*Molecular Physics*, Vol. 103, No. 2–3, 20 January–10 February 2005, 279–308

Taylor & Francis
Taylor & Francis Group

# Development of a CAS-DFT method covering non-dynamical and dynamical electron correlation in a balanced way

JÜRGEN GRÄFENSTEIN and DIETER CREMER*
Department of Theoretical Chemistry, Göteborg University,
Box 460, SE-405 30 Göteborg, Sweden

(*Received 12 August 2004; in final form 14 September 2004*)

CAS-DFT (Complete Active Space Density Functional Theory) is presented as a method that allows an economical, simultaneous treatment of non-dynamical and dynamical correlation effects for electronic systems with multi-reference character. Central problems of CAS-DFT concern the effective coupling between wave function and DFT method, the double counting of dynamical correlation effects, the choice of the proper input quantities for the DFT functional, the balanced treatment of core and active orbital correlation, of equal-spin and opposite-spin correlation effects, and the inclusion of spin polarization to handle closed- and open-shell systems in a balanced way. We present CAS-DFT2(CS,SPP,FOS,DS) (CAS-DFT using level 2 for the distinction of core and active orbital correlations, carried out with the Colle–Salvetti functional, using the Stoll–Pavlidou–Preuss functional for equal-spin correlation corrections, including spin polarization in the scaling procedure, and correcting with the Davidson–Staroverov density for low-spin cases). The method is free of any self-interaction error and size extensive provided the active space is properly chosen. For the three lowest states of methylene, stringent and less stringent tests are used to demonstrate the performance of the new CAS-DFT method for six different active spaces.

## 1. Introduction

Kohn–Sham (KS) density functional theory (DFT) [1–4] in its present approximate form leads to erroneous results when applied to electron systems with strong multi-reference character. There are of course several wave function theory (WFT) methods available that provide satisfactory descriptions of multi-reference problems, however these WFT methods are computationally no longer feasible for larger molecules [5–8]. Ideal for larger systems would be a KS-DFT method that handles both dynamical and non-dynamical correlation effects in a balanced way so that neither the degree of multi-reference character nor the size of the system investigated represents any longer a fundamental obstacle. This has led to a number of attempts of improving or extending KS-DFT in such a way that multi-reference problems can also be described with reasonable accuracy [9–45]. Developments and investigations made in this direction have necessarily been confronted with the question how to distinguish between dynamical and non-dynamical electron correlation effects, which is of general interest for both WFT and DFT [46–50].

* Corresponding author. Email: cremer@theoc.gu.se

Our own efforts have focused on the question how to combine features of a WFT method for treating non-dynamical correlation with those of DFT that treats preferentially dynamical correlation effects (see below). In previous work, we have shown that even in the case of multi-determinantal problems that have no multi-reference character, there is a need to extend KS-DFT by methods well known in WFT. This led to the development of restricted open shell DFT for singlet biradicals (ROSS-DFT) [51]. Early attempts in our group to combine generalized valence bond (GVB) theory [52] with the local spin density approximation (GVB-LSD) [36, 37] extended this work to multi-reference systems and led finally to the development of the more general complete active space (CAS)-DFT approach [41, 42]. In this work, we report on a further development of CAS-DFT where we will focus on the basic requirements for a functioning method and the realization of these requirements. We will develop a CAS-DFT method in our work that provides a balanced description of the different situations (closed-shell, open-shell with high- or low-spin electron coupling) that may be encountered in different systems. We will also show how CAS-DFT can be further developed to meet the requirements of an easy to handle and broadly applicable WFT-DFT method.

*Molecular Physics*
ISSN 0026–8976 print/ISSN 1362–3028 online © 2005
Taylor & Francis Group Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/00268970512331318858

The development of a new CAS-DFT method will be presented in the following way. In section 2, previous attempts of including non-dynamical electron correlation into DFT are shortly reviewed. In section 3, we will discuss the basic requirements of a CAS-DFT method, demonstrate how these can be handled when setting up a method useful for applications, and present CAS-DFT2. Finally, in section 4 we will compare CAS-DFT with other WFT-DFT methods and outline future possible developments of CAS-DFT.

## 2. Covering non-dynamical electron correlation at the KS-DFT and WFT-DFT levels of theory

The exact DFT exchange and correlation (XC) functional should cover all non-dynamical and dynamical correlation contributions [2], however the approximate XC functionals available today are considered to describe only dynamical correlation effects [3, 4]. There were early speculations that the local XC functionals in use also account for some non-specific non-dynamical correlation effects [24, 53–57]. For example, LDA or GGA exchange functionals lead to molecular densities that when compared with suitable reference densities reveal the presence of non-dynamical correlation effects [56, 58–60]. Speculations were confirmed in a quantitative way when investigating the exchange hole provided by local or semilocal exchange functionals [61–63]: *the self-interaction error of approximate exchange functionals mimics long-range correlation effects, which improve the stability of the KS wave function and the performance of DFT methods run with approximate X-functionals*.

Two important corollaries result from these findings: (a) KS-DFT with approximate functionals generally performs much better than Hartree–Fock (HF) in the case of electronic systems with some multi-reference character. It may even reach in some cases the quality of coupled cluster calculations [59]. However, since the coverage of non-dynamical correlation effects is non-specific, *a priori* there is no guarantee whether KS-DFT can really describe a given multi-reference problem in a satisfactory way. (b) Any improvement of DFT aimed at including multi-reference effects in an explicit way has to consider that use of local or semilocal exchange functionals will automatically lead to a double counting of electron correlation effects and in this way give flawed results. This fact is common to most attempts of improving DFT in such a way that it can satisfactorily describe non-dynamical correlation effects.

The simplest improvement of DFT for multi-reference problems is to use broken-symmetry unrestricted DFT (BS-UDFT) [60, 64–67]. This approach leads to reasonable results when applied, e.g. to biradicals with a small singlet–triplet splitting [64]. Analysis of BS-UDFT (carried out over the years by many groups, but here just a few representative citations are given) [60, 64–67] reveals that the BS-UDFT wave function, although a single-determinant wave function, has two-configurational character similar to that of the generalized valence bond approach for one electron pair (GVB(1)). However, the BS-UDFT wave function never takes the explicit form of a GVB(1) function, because it is the admixture of a triplet state that leads to a single-determinantal form. The non-dynamical electron correlation effects accounted for by BS-UDFT are reflected by the form of the $\alpha$- and $\beta$-spin orbitals calculated at the UDFT level of theory [60, 64, 65]. They can be quantitatively determined via the natural orbital occupation numbers of a BS-UDFT calculation.

Due to the two-configurational character of BS-UDFT, the double counting of non-dynamical correlation effects, although mostly ignored, becomes a serious problem in this case: non-dynamical correlation effects explicitly introduced by the wave function are suppressed by non-specific non-dynamical correlation effects invoked by the LDA or GGA exchange functional used in the calculation. Accordingly, it is always better to carry out UDFT calculations with hybrid functionals thus reducing any double counting as much as possible (B3LYP is better than BLYP in the case of a BS-UDFT description) [63].

Gräfenstein and Cremer [42, 60] have introduced a classification of electronic systems in those that are dominantly single-configurational single-determinantal (type 0), single-configurational multi-determinantal (type I), multi-configurational with each configuration represented by a single determinant (type II), and multi-configurational where one or more configurations have to be represented by a multi-determinantal description (type III). KS-DFT already fails to describe type I systems correctly, although in this case no non-dynamical correlation effects are involved. These systems have been described by permuted orbitals (PO) UDFT [60, 64], ROSS-DFT [51] or similar approaches [68, 69] where in these cases knowledge from WFT [60, 70, 71] was used to derive the appropriate formalism. ROKS (restricted open-shell KS) DFT by Filatov and Shaik [69] can be applied for type I and two-configurational type II systems, for which BS-UDFT or TCSCF-DFT [72] are also useful.

The primary objective of the present research is however to describe type I, type II and type III systems in a general way using standard KS-DFT as a starting point. A larger number of attempts and approaches have been made in this direction [9–45] where only a few should be mentioned here.

(1) *Ensemble theory*. The question of how to include multi-reference effects into DFT touches the fundamental problem whether any ground state density (including single- and multi-reference systems) associated with an antisymmetric ground state wave function of a Hamiltonian with the external potential $V(r)$ is *v*-representable. Basic work by Levy [13], Lieb [14], Chayes *et al*. [15], Englisch and Englisch [16] and Kohn and co-workers [17, 19] has led to an extension of the original KS theory, which was formulated for pure-state *v*-representability (i.e. densities representable by a single-determinant ground state wave function of a non-interacting reference system), to ensemble KS theory, which is formulated for *ensemble v-representability* (i.e. densities representable by a weighted sum of pure state wave functions of non-interacting reference systems). It has been formally proven by Englisch and Englisch [16] that any strictly positive density is an ensemble *v*-representable density where the pure-state *v*-representability is a special case. Baerends and co-workers have given numerical examples ($C_2$ and $H_2 + H_2$ description) where, for typical multi-reference cases, the original conjecture of pure-state *v*-representability breaks down [18]. (For related work, see also Savin *et al*. [20] and Morrison [21].)

From the background of ensemble theory, repeated attempts of describing multi-reference systems within KS theory using fractional occupation numbers have been made [9–12]. These approaches have been systemized in the Restricted Ensemble Kohn–Sham (REKS) theory of Filatov and Shaik [22, 23]. In the REKS scheme, each of the single-determinant configurations that would occur, e.g. in a CAS-SCF(2,2) calculation is treated at the DFT level including DFT exchange and correlation. The four configurations are superimposed to account for the non-dynamical correlation and simultaneous optimization of the weighting factors and the KS orbitals is carried out. The REKS approach has the important advantage of being compatible in a DFT context, i.e. REKS energies are directly comparable to standard KS-DFT energies for states without non-dynamical correlation effects. However, dynamical and non-dynamical correlation effects are not cleanly separated in the REKS method.

The REKS method is applicable to (2,2) and (3,3) [22, 23] problems where the derivation of these methods has been guided by the CI handling of the corresponding cases. Further extensions of REKS (e.g. to (4,4) problems) is difficult because the CI expansion contains more variational parameters which can be provided by formal ensemble DFT. Simplifications in the sense of GVB-like schemes [52] have to be applied, the usefulness of which are difficult to foresee, apart from the fact that in any case the simplicity and cost effectiveness of standard KS-DFT is lost. As in the case of BS-UDFT or any other related approach, the problem of double counting of the correlation effects, which is inherent in the method, can however be partly overcome by using hybrid rather than pure XC functionals [63].

Recently, Khait and Hoffmann [73] have suggested a procedure to decompose a spin-adapted density into densities of configuration-state functions. In this way, one can write down densities of, for example, atomic multiplet states in terms of densities of single determinants. It remains to be seen how these densities transform to energies and how this can be formulated within a generally applicable ensemble method. In any case, such an approach will require the development of a new XC functional in order to avoid the double-counting problem.

(2) *Implicit or explicit splitting of the electron–electron interaction operator*. Multi-reference DFT methods that either explicitly or implicitly invoke a partitioning of the electron–electron interaction operator are no longer hampered by the double-counting problem. Savin and co-workers [25–29] decompose the Coulomb interaction potential into a long-ranged smooth and a short-ranged singular part. The effect of the long-ranged interactions is described by a multi-reference wave function, whereas the correlation due to the short-ranged portion is treated with a DFT-type functional. This implies that one has to develop new DFT functionals, which in addition will depend parametrically on the cut-off parameters used to decompose the Coulomb interaction.

Despite some first successes of the approach [29], the method depends on the validity of the assumption that the short-range part of the perturbation will always result in short-range correlation effects only. Especially in cases where non-dynamical correlation effects are important, e.g. in the case of stretched diatomic molecules, long-range correlation effects allow the molecule to decrease its electron–electron interaction energy at a low cost in terms of the

one-particle energy. In this case, the short-range part of the electron–electron interaction can play a role for the formation of long-range correlation effects as well, which cannot be covered properly in this computational scheme.

Panas and co-worker [31] have suggested an implicit partitioning of the electron–electron interaction operator. They use a CAS-SCF reference function to cover the non-dynamical electron correlation effects. For the description of the dynamical correlation effects, they start from the idea that dynamical electron correlation prevents the electrons from being close to each other. The total electron–electron interaction energy for the exact pair distribution, which shows a correlation hole for each position of the reference electron, should therefore be close to that of a model pair distribution without this correlation hole, provided that the electron–electron repulsion potential is truncated for small distances. In the approach by Panas, the two-electron integrals are therefore not calculated with the full Coulomb potential $1/r_{12}$ but with an expression that is truncated for short $r_{12}$ values. This should provide an inclusion of the dynamical electron correlation into the CAS-SCF calculation at no additional computational cost. It is, however, difficult to provide a well-founded rationale how the truncation of the $1/r_{12}$ repulsion should be done. In particular, the extent of the dynamical electron correlation will depend on the local electron density, which is not known in advance. Thus it is questionable whether an *a priori* modification of the two-electron integrals can describe dynamical electron correlation in a general fashion.

