if smaller AO basis sets can be used to achieve sufficient accuracy for a given application, considerable savings of computer resources can be achieved.

5.3.2 Differences in Hartree-Fock accuracy

Tab. 5.10 gives an overview over the accuracy obtained at Hartree-Fock level by the newly introduced basis set combinations. Data is given for both normal density fitted Hartree-Fock calculations, and for Hartree-Fock calculations with perturbative singles corrections applied.

We can see that without the CABS corrections, the VnZ-F12 basis sets offer a significantly improved performance compared to the AVnZ ones (they were constructed to do that [134]). Interestingly, this holds even for the case of electron affinities, despite the fact that the VnZ-F12 basis sets are not fully augmented. The (completely non-augmented) def2-nZVPP basis sets are less accurate than AVnZ and VnZ-F12 throughout, but at the same level of cardinality they are also much smaller (see Tab. 5.9). Not surprisingly, the non-augmented nZVPP sets fail badly when used for calculating of electron affinities.

The accuracy difference between the VnZ-F12 and AVnZ is greatly diminished when the singles correction is applied. The nZVPP sets, however, stay less accurate than both also with singles corrections enabled (at least when using the generated RI sets for the CABS). For def2-SVP, in particular, applying the singles correction is not sufficient for obtaining reasonable Hartree-Fock values.

The RTZVP and RQZVP sets slightly improve on the SVP and TZVPP sets, respectively, but the effect is smaller than I hoped. A full or partial augmentation of the SVP/TZVPP sets might have served better for improving the Hartree-Fock performance.

5.3.3 Effect of geminal exponents on relative energies

For the fixed amplitude ansatz, the quality of the F12 basis set correction depends on the used geminal exponent γ. In order to obtain an idea of how much relative energies are affected by the concrete choice, the parameter space of γ was scanned for AEs, IPs, EAs and REs for all the orbitals sets of Sec. 5.3.1, using the DF-RMP2-F12/3C(FIX) method. The given default combinations of auxiliary basis sets were employed.

The figures 5.7 to 5.10 give a graphical overview over the obtained results. The figures display the root mean square deviation of the F12 treatment compared to conventional RMP2/CBS[56] basis set limit estimates. That means, each single γ-point corresponds to an average of a large number of different, realistic calculations. We can see that there is a sizable dependence on the geminal exponent, as different molecules (and in fact, different regions within the same molecule) require different exponents for an optimal theoretical description. But on the other hand, we also see that sensible compromise values for γ, which work decently for the entire set, can readily be chosen when calculating valence correlation energies.

When judging these basis set combinations, we can draw the following conclusions:

- When calculating electron affinities, the AVnZ sets are clearly superior to both the VnZ-F12 and the (non-augmented) def2-nZVPP sets. The reason for this is that only the AVnZ sets are fully augmented, and thus only they show sufficient flexibility in the diffuse outer regions of molecules. This flexibility is necessary to describe anions accurately. The augmentation appears to be necessary for both low angular momenta (for forming diffuse occupied orbitals) and high angular momenta (for the coarse form of the correlated wave function).

- The AVDZ and the somewhat larger VDZ-F12 basis sets work similarly well for our investigated properties apart from EAs. It is known, however, that the latter show better performance for molecules containing second-row elements [126].

Compared to AVDZ and VDZ-F12, the again somewhat larger def2-TZVPP basis set displays a better performance for AEs and REs and a greatly reduced dependence on γ. Interestingly, in the case of REs, the consistently better performance of def2-TZVPP for correlation energies (left side of Fig. 5.7) does not result in more accurate total energies (right side of Fig. 5.7). A closer look reveals that this is a result of correlation errors and Hartree-Fock errors not compensating each other, for reasons not investigated further. If the RQZVP set is used, which replaces *s* and *p* shells of TZVPP by *s* and *p* shells from QZVP, the Hartree-Fock errors are reduced and then also total energy differences become more accurate (Fig. 5.7, right, especially for REs).

The def2-TZVPP or RQZVP sets may be a good option for elements for which VDZ-F12 orbital sets or AVDZ auxiliary sets are not yet available. For all three basis sets, $\gamma = 0.9$ seems to be the optimal compromise value for valence correlation energies, with $\gamma = 1.0$ (Molpro default value) being only minimally worse.

- From Fig. 5.8 we can see that SVP and RTZVP work much worse than the AVDZ set. While SVP might be still be warranted for some very coarse overview calculations, even with F12 methods quantitative accuracy can not be expected from such basis sets. In particular, in the RTZVP set, which replaces *s* and *p* shells from SVP by *s* and *p* shells from TZVP, basis set deviations are not reduced enough to become competitive with the somewhat larger AVDZ set, which is much more accurate.

- The VTZ-F12 basis is superior to the (somewhat smaller) AVTZ basis set for AEs and REs. For IPs nearly no difference is observable between them. Again $\gamma = 1.0$ seems to be good value of the geminal exponents for all calculations. The again somewhat larger def2-QZVPP basis set offers an accuracy lying somewhere between AVTZ and VTZ-F12. It does require larger γ exponents ($\gamma = 1.2 \ldots 1.4$) to achieve it, which is not surprising [5].

- The plots for AVQZ-sized basis sets in Fig. 5.10 are not very enlightening since the used CBS[56] reference values are apparently not accurate enough to judge these F12 treatments (note the scale).

For comparison, Tab. 5.11 lists the basis deviations for selected benchmark sets and values of γ also for coupled cluster calculations.

[5] The optimum exponents γ are supposed to increase with the basis set cardinality. Because the wave function expansion in terms of Slater determinants then becomes more accurate at short-range, less space must be covered by the F12 configurations.

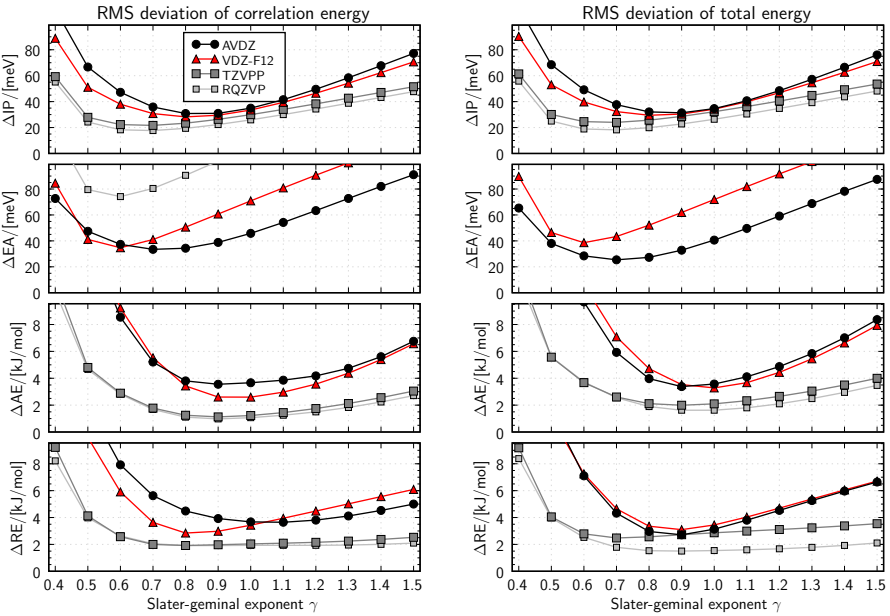

Figure 5.7: γ-dependence of DF-RMP2-F12/3C(FIX) for ionization potentials, electron affinities, atomization energies, and reaction energies. RMS deviations of the correlation energy (left) and of the total energy (right) from RMP2/CBS[56] are shown. The scales in the subplots are equal (1 kJ/mol≈10.36 meV).

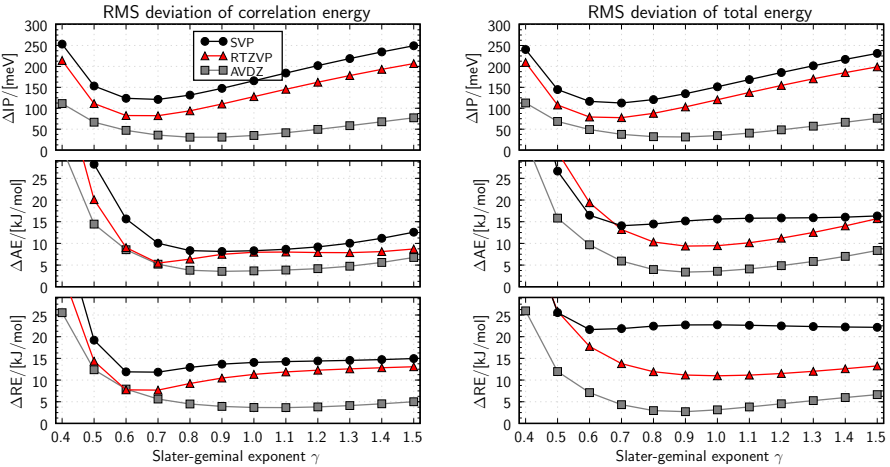

Figure 5.8: γ-dependence of DF-RMP2-F12/3C(FIX) for basis sets smaller than AVDZ.

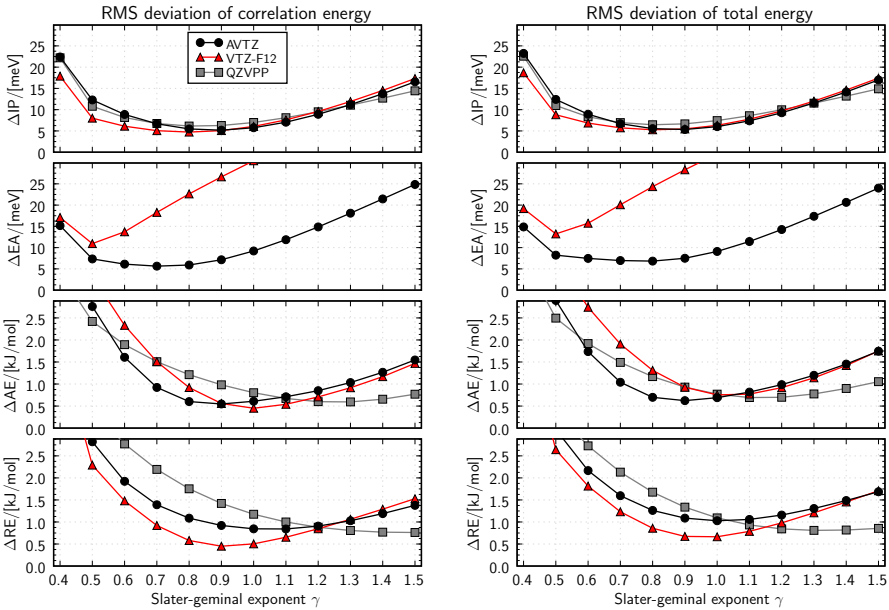

Figure 5.9: γ-dependence of DF-RMP2-F12/3C(FIX) for basis set of size comparable to VTZ-F12.

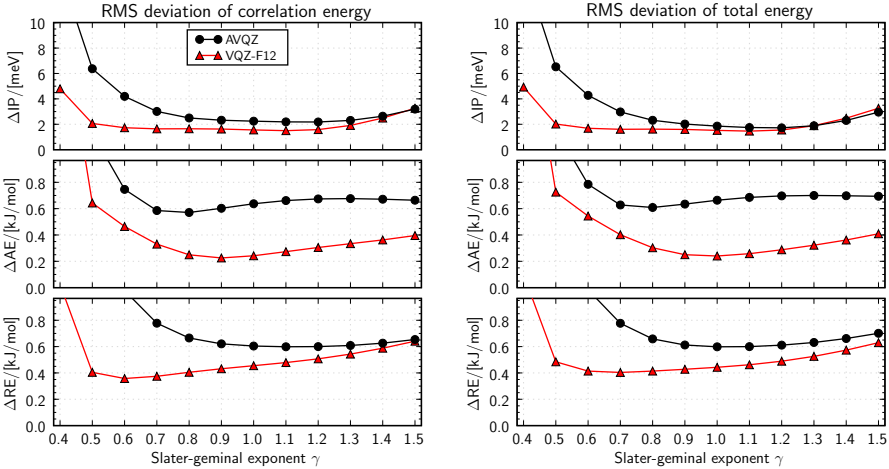

Figure 5.10: γ-dependence of DF-RMP2-F12/3C(FIX) for AVQZ and VQZ-F12 basis sets. *Warning:* The CBS[56] reference values are not much more accurate (if at all) than the AVQZ/VQZ-F12 F12 values themselves [cf. Tab. 5.5].

Table 5.11: RMS basis deviations of CCSD(T)-F12x treatments for selected combinations of basis set and geminal exponent γ. Reactions without CBS[56] reference values were excluded.