(3) *Merging of WFT with DFT*. BS-UDFT can be considered as the simplest example for this approach. As mentioned above, even at this stage the double counting of electron correlation effects becomes a problem and in so far the neat separation of non-dynamical and dynamical correlation effects, i.e. the avoidance of a double counting of correlation effects is a central problem of this approach. Various authors have tried to handle this problem in different ways.

(a) Colle and Salvetti [32] start from a MCSCF reference wave function. In the two-particle density matrix resulting from the wave function, they check locally point by point to what extent the electron–electron repulsion potential is reflected in the electronic structure, i.e. to what extent short-range electron correlation is covered by the MCSCF wave function. The dynamical electron correlation is then accounted for by a generalized version of the original Colle–Salvetti functional [33] where the information from the first step is used to reduce the dynamical correlation energy accordingly. The pointwise assessment of the wave function is expensive, which counteracts the goal of setting up a cheap and efficient computational scheme.

(b1) The simplest way to avoid a double counting of energy contributions is to keep dynamical electron correlation out of the multi-reference wave function, i.e. to choose as small an active space as possible. This approach was pursued first by Lie and Clementi [34], who investigated the dissociation of second-row hydrides and diatomics. In their work, just as many additional configurations were superimposed to the HF state so that the dissociation is described qualitatively correctly. The dynamical correlation is then taken into account by applying a version of the Gombas correlation potential [74] which was reparametrized in [34] to improve the balance in the description of the closed-shell molecules and the open-shell fragments. Lie and Clementi found that the approach required further refinement to become practicable.

(b2) Kraka [36] and Kraka *et al*. [37] followed the same basic idea by combining a GVB wave function with a correlation-energy term calculated with a LDA correlation functional. This approach allowed one to reproduce dissociation energies of a set of 28 small molecules with an accuracy of $4\,\mathrm{kcal\,mol^{-1}}$, as compared to $24\,\mathrm{kcal\,mol^{-1}}$ for the corresponding GVB calculations. In this approach, the proper choice of configurations for the molecules and fragments proved crucial and subtle. The work by Kraka and co-workers [36, 37] demonstrates that a straightforward combination of small-active-space WFT calculations and a DFT correlation functional does allow a simultaneous description of non-dynamical and dynamical correlation effects. However, one cannot guarantee that the choice of a minimal active space in all cases will exclude dynamical electron correlation from the reference function. It should be

mentioned that the idea of this approach has recently been revived by Stoll [75].

(c) All methods described so far start from a theoretical description of the correlation hole. Gusarov *et al*. [45] pursue a complementary line, trying to construct an XC potential and, subsequently a multi-reference DFT correlation-energy functional by studying a set of test molecules and generalizing the features found there. They start by doing high-level *ab initio* calculations for the test molecules, providing reliable reference values for their one-electron densities as well as total energies. Then they attempt to reproduce the densities found by a CAS-SCF calculation with an adjustable extra one-electron potential. This potential, which is a generalized KS correlation potential, will depend not only on the density to be reproduced but also on the size of the active space for the reference function. The energy difference between the high-level *ab initio* and the CAS-SCF result is a generalized KS correlation energy. The correlation energies and potentials obtained are collected in a data base with the goal of establishing general relationships between the electron density of the molecule and the active size on the one hand and the KS correlation potential on the other hand. So far, no strategy has been provided for the evaluation of the compiled data and the systematic construction of a KS potential for a given density and active space.

## 3. Basic requirements for a useful CAS-DFT method

The goal of this work is to develop a calculational scheme that provides a balanced description of non-dynamical and dynamical electron correlation effects by combining a multi-reference wave function approach with a DFT correlation functional for a description of the dynamical correlation. For this purpose, a list of requirements has to be fulfilled.

(1) Any useful WFT + DFT method has to refrain from using spin densities $\varrho_\alpha$ and $\varrho_\beta$ as input quantities because they lead to errors in the description of state multiplets [51, 60]. Other input quantities have to be found.

(2) As one includes specific multi-reference effects explicitly into the calculation, approximate DFT exchange cannot be used any longer because it implicitly mimics non-specific non-dynamical correlation effects [61–63]. Exchange has to be calculated in an exact way by the WFT method to avoid any double counting of non-dynamical correlation effects caused by the exchange functional.

(3) The WFT approach treats, in a specific way, a part of the total correlation energy. This part is larger the larger the active space is. In practice the active space is chosen in such a way that all non-dynamical correlation effects are covered. The DFT correlation functional, however, is insensitive to the size of the active space. It is designed to describe all electron correlation effects in the system, however, in an unspecified way: all correlation effects, including those related to low-lying virtual orbitals, are treated as dynamical ones. Consequently, some correlation-energy contributions in the molecule are treated twice: once in an exact way by the multi-reference wave function and once more in a less specific way by the DFT correlation functional. This becomes actually more problematic the larger the active space is. This double counting of correlation effects has to be avoided [41, 42, 60].

(4) Problem (3) can be reformulated to lead to the more general question: how does the DFT part of any WFT + DFT method 'know', which non-dynamical or dynamical correlation effects are already included by the WFT method so that the DFT method is set up to compensate just for missing correlation effects?

(5) Problems (3) and (4) lead to the necessity of distinguishing between active and inactive (core) orbitals to keep track of electron correlation effects accounted for by the WFT method.

(6) There is also a necessity of considering which orbitals are doubly and which are singly occupied in the active space. Closed shell and open shell systems encounter different situations in terms of opposite and equal spin correlation effects, which has to be considered when merging any WFT method with DFT.

(7) Even if just two open-shell systems are compared, opposite and equal spin correlation effects will be different for a high-spin and a low-spin system. This also has to be considered.

(8) An appropriate choice of the DFT correlation functional has to be made.

(9) The WFT + DFT method to be developed has to be size-consistent (extensive) within the limitations given by the WFT approach (where of course it is useful just to consider size-consistent WFT methods).

Some of these questions have already been discussed in our previous work [41, 42, 60] and therefore, they need only to be considered in a short, summarizing fashion in the following. Also there are some general questions, which have to be considered before starting with the actual development work. In this connection it is also useful to demonstrate the various points discussed for a suitable molecular example, for which the various correlation effects can still be assessed in a simple way (i.e. a relatively small molecular system), but one which is not trivial. For this purpose we have chosen the three lowest states of methylene $CH_2$, the $^3B_1$ ground state, the $^1A_1$ first excited state and the $^1B_1$ second excited state. These states represent typical examples for a high-spin open shell case with no correlation (type 0 system), a closed-shell case with non-dynamical electron correlation (type II system) and a low-spin open shell biradical again with no (or just small) non-dynamical correlation (type I or type III system). The three states present different cases of opposite/equal spin correlation situations and therefore the calculated state energy differences will be sensitive indicators of the usefulness of any WFT-DFT method.

In so far as these calculations are not trivial, they will reveal in a system with no non-dynamical correlation whether CAS-DFT can reduce unwanted correlation effects to zero. This would be difficult to check in a larger system because there doubly-counted correlation effects could still be used to fill up lacking correlation so that any weakness of the method in this respect would be disguised. For the three states of methylene however accurate excitation energies are known and the three different electronic situations of the system provide a basis for a rigorous testing of CAS-DFT.

All methylene calculations discussed in the following were carried out with a 6-311++G(2df,2pd) basis set [76]. For the $^1B_1$ state, the low-spin ROMP2 (OPT2) geometry from [71] was used, the geometry of the other two states was calculated at the R(O)MP2 level with the TZ2P basis set as described in [71] (see table 1). For the calculations, the *ab initio* packages COLOGNE 2004 [77] and Gaussian 98 [78] were used. In figure 1, the energy gain for an increase of the active space relative to the (0,0) space is shown for the $^1A_1$ state of methylene. The corresponding energy differences are given in table 1. In view of the multi-reference character of this state, they will be a measure for the performance of the various methods.

### 3.1. *Choice of the appropriate multi-reference wave function*

Previous experience in our group with the GVB-DFT approach [36, 37] clearly showed that the applicability of GVB as a suitable WFT method for describing multi-reference systems is too limited and a more general MCSCF approach seems to be more desirable. A straightforward procedure to cover non-dynamical electron correlation effects is to include all configuration state functions with similar energy into a self-consistent field calculation optimizing both the orbitals and the coefficients of the configuration state functions as done in the *full optimized reaction space* (FORS) method of Ruedenberg *et al*. [79] or the equivalent (*complete active space*) CAS-SCF approach [80]. The special case of a CAS-SCF(2,2) calculation, i.e. 2 electrons in two active orbitals as needed for describing the dissociation of a single bond, includes the same non-dynamical near-degeneracy correlation effects as provided by GVB(1) [52] or in a related way by BS-UHF.

What should the energy gain for a CAS-DFT method with successively increasing active space look like? For a (0,0) space, i.e. HF + DFT correlation, non-dynamical correlation contributions will be missing in the energy. The DFT correlation term will describe these correlation effects in an unspecified way and, accordingly, underestimate their contribution to the correlation energy. If the orbitals essential for the non-dynamical correlation effects are added to the active space, the energy should decrease. For example, the $^1A_1$ state of methylene represents in a first approximation a (2,2)-active space problem where the active orbitals are the occupied $a_1$-symmetrical $\sigma$ lone pair orbital and the empty $b_1$-symmetrical p$\pi$ orbital. If the active space is increased by adding orbitals that are responsible for dynamical correlation effects, the energy gain in the CAS-SCF wave function should be balanced by a reduction of the DFT correlation energy, i.e. the total energy should remain basically unchanged. For the $^1A_1$ state of methylene, there should be a substantial gain in energy from the (0,0) to the (2,2) and (2,4) active space where in the latter case the $\sigma^*$(CH) orbitals are added, which could be important for obtaining a balanced description of the $^1A_1$, $^3B_1$ and $^1B_1$ states.

Only minor energy gains can be expected when increasing the active space from (2,2) to (6,4) or from (2,4) to (6,6) and (8,7). In the first case the (2,2) space is enlarged by the occupied $\sigma$(CH) orbitals so to include all valence electrons where the CH bond electrons should only undergo dynamical correlation. The same holds for the second case (again addition of the $\sigma$(CH) orbitals), which represents the full valence (6,6) active space. The addition of the 1s core orbital leads to the (8,7) space, which is interesting in so far as the core electrons will definitely encounter just dynamical rather than any non-dynamical electron correlation. Hence, a sensitive test would be provided by the CAS-DFT(6,6) and CAS-DFT(8,7) energies, which

Table 1. Absolute HF(0,0), MP2(0,0), and HF + CS(0,0) references energies of the $^1A_1$ state of $CH_2$. For increasing active space, the lowering of the energy relative to the (0,0) space is given. Values in parentheses denote SPP corrected energies.[a]

| Active space | CASSCF | CAS-MP2 | CAS + LYP | CAS + CS | CAS + $f_{CS}$ CS (CAS + $f_{CS}$ CS(SPP)) | CAS-DFT1 (CS, FCS) (CAS-DFT1(CS,SPP,FCS)) | CAS-DFT2 (CS, FCS) (CAS-DFT2(CS,SPP,FCS)) |
|---|---|---|---|---|---|---|---|
| (0,0) | −38.89206 | −39.05004 | −39.12460 | −39.12309 | −39.12309 (−39.12309) | −39.12309 (−39.12309) | −39.12309 (−39.12309) |
| (2,2) | −14.22 | −3.56 | −14.45 | −10.02 | 4.10 (0.57) | −2.10 (−4.56) | −1.59 (−4.07) |
| (6,4) | −14.63 | −3.68 | −14.88 | −10.38 | 3.83 (0.36) | 2.49 (0.97) | 3.58 (0.25) |
| (2,4) | −18.07 | +1.83 | −18.30 | −12.81 | 23.62 (15.97) | 2.23 (−1.69) | 4.50 (0.33) |
| (6,6) | −38.41 | −5.18 | −38.75 | −28.89 | 9.26 (1.36) | 4.32 (−3.17) | 7.44 (0.37) |
| (8,7) | −38.56 | −4.28 | −38.90 | −29.04 | 9.18 (1.27) | 9.18 (1.27) | 9.18 (1.27) |

[a]Absolute energies in hartree, relative energies in kcal mol. All calculations with the 6-311 ++ G(2df,2pd) basis set. The CS correlation energy leads to an energy lowering of $E(\mathrm{RHF+CS}) - E(\mathrm{RHF}) = -144.98$ kcal mol. Active spaces: (2,2): minimal; (6,4): minimal + all occupied valence orbitals; (2,4): minimal + all virtual valence orbitals; (6,6): full valence; (8,7): full valence + 1s(C). CS: Colle–Salvetti functional; FCS: scaling factor $f$ calculated for a closed shell situatio; SPP: Stoll–Pavlidou–Preuss functional for elimination of equal-spin correlation; levels 1 and 2: different levels of distinguishing between core and active electron correlation.

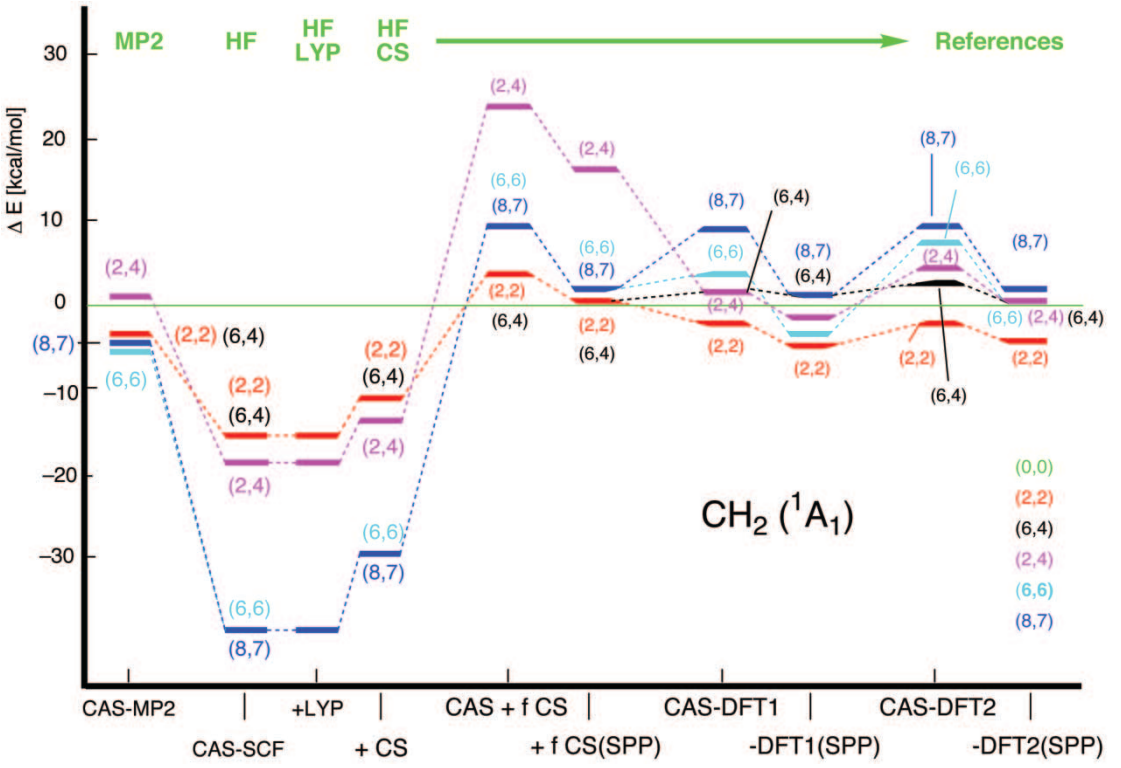


Figure 1. Changes in the correlation energy relative to a reference energy (HF, MP2, HFLYP, HFCS) for an increasing active space as calculated by different methods with a 6-311 + + G(2df,2pd) basis set. CS: Colle–Salvetti functional; FCS: scaling factor $f$ calculated for a closed shell situation; SPP: Stoll–Pavlidou–Preuss functional for elimination of equal-spin correlation; levels 1 and 2: different levels of distinguishing between core and active electron correlation.

should be strictly identical. Although in a practical application one would be interested in just the minimal (2,2) or the extended (2,4) active space, consideration of the three other active spaces provides energy data reflecting the sensitivity of the method to double counting of correlation effects.

Summarizing the following relationships should hold for the absolute CAS-DFT energies of the $^1A_1$ state of methylene provided a double counting of correlation effects can be excluded:

$$E^{\mathrm{CAS-DFT}}(2,4) < E^{\mathrm{CAS-DFT}}(2,2) \ll E^{\mathrm{CAS-DFT}}(0,0), \quad (1a)$$

$$E^{\mathrm{CAS-DFT}}(2,2) \approx E^{\mathrm{CAS-DFT}}(6,4), \quad (1b)$$

$$E^{\mathrm{CAS-DFT}}(2,4) \approx E^{\mathrm{CAS-DFT}}(6,6) = E^{\mathrm{CAS-DFT}}(8,7). \quad (1c)$$

Figure 1 shows the CAS-SCF/6-311 + + G(2df,2pd) energy of the $^1A_1$ state of $CH_2$ for the five different active spaces, relative to the energy for the (0,0) active space (i.e. HF energy). The CAS-SCF energy decreases with the size of the active space. The minimal (2,2) space gives a gain of 14.2 kcal mol$^{-1}$ relative to the HF energy, caused mainly by the non-dynamical correlation effects of the electron lone pair where the two electrons avoid each other by moving in opposite directions out of the molecular plane. Major energy gains are obtained when adding the $\sigma^*$(CH) orbitals to the minimal active space and increasing to the (2,4) space (−3.8 kcal mol$^{-1}$, table 1) or applying the full valence (6,6) space (−20.4 kcal mol$^{-1}$ relative to the minimal active space). The latter energy change is caused by adding dynamical (primarily left–right) electron correlation (involving excitations from the $\sigma$(CH) to the $\sigma^*$(CH) orbitals) whereas in the former case the energy lowering may be due to either non-dynamical or dynamical correlation effects. As becomes obvious from figure 1, neither the extension from the (2,2) to the (6,4) or from the (6,6) to the (8,7) space lead to any significant effect (see also table 1). Excitations of the $\sigma$(CH) electrons into the $\pi^*$(C) orbital do not play any role for non-dynamical or dynamical electron correlation. Also, incorporating the 1s(C) orbital into the active space leads to an energy gain of just 0.2 kcal mol$^{-1}$, which is not astonishing because the unoccupied valence orbitals are not appropriate for describing dynamical correlation between the two 1s(C) electrons.

For CASPT2 or CAS-MP2, the energy relationships of equations (1) should hold. This has to do with the fact that methylene is an example for which pair effects

dominate dynamical electron correlation whereas three- or higher-order electron correlation effects play only a minor role [81]. Accordingly, CAS-MP2 (CASPT2) should provide rather accurate energies. CAS-MP2/6-311++G(2df,2pd) calculations with a finite basis at fixed geometries reveal how the actual energies behave (table 1, figure 1). The CAS-MP2 energies for the (2,2), (6,4) and (6,6) spaces fulfill equations (1) confirming that with the minimal active space not all non-dynamical electron correlation is introduced. However, this is also not included when using the (2,4) space, which seems to be unbalanced pushing the correlation energy above that of the MP2(0,0) calculation. This is not surprising in so far as MP2 is known to exaggerate pair correlation effects [81] and provides a less appropriate reference than HF does. In any case, the inclusion of the $\sigma^*$(CH) orbitals into the (2,2) space to yield the (2,4) space reduces the effectiveness of non-dynamical electron correlation without introducing new effects. In this respect, the (6,6) space with the possibility of left–right correlation for the bond electrons seems to add both dynamical and non-dynamical correlation.

These observations show that it is misleading to make presumptuous classifications along the lines that left–right correlation is always non-dynamical and angular (or in–out) correlation effects are always dynamical. The classification of correlation effects will have to change from electronic system to electronic system and their nature should be assessed by quantitative means rather than with a rigid classification scheme.

At this point it is useful to understand how in CASPT2 [5] or CAS-MP2 [6] lacking correlation effects are included by perturbation theory. Besides dynamical correlation effects, CASPT2/CAS-MP2 does include non-dynamical correlation effects in its correction term and can thus balance an inconsistent choice of the active space to some extent: the CASPT2 or CAS-MP2 correlation energy can be found by expanding the expression

$$\Delta E = \left\langle \Psi_{\mathrm{CAS-SCF}} \middle| \hat{H}\hat{Q}\frac{1}{E_{\mathrm{CAS-SCF}} + \Delta E - \hat{Q}\hat{H}\hat{Q}}\hat{Q}\hat{H} \middle| \Psi_{\mathrm{CAS-SCF}} \right\rangle, \tag{2}$$

which gives in second order

$$\Delta E_{\mathrm{CASPT2}} = \left\langle \Psi_{\mathrm{CAS-SCF}} \middle| \hat{V}\hat{Q}\frac{1}{E_{\mathrm{CAS-SCF}} - \hat{Q}\hat{H}_0\hat{Q}}\hat{Q}\hat{V} \middle| \Psi_{\mathrm{CAS-SCF}} \right\rangle. \tag{3}$$

Equation (3) shows the sensitivity of CASPT2 or CAS-MP2 to external configurations (outside the active space) rather than just orbitals or densities: the operator $\hat{V}\hat{Q}$ first applies the perturbation (i.e. the electron–electron interaction) onto the CAS-SCF reference function and then projects out that part which is outside the active space. The weight of the part is determined by the energy of the configurations in the subspace orthogonal to the active space: if there are low-lying ones, the expectation value of $\hat{Q}\hat{H}_0\hat{Q}$ will become close to $E_{\mathrm{CAS-SCF}}$, and the corresponding configuration will make a large contribution to the CAS-MP2 (CASPT2) energy. That is, the structure of the energy expression (3) scans the orthogonal space for configurations that (i) are coupled to $\Psi_{\mathrm{CAS-SCF}}$ via the interaction operator and (ii) have an energy close to the CAS-SCF energy. In this way, equation (3) generates large energy contributions for such configurations. The occurrence of this kind of configuration actually indicates that the active space should be extended by active orbitals leading to these configurations.

Clearly, the perturbation method that includes dynamical electron correlation 'knows' what correlation effects are already accounted for by the multi-reference method so that the correlation energy can be further optimized. The result of the CASPT2 (CAS-MP2) calculation is in this way optimal. For a CAS-DFT approach the coupling between the two methods proceeds through the CAS electron density, which is the input for the DFT correlation functional. If the latter is 'blind' to any changes in the CAS density, then DFT will not 'know' what CAS-SCF is doing. But even in the case that the correlation functional responds to the features of the CAS density, an optimization of the coupling effects can only be achieved by a self-consistent procedure. Experiences made by Kraka, who developed a SCF GVB-DFT method [36, 37] indicate that self-consistency corrections (needed for analytical energy derivatives) are small and that the sensitivity of the correlation functional to the properties of the active space of the WFT method is low. Techniques (borrowed from WFT) to improve this sensitivity are in conflict with the goal of developing a simple, largely applicable CAS-DFT method.

Several authors have advanced [22, 23, 34, 36, 37] what is considered (to our opinion erroneously) as a basic rule: methods such as GVB-DFT or REKS are successful because they operate with the smallest active space possible (*minimal active space*) to include just non-dynamical rather than any dynamical electron correlation. Therefore, CAS-DFT calculations should be carried out with the minimal active space to keep the double-counting problem as small as possible. The sensitivity of the correlation functional is less serious for a minimal active space calculation. However, as we will show, in this work *even for minimal active space*

*calculations, double counting of correlation effects is always a problem*. Only, that this double-counting of correlation effects is often disguised (correlation is largely underestimated when using minimal active spaces) and may even lead to better results (*right results for the wrong reason*).

There are many situations (see above) where a minimal active space description is not sufficient and therefore any WFT + DFT theory based on a minimal active space rule cannot offer (apart from cost considerations) any advantages. We will use extended (rather than minimal) active spaces to keep the flexibility of the CAS-SCF description and, because of this, we will focus on appropriate means to suppress the double counting of correlation effects. Clearly such a means must become most sensitive when large active spaces are used. For the present example of the correlation energies of the $^1A_1$ state of methylene, we can reformulate the rules of equation (1) (applicable to the ideal case of a very large basis set) in the following way.

(a) CAS-DFT absolute energies should be separated from each other by small energies (for methylene just 5 kcal mol$^{-1}$ or less). The magnitude of the energy range is a direct reflection of the sensitivity of the correlation functional to include missing correlation effects.

(b) By avoiding any stronger deviations for the (2,4) or (8,7) space (spaces, for which CAS-MP2 turns out to be unstable), CAS-DFT should prove to be stable against variations in the active space.

(c) CAS-DFT should be located below the energy of the HF + CS reference, reflecting the non-dynamical correlation energy missing in the HF + CS description. Any reordering of CAS-DFT energies relative to that given in equation (1) (figure 1) indicates deficiencies of the method chosen.