	UCCSD(T)-F12a			UCCSD(T)-F12b		
Basis/γ	RE[a]	AE	IP	RE	AE	IP
SVP/1.0	21.13	16.85	188.2	20.60	14.81	204.4
(RTZVP/1.0)	11.70	9.82	161.7	12.06	13.65	176.5
AVDZ/0.9	3.80	9.71	72.4	4.43	15.15	95.5
AVDZ/1.0	3.63	8.99	73.6	4.07	14.28	94.8
VDZ-F12/0.9	4.71	7.63	73.1	5.72	12.17	93.7
VDZ-F12/1.0	4.59	7.00	74.8	5.45	11.43	93.5
TZVPP/1.0	2.45	2.52	44.3	3.47	6.07	62.5
(RQZVP/1.0)	1.40	1.95	37.8	2.13	5.48	55.4
AVTZ/1.0	1.26	1.30	13.1	2.00	3.91	30.0
VTZ-F12/1.0	1.50	1.39	12.2	2.39	4.24	27.6
QZVPP/1.0	1.52	1.30	5.1	1.14	1.62	17.9
QZVPP/1.3	1.14	0.93	7.1	0.98	1.73	18.4
VQZ-F12/1.0	0.71	1.28	5.4	0.75	1.42	7.9
VQZ-F12/1.3	0.83	1.47	5.6	0.79	1.13	6.5
AVQZ/1.0	1.00	1.35	4.9	0.93	0.98	8.9
AVQZ/1.3	1.02	1.41	5.4	0.87	0.82	6.8

(a) REs and AEs in kJ/mol, IPs and EAs in meV

5.4 Exploratory calculations

While Sec. 5.2 established the basic basis set convergence of the proposed electronic structure methods for routine calculations, this section deals with additional application scenarios in other cases. Within this section, the coupled cluster methods will also be compared to accurate experimental or theoretical reference data, and their potential for combination with other high-level electronic structure methods will be explored.

5.4.1 Diatomic molecules: Basis set convergence for vibrational spectroscopic constants

In order to investigate how the F12 corrections influence potential energy surfaces (PES), we benchmarked the basis set convergence of the vibrational spectroscopic constants of several diatomic molecules. The investigated systems and corresponding

Table 5.12: Investigated diatomic molecules and some of their experimentally derived spectroscopic constants [145]. r_e are given in pm, ω_e and $x_e\omega_e$ are given in cm^{-1}. mad(exp) and rms(exp) give deviations from the experimental values.

	Experimental [145]			UCCSD(T)/CBS[56]		
Molecule	r_e	ω_e	$x_e\omega_e$	r_e	ω_e	$x_e\omega_e$
HF	91.68	4138.3	89.9	91.73	4142.6	89.9
N_2	109.77	2358.6	14.3	109.89	2362.3	13.7
CO	112.83	2169.8	13.3	113.02	2166.8	13.2
BF	126.25	1402.1	11.8	126.65	1395.0	11.6
F_2	141.19	916.6	11.2	140.92	931.4	11.5
C_2 ($^1\Sigma_g^+$)	124.25	1854.7	13.3	124.40	1861.2	12.4
OH ($^2\Pi$)	96.97	3737.8	84.9	97.01	3743.1	84.2
NH ($^3\Sigma^-$)	103.62	3282.3	78.3	103.68	3286.2	76.7
CH ($^2\Pi$)	111.99	2858.5	63.0	111.96	2857.9	62.7
CN ($^2\Sigma^+$)	117.18	2068.6	13.1	117.31	2070.5	12.7
NO ($^2\Pi$)	115.08	1904.2	14.1	115.13	1915.3	13.4
O_2 ($^3\Sigma_g^-$)	120.75	1580.2	12.0	120.60	1605.8	10.9
CF ($^2\Pi$)	127.18	1308.1	11.1	127.38	1308.5	10.6
mad(exp)				0.14	6.8	0.6
rms(exp)				0.17	9.5	0.7

experimental values and CCSD(T)/CBS[56] reference values are listed in Tab. 5.12. The set includes several difficult cases, such as the triply-bonded N_2, the C_2 and CN molecules, which contain low-lying excited states, and the F_2 molecule, which is unbound at Hartree-Fock level.

For each molecule, a grid of ten energy points within $\pm 0.15 a_0$ of the equilibrium distance was calculated and the equilibrium harmonic frequency ω_e, equilibrium distance r_e and anharmonicity constant $x_e\omega_e$ were determined from an eighth order polynomial fit performed with the Polfit program of Rosmus and Werner. The experimental and calculated spectra always refer to the most abundant isotopes (e.g., ^{1}H, ^{9}F). The values were calculated with bare nuclear masses; the weight of the electrons was neglected.

The tables 5.13 and 5.14 give an overview over the basis set convergence of F12 methods with respect to UCCSD(T)/CBS[56] reference values. F12 methods have been applied with $\gamma = 1.0$ unless noted otherwise and with the same basis set combinations as introduced in Sec. 5.3.

In total we can see that F12 methods applied with AVnZ or VnZ-F12 basis sets work excellently for this purpose of determining spectroscopic constants. When comparing with Tab. 5.12, we see that for the CCSD(T)-F12x methods, already AVDZ,

Table 5.13: RMS deviations of ω_e/[cm^{-1}] from CBS[56] reference values for the 13 diatomic molecules. "Conv" denotes conventional calculations, "Co+S." denotes RHF with CABS correction.

	δRHF		δMP2		δCCSD(T)		
Basis	Conv	Co+S.	Conv	F12	Conv	F12a	F12b
AVDZ	45.5	3.3	69.9	6.6	75.7	5.7	5.1
AVTZ	5.7	0.9	19.8	1.3	21.0	3.0	1.7
AVQZ	1.2	0.2	5.8	0.9	6.5	0.7	1.5
AV5Z	0.1		2.8		2.9		
CBS[23]			8.8		9.8		
CBS[34]			3.7		3.0		
CBS[45]			1.4		0.9		
VDZ-F12	8.2	2.5	33.1	4.0	35.2	5.5	2.8
VTZ-F12	3.5	0.4	12.7	1.1	16.4	3.4	2.1
VQZ-F12	0.3	0.1	6.4	0.8	7.3	1.1	0.7
SVP	47.7	25.3	44.8	48.7	43.2	34.1	37.6
TZVPP	2.8	1.6	10.4	4.7	12.4	4.3	4.7
QZVPP	1.2	0.3	8.0	4.1	7.8	4.1	5.4

Table 5.14: RMS deviations of r_e/[pm] for the 13 diatomic molecules.

	δRHF		δMP2		δCCSD(T)		
Basis	Conv	Co+S.	Conv	F12	Conv	F12a	F12b
AVDZ	1.36	0.10	2.28	0.11	2.48	0.20	0.15
AVTZ	0.21	0.02	0.55	0.04	0.59	0.08	0.06
AVQZ	0.03	0.00	0.16	0.01	0.19	0.02	0.01
AV5Z	0.00		0.07		0.08		
CBS[23]			0.30		0.31		
CBS[34]			0.05		0.04		
CBS[45]			0.02		0.02		
VDZ-F12	0.35	0.06	0.93	0.15	0.97	0.17	0.11
VTZ-F12	0.07	0.01	0.31	0.04	0.39	0.09	0.07
VQZ-F12	0.01	0.00	0.14	0.01	0.16	0.02	0.01
SVP	0.60	0.24	0.91	0.70	1.16	0.54	0.58
TZVPP	0.17	0.04	0.37	0.04	0.45	0.09	0.07
QZVPP	0.02	0.01	0.12	0.02	0.16	0.03	0.03

Table 5.15: RMS deviations of ω_e/[cm^{-1}] from CBS[56] reference values for the 13 diatomic molecules: Influence of geminal exponent γ.

	MP2-F12			CCSD(T)-F12a			CCSD(T)-F12b		
Basis	1.0	1.2	1.4	1.0	1.2	1.4	1.0	1.2	1.4
AVDZ	6.6	9.3	11.4	5.7	5.0	5.6	5.1	6.3	7.3
AVTZ	1.3	1.4	2.0	3.0	2.6	2.2	1.7	1.5	1.3
AVQZ	0.9	1.0	1.2	0.7	0.7	0.7	1.5	1.3	1.2
VDZ-F12	4.0	2.9	4.7	5.5	3.4	2.3	2.8	2.4	2.9
VTZ-F12	1.1	1.1	1.7	3.4	2.8	2.4	2.1	1.6	1.3
VQZ-F12	0.8	0.8	1.0	1.1	1.0	0.9	0.7	0.6	0.6
SVP	48.7	58.5	65.4	34.1	42.7	49.0	37.6	45.7	51.5
TZVPP	4.7	7.4	9.9	4.3	4.8	6.3	4.7	5.9	7.3
QZVPP	4.1	3.9	4.1	4.1	3.6	3.4	5.4	4.8	4.4

VDZ-F12, or def2-TZVPP basis sets are sufficient to obtain basis set deviations comparable to or smaller than the intrinsic accuracy of the CCSD(T) method itself. That means that unless other higher level effects beyond CCSD(T) are considered explicitly (which are very expensive to calculate, see Sec. 5.4.2), the accuracy with respect to experimental values is nearly optimal already with such small sets. Already AVQZ or VQZ-F12 basis sets are sufficient to obtain basis set deviations from CBS[56] which are smaller than the ones from CBS[45]; so the question must be asked whether the CBS[56] reference values are actually more accurate than these QZ-level F12 treatments. The F12 methods can thus serve as a valuable tool for both practical calculations (when applied with small basis sets) and for establishing basis set limits (when applied with large basis sets). This latter aspect is particularly important because conventional basis set extrapolation does not actually work well for determining vibrational constants (e.g., in Tab. 5.13, for CCSD(T)/AVQZ, $\delta\omega_e = 6.5\,\mathrm{cm}^{-1}$ while for CBS[34] it is still nearly half as large at $\delta\omega_e = 3.0\,\mathrm{cm}^{-1}$); here the F12 methods offer a practical alternative.

In the current case the def2-nZVPP basis sets, maybe except TZVPP, do not produce good results. Diffuse augmentation of the sets or a more careful construction of the RI basis sets (particularly for SVP) might bring improvement in this regard, but here this topic is not further investigated.

Tab. 5.15 illustrates the effect of the geminal exponent γ on the calculation results for the different basis sets and F12 methods. The most reliable set of columns here is the one for the non-approximate RMP2-F12 method. These columns exhibit a rather weak dependence on γ, but it must be noted that the set of molecules is too small to draw firm conclusions or determine optimal γ parameters. At first glance,

the rows for AVQZ and VQZ-F12 in Tab. 5.15 seem to indicate that no improvement can be obtained by going from AVTZ or VTZ-F12 to AVQZ or VQZ-F12. However, this impression is likely misleading because it is quite doubtful whether the CBS[56] reference values (and maybe also the fitting process) are actually accurate to more than 1–2 cm^{-1}.

5.4.2 Diatomic molecules: Accurate calculations with high-level composite methods

CCSD(T) calculations offer an excellent accuracy for many kinds of applications, but in some cases one might really be interested in getting more accurate results. In order to obtain them, several physical effects need to be considered which are neglected in a default application of the CCSD(T) method. For molecules composed of light atoms, the following such effects play dominant roles [83]:

Core-core and core-valence correlation: Typically, only valence orbitals are subjected to the correlation treatment, and the core orbitals are frozen. For the core electron space, the wave function is thus not corrected from its description in the reference function. The reason for this is not that core correlation is a small effect—quite on the contrary, especially core-core correlation energies are very large. However, the core shells are hardly influenced by the chemical environment; so for relative energies the neglect of core correlation is usually a very good approximation. But it can still account for several kJ/mol in atomization energies, for example.

Higher excitations in the wave function: CCSD produces wave functions which exactly fulfill the many-body Schrödinger equation [in the given AO basis] within the Fock space spanned by the reference function and its singly and doubly excited configurations. For the subspaces of higher excitations the wave function can, in principle, be arbitrarily wrong. In some applications a more accurate treatment of many-body effects is desirable; this can be achieved, for example, by an explicit treatment of higher excitations with the CCSDT(Q) or CCSDTQ methods.

Relativistic effects: The Schrödinger equation is non-relativistic. For first- and second-row elements this is not a large problem because both core- and valence electrons move much slower than the speed of light. But even then the modified kinetic energy operator and electron/nucleus contact terms might lead to relative energy differences in the 1 kJ/mol range. Similarly, also the non-scalar relativistic effect of spin-orbit coupling may need to be considered in the sub-kJ/mol range (for heavy elements LS coupling is essential, but for light elements it is only another small effect).

Quantum nature of nuclei, electronic coupling: Almost all electronic structure calculations employ the Born-Oppenheimer approximation in its clamped nuclei variant: The electrons see the nuclei only as fixed point charges in space, which enter the electronic calculation merely by the Coulomb potential v_{nuc} they emit. On the electronic side, this is a very good approximation, at least in the absence of vertical (near-)degeneracy of electronic states. Nevertheless, correction terms for the coupling between electron motion and nuclear motion should be accounted for in sub-kJ/mol accuracy calculations when very light atomic cores, like hydrogen, are present. This is usually done in the form of the diagonal Born-Oppenheimer correction (DBOC).