These rules are more stringent than those actually needed for a consistent description of the various states of a molecule or the different energies of a reaction complex along the reaction path. Nevertheless, they provide in the following a guidance on how to improve any CAS-DFT method.

### 3.2. *Choice of the input quantities*

As was pointed out previously [41, 42], spin densities $\varrho_\alpha$ and $\varrho_\beta$ are no longer useful input quantities for CAS-DFT. One has to replace these densities, similarly as discussed by Perdew *et al*. [82] in the case of UDFT (see also [40, 60]), by the total density $\varrho(\mathbf{r})$ and the on-top pair density $P(\mathbf{r},\mathbf{r})$ as input quantities for the correlation functional. As a two-particle quantity, $P(\mathbf{r},\mathbf{r})$ can distinguish between states with different multiplicity and, contrary to the spin densities $\varrho_\alpha, \varrho_\beta$, the on-top density is identical for the components of a state multiplet such as T($M=-1$), T($M=0$), and T($M=1$). Equation (4) can be used to convert existing correlation functionals so that they depend on total density $\varrho(\mathbf{r})$ and on-top pair density $P(\mathbf{r},\mathbf{r})$ rather than spin densities $\varrho_\alpha(\mathbf{r})$ and $\varrho_\beta(\mathbf{r})$ [35, 41, 42];

$$\varepsilon_C^{\text{on-top}}(\varrho,P)=\varepsilon_C\left(\frac{\varrho}{2}+\left[\left(\frac{\varrho}{2}\right)^2-P\right]^{1/2},\frac{\varrho}{2}-\left[\left(\frac{\varrho}{2}\right)^2-P\right]^{1/2}\right), \tag{4}$$

where $\varepsilon_C$ is the correlation energy.

### 3.3. *Choice of the appropriate correlation functional: CAS-SCF + LYP versus CAS-SCF + CS*

CAS-SCF + LYP calculations, in which the LYP correlation energy is simply added to the CAS-SCF energy, as suggested by several authors [38, 48], are also included in table 1 and figure 1. Clearly, CAS + LYP does not lead to a change in the relative CAS-SCF energies because the LYP correlation functional is insensitive to the changes in the CAS pair density when calculating the three different states. CAS + LYP follows the overall changes in the active space energies of CAS-SCF. We conclude that a simple addition of the LYP functional (calculated with the CAS-SCF densities) does not lead to any results better than those of CAS-SCF.

This changes almost dramatically when the Colle–Salvetti (CS) functional [33] is used. We have previously [41, 42] chosen this functional for CAS-DFT calculations for the following reasons.

(i) The CS functional has been developed for the subsequent correction of Hartree–Fock wave functions [33].

(ii) The CS functional does not use the spin-resolved densities but the total density and pair density as input quantities. It is thus more sensitive to the features of the electronic structure that are reflected in the CAS-SCF wave function.

(iii) The CS functional contains gradient corrections and therefore is superior to LDA functionals.

In table 1 and figure 1, the results of the CAS + CS calculations are given. The CS functional is, in distinction to the LYP functional [3] (actually derived from the CS functional), sensitive to the quality of the reference wavefunction because CS depends not only on the total

density but also on the on-top density and the curvature of the pair distribution function around the position of the reference electron. Both of the latter quantities will change (on-top density down, curvature up) if the quality of the CAS reference wave function increases. In LYP however, the on-top density and curvature are expressed by the total density and its gradient, implicitly assuming a wave function of HF quality.

Figure 1 reveals that the CS functional introduces the expected (and needed) sensitivity to the changes in the CAS density by decreasing the energy gains with increasing active space (by 4.4, 5.5 and 9.9 kcal mol$^{-1}$; table 1), i.e. it scales back correlation effects so that more of them are included in the active space. The back-scaling (coupling via the CAS density) is however too weak so that the double counting of electron correlation effects experienced at the CAS + LYP level largely remains. This double counting problem will be investigated in section 4.

## 4. A balanced description of non-dynamical and dynamical correlation in CAS-DFT

The basic problem of a simple additive CAS-SCF + DFT approach is illustrated in figure 2. The CAS-SCF reference wave function contains the correlation effects related to excitations into low-lying unoccupied orbitals and treats them exactly. For the purpose of understanding what correlation effects are accounted for by the DFT correlation functional, we revert to the homogeneous electron gas (HEG) and a LDA functional. The LDA correlation functional models all correlation effects, no matter whether high-lying or low-lying virtual orbitals are involved. Similar arguments apply to a GGA functional such as the CS functional. By simply adding the CAS-SCF and DFT energy contributions, one counts the correlation effects arising from low-lying excitations twice, however in the form of dynamical rather than non-dynamical correlation, i.e. just a part of the non-dynamical correlation is counted twice (assuming that the non-dynamical correlation energy associated with the excitations in question is larger than the dynamical correlation energy). This becomes more of a problem the larger the active space is, i.e. the higher the quality of the underlying CAS-SCF calculation is.

A prospective approach to tackle the double counting of correlation effects was suggested by Savin [39], worked out later by Miehlich, Savin and Stoll (MSS) for atom calculations [40], and then developed to a generally applicable method by Gräfenstein and Cremer [41, 42]. According to this approach, the correlation effects related to excitations in low-lying virtual orbitals are excluded from the DFT correlation energy term by using equation (5);

$$E^{\mathrm{CAS-DFT}} = E^{\mathrm{CAS-SCF}} + \int \mathrm{d}^3 r \varepsilon_{\mathrm{c}}^{\mathrm{DFT}}(\mathbf{r}), \qquad (5\,a)$$

$$\varepsilon_{\mathrm{c}}^{\mathrm{DFT}}(\mathbf{r}) = \varepsilon_{\mathrm{c}}^{\mathrm{DFT,\,std}}(\mathbf{r}) - \varepsilon_{\mathrm{c}}^{\mathrm{DFT,\,trunc}}(\mathbf{r}), \qquad (5\,b)$$

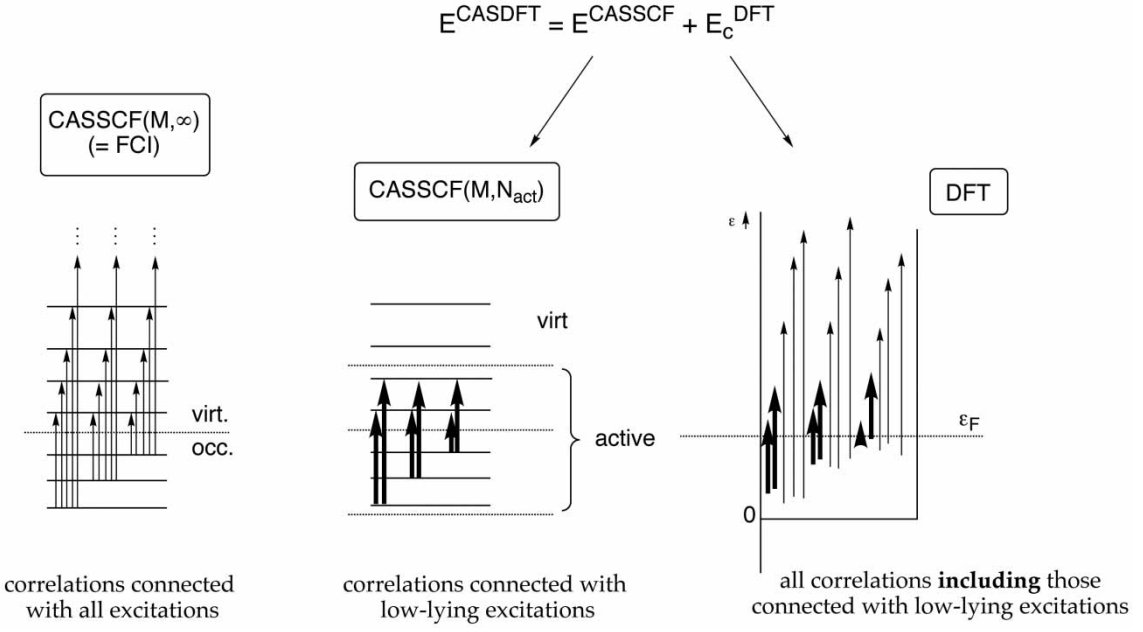


Figure 2. Double counting of correlation effects in a simple CAS + DFT approach. The electron correlation effects accounted for by all excitations of a FCI (left), those of CAS-SCF with $M$ electrons and $N_{\mathrm{act}}$ orbitals (middle) and those of a DFT calculation (right) are schematically indicated by arrows. Bold arrows denote those excitations (correlation effects) counted in CAS-DFT twice. virtual: virtual orbitals; occupied: occupied orbitals; $\epsilon_{\mathrm{F}}$: Fermi energy.

where 'std' refers to a standard DFT correlation energy (i.e. the density of the CAS-SCF wave function is simply inserted into a DFT correlation functional, in our case the CS functional) and 'trunc' means a truncated DFT correlation energy related to low-lying excitations only. The larger the active space in the CAS-SCF wave function is the higher the limit energy in the truncated DFT part has to be chosen. Figure 3 shows schematically the partitioning of the correlation energy between the CAS-SCF reference function and the DFT correlation functional.

The realization of this idea raises two questions. (a) How can the limiting energy for the DFT excitations be determined? The orbital energy spectrum is discrete for the real system, whereas it is continuous for the HEG that is used as a model to calculate the DFT correlation energy. In addition, orbital energies are not well-defined quantities in CAS-SCF. A measure for the DFT limiting energy that can be directly transferred to the size of the active space can be obtained by considering local densities. For this purpose, one calculates for the real system both the real density $\varrho(\mathbf{r})$ and a reference density $\varrho_{\mathrm{ref}}(\mathbf{r})$ defined by

$$\varrho_{\mathrm{ref}}(\mathbf{r}) = \sum_{k}^{\mathrm{act.}} 2[\varphi_k(\mathbf{r})]^2, \tag{6}$$

i.e. as that density, one would determine if all active orbitals were doubly occupied. The ratio

$$\eta_{\mathrm{ref}}(\mathbf{r}) = \left[\frac{\varrho_{\mathrm{ref}}(\mathbf{r})}{\varrho(\mathbf{r})}\right]^{1/3} \tag{7}$$

is then a local measure for the size of the active space. For a HF wave function with no non-dynamical correlation at all, $\eta_{\mathrm{ref}}(\mathbf{r}) = 1$; with increasing active space, $\eta_{\mathrm{ref}}(\mathbf{r})$ becomes larger and tends to infinity for the limit of a full CI (FCI) calculation with a complete basis set. It should be noted that $\eta_{\mathrm{ref}}(\mathbf{r})$ can be equal or close to 1 locally even for a large active space if the weakly populated active orbitals are concentrated in other regions of the molecule. The ratio $\eta_{\mathrm{ref}}(\mathbf{r})$ is thus a measure describing how many weakly populated active orbitals are available locally to describe correlation effects in the real system. Equivalently, $\eta_{\mathrm{ref}}(\mathbf{r})$ can be seen as a measure for the size of the local correlation hole present in the reference wavefunction, which is a view that fits well with the KS-DFT approach.

(b) The second question concerns the calculation of the truncated DFT correlation energy. This is straightforward at the LDA level. The correlation energy for the HEG is recalculated with the excitations limited to low-lying virtual orbitals up to a certain truncation energy. This truncation energy is defined in a way that the reference density for occupied and low-lying virtual orbitals just equals $\varrho_{\mathrm{ref}}(\mathbf{r})$ of the real system. One recalculates then the correlation energy of the HEG with this truncation, providing an expression for the truncated LDA correlation energy $\varepsilon_{\mathrm{c}}^{\mathrm{LDA}}(\varrho, \eta_{\mathrm{ref}})$, where the conventional LDA correlation energy density is given by the special case $\varepsilon_{\mathrm{c}}^{\mathrm{LDA}}(\varrho) = \varepsilon_{\mathrm{c}}^{\mathrm{LDA}}(\varrho, \infty)$. With this truncated LDA energy, one can define the DFT correlation terms in equation (4) as

$$\varepsilon_{\mathrm{c}}^{\mathrm{DFT,\,std}}(\mathbf{r}) = \varepsilon_{\mathrm{c}}^{\mathrm{LDA}}(\varrho)\Big|_{\mathbf{r}}, \tag{8 a}$$

$$\varepsilon_{\mathrm{c}}^{\mathrm{DFT,\,trunc}}(\mathbf{r}) = \varepsilon_{\mathrm{c}}^{\mathrm{LDA}}(\varrho, \eta_{\mathrm{ref}})\Big|_{\mathbf{r}}. \tag{8 b}$$

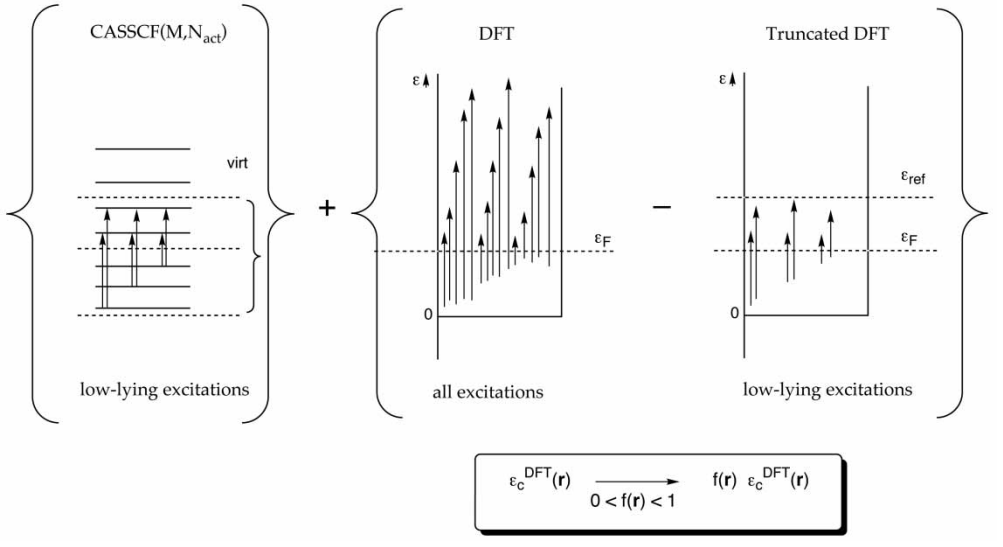


Figure 3. Partitioning of the correlation energy in a CAS-DFT calculation according to Miehlich *et al.* [40]. For explanations, see figure 2 and text.

This corresponds to an LDA description of the correlation energy. Of course, the accuracy of LDA is insufficient for reliable quantum-chemical calculations. On the other hand, it is not clear how the truncated DFT correlation energy can be defined in a GGA or other more sophisticated approaches. Therefore, one uses expressions (9) rather than equations (8 *a*) and (8 *b*):

$$\varepsilon_c^{\mathrm{DFT,\,std}}(\mathbf{r}) = \varepsilon_c^{\mathrm{GGA}}[\varrho]\Big|_{\mathbf{r}}, \tag{9 a}$$

$$\varepsilon_c^{\mathrm{DFT,\,trunc}}(\mathbf{r}) = \frac{\varepsilon_c^{\mathrm{LDA}}(\varrho,\eta_{\mathrm{ref}})}{\varepsilon_c^{\mathrm{LDA}}(\varrho,\infty)}\varepsilon_c^{\mathrm{GGA}}[\varrho]\Big|_{\mathbf{r}}, \tag{9 b}$$

i.e. the absolute correlation energy is calculated at the GGA (or another elaborate) level of DFT, only the relative effect of the truncation is determined at the LDA level. Technically, one calculates the scaling factor

$$f(\varrho,\varrho_{\mathrm{ref}}) = 1 - \frac{\varepsilon_c^{\mathrm{LDA}}(\varrho,\eta_{\mathrm{ref}})}{\varepsilon_c^{\mathrm{LDA}}(\varrho,\infty)} \tag{10}$$

with dependence on $\varrho$ and $\eta_{\mathrm{ref}}$ once and forever, then parametrizes it in an analytical form and determines then the total energy according to

$$E^{\mathrm{CAS-DFT}} = E^{\mathrm{CAS-SCF}} + \int \mathrm{d}^3 r f(\varrho,\varrho_{\mathrm{ref}})\varepsilon_c^{\mathrm{GGA}}[\varrho]\Big|_{\mathbf{r}}. \tag{11}$$

In the following, this approach [40–42] will be called CAS + f DFT with reference to the scaling factor *f*. In this work, we use for all calculations in this connection CAS + f CS as explained above.

Both the test calculations by MSS [40] and our own test calculations [41, 42] suggest that CAS + f DFT is a reasonable approach for incorporation of dynamical correlation effects into the active space. Figure 1 and table 1 reveal however that the approach suffers from some serious shortcomings.

The CAS + f CS energies for the (2,2) and (6,4) active spaces of $CH_2$ ($^1A_1$) are about 4 kcal mol$^{-1}$, those for (6,6) and (8,7) spaces about 9.2 kcal mol$^{-1}$, and that for the (2,4) active space even 23.6 kcal mol$^{-1}$ above the HF + CS reference energy. Although the dependence of the energy on the active space has been reduced by 10 kcal mol$^{-1}$, the relative CAS + f CS energies are far from reaching CAS-MP2 quality (even if the (2,4) value is excluded). In addition, the energy ordering is incorrect for CAS-SCF + f DFT: the (6,6) and (8,7) energies are above rather than below the (2,2) and (6,4) ones, which suggests that the scaling factor exaggerates the elimination of doubly counted correlation effects. This error increases with the size of the active space where for the (2,4) space CAS + f CS even enlarges the unbalanced description of non-dynamical and dynamical correlation effects observed at the CAS-MP2 level. Hence, there is a need to analyse the failures of the scaling procedure inherent in the CAS + f CS approach, which will be done in the next subsection.

### 4.1. *Improved determination of the scaling factor by considering equal-spin and opposite-spin correlation effects*

An obvious limitation of the CAS + f DFT approach is the calculation of the scaling factor at the LDA level, which is known to have limited accuracy. An improved treatment of the truncated DFT energy may resolve at least some of the problems encountered for CAS + f DFT. Stoll, Pavlidou and Preuss (SPP) [84] analysed the performance of LDA in standard DFT calculations and pointed out that an important reason for the limited performance of LDA is that equal-spin correlation effects are overestimated relative to opposite-spin correlation effects. This becomes manifested in the large self-interaction error (SIE) that occurs for LDA correlation functionals, i.e. the fact that LDA incorrectly predicts a non-vanishing correlation energy for a one-electron system. This overestimation is due to the difference between the HEG and finite systems.

(a) In the HEG, the exchange hole is more diffuse the lower the density is and shows long-range oscillations, which are an artefact related to the vanishing HOMO–LUMO energy difference. An important portion of the equal-spin LDA correlation hole just fills the minima in the oscillations of the exchange hole, which results in a considerable amount of equal-spin correlation energy.

(b) In the real system, in contrast, the size of the exchange hole is limited, and there is in particular no counterpart to the long-range oscillations one finds in the HEG. As a consequence, equal-spin correlation effects are much weaker than in the case of the HEG.

(c) For opposite-spin electrons there is no exchange and thus the inconsistency in the LDA description of electron correlation vanishes.

The exaggeration of equal-spin correlation is an artefact of the HEG itself and occurs even if this model system is treated exactly. This artefact reflects an essential difference between the HEG and the real system, which has been pointed out by Dobson [85]: in the HEG, each electron has a large number of other equal-spin electrons in its vicinity, allowing for substantial energy gain by equal-spin correlation.

A similar situation is found in metals, which rationalizes the success of LDA for the description of metals. In molecules or covalently bonded solids, in contrast, the equal-spin electrons are more strongly separated. It is unlikely, e.g. to find two equal-spin electrons in the region of one covalent bond. Consequently, equal-spin correlation effects are weak and contribute only little to the total energy. In a one-electron system, there are no other equal-spin electrons at all, but the LDA description still assumes nearby equal-spin electrons for each reference point. This accounts for the large SIE of the LDA correlation.

In the CAS + f DFT approach, the LDA correlation energy is calculated with the random-phase approximation (RPA), which further increases the exaggeration of equal-spin correlation effects.

One might expect that this inconsistency of LDA is problematic mainly for open-shell systems. However, even for a closed-shell system this inconsistency can influence the full and truncated DFT correlation energy differently and thus influence the value of scaling factor $f$. The long-range oscillating part of the LDA correlation hole is mainly due to low-lying excitations, i.e. excitations involving orbitals close to the Fermi level. This means eventually that an RPA calculation of $f$ will give too small values of $f$, i.e. tends to reduce the DFT correlation energy too strongly.

The solution suggested by SPP [84] is to completely eliminate equal-spin correlations from the DFT correlation functional. This implies that the DFT correlation energy is approximated by

$$E_c^{SPP}[\varrho_\alpha, \varrho_\beta] = E_c^{LDA}[\varrho_\alpha, \varrho_\beta] - E_c^{LDA}[\varrho_\alpha, 0] - E_c^{LDA}[0, \varrho_\beta]. \quad (12)$$

This expression is self-interaction free by construction, no matter which approximation is used for the correlation functional. Here $E_c^{LDA}[\varrho_\alpha, \varrho_\beta]$ is the standard LDA correlation energy for the spin-resolved density $\varrho_\alpha, \varrho_\beta$. As SPP showed [84], the modification of the DFT functional along these lines will improve accuracy, especially if the SPP functional is used together with a HF description of the exchange energy. Even though the complete suppression of equal-spin correlation effects is physically incorrect, the error of this approximation is considerably smaller than the exaggeration of equal-spin correlation in conventional LDA. It should be noted in this connection that the GVB-LDA approach by Kraka [36, 37] used a SPP-corrected correlation functional.

The idea of the SPP correction is easy to transfer to the truncated DFT used in the CAS + f DFT approach. For this purpose, the truncated DFT energy is calculated both for a spin-unpolarized HEG with the total density $\varrho$ (i.e. $\varrho_\alpha = \varrho_\beta = \varrho/2$) and a completely spin polarized electron gas with $\varrho_\alpha = \varrho/2$, $\varrho_\beta = 0$ with the same limiting energy for the virtual orbitals. Then, the truncated correlation energy is calculated in analogy to equation (12).

The SPP-corrected CAS-SCF + f DFT energies for $^1A_1$ $CH_2$ are shown in figure 1. The (2,2), (6,4), (6,6) and (8,7) energies are at this level less than 1.5 kcal mol$^{-1}$ above the (0,0) reference value, i.e. the inclusion of the SPP corrections improves the balance between dynamical and non-dynamical correlation effects for different active spaces.

The CAS + f DFT(SPP) energies collected in figure 1 and table 1 show that there is indeed a significant improvement caused by the SPP functional: the energies of four of the active spaces considered are now found within 1.4 kcal mol$^{-1}$ above the reference energy, which corresponds to an improvement factor larger than 5. However, the energy ordering of the states is still incorrect (energy for (2,2) below (6,6), figure 1) similarly as in the original CAS + f DFT method. The (2,4) state is 16 kcal mol$^{-1}$ above the (0,0) reference state. This is a strong indication that there are more inconsistencies to be corrected in the original CAS + f DFT method.

### 4.2. *The distinction between active and inactive orbitals: CAS-DFT1*

The CAS-SCF + f DFT method assumes implicitly that all orbitals are either active or virtual, i.e. that there is no inactive (core) space. This applies to all examples investigated by MSS [40] but is not typical for normal applications, where one has to limit the active space to those orbitals involved in the chemical processes of bond breaking and bond forming for computational reasons [41, 42].

Figure 4 shows the major inconsistency in the treatment of correlation effects in a CAS + f DFT calculation that does not differentiate between active and inactive orbitals: CAS + f DFT corrects for the double-counting occurring in CAS + DFT but suppresses incorrectly excitations from inactive into weakly occupied active orbitals. The resulting errors will be especially large when the inactive and the weakly occupied active orbitals are deformed during a chemical process (chemical reactions, excitations, etc.).

The determination of the scaling factor $f$ has to be generalized so that it accounts for the size of both the active and the inactive space. The basic idea of this generalized approach is shown schematically in figure 5. In the truncated DFT calculation for CAS-DFT one excludes not only all excitations into high-lying virtual orbitals but also those from low-lying occupied orbitals. The limiting energy for the occupied

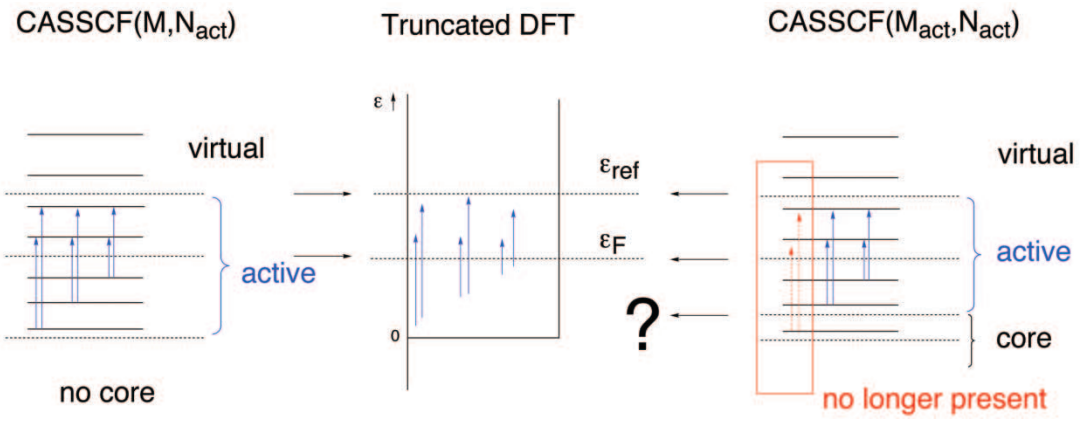


Figure 4. The missing excitations (correlation effects) in the CAS + f DFT approach when introducing an inactive (core) space of finite size (right, dashed arrows). Despite the fact that they are not present in the reference wave function, they are scaled-back via the truncated DFT. For further explanations, see figure 2 and text.