Quantum treatment of nuclei: A more serious problem is an accurate treatment of the zero-point motion of the nuclei themselves. Scaled zero point vibrational energies (ZPVEs) calculated from DFT, Hartree-Fock, or MP2 harmonic frequencies are typically sufficient to get relative vibrational energies accurate to less than 1 kcal/mol. However, in calculations with higher accuracy standards, it may be necessary to (i) go beyond the harmonic approximation and actually calculate core wave functions (e.g., by employing a vibrational CI method [146]) *and* to (ii) use more accurate electronic structure methods as basis for the potential energy surface on which the core wave function is adjusted. These issues can get very problematic when floppy molecules need to be treated.

Ref. [83] gives a very good overview over the relative importance and the interplay of the various physical effects for molecules composed of light elements. Of special importance is the fact that corrections for the two most important approximations—namely, core-valence correlation and neglect of higher valence excitations—have the unfortunate tendency of having a comparable size but opposite signs. Therefore *both* need to be considered in order to improve relative energies. A similar statement holds for the higher correlation classes: CCSDT tends to be *less* accurate than CCSD(T), and CCSDT(Q) is generally the first many-body method which reliably improves on CCSD(T). This is bad because CCSDT(Q) is a very complicated method, and features a canonical scaling of $O(N^9)$ where N is some measure of the molecule size (CCSD(T) has $O(N^7)$). Thus, it generally is only tractable for very small molecules and small basis sets. Higher coupled cluster methods fare even worse.

Previous studies [83, 84] have established that very high accuracy can be obtained with first principles composite methods. Such methods assume that the various small physical effects due to different sources have negligible quantitative interplay among each other and can thus be combined additively to a good approximation. In the same spirit we now perform some calculations on diatomic molecules to investigate how augmentations of F12 calculations by comparatively simple correction terms can improve agreement of theoretical data with experimental data. The potential

energy curve is now calculated with energies composed as

$$E = E_{\text{val}} + \Delta E_{\text{core-val}} + \Delta E_{\text{rel}} + \Delta E_{\text{HL}}, \tag{5.1}$$

where E_{val} is calculated with CCSD(T)-F12b/VQZ-F12/γ=1.1 employing the (default) frozen core approximation; $\Delta E_{\text{core-val}}$ is a correction for core-valence correlation, calculated as

$$\Delta E_{\text{core-val}} = \text{CCSD(T)-F12b[all elec.]} - \text{CCSD(T)-F12b[frozen core]} \tag{5.2}$$

at aug-cc-pwCVTZ [147] level [with γ=1.4 and QZVPP/JKFIT [142] for JKFIT and AVTZ+d/OPTRI [135] for RI]; ΔE_{rel} is a relativistic correction (consisting of the expectation values of the mass-velocity term and one-electron Darwin term; called Cowan-Griffin operator [148] in the Molpro manual), calculated at conventional CISD/VQZ-F12 level; and ΔE_{HL} is a correction for higher level correlation,

$$\Delta E_{\text{HL}} = \text{CCSDT(Q)} - \text{CCSD(T)} \tag{5.3}$$

calculated at frozen core cc-pVDZ or cc-pVTZ level. The CCSDT(Q) [149] calculations were performed with the MRCC program of Kállay [150, 151] as interfaced into Molpro. Regarding the computational cost, E_{val} is somewhat more expensive than $E_{\text{core-val}}$, while ΔE_{rel} is cheaper, and ΔE_{HL} at VDZ level is similarly expensive as the other ones combined, despite being basically the least expensive sensible correction for higher level correlation effects that is imaginable.[6]

Tab. 5.16 gives a statistical overview over the results (compared to experiment) which are obtained with this energy formula. Tab. 5.17 gives data for ω_e for the individual molecules. The most striking feature of Tab. 5.16 is that the intrinsic accuracy of frozen core CCSD(T) ("E_{val} only" row) with respect to experiment can only be surpassed when corrections for *both*, core-valence correlation and higher level correlation are considered. Including only corrections for core-valence correlation and/or relativistic effects, for example, deteriorates the results instead of improving them (compare the rows "$E_{\text{val}} + \Delta E_{\text{core-val}}$" and "$E_{\text{val}}$ only").

This dependence on corrections for higher level correlation effects is most unfortunate, because—due to their computational cost—the HL corrections can only be evaluated for very small molecules. In fact, even for the diatomic molecules presented here, the crossover point between the F12 calculations and the CCSDT(Q)/VDZ calculations is already reached, even though our current implementation of F12 methods is not very efficient for small molecules with high symmetry. I would like to state explicitly that this high computational cost of the CCSDT(Q) method is caused by

[6] For example, one energy point of the O_2 molecule required 110s for ΔE_{val}, 7s for ΔE_{rel}, (59+40)s for $\Delta E_{\text{core-val}}$, and 171s for ΔE_{HL} at VDZ level or 7137 s at VTZ level, all on an Opteron 2.6 GHz processor. The MP2-F12 parts of the F12 calculations did not even take molecular symmetry into account (which could have accelerated them by a factor of 10–60).

Table 5.16: RMS deviations from experimental values in Tab. 5.12 for 11 of the 13 molecules [NO and CN were excluded due to difficulties in converging UHF]. The rows give the results obtained when only the indicated components of the energy in eq. (5.1) are considered. r_e are given in pm, ω_e and $x_e\omega_e$ in cm^{-1}.

	ΔE_{HL} with VDZ			ΔE_{HL} with VTZ		
Components	ω_e	r_e	$\omega_e x_e$	ω_e	r_e	$\omega_e x_e$
all	3.0	0.08	0.3	2.4	0.08	0.3
$E_{\mathrm{val}} + \Delta E_{\text{core-val}} + \Delta E_{\mathrm{rel}}$	12.3	0.16	0.6	12.3	0.16	0.6
$E_{\mathrm{val}} + \Delta E_{\text{core-val}}$	13.5	0.16	0.6	13.5	0.16	0.6
E_{val} only	9.6	0.19	0.8	9.6	0.19	0.8
$E_{\mathrm{val}} + \Delta E_{\mathrm{HL}}$	7.8	0.26	0.4	7.3	0.25	0.3

Table 5.17: Influence of individual corrections of eq. (5.1) on harmonic frequencies $\omega_e/[\mathrm{cm}^{-1}]$. "Exp." denotes the experimental values, "Tot." the theoretical values obtained with the full formula (5.1), "Val." the values obtained from a E_{val} potential curve. The Δ_X values give the change in ω_e when using a $E_{\mathrm{val}} + \Delta E_X$ potential instead of E_{val} alone. HL corrections were computed with the VDZ basis.

Molecule	Exp.	Tot.	Val.	$\Delta_{\text{core-val}}$	Δ_{HL}	Δ_{rel}
HF	4138.3	4137.3	4141.9	+4.7	-6.9	-2.5
N_2	2358.6	2355.4	2363.0	+8.8	-15.7	-0.9
CO	2169.8	2165.4	2167.0	+9.1	-9.9	-0.8
BF	1402.1	1398.8	1394.8	+9.3	-4.9	-0.5
F_2	916.6	917.6	930.3	+1.1	-13.0	-0.7
C_2 ($^1\Sigma_g^+$)	1854.7	1856.9	1861.4	+10.3	-14.5	-0.6
OH ($^2\Pi$)	3737.8	3738.5	3743.3	+4.9	-7.5	-2.2
NH ($^3\Sigma^-$)	3282.3	3281.6	3286.5	+4.6	-7.8	-1.7
CH ($^2\Pi$)	2858.5	2860.9	2858.6	+7.3	-4.0	-1.1
O_2 ($^3\Sigma_g^-$)	1580.2	1582.7	1605.4	+5.0	-25.5	-2.3
CF ($^2\Pi$)	1308.1	1301.8	1308.9	+5.8	-11.9	-1.2

the method's complexity itself and is not an artificial feature caused by being evaluated with a bad program—quite on the contrary, Kállay's ingenious MRCC program is very efficient for high-order coupled cluster methods. Nevertheless, for larger molecules, CCSDT(Q) quickly becomes intractable. For this reason, the intrinsic accuracy of CCSD(T) is currently an upper bound to the accuracy possibly achievable for chemical systems for which CCSDT(Q) calculations cannot be afforded. And unless electronic structure methods are discovered which far surpass the power of the coupled cluster hierarchy, this is unlikely to change any time soon.

5.4.3 Atomic dipole polarizabilities: Basis set convergence

As a first test of derived properties, we investigated static dipole polarizabilities of first-row atoms. These were calculated as numerical derivatives of the atom energy with respect to the field strength of a small dipole field:

$$\alpha = \left.\frac{\partial^2}{\partial h^2}\right|_{h=0} \left\langle \Psi \,\middle|\, \hat{H} + h \cdot \hat{d}_z \,\middle|\, \Psi \right\rangle, \tag{5.4}$$

where H is the Hamiltonian and $\hat{d}_z$ is the dipole field operator centered at the atomic nucleus. (This relation derives from *Hellman-Feynman theorem* for general properties; in practice it is realized by adding finite dipole fields to the core Hamiltonian in (2.14), and for each field strength, performing both RHF and correlation calculation with this modification).

Atoms which are not in a S state (i.e., which are not spherically symmetric) have different polarizabilities depending on the relative orientation of the dipole field to the wave function polarization. For these atoms in P states we define the average polarizability $\bar{\alpha}$ and the anisotropy $\Delta\alpha$ as

$$\bar{\alpha} = \tfrac{1}{3}\alpha_{\parallel} + \tfrac{2}{3}\alpha_{\perp} \tag{5.5}$$

$$\Delta\alpha = \alpha_{\perp} - \alpha_{\parallel}, \tag{5.6}$$

where $\alpha_{\parallel}$ is the component for dipole field and wave function polarization being parallel, and $\alpha_{\perp}$ is the component for both being perpendicular. These definitions are consistent with Ref. [152]. The electronic polarizabilities are given in atomic units: 1 a.u. $= e^2 a_0^2 E_h^{-1} \approx 1.64877727 \cdot 10^{-41}$ $\mathrm{C^2 m^2 J^{-1}}$ [152], corresponding to gas volume polarizability $\frac{\alpha=1\,\text{a.u.}}{4\pi\varepsilon_0} = 1.481847 \cdot 10^{-31}$ $\mathrm{m^3}$.

The numerical second derivatives are determined from the five point formula

$$\begin{aligned}\alpha &\approx \tfrac{1}{h^2}\big[-\tfrac{1}{12}E(2h) + \tfrac{16}{12}E(h) - \tfrac{30}{12}E(0) + \tfrac{16}{12}E(-h) - \tfrac{1}{12}E(-2h)\big] \\ &= \tfrac{1}{h^2}\big[-\tfrac{2}{12}\big[E(2h) - E(0)\big] + \tfrac{32}{12}\big[E(h) - E(0)\big]\big],\end{aligned} \tag{5.7}$$

where h is the finite field strength and $E(x)$ is the system energy when a dipole field of strength x is applied. Numerical experimentation indicated that for the first row atoms a field strength of $h = 1.0 \cdot 10^{-3}$ a.u. is appropriate, which results in approximately five to six significant digits in the polarizabilities (see Tab. 5.18) when UCCSD(T) is converged to 10^{-14} $\mathrm{E_h}$[7].

Tab. 5.19 gives an overview over the basis set convergence obtained with conventional methods and F12 methods. It lists regular augmented orbital basis sets

[7] This criterion proved to be unproblematic even for TAV6Z basis sets, so I never tried a different one.

Table 5.18: α from formula (5.7) for UCCSD(T)/DAVTZ with different field strengths h for the largest (B) and smallest (Ne) αs occurring in the test set. Actual calculations were done with $h = 1.0 \cdot 10^{-3}$, leading to 5–6 significant digits expected accuracy.

h/[a.u.]	α(Ne)/[a.u.]	$\alpha_\parallel$(B)/[a.u.]	$\alpha_\perp$(B)/[a.u.]
$3.20 \cdot 10^{-2}$	2.68843451	(no conv.)	(no conv.)
$1.00 \cdot 10^{-2}$	2.68884939	24.57367415	18.65665027
$3.20 \cdot 10^{-3}$	2.68885288	24.68584039	18.66098439
$1.00 \cdot 10^{-3}$	2.68885213	24.68653895	18.66102302
$3.20 \cdot 10^{-4}$	2.68884234	24.68653333	18.66100998
$1.00 \cdot 10^{-4}$	2.68876098	24.68641895	18.66090050
$3.20 \cdot 10^{-5}$	2.68795171	24.68527350	18.65979015

(the AVnZ sets defined before) and doubly augmented sets (DAVnZ). Calculations with triply augmented basis sets (TAVnZ) were also performed, but are not shown here because their results did not differ significantly from the DAVnZ results. The DAVnZ and TAVnZ orbital series are obtained by adding sets of even tempered diffuse functions to the AVnZ sets. Extrapolated results are obtained by l^3 extrapolation as described in Sec. 2.3; in order to indicate the augmentation level of the basis set, they are labeled for example as DAV[xy]Z for a value extrapolated from DAVxZ and DAVyZ calculations. F12 methods were applied with AVnZ/JKFIT as basis sets for RI and JKFIT and AVnZ/MP2FIT for MP2FIT; the MP2-F12 method used was DF-RMP2-F12 with $\gamma = 1.0$ and ansatz 3C(FIX) or, for the marked rows, 3C(D). For CCSD(T)-F12 calculations the latter choice meant that frozen diagonal T^{ij}_{mn} amplitudes from MP2-F12 were used in CCSD-F12.