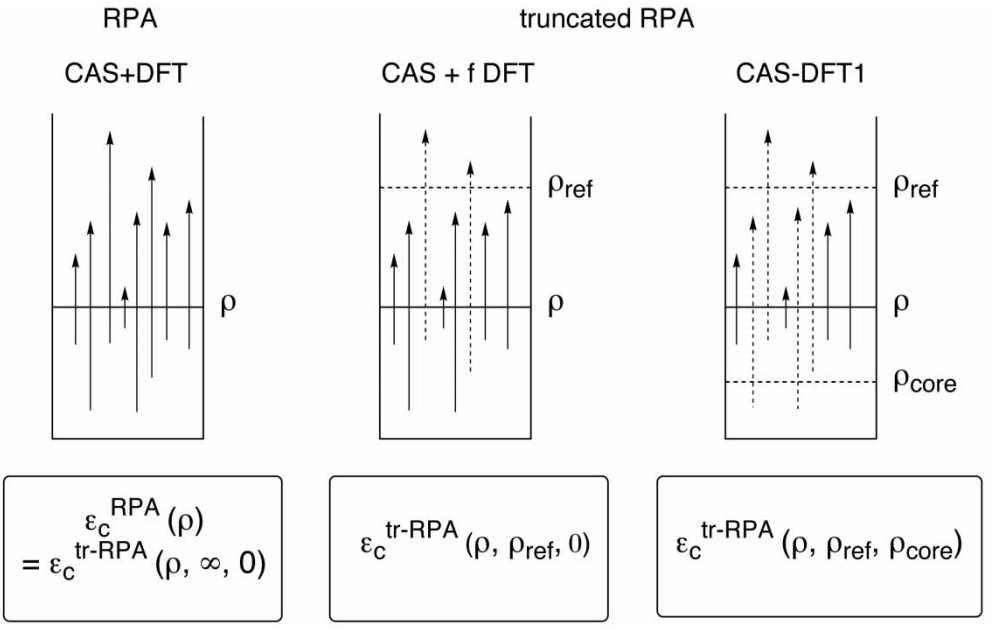


Figure 5. Schematic description of a more rigorous treatment of scaling active and core excitations (correlation effects) in the CAS-DFT1 method using the Random Phase Approximation (RPA) in a truncated form. The dashed arrows indicate excitations that are eliminated when a finite size of the active space is considered (middle) and in addition a finite size of the inactive (core) space (right). For further explanations, see figure 2 and text.

orbitals is adjusted in a way that the total density of all low-lying (i.e. inactive) occupied orbitals just equals $\varrho_{\text{core}}$ in the real system. This means that the scaling factor $f$ is generalized to depend on both the actual, the reference, and the core density. The DFT correlation energy calculated at this level takes the form of equation (13):

$$\varepsilon_{\text{c}}^{\text{DFT}}(\mathbf{r}) = f(\varrho, \varrho_{\text{ref}}, \varrho_{\text{core}})\varepsilon_{\text{c}}^{\text{GGA}}[\varrho]\Big|_{\mathbf{r}}. \qquad (13)$$

The calculation of $\varepsilon_{\text{c}}^{\text{LDA}}(\varrho, \eta_{\text{ref}}, \eta_{\text{core}})$, referred to as 'truncated RPA', is described in Appendix A. The results were parametrized for the use in the CAS-DFT code incorporated into COLOGNE 2004 [77] as described in Appendix B. Energies obtained with the general scaling factor are denoted as CAS-DFT1.

The parametrization suggested by MSS [34] is inappropriate for an extension to the CAS-DFT1 approach. We constructed therefore a new model function for the scaling factor $f$, which incorporates some known features of the scaling factor from the beginning. This model function allowed us to parametrize the scaling factor with a sufficient accuracy. The analytical expression found from the parametrization contains the scaling factor for CASSCF + f DFT

as a special case and will be used in this case as well instead of the expression suggested in [34].

The CAS-DFT1 results for the $^1A_1$ state of $CH_2$ (figure 1) reveal an equally directed, but even stronger improvement than obtained with the SPP functional. The relative energies are spread over 11 rather than 19 kcal mol$^{-1}$ as in the case of CAS + f DFT (table 1). The energies of the active spaces are ordered in the way that the minimal active space has the lowest, the full valence + 1s(C) active space has the highest energy, which is just opposite to the CAS-SCF ordering. Introduction of the SPP functional thus leading to CAS-DFT1(SPP) helps to improve results further. Energies are spread over just 6 kcal mol$^{-1}$ reaching almost CAS-PT2 quality. However, the ordering of states is still wrong suggesting that with the number of electrons in the active space the scaling factor exaggerates correlation effects covered by DFT thus leading to a stronger reduction of the DFT correlation energy as it should.

The discrepancies suggest that electron correlation effects involving low-lying occupied orbitals are described in an essentially different way in the HEG than in the real system. In the HEG, the lowest occupied orbitals correspond to plane waves with low wave numbers, which vary in space very slowly. This means that the corresponding correlation hole may have long-range tails. Excitations from the occupied orbitals into all virtual orbitals, including those closely above the Fermi energy, may make substantial contributions to the correlation hole. In the real system, in contrast, the lowest-lying occupied orbitals (1s orbitals) are concentrated in space, and the size of the correlation hole is limited by the extension of the orbital. The virtual orbitals needed to describe the correlation effects between 1s orbitals have high orbital energies; excitations from the 1s orbitals into the low-lying valence orbitals make only small contributions to the correlation energy. Overall, the lowest occupied orbital makes a larger relative contribution to the correlation energy in the HEG than in the real system, and one should expect that CAS-DFT1 largely overcompensates the decrease in the CAS-SCF energy one obtains by including inner core orbitals into the active space due to large back-scaling appropriate for the HEG but not for the real system. This is exactly what one observes for the (6,6) and (8,7) CAS-DFT1 energies for the $^1A_1$ state of $CH_2$.

### 4.3. *The distinction between active and inactive orbitals: CAS-DFT2*

As was pointed out in the previous subsection, the CAS-DFT1 method generates too extended correlation holes in connection with the inner core orbitals and exaggerates thus the correlation-energy contributions from these orbitals. This problem can be remedied if the core correction of the correlation energy is based on the core density alone rather than the total density. This is the basic idea for an improved CAS-DFT approach, which will be denoted as CAS-DFT2 in the following.

In CAS-DFT2, a correction term is added to the calculation of the correlation energy, which accounts for excitations from the inactive into the weakly populated active orbitals:

$$\varepsilon_c^{\mathrm{DFT}} = \varepsilon_c^{\mathrm{DFT,\,std}} - \varepsilon_c^{\mathrm{DFT,\,trunc}} + \varepsilon_c^{\mathrm{DFT,\,core\to weak}}. \quad (14)$$

The correction term in turn is approximated by

$$\varepsilon_c^{\mathrm{DFT,\,core\to weak}} = \varepsilon_c^{\mathrm{DFT,\,core\to act}} - \varepsilon_c^{\mathrm{DFT,\,core\to strong}}, \quad (15)$$

where DFT, core $\to$ act denotes excitations from the inactive orbitals into all active orbitals and DFT, core $\to$ strong, excitations from the inactive orbitals into all strongly occupied active orbitals. Each term in equation (14) is equivalent to $\varepsilon_c^{\mathrm{DFT,\,trunc}}$ from equation (5*b*) with suitably substituted actual and reference densities. One can thus write

$$\varepsilon_c^{\mathrm{DFT,\,core\to act}} = [1 - f(\varrho_{\mathrm{core}}, \varrho_{\mathrm{ref}})]\varepsilon^{\mathrm{GGA}}[\varrho_{\mathrm{core}}]\Big|_{\mathbf{r}}, \quad (16a)$$

$$\varepsilon_c^{\mathrm{DFT,\,core\to strong}} = [1 - f(\varrho_{\mathrm{core}}, \varrho)]\varepsilon^{\mathrm{GGA}}[\varrho_{\mathrm{core}}]\Big|_{\mathbf{r}}, \quad (16b)$$

and altogether

$$\varepsilon_c^{\mathrm{DFT}} = f(\varrho, \varrho_{\mathrm{ref}})\varepsilon^{\mathrm{GGA}}[\varrho] + [f(\varrho_{\mathrm{core}}, \varrho) - f(\varrho_{\mathrm{core}}, \varrho_{\mathrm{ref}})]\varepsilon^{\mathrm{GGA}}[\varrho_{\mathrm{core}}]\Big|_{\mathbf{r}}. \quad (17)$$

The expression $\varepsilon^{\mathrm{GGA}}[\varrho_{\mathrm{core}}]$ utilizes a correlation hole, the extent of which is adjusted to the extent of the core orbitals, that is, the artificial long-range core correlations in CAS-DFT1 are suppressed from the beginning. This method was suggested previously [41, 42] to overcome the major shortcoming of the CAS + f DFT approach. In this work, we reintroduce this approach as CAS-DFT2 where we take advantage of the new scale factor determination described in the Appendix, the equal-spin correction, and other features described in the next section.

Inspection of figure 1 and table 1 does not reveal any improvement of the active space energies calculated with CAS-DFT2 relative to those obtained with CAS-DFT1. However, CAS-DFT2(SPP) energies are clearly better than CAS-DFT1 or CAS-DFT2 energies. (a) They reach CAS-MP2 quality (spreading over just 4.4 kcal mol$^{-1}$

when excluding the (8,7) case). (b) Absolute energies for the (2,4), (6,4) and (6,6) spaces are almost identical, i.e. the (2,4) active space is handled by CAS-DFT2(SPP) on an equal footing contrary to the failure of the CASPT2 (CAS-MP2) method in this case.

The improvement caused by CAS-DFT2(SPP) becomes especially clear when comparing the energy difference between the (6,6) and (8,7) spaces. Ideally, this energy difference should be zero (see equation (1)). For the $^1A_1$ state, the value is decreased from 4.4 kcal mol$^{-1}$ at CAS-DFT1(SPP) to 0.9 kcal mol$^{-1}$ at CAS-DFT2(SPP). For the $^3B_1(^1B_1)$ state, the corresponding values are 0.9 and 3.6 (0.9 and 3.7) kcal mol$^{-1}$ thus underlining the improvement caused by CAS-DFT2(SPP). This comparison reveals also that CAS-DFT2(SPP) seems to provide a cancellation of errors where CAS-DFT1(SPP) introduces a variation of 0.7 to 0.8 kcal mol$^{-1}$ in the relative energies.

There is still an inconsistency in that the energy for the four largest active spaces is 4 kcal mol$^{-1}$ above that of the (2,2) space. Clearly, this indicates that the scaling factor still is too small for larger active spaces (too strong cut-back of correlation effects). In particular, the (6,4) energy is 4 kcal mol$^{-1}$ above the (2,2) energy even though the CAS-SCF value indicates that the two CAS-SCF wave functions are nearly identical. The $f$ scaling factor is obviously not sensitive to situations where an increase of the active space does not lead to additional correlation effects in the reference wave function. Besides, the energies for the largest active spaces are not below the (0,0) energy. This discrepancy can be traced back to the interplay between the CS correlation functional and the scaling factor: the CS functional is to some extent sensitive to the quality of the reference wavefunction whereas in the derivation of the scaling factor it is assumed that the DFT correlation functional does not possess such sensitivity. Therefore, for large active spaces, parts of the correlation energy may be eliminated twice: once by a decrease of the total CS correlation energy and once more by a decrease of the scaling factor $f$.

It was mentioned above that the tests made for the data in table 1 and figure 1 represent stringent tests where changes in the absolute energies of a single electronic system are directly used. We refrain from making similar tests for the $^3B_1$ and $^1B_1$ states because we have to consider now the chemically more interesting energies. Any real electronic system experiences a variety of different electron correlation effects (left–right, in–out, angular pair correlation effects; different types of three-, four-, $n$-electron correlation effects; short-range, long-range correlation effects; equal-spin, opposite-spin correlation effects; etc.), which lead to correlation energy contributions of significantly varying magnitude. One cannot expect that a scaling factor based on the HEG can account for all these variations.

There is however no need for such sensitive scaling factors. For real electronic systems, differences of energies have to be discussed rather than absolute energies. Because of this one can expect that for relative energies most of the smaller inconsistencies caused by variations in the correlation effects cancel each other out. This will be considered in the next section.

## 5. Differences in electron correlation for closed- and open-shell systems

In table 2 and figure 6, the singlet–triplet (S–T) splitting energy for $CH_2$, i.e. the excitation energy $E(^1A_1) - E(^3B_1)$ is shown where again a stepwise improvement of the CAS-SCF to the CAS-DFT2(SPP) description is followed for the different active spaces considered.

Table 2. Energy differences $E(^1A_1) - E(^3B_1)$ calculated for $CH_2$ at various level of theory.[a]

| | | | | CAS + f CS | | | CAS-DFT2(cs) | |
|---|---|---|---|---|---|---|---|---|
| Active space | CAS-SCF | CAS-MP2 | CAS + CS | $f_{CS}$ ($f_{CS}$ + SPP) | $f_{OS}$ ($f_{OS}$ + SPP) | CAS-DFT1(CS) FCS (FCS,SPP) | FCS (FCS,SPP) | FOS (FOS,SPP) |
| (0,0) | 24.70 | 15.30 | 15.82 | 15.82 (15.82) | 15.82 (15.82) | 15.82 (15.82) | 15.82 (15.82) | 15.82 (15.82) |
| (2,2) | 10.48 | 9.65 | 5.80 | 5.77 (6.48) | 13.41 (10.64) | 4.53 (5.80) | 7.01 (7.79) | 14.64 (11.94) |
| (6,4) | 10.07 | 1.63 | 5.44 | 5.53 (6.19) | 13.17 (10.37) | 5.18 (5.97) | 5.81 (6.53) | 13.45 (10.71) |
| (2,4) | 6.62 | 18.22 | 3.01 | 1.08 (1.76) | 6.21 (4.75) | 3.98 (4.89) | 5.90 (6.95) | 11.03 (9.95) |
| (6,6) | 10.36 | 9.04 | 5.20 | 6.63 (6.74) | 11.84 (9.74) | 5.65 (5.86) | 6.60 (6.72) | 11.81 (9.72) |
| (8,7) | 10.36 | 8.94 | 5.20 | 6.62 (6.73) | 11.83 (9.73) | 6.62 (6.73) | 6.62 (6.73) | 11.83 (9.73) |

[a]Relative energies in kcal mol$^{-1}$. All calculations with the 6-311++G(2df,1pd) basis set. The (0,0) space denotes HF, MP2, HFLYP, HFCS references energies. Active spaces: (2,2): minimal; (6,4): minimal + all occupied valence orbitals; (2,4): minimal + all virtual valence orbitals; (6,6): full valence; (8,7): full valence + 1s(C). CS: Colle–Salvetti functional; FCS and FOS: scaling factor $f$ calculated for a closed- or open-shell situation, respectively; SPP: Stoll–Pavlidou–Preuss functional for elimination of equal-spin correlation; levels 1 and 2: different level of distinguishing between core and active electron correlation.

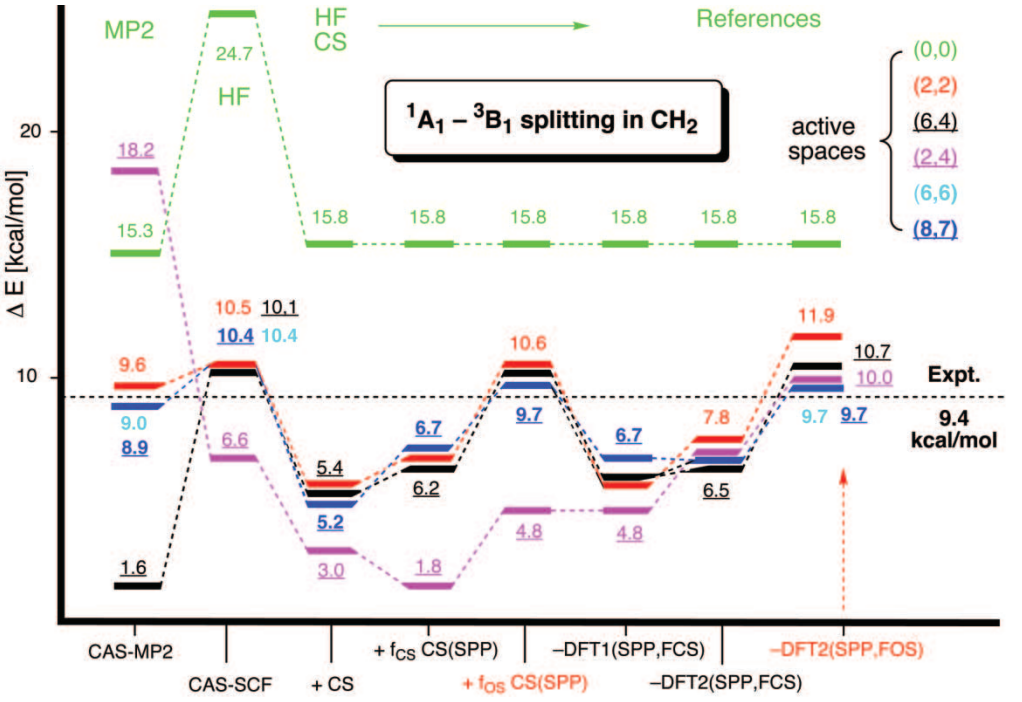


Figure 6. Changes in the excitation energy $E(^1A_1) - E(^3B_1)$ of $CH_2$ for an increasing active space calculated at different levels of theory with a 6-311 + + G(2df,2pd) basis set. CS: Colle–Salvetti functional; FCS and FOS: scaling factor *f* calculated for a closed-shell and open-shell situation, respectively; SPP: Stoll–Pavlidou–Preuss functional for elimination of equal-spin correlation; levels 1 and 2: different levels of distinguishing between core and active electron correlation, as calculated by different methods.

Since HF misses the non-dynamical correlation effects involving the lone-pair electrons of the $^1A_1$ state, an $^1A_1$–$^3B_1$ splitting of 24.7 kcal mol$^{-1}$ is predicted, about 15 kcal mol$^{-1}$ above the experimental value of 9.4 kcal mol$^{-1}$ derived from the experimental enthalpy value of 9.0 kcal mol$^{-1}$ [86] by considering thermal corrections [42]. CAS-SCF, in contrast, gives an S–T splitting between 10.1 and 10.5 kcal mol$^{-1}$, i.e. less than 1 kcal mol$^{-1}$ off the experimental value. An exception is found in the case of the (2,4) space, which underestimates the splitting by about 3 kcal mol$^{-1}$ (figure 6). The $\sigma^*$(CH) orbitals in this active space are relevant for correlations between the two $\sigma$ lone-pair electrons in the $^1A_1$ state, and adding them to the active space decreases the CASSCF energy. For the $^3B_1$ state, in contrast, the correlations between the two unpaired electrons are weak in the first place because of the high-spin coupling. Besides, the correlations between the $\sigma$ and the $\pi$ electron involve out-of-plain movements of the $\sigma$ electron, which cannot be described by excitations into the two $\sigma^*$(CH) orbitals. Therefore, the (2,4) active space gives nearly the same CASSCF energy as the (2,2) space for the $^3B_1$ state. Accordingly, the (2,4) space stabilizes the $^1A_1$ state artificially.

MP2/6-311 + + G(2df,2pd) calculations improve the HF splitting to 15.3 kcal mol$^{-1}$. CAS-MP2 gives much better results for the (2,2) (9.6 kcal mol$^{-1}$, table 2) and the (6,6) or (8,7) spaces (9.0, 8.9 kcal mol$^{-1}$, table 1) all being close to the experimental value. CAS-MP2 fails however when using the (2,4) space (18.2 kcal mol$^{-1}$ due to an unbalanced stabilization of the $^3B_1$ state) or the (6,4) space (1.6 kcal mol$^{-1}$ due to an unbalanced stabilization of the $^1A_1$ state). In agreement with the observations made for the absolute energies of the $^1A_1$ state, we find that CAS-MP2 (CASPT2), despite the strong coupling of the perturbation correction to the CAS-SCF wave function, is partly unstable.

Adding dynamical electron correlation effects via the DFT correlation functional reduces the S–T splitting to about 3–6 kcal mol$^{-1}$, i.e. deteriorates the accuracy below that of CAS-SCF or DFT itself (B3LYP leads to a value of 10.5 kcal mol$^{-1}$, see, e.g. [64]). In view of the fact that the absolute energies for the $^1A_1$ state are reasonable (figure 1, table 1), the (lower-lying) open-shell state must be destabilized relative to the closed-shell state. According to calculated S–T splittings (figure 6, table 2), this destabilization is also acting, although slightly reduced, when scaling at the CAS + f CS(SPP) level is introduced and in addition active and inactive orbitals are distinguished (CAS-DFT1 or CAS-DFT2).

For the purpose of understanding the reason for this failure, we consider the CASF + f DFT calculations with a (2,2) active space for the $^1A_1$ and $^3B_1$ states.

For the $^1A_1$ state, a (2,2) calculation implies that the ground-state configuration and a doubly excited configuration are superimposed, which means that some electron correlation is covered by the CAS-SCF wave function and has to be excluded from the DFT correlation energy. This is reflected correctly by $\eta_{\text{ref}} > 1$, hence $f < 1$. For the $^3B_1$ state, CAS-SCF(2,2) is equivalent to restricted open-shell (RO) HF, i.e. no correlation is contained in the wave function and $f = 1$ must hold. However, as $\eta_{\text{ref}} > 1$, one obtains $f < 1$ in this case as well, and the correlation energy is scaled down. This leads to an artificial destabilization of the open-shell state, which is carried through all active spaces and CAS-DFT procedures whereas CAS-MP2 (CASPT2) is of course immune against this error because the form of the perturbation operator rather than a general scaling factor decide on the added dynamical electron correlation.

### 5.1. *Handling of closed- and open-shell systems in a balanced way*

The current form of CAS-DFT is not sensitive to the differences in electron correlation for open-shell and closed-shell states. This can be remedied if the scaling factor $f$ is derived for the general case of a spin-polarized system, i.e. if the calculation of $\varepsilon_c^{\text{LDA}}$ is done for the case of a spin-polarized electron gas.

The expression for the correlation energy in this case is given in Appendix A. A subtle problem is the parametrization of the resulting energy expressions, which for CAS-DFT1 depend nonlinearly on four parameters, namely $\varrho$, $\varrho_{\text{ref}}$, $\varrho_{\text{core}}$ and the spin polarization $\zeta$. The calculation of the scaling factor is straightforward in this case, however the fitting of calculated $f$-values to a four-dimensional function turns out to be difficult. For the time being, there is no parametrization of $f$ for the general case of CAS-DFT1 available. As open-shell systems are often considered in direct comparison to closed-shell systems, which gives a partial compensation of core-orbital effects, this lack is not too severe. CAS-DFT2 provides a reasonable approximation for the open-shell correlation energy without the need to perform a four-parameter nonlinear fit for the scaling factor:

$$E_c^{\text{CAS-DFT2, open}} = \int d^3r \varepsilon_c^{\text{DFT}}[\varrho]\ f(\varrho, \varrho_{\text{ref}}, \zeta) + \int d^3r\, \varepsilon_c^{\text{DFT}}[\varrho_{\text{core}}]\ [f(\varrho_{\text{core}}, \varrho, \zeta = 0) - f(\varrho_{\text{core}}, \varrho_{\text{ref}}, \zeta = 0)]. \tag{18}$$

Here the leading term of the correlation energy is corrected for the spin polarization whereas the correction term for the inactive orbitals is treated in a closed-shell approximation, i.e. as if the total electron density were 50% $\alpha$ and $\beta$ spin each (compare with figure 7). We will denote these methods using equation (18) as methods based on an open-shell corrected scaling procedure, abbreviated as FOS ($f$ for open-shell problems) or $f_{\text{OS}}$ DFT as opposed to FCS ($f$ for closed-shell problems) or $f_{\text{CS}}$ DFT.