In Tab. 5.19 we first see that double augmentation of the orbital basis sets is generally necessary in order to obtain accurate polarizabilities; this applies to all methods, pure conventional, extrapolated methods, and F12 methods. However, there are significant differences with respect to the considered atom (not shown): For boron, values from singly augmented calculations are fine (both for conventional and for F12 calculations), while for neon even the AVDZ Hartree-Fock polarizability deviates by 23% from the basis set limit (1.832 a.u. versus 2.376 a.u.).

The next striking feature is that the convergence improvements obtained by F12 methods is rather unimpressive compared to the ones found in the previous sections. In particular, CCSD(T)-F12b offers only minor advantages over conventional CCSD(T) for absolute values (the relative values look better, but are still unimpressive). But also RMP2-F12, which is exactly implemented, appears to have trouble with polarizabilities. These problems might be a result of

- The small impact of the correlation treatment on polarizabilities: The small dipole fields mainly lead to a small displacement of the electron density in

diffuse outer regions of the molecule. Since the electron density is low there, electrons spend less time near each other compared with the proper valence space. As a result, correlation itself is not so important (even Hartree-Fock fares comparatively well), and furthermore, also the basis set requirements for correlating higher angular momentum functions appear to be rather low. Both aspects facilitate accurate conventional calculations.

- The relevant correlation effects occurring on multiple different length scales. Sec. 3.2 noted that the size of the correlation hole is related to the electron density. For the diffuse outer regions of molecules, the single Slater function we employ might simply not be a good representation of the correlation hole. This impression is reinforced by the large convergence improvements obtained in MP2-F12 by going from ansatz 3C(FIX) to 3C(D). To a certain degree, the variable amplitudes can correct for the length scale effect by amplifying the F12 contributions for diffuse occupied orbitals. While in most calculations choosing either 3C(D) or 3C(FIX) has only minor impact on calculation results [Sec. 5.2.2], in this particular case the effects are large; for example, the MP2-F12/DAVQZ deviations go down from 0.06% for 3C(FIX) to 0.01% for 3C(D). Also F12b results are significantly improved by the ansatz change. The effects are likely less pronounced there because the used T_{mn}^{ij} amplitudes were optimized for MP2-F12, and not for CCSD-F12.

Compared to other kinds of calculations, the increased importance of one-particle relaxation also has the effect that double-ζ-level results can be improved by evaluating complete RMP2 singles in the combined MO+CA basis instead of applying only the spin-free singles part for RHF relaxation [Sec. 5.2.1]. For larger basis sets either choice has no noteworthy effect. As a last observation, we might note that in this case of polarizabilities the F12a method up to DAVQZ appears to work better than F12b in absolute terms; for relative improvements there is not much difference. Similar behavior was found in some benchmarks in Sec. 5.2.3.

5.4.4 Atomic dipole polarizabilities: Accurate calculations

Interestingly, despite being hard to measure experimentally [153] and despite their importance for constructing asymptotic interaction potentials in various application scenarios (e.g., Refs. [154–156]), accurate calculations of such polarizabilities are sparse. The PNO-CEPA calculations of Werner and Meyer [157] apparently still are the most accurate systematic investigations of polarizabilities for first row atoms[8]. A

[8]There are some CCSD(T) calculations, but in the absence of strong correlations there is no reason to believe that CCSD(T) alone is more accurate than the fully relaxed PNO-CEPA with carefully selected basis sets.

Table 5.19: Basis set convergence of dipole polarizabilities α for the set [α from Ne and N; $\alpha_\parallel$ and $\alpha_\perp$ of B, C, O, F]. (top): rms deviations from conventional TAV[56]Z reference values in a.u.; (bottom): rms of relative deviations $[\frac{\text{val}-\text{ref}}{\text{ref}} \cdot 100\%]$ in %. "Conv" denotes conventional calculations, "Co+S." denotes RHF+CABS correction.

	δRHF		δMP2		δCCSD(T)		
Basis	Conv	Co+S.	Conv	F12	Conv	F12a	F12b
AVDZ	0.520	0.134	0.562	0.255	0.591	0.303	0.291
AVTZ	0.132	0.034	0.185	0.093	0.168	0.095	0.108
AVQZ	0.027	0.010	0.063	0.031	0.046	0.031	0.040
AV[23]Z			0.145		0.147		
AV[34]Z			0.019		0.019		
AV[45]Z			0.010		0.008		
DAVDZ	0.081	0.010	0.269	0.064	0.151	0.099	0.086
DAVTZ	0.015	0.002	0.108	0.020	0.061	0.016	0.058
DAVQZ	0.008	0.003	0.058	0.009	0.039	0.008	0.026
DAVDZ[a]		0.020		0.039		0.106	0.078
DAVTZ[a]		0.002		0.010		0.016	0.037
DAVQZ[a]		0.003		0.002		0.008	0.013
DAV[23]Z			0.029		0.042		
DAV[34]Z			0.011		0.011		
DAV[45]Z			0.003		0.004		

	δ%RHF		δ%MP2		δ%CCSD(T)		
Basis	Conv	Co+S.	Conv	F12	Conv	F12a	F12b
AVDZ	12.35	3.13	14.00	6.14	13.89	6.10	6.28
AVTZ	3.82	0.96	4.74	2.47	4.58	2.39	2.51
AVQZ	0.87	0.28	1.23	0.84	1.13	0.81	0.87
AV[23]Z			4.03		3.92		
AV[34]Z			0.58		0.56		
AV[45]Z			0.17		0.17		
DAVDZ	0.78	0.15	1.61	0.38	1.97	0.91	0.58
DAVTZ	0.08	0.02	0.74	0.15	0.89	0.26	0.33
DAVQZ	0.11	0.02	0.47	0.06	0.53	0.11	0.14
DAVDZ[a]		0.19		0.35		0.92	0.61
DAVTZ[a]		0.02		0.08		0.24	0.21
DAVQZ[a]		0.03		0.01		0.11	0.08
DAV[23]Z			0.38		0.53		
DAV[34]Z			0.20		0.18		
DAV[45]Z			0.05		0.03		

(a) Ansatz 3C(D) instead of 3C(FIX) and F12-singles instead of RHF-relax [Sec. 5.2.1]

Table 5.20: Average polarizabilities $\bar{\alpha}$/[a.u.] and anisotropies $\Delta\alpha$/[a.u.] computed with the high level composite method. "Tot." denotes values obtained with the full formula (5.8), "Val." values obtained from a E_{val} potential curve. The Δ_X values give the change in α when using a $E_{\text{val}} + \Delta E_X$ potential instead of E_{val} alone.

Atom	Tot.	Val.	$\Delta_{\text{core-val}}$	Δ_{HL}	Δ_{rel}
$\bar{\alpha}$(B)	20.446	20.619	−0.169	+0.000	−0.005
$\bar{\alpha}$(C)	11.711	11.755	−0.075	+0.035	−0.004
α(N)	7.267	7.292	−0.036	+0.014	−0.004
$\bar{\alpha}$(O)	5.231	5.231	−0.021	+0.018	+0.002
$\bar{\alpha}$(F)	3.685	3.685	−0.011	+0.008	+0.004
α(Ne)	2.664	2.665	−0.007	+0.003	+0.004
$\Delta\alpha$(B)	−5.803	−5.981	+0.192	+0.000	−0.014
$\Delta\alpha$(C)	2.092	2.142	−0.053	−0.004	+0.007
$\Delta\alpha$(O)	−0.901	−0.884	+0.011	−0.017	−0.011
$\Delta\alpha$(F)	0.409	0.403	−0.005	+0.004	+0.007

recent investigation of static polarizabilities for second-row atoms can be found in Ref. [152]; this article also contains pointers to specific problems occurring in their calculation.

Although F12 methods did not fare too well in the calculation of atomic polarizabilities, we here repeat the calculation of Sec. 5.4.3 using a composite method like discussed in Sec. 5.4.2. Again, the energy formula is

$$E = E_{\text{val}} + \Delta E_{\text{core-val}} + \Delta E_{\text{rel}} + \Delta E_{\text{HL}}, \tag{5.8}$$

but now E_{val} is calculated with CCSD(T)-F12b/DAVQZ/γ=1.1 and ansatz 3C(D), $\Delta E_{\text{core-val}}$ is calculated as with d-aug-cc-pwCVTZ [with γ=1.4, QZVPP/JKFIT for JKFIT and AVTZ/JKFIT for RI, and also ansatz 3C(D)]; ΔE_{rel} is calculated at CISD/DAVQZ level and and ΔE_{HL} with CCSDT(Q) and the DAVDZ basis sets, all using the same programs and techniques as before.

Tab. 5.20 reports the obtained results. Here we can see that, contrary to the case of spectroscopic constants, neither the higher-level correlation effects (expressed through ΔE_{HL}) nor the relativistic effects (expressed through ΔE_{rel}) do very much at all. For some atoms, core-valence correlation has a significant influence on the results, though.

As noted in Ref. [152], accurate experimental data for such polarizabilities are hard to come by, save for rare gases. For this reason no in-depth comparison to experiment is performed. These calculations therefore have the character of (uncertain) predictions, although I do not claim that they are the most accurate calculations that can be performed (each single energy point required less than five minutes of

CPU time). Nevertheless, for the neon atom the value of $\alpha = 2.664$ a.u. found is in decent agreement with the experimental values $\alpha = 2.669$ a.u. of Ref. [158] and $\alpha = (2.678 \pm 0.003)$ a.u. of Ref. [159] (especially when considering that the latter reference also gives a value of $\alpha = (0.139\,07 \pm 0.000\,10)$ a.u. for helium, which lies way beyond the error boundaries of the accepted theoretical value of 0.138 376 1 a.u. [160]).

5.4.5 Reaction barrier heights

Reaction barrier heights are of considerable interest for users and developers of electronic structure methods alike. The development interest rests on the fact that most density functional theory (DFT) methods perform poorly for this task [161–164], even if they are able to capture reaction energies sufficiently well. Some DFT methods were re-parameterized for reaction barrier heights [165–167], but this domain specific fitting usually deteriorated thermochemical results in the process [168, 169]. Some authors of highly parameterized empirical DFT methods claim to have achieved a balanced description of both thermochemistry and barrier heights [170], but to this day no such "density functional" has seen universal acceptance. The ubiquitous (also empirically fitted) B3LYP functional remains the most often applied DFT method for the task, despite being known to produce poor energy barriers.

Here we probe our explicitly correlated coupled cluster methods for the DBH24 benchmark set [172] of reaction barrier heights from the Truhlar group. This is a set of 22 reaction barrier heights[9] for which accurate theoretical reference data is available [171]. We performed RHF-UCCSD(T)/CBS[45] reference calculations and UCCSD(T)-F12x (x=a,b) calculations (using default parameters) using the original QCISD/MG3 geometries of Ref. [172]. Tab. 5.21 lists the DBH24 reactions. For their forward and reverse barriers, W4 reference values from Ref. [171], the original W1 reference values from Ref. [172], and our own UCCSD(T)-F12b and CBS[45] values are given (the F12a results are not very different from F12b). The W4 [83] reference values include corrections for higher order correlation, core-valence correlation, relativistic effects, and non-adiabatic effects, but vibrational zero point energies were omitted (i.e., we compare only the electronic barriers at 0K). Thus, the difference between the CCSD(T)/CBS[45] values and the W4 values reflects these high level effects, which are expensive to calculate, and which normally cannot be accounted for in a routine calculation.

We can see in Tab. 5.21 that the F12 values are very accurate—not only with

[9] Two of the barrier heights belong to symmetric reactions and occur as *both* forward and reverse barriers. In statistics we will count them as independent variables (i.e., twice), because this is done in other articles dealing with DBH24. For the same reason, all DBH24-related values are given in kcal/mol instead of kJ/mol.

Table 5.21: Barrier heights of the DBH24 set. Given are UCCSD(T)-F12b values, the very accurate W4 theory [83] reference data from Ref. [171] and the reference values Zheng et al. originally suggested for the set [172]. The last four lines give MAD and RMS deviations from W4 and UCCSD(T)/CBS[45] values, respectively.

Reaction	F12b[a]	F12b[b]	W4[c]	W1[d]	CBS
Forward barrier heights [kcal/mol]					
$H + N_2O \rightarrow OH + N_2$	17.47	17.49	17.23	18.14	17.47
$H + ClH \rightarrow HCl + H$	18.46	17.82	17.47	18.00	17.70
$CH_3 + FCl \rightarrow CH_3F + Cl$	7.20	6.86	6.77	7.43	6.99
$Cl^- \cdots CH_3Cl \rightarrow ClCH_3 \cdots Cl^-$	14.19	13.47	13.42	13.61	13.38
$F^- \cdots CH_3Cl \rightarrow FCH_3 \cdots Cl^-$	3.89	3.38	3.45	2.89	3.34
$OH^- + CH_3F \rightarrow HOCH_3 + F^-$	-2.44	-2.71	-2.44	-2.78	-2.42
$H + N_2 \rightarrow HN_2$	14.90	14.47	14.52	14.69	14.50
$H + C_2H_4 \rightarrow CH_3CH_2$	2.05	1.91	1.84	1.72	1.97
$HCN \rightarrow HNC$	47.20	47.69	48.10	48.16	47.69
$OH + CH_4 \rightarrow CH_3 + H_2O$	5.95	6.14	6.54	6.70	6.34
$H + OH \rightarrow O + H_2$	10.36	10.57	10.74	10.70	10.81
$H + H_2S \rightarrow H_2 + HS$	4.13	3.85	3.76	3.60	3.87
Reverse barrier heights [kcal/mol]					
$H + N_2O \rightarrow OH + N_2$	82.96	82.81	82.56	83.22	83.08
$H + ClH \rightarrow HCl + H$	18.46	17.82	17.47	18.00	17.70
$CH_3 + FCl \rightarrow CH_3F + Cl$	59.01	59.64	60.04	61.01	59.49
$Cl^- \cdots CH_3Cl \rightarrow ClCH_3 \cdots Cl^-$	14.19	13.47	13.42	13.61	13.38
$F^- \cdots CH_3Cl \rightarrow FCH_3 \cdots Cl^-$	30.10	29.57	29.42	29.62	29.56
$OH^- + CH_3F \rightarrow HOCH_3 + F^-$	17.32	17.11	17.69	17.33	17.57
$H + N_2 \rightarrow HN_2$	10.61	10.87	10.73	10.72	10.79
$H + C_2H_4 \rightarrow CH_3CH_2$	41.87	42.19	41.83	41.75	42.11
$HCN \rightarrow HNC$	32.40	32.84	32.86	33.11	32.78
$OH + CH_4 \rightarrow CH_3 + H_2O$	19.13	19.29	19.51	19.60	19.63
$H + OH \rightarrow O + H_2$	12.92	13.15	13.18	13.10	13.28
$H + H_2S \rightarrow H_2 + HS$	16.72	17.32	17.83	17.30	17.32
Statistical deviations [kcal/mol]					
mad(W4)	0.51	0.22	(0)	0.33	0.19
rms(W4)	0.60	0.28	(0)	0.43	0.25
mad(CBS[45])	0.40	0.13	0.19	0.32	(0)
rms(CBS[45])	0.47	0.17	0.25	0.43	(0)

(a) UCCSD(T)-F12b, AVDZ

(b) UCCSD(T)-F12b, AVTZ

(c) W4 theory [83], excluding zero point vibrational effects, from Ref. [171]

(d) Original W1 reference values from the Truhlar group [172]

respect to the CBS[45] basis set limit estimates, but also with respect to W4 reference values which account for all dominant effects beyond CCSD(T). Even with the small AVDZ basis, CCSD(T)-F12b is not only more accurate than every single method investigated by Zheng et al. in Ref. [161] (including their own empirical Minnesota density functionals, which were fitted on a benchmark set related to the DBH24 reactions), but also nearly as accurate as the *reference values* given in Ref. [172]. F12b/AVTZ clearly surpasses this accuracy.

If we compare the CBS[45] values to the W4 values for this set of reaction barriers, we see a RMS deviation of $\approx$ 0.25 kcal/mol $\approx$ 1 kJ/mol. From that we can conclude that the intrinsic accuracy of CCSD(T) here lies in the same order of magnitude. Thus, in contrast to DFT methods, for CCSD(T) methods—including our F12 variants—reaction barrier heights apparently are not more difficult to calculate than reaction energies or similar quantities.

5.5 Automated generation of RI basis sets

In most applications reported so far, employing JKFIT sets as RI basis sets worked rather well, despite them being optimized for rather different purposes. For the def2-nZVPP basis sets, this was not the case. The universal JKFIT sets [142] were not diffuse enough, and were highly redundant when combined with the AO basis. They are also quite large. The non-existing diffuseness in the RI basis is a more serious problem for the def2-nZVPP sets than for the AVnZ or VnZ-F12 sets, because the latter ones are rather diffuse on their own, so only a smaller part of the one-particle space must be covered by the RI set.

Therefore new RI basis sets for def2-nZVPP were generated. The goal of the following construction is to obtain a set of functions which extends the one-particle space described by the AO basis with functions "similar" to the AO functions. The new RI sets consist of uncontracted Gaussian primitives exclusively. The following procedure was employed for each chemical element and each orbital basis set:

- For each angular momentum l, a set of Gaussian primitive exponents was determined which roughly describes the range of exponents covered by the basis set. For this purpose, for each contracted radial function $f_n(r) = r^l \sum_i c_{ni} \exp(-\zeta_i r^2)$ [see eq. (2.4). The r^l originates from the solid harmonic S_{lm}], a primitive function $f_\zeta(r) = r^l \exp(-\zeta r^2)$ was determined such that

$$\operatorname*{argmax}_{r\in[0,\infty)} r^2 f_n(r) = \operatorname*{argmax}_{r\in[0,\infty)} r^2 f_\zeta(r)$$

  (i.e., ζ was assigned such that the maximum $r_{\max}$ of the primitive and of the contracted function coincide when multiplied by r^2. The r^2 is supposed to account for the volume element $\mathrm{d}V = r^2 \sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi\,\mathrm{d}r$). For non-contracted functions, their primitive exponent is used directly. The set of all average exponents obtained for a single angular momentum in this fashion is called the *cover set* C_l of this angular momentum.

- For each l, an extended cover set E_l is created by extending C_l by one sharp and one diffuse even tempered exponent (e.g., $\zeta_{\min}^2/\zeta_{\min-1}$ on the diffuse side, when $\zeta_{\min}$ is the most diffuse and $\zeta_{\min-1}$ is the second-to most diffuse exponent). If there is only one exponent ζ in C_l, then $\zeta/2.5$ and $\zeta * 2.5$ are added instead. If C_l is empty, E_l is left empty. E_l is sorted by exponent.

- If $l \leq l_{\max}$, where $l_{\max}$ is the largest angular momentum present in the AO basis for the current element, then the RI set is constructed in such a way that additional primitives are added between the exponents already covered in E_l. For each two consecutive exponents ζ_i and ζ_{i+1} of E_l, an exponent $\sqrt{\zeta_i \cdot \zeta_{i+1}}$ is added to the RI set, unless it would lie too near to an exponent already present

in E_l. A new exponent ζ is considered too near if it lies within a factor of 1.4 from any exponent in E_l. This omission is done to prevent near-linear dependencies between the AO basis and the new RI basis.

- If $l = l_{\max} + 1$, then a new RI function is added for every exponent in C_{l-1} (i.e., in the previous angular momentum). This exponent ζ_{new} is determined in such a way that the maxima of $r^{l+2}\exp(-\zeta_{\text{new}} r^2)$ and $r^{l+1}\exp(-\zeta_{\text{old}} r^2)$ coincide, where ζ_{old} is the exponent in C_{l-1}. Again, the factor r^2 is supposed to account for the r^2 in the volume element.

- This was the general process. The following special rules are applied:
  - Prior to any other considerations, sharp exponents are removed from the cover set C_l as long as the two sharpest exponents lie within a factor of 1.4 of each other. Such close exponents commonly appear in the QZVPP set, apparently resulting from their split valence nature.
  - For H and He no sharp even tempered functions are added to the extended cover set, only diffuse ones.
  - For the s group elements (H, He, Li, Be, Na,...) no RI functions are added at level $l_{\max} + 1$.
  - For the elements from atomic number 11 to 18, one core s function is deleted from E_l. For atomic number 19 to 36, two core s functions and one core p function are deleted. For higher elements no core functions are deleted, since the core is represented implicitly by an effective core potential.

While some care has been taken in the construction of these RI sets, no claim is made that this concrete process is the most reasonable one imaginable. The generated sets appear to work significantly better than universal JKFIT sets for RI purposes, but further experimentation is clearly required. For example, it is by no means clear what exactly is the most important part of the one-particle space that must be represented by the RI expansions; here we simply assumed that the critical part is covered by the AO set itself, and that we simply need to be extend it somewhat. Also it is not clear if the chosen way of assigning primitive exponents to contracted functions for the cover generation is the best way of approaching this problem (for example, considering $r^2 f(r)^2$ instead of $r^2 f(r)$ as the basic quantity of the radial part might have been a better choice, because in matrix elements with themselves, wave functions occur in squared form; or one could demand that different quantities than $r_{\max}$ coincide for determining the "closest" primitive). Nevertheless, I am quite optimistic that some way of generating decent RI basis sets automatically can be found, and that the algorithm proposed here is a viable starting point.

Table 5.22: Overview of Accuracy: RMS deviations from UCCSD(T)/CBS[56] of F12a and conventional calculations for the benchmark sets presented previously.

	F12a/AVTZ	Conv/AVTZ
Reactions (open-shell)	1.3 kJ/mol	19.1 kJ/mol
Atomization Energies	1.3 kJ/mol	26.7 kJ/mol
Electron Affinities	12.4 meV	83.0 meV
Ionization Potentials	13.1 meV	105.3 meV
Equilibrium Distances	0.08 pm	0.59 pm
Harmonic Frequencies	3.0 cm^{-1}	21.0 cm^{-1}

5.6 Summary & Outlook: Calculations

Method summary

- We have shown that F12 corrections lead to tremendous improvements in basis set convergence for both RMP2 and RHF-UCCSD(T).

  Also the proposed corrections for RHF treatments lead to large improvements. This is important because our calculations have shown that considering some kind of Hartree-Fock correction is absolutely mandatory for small and medium-sized basis sets. Otherwise no accurate total energy differences can be expected.

  The proposed crude (T*) triples correction should be used with care, and usually only with F12b and large basis sets. It is not to be used for calculating potential energy surfaces. Nevertheless, in certain cases the (T*) correction is an attractive choice.

- The approximate, simplified CCSD(T)-F12x (x=a,b) methods are not quite as reliable for obtaining complete basis set limits for CCSD(T) as RMP2-F12 is for RMP2. That is a very high standard, however, and by all means, the F12x (a,b) methods do work very well. An overview for UCCSD(T)-F12a, when applied in a black-box fashion with default parameters only, is given in Tab. 5.22. All of the F12/AVTZ values reported in Tab. 5.22 are so close to the basis set limit that the remaining basis set deviation is unlikely to be the largest source of inaccuracy in routine calculations.

- At the CCSD(T) level, the F12 corrections have only minor impact on the calculation cost due to their favorable computational scaling. Thus, F12 variants should be preferred to conventional CCSD(T) in almost all cases.

- Already with medium sized basis sets (AVTZ), CCSD-F12x methods reach the intrinsic accuracy of the CCSD(T) method: There is a good balance of the

remaining errors due to different sources (perturbative triples and neglect of higher excitations, core-valence correlation, relativistic effects, non-adiabatic effects etc.)

⇒ Unless *all* these effects were accounted for in separate calculations, a further improvement in basis set convergence would not lead to better total energies.

⇒ That means that more complicated CCSD-F12 treatments than F12x (x=a,b) are not necessary for any kind of routine calculation, in which these extra calculations are not done.

Basis-set and calculation task summary

- Regarding the tested basis sets, the AVnZ series is a good choice for F12 calculations, as is, quite naturally, the VnZ-F12 series. The non-augmented def2-nZVPP sets were found to not work too impressively, although they certainly performed much better than the non-augmented cc-pVnZ sets (not shown). def2-TZVPP calculations may be a good choice when facing chemical elements for which auxiliary basis sets for the AVnZ or VnZ-F12 series are not yet available. Augmenting the def2-sets with (even tempered) diffuse functions would likely also improve their F12 performance significantly, but this has not been tested here.

- Regarding the tested calculation types, F12 methods work best for energy differences between like species with the same number of electrons (e.g., reaction energies, atomization energies, reaction barrier heights) and (geometric) potential energy surfaces (e.g., for spectroscopic constants of molecules).

  Similar results were also found for interaction energies of weakly interacting dimers [136, 137] and both harmonic and anharmonic vibrational spectroscopic constants of polyatomic molecules [173, 174]. In a benchmark study of many vibrational modes of different molecules, Ref. [173] found CCSD(T)-F12a/AVTZ anharmonic frequencies to have a mean absolute deviation of only $4\,\mathrm{cm}^{-1}$ from experimentally measured frequencies.

Outlook

- All of the molecules considered in the calculations were composed of only hydrogen and first- and second-row main-group elements. If/how well the F12 methods fare for chemical elements with more electrons remains to be seen. A preliminary support for effective core potentials (ECPs) was recently added to Molpro's F12 programs, but no extensive benchmarks have been performed yet. In particular, CCSD(T)-F12 applicability for transition metal chemistry must be investigated.

- One problem the F12 methods still have is that, additionally to the orbital basis sets, also auxiliary basis sets are required for the DF and the RI approximations. At this moment in time, such auxiliary basis sets exist neither for all potentially interesting chemical elements nor for all kinds of application scenarios (e.g., for core-valence correlation). Doubtlessly, over time such sets will be developed; the Peterson group is already working on specialized basis sets for F12 calculations. As an intermediate solution, the route of automatically generating adequate RI sets might also be followed further.

- The calculations in this thesis focused on valence correlation energies. While the F12 theory itself is just as valid for core-correlation and core-valence correlation, such treatments of core correlation might pose problems to the F12 methods because the relevant correlation effects occur at multiple different length scales. Such multiple length scales might not be well-described by a single Slater function. The Peterson group is currently investigating the matter.

- Basis set extrapolation schemes for F12 methods were recently developed [175], and these may be used to further increase the accuracy of F12 methods for high level calculations.

6 Closing Words

During this project, a great amount of time was spent on purely practical development issues, which were basically unrelated to the actual theories being developed. *Instead of making a bad, theoretically 102% justified CCSD-F12 program, we wanted to make a good approximate CCSD(T)-F12 program.* Because development time is a limited resource, sometimes this made it necessary to sacrifice a possibly more elaborate theoretical treatment for issues of greater concern in real applications. For examples, large amounts of time were spent with

- Synchronizing the open- and closed-shell programs, in order to make them obtain equal values when applied to closed shell molecules. Achieving this rather innocent looking goal—in all typical application scenarios—turned out to be one of the hardest problems encountered during this work. It could only be solved by the combined efforts of both Prof. Werner and me.

- Detecting and repairing problems with numerical stability. Originally, the (closed-shell) lmp2-f12 program and the rmp2-f12 program used very different approaches for calculating their intermediates. If applied with realistic basis sets, they produced significantly different results despite implementing exactly the same theory. This not only made comparisons of intermediate results difficult, it was also unclear *which* program was right. We identified the main reason for the discrepancies to lie in a particularly ill-conditioned CABS projector transformation in lmp2-f12, which has since been updated to the more stable version described in Ref. [103].

- Adjusting the programs to allow execution when molecular symmetry was present. This was a major problem because Molpro's density fitting integral evaluation subroutines were designed for local correlation methods and would refuse to run unless symmetry was disabled.

- The smooth integration of the rmp2-f12 program into the Molpro program system (e.g., recognizing the numerous input data, options, and output variables; converting data and calculation intermediates between different programs, etc.).

- Improving support for calculations of potential energy surfaces and derivative properties by allowing the auxiliary basis sets to be frozen at some reference geometry.

- Detecting and repairing program errors which occurred only on certain compiler- and hardware platforms, and working around compiler bugs.

While not exactly scientific in nature, such developments, and the corresponding practical assessments, were necessary in order to create a *production level* program which is both stable and efficient in real scientific applications. The benchmark results presented in Sec. 5 demonstrate that we (Prof. Werner, I, and Thomas Adler) succeeded in this goal.

7 Acknowledgments

Foremost I would like to express my deep gratitude to Prof. Hans-Joachim Werner, who supported me not only with his vast scientific knowledge, but also with his tolerance for my somewhat unconventional style of work- and time scheduling. I thank the professors Hermann Stoll, Guntram Rauhut, and László von Szentpály for always keeping an open ear to my inquiries, Christof Hättig for pointing out some errors in the manuscript, and Peter Fulde for introducing me into the world of the electron. My thanks go to K. R. Shamasundar, Thomas B. Adler, Kirk A. Peterson and J. Grant Hill, for all the collaborations, feedback and discussions, which proved most valuable. I thank all my labmates at the institute for keeping such a good atmosphere, and especially Oliver Marchetti for all the sport, talk, and coffee :]. I am in indebted to Stefan Jagiella, Johannes Dietrich, Klaus Pflüger and Michael Neff for all the effort they spent on shepherding our computers and, particularly, for occasionally rescuing my data.

This project was funded by the Deutsche Forschungsgemeinschaft (DFG) in the priority program SPP 1145 (*Modern and universal first-principles methods for many-electron systems in chemistry and physics*). Apart from the financial support, I am also deeply grateful for the opportunity to meet many colleagues in the SPP-organized symposia and workshops, and for numerous interesting discussions on the fascinating topic of electronic structure theory.

Additional resources and programs used in this work:

- Scripting for setup, execution and evaluation of benchmark calculations was done with Python 2.4.2 (http://www.python.org/).
- The `scipy` 0.5.2 package (http://www.scipy.org/) was employed for numerical scripting work. Most data fitting was done with its `fmin` downhill simplex optimizer.
- The `matplotlib` 0.91 package (http://matplotlib.sourceforge.net/) was used for all data plots (in a slightly modified local version). The `xfig` 3.2 program was used for all other graphics.
- Fig. 1.1 was based on a figure in Wikipedia (http://en.wikipedia.org/wiki/File:MOO2.png)

Bibliography

[1] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, *Gaussian-2 theory for molecular energies of first and second-row compounds,* J. Chem. Phys. **94**, 7221 (1991).

[2] E. A. Hylleraas, *Neue Berechnung der Energie des Heliums im Grundzustande, sowie des tiefsten Terms von Ortho-Helium,* Z. Phys. **54**, 347 (1929).

[3] M. Stanke, J. Komasa, D. Kedziera, S. Bubin, and L. Adamowicz, *Three lowest S states of $^9Be^+$ calculated with including nuclear motion and relativistic and QED corrections,* Phys. Rev. A **77**, 062509 (2008).

[4] S. Bubin and L. Adamowicz, *Calculations of low-lying ^1P states of the beryllium atom,* Phys. Rev. A **79**, 022501 (2009).

[5] W. Cencek and K. Szalewicz, *Ultra-high accuracy calculations for hydrogen molecule and helium dimer,* Int. J. Quantum Chem. **108**, 2191 (2008).

[6] D. Bressanini and G. Morosi, *A compact boundary-condition-determined wavefunction for two-electron atomic systems,* J. Phys. B **41**, 14501 (2008).

[7] H. Nakatsuji, *General method of solving the Schrödinger equation of atoms and molecules,* Phys. Rev. A **72**, 062110 (2005).

[8] H. Stoll and A. Savin, *Density functionals for correlation energies of atoms and molecules;* In: Density Functional Methods in Physics; R. Dreizler, J. da Providencia (eds.), 177. New York: Plenum, (1985).

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

[10] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao, *A long-range correction scheme for generalized-gradient-approximation exchange functionals,* J. Chem. Phys. **115**, 3540 (2001).

[11] J. Toulouse, F. Colonna, and A. Savin, *Long-range/short-range separation of the electron-electron interaction in density-functional theory,* Phys. Rev. A **70**, 062505 (2004).

[12] J. G. Ángyán, I. C. Gerber, A. Savin, and J. Toulouse, *van der Waals forces in density functional theory: Perturbational long-range electron-interaction corrections,* Phys. Rev. A **72**, 012510 (2005).

[13] E. Goll, H.-J. Werner, and H. Stoll, *A short-range gradient-corrected density functional in long-range coupled-cluster calculations for rare gas dimers,* Phys. Chem. Chem. Phys. **7**, 3917 (2005).

[14] E. Goll, H.-J. Werner, H. Stoll, T. Leininger, P. Gori-Giorgi, and A. Savin, *A short-range gradient-corrected spin density functional in combination with long-range coupled-cluster methods: Application to alkali-metal rare-gas dimers,* Chem. Phys. **329**, 276 (2006).

[15] I. C. Gerber and J. G. Ángyán, *London dispersion forces by range-separated hybrid density functional with second order perturbational corrections: The case of rare gas complexes,* J. Chem. Phys. **126**, 044103 (2007).

[16] E. Goll, M. Ernst, F. Moegle-Hofacker, and H. Stoll, *Development and assessment of a short-range meta-GGA functional,* J. Chem. Phys. **130**, 234112 (2009).

[17] B. G. Janesko, T. M. Henderson, and G. E. Scuseria, *Long-range-corrected hybrids including random phase approximation correlation,* J. Chem. Phys. **130**, 081105 (2009).

[18] W. Kutzelnigg, *r_{12}-Dependent terms in the wave function as closed sums of partial wave amplitudes for large l,* Theor. Chim. Acta **68**, 445 (1985).

[19] W. Kutzelnigg and W. Klopper, *Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. I. General Theory,* J. Chem. Phys. **94**, 1985 (1991).

[20] V. Termath, W. Klopper, and W. Kutzelnigg, *Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. II. Second-order Møller- Plesset(MP2-R12) calculations on closed-shell atoms,* J. Chem. Phys. **94**, 2002 (1991).

[21] J. Noga, W. Kutzelnigg, and W. Klopper, *CC-R12, a correlation cusp corrected coupled-cluster method with a pilot application to the Be_2 potential curve,* Chem. Phys. Lett. **199**, 497 (1992).

[22] J. Noga and W. Kutzelnigg, *Coupled cluster theory that takes care of the correlation cusp by inclusion of linear terms in the interelectronic coordinates,* J. Chem. Phys. **101**, 7738 (1994).

[23] R. J. Gdanitz, *Accurately solving the electronic Schrödinger equation of atoms and molecules using explicitly correlated (r_{12}-)MR-CI. II. Ground-state energies of first-row atoms and positive atomic ions,* J. Chem. Phys. **109**, 9795 (1998).

[24] R. J. Gdanitz, *Accurately solving the electronic Schrodinger equation of atoms and molecules using explicitly correlated (r(12)-) multireference configuration interaction. III. Electron affinities of first-row atoms,* J. Chem. Phys. **110**, 706 (1999).

[25] J. R. Flores and R. J. Gdanitz, *Accurately solving the electronic Schrodinger equation of small atoms and molecules using explicitly correlated (r(12)-)MR-CI. VIII. Valence excited states of methylene (CH2),* J. Chem. Phys. **123**, 144316 (2005).

[26] W. Cardoen, R. J. Gdanitz, and J. Simons, *Transition-state energy and geometry, exothermicity, and van der Waals wells on the F+H_2 → FH+H ground-state surface calculated at the r(12)-ACPF-2 level,* J. Phys. Chem. A **110**, 564 (2006).

[27] J. Noga, S. Kedžuch, J. Šimunek, and S. Ten-no, *Explicitly correlated coupled cluster F12 theory with single and double excitations,* J. Chem. Phys. **128**, 174103 (2008).

[28] J. Noga, S. Kedžuch, J. Šimunek, and S. Ten-no, *Erratum: "Explicitly correlated coupled cluster F12 theory with single and double excitations" [J. Chem. Phys. 128, 174103 (2008)],* J. Chem. Phys. **130**, 029901 (2009).

[29] W. Klopper and C. C. M. Samson, *Explicitly correlated second-order Møller-Plesset methods with auxiliary basis sets,* J. Chem. Phys. **116**, 6397 (2002).

[30] E. F. Valeev, *Improving on the resolution of the identity in linear R12 ab initio theories,* Chem. Phys. Lett. **395**, 190 (2004).

[31] S. Ten-no and F. R. Manby, *Density fitting for the decomposition of three-electron integrals in explicitly correlated electronic structure theory,* J. Chem. Phys. **119**, 5358 (2003).

[32] F. R. Manby, *Density fitting in second-order linear-r_{12} Møller-Plesset perturbation theory,* J. Chem. Phys. **119**, 4607 (2003).

[33] S. Ten-no, *Initiation of explicitly correlated Slater-type germnal theory,* Chem. Phys. Lett. **398**, 56 (2004).

[34] A. J. May, E. Valeev, R. Polly, and F. R. Manby, *Analysis of the errors in explicitly correlated electronic structure theory,* Phys. Chem. Chem. Phys. **7**, 2710 (2005).

[35] S. Ten-no, *Explicitly correlated second order perturbation theory: Introduction of a rational generator and numerical quadratures,* J. Chem. Phys. **121**, 117 (2004).

[36] H.-J. Werner, T. B. Adler, and F. R. Manby, *General Orbital Invariant MP2-F12 Theory,* J. Chem. Phys. **126**, 164102 (2007).

[37] J. Noga, S. Kedžuch, and J. Šimunek, *Second order explicitly correlated R12 theory revisited: A second quantization framework for treatment of the operators' partitionings,* J. Chem. Phys. **127**, 034106 (2007).

[38] H. Fliegl, W. Klopper, and C. Hättig, *Coupled-cluster theory with simplified linear-r_{12} corrections: The CCSD(R12) model,* J. Chem. Phys. **122**, 084107 (2005).

[39] H. Fliegl, C. Hättig, and W. Klopper, *Inclusion of the (T) Triples Correction into the Linear-r12 Corrected Coupled-Cluster Model CCSD(R12),* Int. J. Quantum Chem. **106**, 2306 (2006).

[40] D. P. Tew, W. Klopper, C. Neiss, and C. Hättig, *Quintuple-zeta quality coupled-cluster correlation energies with triple-zeta basis sets,* Phys. Chem. Chem. Phys. **9**, 1921 (2007).

[41] T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, *Explicitly correlated coupled-cluster singles and doubles method based on complete diagrammatic equations,* J. Chem. Phys. **129**, 071101 (2008).

[42] T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, *Equations of explicitly-correlated coupled-cluster methods,* Phys. Chem. Chem. Phys. **10**, 3358 (2008).

[43] A. Köhn, G. W. Richings, and D. P. Tew, *Implementation of the full explicitly correlated coupled-cluster singles and doubles model CCSD-F12 with optimally reduced auxiliary basis dependence,* J. Chem. Phys. **129**, 201103 (2008).

[44] A. Köhn, *Explicitly correlated connected triple excitations in coupled-cluster theory,* J. Chem. Phys. **130**, 131101 (2009).

[45] C. Neiss, C. Hättig, and W. Klopper, *Extensions of r12 corrections to CC2-R12 for excited states,* J. Chem. Phys. **125**, 064111 (2006).

[46] A. Köhn, *A modified ansatz for explicitly correlated coupled-cluster wave functions that is suitable for response theory,* J. Chem. Phys. **130**, 104104 (2009).

[47] H.-J. Werner and F. R. Manby, *Explicitly correlated second-order perturbation theory using density fitting and local approximations,* J. Chem. Phys. **124**, 054114 (2006).

[48] F. R. Manby, H.-J. Werner, T. B. Adler, and A. J. May, *Explicitly correlated local second-order perturbation theory with a frozen geminal correlation factor,* J. Chem. Phys. **124**, 094103 (2006).

[49] H.-J. Werner, *Eliminating the domain error in local explicitly correlated second-order Møller-Plesset perturbation theory,* J. Chem. Phys. **129**, 101103 (2008).

[50] T. B. Adler, H.-J. Werner, and F. R. Manby, *Local explicitly correlated second-order perturbation theory for the accurate treatment of large molecules,* J. Chem. Phys. **130**, 054106 (2009).

[51] T. B. Adler and H.-J. Werner, *Local explicitly correlated coupled-cluster methods: Efficient removal of the basis set incompleteness and domain errors,* J. Chem. Phys. **130**, 241101 (2009).

[52] J. Noga, P. Valiron, and W. Klopper, *The accuracy of atomization energies from explicitly correlated coupled-cluster calcualtions,* J. Chem. Phys. **115**, 2022 (2001).

[53] J. Noga and P. Valiron, *Explicitly correlated R12 coupled-cluster calculations for open shell systems,* Chem. Phys. Lett. **324**, 166 (2000).

[54] S. Ten-no, *A simple F12 geminal correction in multi-reference perturbation theory,* Chem. Phys. Lett. **447**, 175 (2007).

[55] D. P. Tew, W. Klopper, M. Heckert, and J. Gauss, *Basis Set Limit CCSD(T) Harmonic Vibrational Frequencies,* J. Phys. Chem. A **111**, 11242 (2007).

[56] R. A. Bachorz, W. Klopper, and M. Gutowski, *Coupled-cluster and explicitly correlated perturbation-theory calculations of the uracil anion,* J. Chem. Phys. **126**, 085101 (2007).

[57] E. F. Valeev and T. D. Crawford, *Simple coupled-cluster singles and doubles method with perturbative inclusion of triples and explicitly correlated geminals: The CCSD(T)$_{\overline{R12}}$ model,* J. Chem. Phys. **128**, 244113 (2008).

[58] W. Klopper, R. A. Bachorz, D. P. Tew, J. Aguilera-Iparraguire, Y. Carissan, and C. Hättig, *Accurate Coupled Cluster Calculations of the Reaction Barrier Heights of Two CH_3·+CH_4 Reactions,* J. Phys. Chem. A **113**, 11679 (2009).

[59] P. J. Knowles, C. Hampel, and H.-J. Werner, *Coupled cluster theory for high spin, open shell reference wave functions,* J. Chem. Phys. **99**, 5219 (1993).

[60] P. J. Knowles, C. Hampel, and H.-J. Werner, *Erratum: Coupled cluster theory for high spin, open shell reference wave functions,* J. Chem. Phys. **112**, 3106 (2000).

[61] M. J. O. Deegan and P. J. Knowles, *Perturbative corrections to account for triple excitations in closed and open shell coupled cluster theories,* Chem. Phys. Lett. **227**, 321 (1994).

[62] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory*. Chichester: Wiley, 2000.

[63] S. M. Cybulski and R. R. Toczy lowski, *Ground state potential energy curves for He_2, Ne_2, Ar_2, He–Ne, He–Ar, and Ne–Ar: A coupled-cluster study,* J. Chem. Phys. **111**, 10520 (1999).

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

[65] K. L. Bak, P. Jørgensen, J. Olsen, T. Helgaker, and W. Klopper, *Accuracy of atomization energies and reaction enthalpies in standard and extrapolated electronic wave function/basis set calculations,* J. Chem. Phys. **112**, 9229 (2000).

[66] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Basis-set convergence in correlated calculations on Ne, N_2, and H_2O,* Chem. Phys. Lett. **286**, 243 (1998).

[67] T. Helgaker, T. A. Ruden, P. Jørgensen, J. Olsen, and W. Klopper, *A priori calculation of molecular properties to chemical accuracyy,* J. Phys. Org. Chem. **17**, 913 (2004).

[68] F. Jensen, *Estimating the Hartree–Fock limit from finite basis set calculations,* Theor. Chim. Acta **113**, 267 (2005).

[69] O. Vahtras, J. Almlöf, and M. W. Feyereisen, *Integral approximations for LCAO-SCF calculations,* Chem. Phys. Lett. **213**, 514 (1993).

[70] W. Kutzelnigg, *Quantum chemistry in Fock space. I. The universal wave and energy operators,* J. Chem. Phys. **77**, 3081 (1982).

[71] W. Kutzelnigg and S. Koch, *Quantum chemistry in Fock space. II. Effective Hamiltonians in Fock space,* J. Chem. Phys. **79**, 4315 (1983).

[72] W. Kutzelnigg, *Quantum chemistry in Fock space. III. Particle-hole formalism,* J. Chem. Phys. **80**, 822 (1984).

[73] W. Kutzelnigg and D. Mukherjee, *Normal order and extended Wick theorem for a multiconfiguration reference wave function,* J. Chem. Phys. **107**, 432 (1997).

[74] M. Nooijen and R. J. Bartlett, *General spin adaption of open-shell coupled cluster theory,* J. Chem. Phys. **104**, 2652 (1996).

[75] H.-J. Werner and P. J. Knowles, *A second order multiconfiguration SCF procedure with optimum convergence,* J. Chem. Phys. **82**, 5053 (1985).

[76] P. J. Knowles, J. S. Andrews, R. D. Amos, N. C. Handy, and J. Pople, *Restricted Møller-Plesset theory for open-shell molecules,* Chem. Phys. Lett. **186**, 130 (1991).

[77] C. W. Murray and N. C. Handy, *Comparison and assessment of different forms of open shell perturbation theory,* J. Chem. Phys. **97**, 6509 (1992).

[78] M. W. Wong and L. Radom, *Radical Addition to Alkenes: An assessment of Theoretical Procedures,* J Phys. Chem. **99**, 8582 (1995).

[79] C. J. Parkinson, P. M. Mayer, and L. Radom, *Cyanovinyl radical: an illustration of the poor performance of unrestricted perturbation theory and density functional theory procedures in calculating radical stabilization energies,* Theor. Chem. Acc. **102**, 92 (1999).

[80] Y.-Y. Chuang, E. L. Coitiño, , and D. G. Truhlar, *How Should We Calculate Transition State Geometries for Radical Reactions? The Effect of Spin Contamination on the Prediction of Geometries for Open-Shell Saddle Points,* J. Phys. Chem. A **104**, 446 (2000).

[81] N. C. Handy, P. J. Knowles, and K. Somasundram, *On the convergence of the Møller-Plesset perturbation series,* Theor. Chim. Acta **68**, 87 (1985).

[82] M. L. Leininger, W. D. Allen, H. F. Schaefer III, and C. D. Sherrill, *Is Møller–Plesset perturbation theory a convergent ab initio method?,* J. Chem. Phys. **112**, 9213 (2000).

[83] A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, *W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions,* J. Chem. Phys. **125**, 144108 (2006).

[84] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, *HEAT: High accuracy extrapolated ab initio thermochemistry,* J. Chem. Phys. **121**, 11599 (2004).

[85] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *A fifth-order perturbation comparison of electron correlation theories,* Chem. Phys. Lett. **157**, 479 (1989).

[86] P. Constans, P. Y. Ayala, and G. E. Scuseria, *Scaling reduction of the perturbative triples correction (T) to coupled cluster theory via Laplace transform formalism,* J. Chem. Phys. **113**, 10451 (2000).

[87] M. Nooijen, K. Shamasundar, and D. Mukherjee, *Reflections on size-extensivity, size-consistency and generalized extensivity in many-body theory,* Mol. Phys. **103**, 2277 (2005).

[88] C. Hampel, K. A. Peterson, and H.-J. Werner, *A comparison of the efficiency and accuracy of the quadratic configuration interaction (QCISD), coupled cluster (CCSD), and Brueckner coupled cluster (BCCD) methods,* Chem. Phys. Lett. **190**, 1 (1992).

[89] M. Heckert, O. Heun, J. Gauss, and P. G. Szalay, *Towards a spin-adapted coupled-cluster theory for high-spin open-shell states,* J. Chem. Phys. **124**, 124105 (2006).

[90] T. Kato, *On the eigenfunctions of many-particle systems in quantum mechanics,* Comm. Pure Appl. Math. **10**, 151 (1957).

[91] W. Kutzelnigg and J. D. Morgan, III, *Rates of convergence of the partial-wave expansions of atomic correlation energies,* J. Chem. Phys. **96**, 4484 (1992).

[92] A. J. May and F. R. Manby, *An explicitly correlated second order Møller-Plesset theory using a frozen Gaussian geminal,* J. Chem. Phys. **121**, 4479 (2004).

[93] D. P. Tew and W. Klopper, *New correlation factors for explicitly correlated electronic wave functions,* J. Chem. Phys. **123**, 074101 (2005).

[94] G. Knizia, *Evaluation of Observables in the Cumulant Expectation Value Formalism,* Diplomarbeit, Institut für Theoretische Physik, TU-Dresden, und Max-Planck Institut für Physik komplexer Systeme, Dresden, (2006).

[95] J. P. Perdew and Y. Wang, *Pair-distribution function and its coupling-constant average for the spin-polarized electron gas,* Phys. Rev. B **46**, 12947 (1992).

[96] P. Gori-Giorgi and J. P. Perdew, *Short-range correlation in the uniform electron gas: Extended Overhauser model,* Phys. Rev. B **64**, 155102 (2001).

[97] S. Höfener, D. P. Tew, W. Klopper, and T. Helgaker, *The geminal basis in explicitly correlated wave functions,* Chem. Phys. **356**, 25 (2009).

[98] R. T. Pack and W. Byers Brown, *Cusp Conditions for Molecular Wavefunctions,* J. Chem. Phys. **45**, 556 (1966).

[99] W. Klopper, *Orbital-invariant formulation of the MP2-R12 method,* Chem. Phys. Lett. **186**, 583 (1991).

[100] D. P. Tew and W. Klopper, *A comparison of linear and non-linear correlation factors for the basis set limit Møller-Plesset second order binding energies and structures of He_2, Be_2, and Ne_2,* J. Chem. Phys. **125**, 094302 (2006).

[101] W. Klopper, *A hybrid scheme for the resolution-of-the-identity approximation in second-order Møller-Plesset linear-r_{12} perturbation theory,* J. Chem. Phys. **120**, 10890 (2004).

[102] T. B. Adler, G. Knizia, and H.-J. Werner, *A simple and efficient CCSD(T)-F12 approximation,* J. Chem. Phys. **127**, 221106 (2007).

[103] G. Knizia, T. B. Adler, and H.-J. Werner, *Simplified CCSD(T)-F12 methods: Theory and benchmarks,* J. Chem. Phys. **130**, 054104 (2009).

[104] D. Bokhan, S. Ten-no, and J. Noga, *Implementation of the CCSD(T)-F12 method using cusp conditions,* Phys. Chem. Chem. Phys. **10**, 3320 (2008).

[105] H. Koch and P. Jørgensen, *Coupled cluster response functions,* J. Chem. Phys. **93**, 3333 (1990).

[106] H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, *Excitation energies from the coupled cluster singles and doubles linear response function (CCSDLR). Applications to Be, CH^+, CO, and H_2O,* J. Chem. Phys. **93**, 3345 (1990).

[107] H. Fliegl, C. Hättig, and W. Klopper, *Coupled-cluster response theory with linear-r12 corrections: The CC2-R12 model for excitation energies,* J. Chem. Phys. **124**, 044112 (2006).

[108] M. Hanauer and A. Köhn, *Response properties with explicitly correlated coupled-cluster methods using a Slater-type correlation factor and cusp conditions,* J. Chem. Phys. **131**, 124118 (2009).

[109] T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, *Higher-order explicitly correlated coupled-cluster methods,* J. Chem. Phys. **130**, 054101 (2009).

[110] T. Shiozaki, E. F. Valeev, and S. Hirata, *Explicitly correlated combined coupled-cluster and perturbation methods,* J. Chem. Phys. **131**, 044118 (2009).

[111] H. B. Schlegel, *Møller-Plesset Perturbation Theory with Spin Projection,* J. Phys. Chem. **92**, 3075 (1988).

[112] W. J. Lauderdale, J. F. Stanton, J. Gauss, J. D. Watts, and R. J. Bartlett, *Many-body perturbation theory with a restricted open-shell Hartree-Fock reference,* Chem. Phys. Lett. **187**, 21 (1991).

[113] W. J. Lauderdale, J. F. Stanton, J. Gauss, J. D. Watts, and R. J. Bartlett, *Restricted open-shell Hartree-Fock-based many-body perturbation theory: Theory and application of energy and gradient calculations,* J. Chem. Phys. **97**, 6606 (1992).

[114] S. Kedžuch, M. Milko, and J. Noga, *Alternative formulation of the matrix elements in MP2-R12 theory,* Int. J. Quantum Chem. **105**, 929 (2005).

[115] P. Pulay, *Convergence acceleration of iterative sequences. The case of SCF iteration.,* Chem. Phys. Lett. **73**, 393 (1980).

[116] P. Pulay, *Improved SCF convergence acceleration,* J. Chem. Phys. **3**, 556 (1982).

[117] R. J. Harrison, *Krylov subspace accelerated inexact Newton method for linear and nonlinear equations,* J. Comput. Chem. **25**, 328 (2003).

[118] P. Valiron, S. Kedžuch, and J. Noga, *Avoiding numerical instabilities in R12 calculations through universal R12 consistent basis sets,* Chem. Phys. Lett. **367**, 723 (2003).

[119] D. P. Tew, W. Klopper, and C. Hättig, *A diagonal orbital-invariant explicitly correlated coupled-cluster method,* Chem. Phys. Lett. **452**, 326 (2008).

[120] W. Kutzelnigg, *Note on the perturbation theory of electron correlation,* Chem. Phys. Lett. **35**, 283 (1975).

[121] E. F. Valeev, *Coupled-cluster methods with perturbative inclusion of explicitly correlated terms: a preliminary investigation,* Phys. Chem. Chem. Phys. **10**, 106 (2008).

[122] T. H. Dunning, Jr. and K. A. Peterson, *Approximating the basis set dependence of coupled cluster calculations: Evaluation of perturbation theory approximations for stable molecules,* J. Chem. Phys. **113**, 7799 (2000).

[123] D. Bakowies, *Extrapolation of electron correlation energies to finite and complete basis set targets,* J. Chem. Phys. **127**, 084105 (2007).

[124] W. Meyer, *Theory of self-consistent electron pairs. An iterative method for correlated many-electron wavefunctions,* J. Chem. Phys. **64**, 2901 (1976).

[125] M. Schütz, R. Lindh, and H.-J. Werner, *Integral-direct electron correlation methods,* Mol. Phys. **96**, 719 (1999).

[126] D. P. Tew, C. Hättig, R. A. Bachorz, and W. Klopper, *Explicitly correlated coupled-cluster theory;* In: Recent Progress in Coupled Cluster Methods: Theory and Applications; P. Cársky, J. Paldus, J. Pittner (eds.). Springer, (2010).

[127] G. Knizia and H.-J. Werner, *Explicitly correlated RMP2 for high-spin open-shell reference states,* J. Chem. Phys. **128**, 154103 (2008).

[128] K. Pflüger and H.-J. Werner, *On the Selection of Domains and Orbital Pairs in Local Correlation Treatments,* Annu. Rep. Comp. Chem. **2**, 53 (2006).

[129] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions,* J. Chem. Phys. **96**, 6796 (1992).

[130] T. H. Dunning, Jr., K. A. Peterson, and A. K. Wilson, *Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited,* J. Chem. Phys. **114**, 9244 (2001).

[131] F. Weigend, *A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency,* Phys. Chem. Chem. Phys. **4**, 4285 (2002).

[132] F. Weigend, A. Köhn, and C. Hättig, *Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations,* J. Chem. Phys. **116**, 3175 (2002).

[133] C. Hättig, *Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core-valence and quintuple- basis sets for H to Ar and QZVPP basis sets for Li to Kr,* Phys. Chem. Chem. Phys **7**, 59 (2005).

[134] K. A. Peterson, T. B. Adler, and H.-J. Werner, *Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B-Ne, and Al-Ar,* J. Chem. Phys. **128**, 084102 (2008).

[135] K. E. Yousaf and K. A. Peterson, *Optimized auxiliary basis sets for explicitly correlated methods,* J. Chem. Phys. **129**, 184108 (2008).

[136] O. Marchetti and H.-J. Werner, *Accurate calculations of intermolecular interaction energies using explicitly correlated wave functions,* Phys. Chem. Chem. Phys. **10**, 3400 (2008).

[137] O. Marchetti and H.-J. Werner, *Accurate Calculations of Intermolecular Interaction Energies Using Explicitly Correlated Coupled Cluster Wave Functions and a Dispersion-Weighted MP2 Method,* J. Phys. Chem. A **113**, 11580 (2009).

[138] F. Weigend and R. Ahlrichs, *Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy,* Phys. Chem. Chem. Phys. **7**, 3297 (2005).

[139] F. Weigend, F. Furche, and R. Ahlrichs, *Gaussian basis sets of quadruple zeta valence quality for atoms H–Kr,* J. Chem. Phys. **119**, 12753 (2003).

[140] F. A. Bischoff, S. Wolfsegger, D. P. Tew, and W. Klopper, *Assessment of basis sets for F12 explicitly-correlated molecular electronic-structure methods,* Mol. Phys. **107**, 963 (2009).

[141] K. E. Yousaf and K. A. Peterson, *Optimized complementary auxiliary basis sets for explicitly correlated methods: aug-cc-pVnZ orbital basis sets,* Chem. Phys. Lett. **476**, 303 (2009).

[142] F. Weigend, *Hartree-Fock exchange fitting basis sets for H to Rn,* J. Comput. Chem. **29**, 167 (2007).

[143] A. Hellweg, C. Haettig, S. Hoefener, and W. Klopper, *Optimized accurate auxiliary basis sets for RI-MP2 and RI-CC2 calculations for the atoms Rb to Rn,* Theor. Chem. Acc. **117**, 587 (2007).

[144] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, *RI-MP2: optimized auxialiary basis sets and demonstration of efficiency,* Chem. Phys. Lett. **294**, 143 (1998).

[145] K. P. Huber and G. Herzberg, *Molecular spectra and molecular structure. IV. Constants of diatomic molecules*. Van Nostrand Reinhold, 1979.

[146] G. Rauhut, *Configuration selection as a route towards efficient vibrational configuration interaction calculations,* J. Chem. Phys. **127**, 184109 (2007).

[147] K. A. Peterson and T. H. Dunning, Jr., *Accurate correlation consistent basis sets for molecular core–valence correlation effects: The second row atoms Al–Ar, and the first row atoms B–Ne revisited,* J. Chem. Phys. **117**, 10548 (2002).

[148] R. D. Cowan and D. C. Griffin, *Approximate relativistic corrections to atomic radial wave functions,* J. Opt. Soc. Am. **66** (1976).

[149] M. Kállay and J. Gauss, *Approximate treatment of higher excitations in coupled-cluster theory,* J. Chem. Phys. **123**, 214105 (2005).

[150] M. Kállay, *"MRCC, a string-based quantum chemical program suite written by M. Kállay. See also M. Kállay, P. R. Surján, J. Chem. Phys. 115 2945 (2001) as well as: www.mrcc.hu.",* (2001).

[151] M. Kállay and P. R. Surján, *Higher excitations in coupled-cluster theory,* J. Chem. Phys. **115**, 2945 (2001).

[152] C. Lupinetti and A. J. Thakkar, *Polarizabilities and hyperpolarizabilities for the atoms Al, Si, P, S, Cl and Ar: Coupled cluster calculations,* J. Chem. Phys. **122** (2005).

[153] U. Hohm, *Polarizabilities of atoms and molecules, new insights into an old subject,* Vacuum **58**, 117 (2000).

[154] S. Y. Ovchinnikov, P. S. Krstic, and J. H. Macek, *Oscillations in the integral elastic cross sections for scattering of protons at inert gases,* Phys. Rev. A **74**, 042706 (2006).

[155] E. Bichoutskaia and N. C. Pyper, *A theoretical study of the cohesion of noble gases on graphite,* J. Chem. Phys. **128**, 024709 (2008).

[156] E. Bichoutskaia and N. C. Pyper, *Theoretical study of the structures and electronic properties of all-surface KI and CsI nanocrystals encapsulated in single walled carbon nanotubes,* J. Chem. Phys. **129**, 154701 (2008).

[157] H.-J. Werner and W. Meyer, *Finite perturbation calculations for the static dipole polarizabilities of the first-row atoms,* Phys. Rev. A **13**, 13 (1976).

[158] R. P. Saxon, *Neon polarizability and C_6 coefficient from oscillator strength distribution,* J. Chem. Phys. **59**, 1539 (1973).

[159] U. Hohm and K. Kerl, *Interferometric measurements of the dipole polarizability α of molecules between 300K and 1100K,* Mol. Phys. **69**, 803 (1990).

[160] G. Lach, B. Jeziorski, and K. Szalewicz, *Radiative Corrections to the Polarizability of Helium,* Phys. Rev. Lett. **92**, 233001 (2004).

[161] J. Zheng, Y. Zhao, and D. G. Truhlar, *The DBH24/08 Database and Its Use to Assess Electronic Structure Model Chemistries for Chemical Reaction Barrier Heights,* J. Chem. Theory Comput. **5**, 808 (2009).

[162] S. Parthiban, G. de Oliveira, , and J. M. L. Martin, *Benchmark ab Initio Energy Profiles for the Gas-Phase SN2 Reactions $Y^- + CH_3X \rightarrow CH_3Y + X^-$ (X,Y = F,Cl,Br). Validation of Hybrid DFT Methods,* J. Phys. Chem. A **105**, 895 (2001).

[163] Y. Kanai, X. Wang, A. Selloni, and R. Car, *Testing the TPSS meta-generalized-gradient-approximation exchange-correlation functional in calculations of transition states and reaction barriers,* J. Chem. Phys. **125**, 234104 (2006).

[164] A. P. Bento, M. Solá, and F. M. Bickelhaupt, *E2 and S_N2 Reactions of $X^- + CH_3CH_2X$ (X = F, Cl); an ab Initio and DFT Benchmark Study,* J. Chem. Theory Comput. **4**, 929 (2008).

[165] A. D. Becke, *A new mixing of Hartree–Fock and local density-functional theories,* J. Chem. Phys. **98**, 1372 (1993).

[166] A. D. Boese and J. M. L. Martin, *Development of density functionals for thermochemical kinetics,* J. Chem. Phys. **121**, 3405 (2004).

[167] Y. Zhao, N. E. Schultz, and D. G. Truhlar, *Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions,* J. Chem. Theory Comput. **2**, 364 (2006).

[168] D. H. Ess and K. N. Houk, *Activation Energies of Pericyclic Reactions: Performance of DFT, MP2, and CBS-QB3 Methods for the Prediction of Activation Barriers and Reaction Energetics of 1,3-Dipolar Cycloadditions, and Revised Activation Enthalpies for a Standard Set of Hydrocarbon Pericyclic Reactions,* J. Phys. Chem. A **109**, 9542 (2005).

[169] S. Grimme, *Semiempirical hybrid density functional with perturbative second-order correlation,* J. Chem. Phys. **124**, 034108 (2006).

[170] Y. Zhao and D. G. Truhlar, *Exploring the Limit of Accuracy of the Global Hybrid Meta Density Functional for Main-Group Thermochemistry, Kinetics, and Noncovalent Interactions,* J. Chem. Theory Comput. **4**, 1849 (2008).

[171] A. Karton, A. Tarnopolsky, J.-F. Lamre, G. C. Schatz, and J. M. L. Martin, *Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical Kinetics,* J. Phys. Chem. A **112**, 12868 (2008).

[172] J. Zheng, Y. Zhao, and D. G. Truhlar, *Representative Benchmark Suites for Barrier Heights of Diverse Reaction Types and Assessment of Electronic Structure Methods for Thermochemical Kinetics,* J. Chem. Theory Comput. **3**, 569 (2007).

[173] G. Rauhut, G. Knizia, and H.-J. Werner, *Accurate calculation of vibrational frequencies using explicitly correlated coupled-cluster theory,* J. Chem. Phys. **130**, 054105 (2009).

[174] G. Rauhut and B. Hartke, *Modeling of high-order many-mode terms in the expansion of multidimensional potential energy surfaces: Application to vibrational spectra,* J. Chem. Phys. **131**, 014108 (2009).

[175] J. G. Hill, K. A. Peterson, G. Knizia, and H.-J. Werner, *Extrapolating MP2 and CCSD explicitly correlated correlation energies to the complete basis set limit with first and second row correlation consistent basis sets,* J. Chem. Phys. **131**, 194105 (2009).