Including the open-shell corrections given by equation (18), one obtains the S–T splittings given in figure 6 at the CAS + $f_{\text{OS}}$ CS(SPP) and CAS-DFT2 (SPP,FOS) level, which are compared with the corresponding CAS + $f_{\text{CS}}$ CS(SPP), CAS-DFT1(SPP, FCS) and CAS-DFT2(SPP,FCS) values. There is a substantial

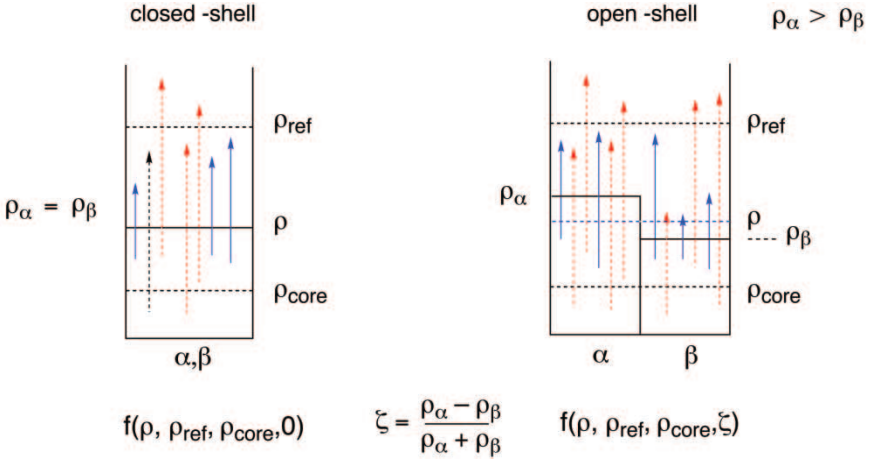


Figure 7. Schematic description of the different scaling procedures used in closed- and open-shell cases. Dashed arrows indicate excitations eliminated from the truncated DFT treatment because they must not be back-scaled (they are not accounted for by CAS-SCF). They are different for the $\alpha$ and the $\beta$ space of an open-shell system with finite spin polarization $\zeta$. For further explanations, see figure 2 and text.

improvement leading to a stabilization of the $^3B_1$ state by 3 to 7 kcal mol$^{-1}$ depending on the size of the active space (larger spaces lead to smaller stabilization of the triplet; the scaling procedures explicitly handling core orbitals lead to larger corrections, table 2). In figure 6, only the S–T splittings obtained with the SPP functional in combination with the FOS correction and CAS + f DFT or CAS-DFT2 are shown because CAS-DFT1 leads to similar results as given for CAS-DFT2.

Contrary to CAS-MP2 (CASPT2), CAS-DFT2 (SPP,FOS) leads to reasonable S–T splittings for all active spaces, where the excitation energies differ by just 0.3 to 2.5 kcal mol$^{-1}$ from the experimental value of 9.4 kcal mol$^{-1}$ (figure 6). In this respect, CAS-DFT2 (as well as CAS-DFT1) is clearly better than CAS + f DFT although the calculated S–T splittings for the smaller active spaces agree better with the experimental value (table 2, figure 6). However, the result for the (2,4) space is unbalanced for CAS + f DFT whereas it is reasonable for CAS-DFT2 (CAS-DFT1, see figure 6).

CAS-DFT2(SPP,FOS) provides for the first time an ordering of results as one should expect it from the design of the active spaces. The (2,2) space leads to a value (11.9 kcal mol$^{-1}$, table 2) below the accuracy of the CAS-SCF result because, as suggested before, important non-dynamical correlation effects are still missing. Also, the simplifications made concerning the spin polarization of the core electrons seem to be another reason for the relatively large discrepancy of 2.5 kcal mol$^{-1}$ from the experimental value. As soon as the number of core electrons is reduced as in the (6,4) space, the S–T splitting is improved to 10.7 kcal mol$^{-1}$. By adding missing non-dynamical electron correlation in the (2,4) and (6,6) space, values of 10.0 and 9.7 kcal mol$^{-1}$ are obtained. It is especially satisfying that the (8,7) space leads to a reproduction of the (6,6) result confirming that the 1s(C) electrons do not play any role.

Clearly, by inspecting relative rather than absolute energies the problems with the scaling factor increasing the cut-back of electron correlation with the number of active electrons largely vanish. However, a balanced description of open-shell cases in relation to closed-shell systems must be guaranteed where the situation will become even more complicated if one has to compare both high-spin and low-spin open-shell systems with closed-shell cases.

### 5.2. *Handling of high-spin and low-spin open-shell states in a balanced way*

The second excited singlet state of $CH_2$ is the low-spin open-shell $^1B_1$ state. The treatment of low-spin open-shell states poses a new problem for the CAS-DFT approach, which can be explained for a (2,2) treatment of the $^1B_1$–$^3B_1$ splitting in methylene. A (2,2) CAS-SCF description is equivalent to a ROHF description of the $^3B_1$ state and to a low-spin ROHF description in the case of the $^1B_1$ state. This means that both states are described at the HF level, the reference wave function contains no non-dynamical correlation effects, and the DFT correlation energy should not be scaled down.

CAS-DFT(FOS) will handle the $^3B_1$ state correctly. In the case of the $^1B_1$ state, in contrast, the following situation is faced: the spin polarization function $\zeta(\mathbf{r})$ introduced in the previous section is identically zero everywhere, wrongly suggesting that the $^1B_1$ state has closed-shell character. Furthermore, $\eta_{\mathrm{ref}} > 1$ (there are partly unoccupied orbitals in the active space) and consequently $f < 1$. The DFT correlation energy is incorrectly scaled down, and the $^1B_1$ state is destabilized relative to the $^3B_1$ state. Clearly, to assess the open-shell character of the $^1B_1$ state, one has to treat the two singly occupied orbitals as if they both were $\alpha$ orbitals in the high-spin state. This problem has been noticed by a number of authors in other contexts [65, 70, 71, 87, 88].

This lack of balance implies that CAS-DFT(FOS) is not size-consistent. This is easily seen for the dissociation of $N_2$. For a large distance between the two N atoms, the energy of the supermolecule should be twice the energy of a N atom. CAS-SCF with a minimal active space (i.e. (6,6) for the molecule, (3,3) for the atom) reflects this correctly. CAS-DFT(FOS), in contrast, describes the quartet state of the N atoms as high-spin open-shell states where no scaling of the CS correlation energy takes place, whereas the singlet state of the molecule is treated as a closed-shell state with a scaling of the CS correlation energy. Eventually, the $N_2$ supermolecule is predicted to be higher in energy than two independent N atoms. Only for a high-spin (i.e. septet) coupling between the two N atoms is CAS-DFT(FOS) size-consistent.

The spin-resolved density does not always give a proper account of the open-shell or closed-shell character of a state. It gives the global probability of finding electrons in a certain $\alpha$ or $\beta$ spin orbital, while the open- or closed-shell character is governed by the conditional probability of finding an electron in a $\beta$ spin orbital *provided* that the corresponding $\alpha$ spin orbital is occupied and vice versa. To follow this concept strictly, one would need the spin-resolved electron pair distribution. A simpler approach is available by considering the occupation of the (spin-unresolved) natural orbitals. Natural orbitals with an occupation number neither close to 0 or to 2 indicate an open-shell character

of the molecule. Based on this idea, Davidson and co-workers [87–89] define the effective density of unpaired electrons

$$D(\mathbf{r}) = \sum_i n_i(2 - n_i)\varphi_i(\mathbf{r})\varphi_i(\mathbf{r}), \tag{19}$$

where $\varphi_i(\mathbf{r})$ are the natural orbitals and $n_i$ are their occupation numbers. It can be shown [89] that $D(\mathbf{r})$ agrees with the conventional spin density for a high-spin ROHF wave function and gives just the density related to the open-shell electrons for a low-spin ROHF wave function.

A shortcoming of equation (19) is that $D(\mathbf{r})$ may exceed the total density, which makes it impossible to use $D(\mathbf{r})$ in lieu of the conventional spin density. This problem however can be avoided by using the quantity

$$\tilde{D}(\mathbf{r}) = \sum_i n_i^2(2 - n_i)\varphi_i(\mathbf{r})\varphi_i(\mathbf{r}), \tag{20}$$

which never exceeds the total density and thus is a proper substitute for the spin density in DFT correlation functionals. We denote the method based on equation (20) to describe a low-spin open-shell system as the DS (Davidson–Staroverov) [89] corrected CAS-DFT method and obtain in this way CAS + $f_{OS}$ CS(SPP,DS), CAS-DFT1(SPP,FOS,DS) and CAS-DFT2(SPP,FOS,DS) as new methods. The DS correction restores the size-consistency of CAS-DFT(FOS). For instance, in the above example, the energy of the $N_2$ supermolecule is calculated as if the molecule were a septet. On the other hand, the DS correction to closed-shell states, such as $N_2$ at its equilibrium geometry, is minimal.

Table 3 and figure 8 report the $^1B_1$–$^3B_1$ splitting in $CH_2$ at the various levels of theory considered. CAS-SCF overestimates the splitting, value $^1B_1$–$^3B_1$, which was determined to be 33.4 kcal mol$^{-1}$ by Bauschlicher [90] by extrapolating CISD + Q energies to the complete basis set limit. Noteworthy is that this overestimation is larger for a full-valence active space (41.2 kcal mol$^{-1}$), than for a minimal active space (38.7 kcal mol$^{-1}$; this result is actually the ROHF value). CAS-MP2 provides only in the case of the (2,2) space a reasonable result (35.3 kcal mol$^{-1}$ corresponding to the ROMP2 result, figure 8, table 2) whereas all other excitation energies are either too small (28.8–30.2 kcal mol$^{-1}$) or too large ((2,4) space: 46 kcal mol$^{-1}$), i.e. the ROMP2 reference value is actually better than any of the CAS-MP2 values with larger active space. One could argue that the two open-shell cases are no longer typical CAS problems. However, CAS-MP2 should compensate for any deficiencies of the active space, which it obviously fails to do.

We note that care has to be taken with the CAS-MP2 treatment of the $^1B_1$ state. A CAS-SCF treatment of the excited state is facilitated by imposing a symmetry constraint, i.e. by excluding the closed-shell singlet configurations from the active space. If this symmetry restriction is not available, one can solve the problem (as done in this work) by first performing a symmetry-restricted CAS-SCF calculation and then restarting the CAS-DFT calculation from the orbitals obtained, with any further orbital optimization suppressed.

The inclusion of DFT correlation energy $E_c$(CS) at the CAS + CS level of theory decreases all CAS-SCF splitting values by about 9 to 30–32 kcal mol$^{-1}$ below but accidently close to the reference value. Scaling does increase the excitation energies so that the full valence space and the (8,7) results are above, with the other below the reference value. Inspection of figure 8 shows that the CAS-DFT1 or CAS-DFT2 results are better than the CAS + f DFT results. CAS-DFT(SPP,FOS,DS) leads to a significant reduction of the CAS-DFT (SPP,FCS) values, which becomes especially obvious when the SPP functional is applied (table 3).

The best values are obtained with the CAS-DFT2(SPP,FOS,DS) method when using the full valence active space (or the (8,7) space). The CAS-DFT2(SPP,FCS) value of 34.5 kcal mol$^{-1}$ is reduced by 0.7 kcal mol$^{-1}$ to 33.8 kcal mol$^{-1}$, just 0.4 kcal mol$^{-1}$ above the reference value. The importance of the DS correction becomes obvious when comparing with the corresponding CAS-DFT2(SPP,FOS) value of 37.7 kcal mol$^{-1}$ (table 3). Clearly, the low-spin open-shell correction is of utmost importance as reflected by a 10% correction of the $^1B_1$–$^3B_1$ splitting energy obtained at the CAS-DFT2(SPP,FOS,DS) level of theory.

It is interesting to note that a CAS-DFT2(SPP,FCS) treatment of both open-shell states leads to a partial error compensation, which will be broken if one of the states is treated properly. Obviously, the error corrections for high-spin and low-spin cases are of comparable magnitude so that reliable relative energies will only be obtained if they are both included.

## 6. Conclusions and outlook

This work has demonstrated that CAS-DFT cannot be set up as a simple combination of CAS-SCF and a DFT correlation functional. The coupling of the WFT and the DFT method in an appropriate way via the WFT density is essential for any WFT-DFT method. Other problems are the avoidance of a double counting of correlation effects and the balanced treatment of core and active orbital correlation effects, equal-spin and opposite-spin correlation effects as well as electron correlation in low-spin open-shell systems.

Table 3. Energy differences $E(({}^1B_1) - E(({}^3B_1)$ calculated $CH_2$ at various levels of theory.[a]

| | | | | CAS + f CS | | | CAS-DFT2(CS) | |
|---|---|---|---|---|---|---|---|---|
| Active space | CAS-SCF | CAS-MP2 | CAS + CS | $f_{CS}$ ($f_{CS}$ + SPP) | $f_{OS}$[FOS + spp] ($f_{OS}$ + SPP,DS) | CAS-DFT1(CS) FCS (FCS,SPP) | FCS (FCS,SPP) | FOS[FOS,SPP] (FOS,SPP,DS) |
| (0,0) | 38.71 | 35.32 | 30.20 | 30.20 (30.20) | 30.20 [30.20] (30.20) | 30.20 (30.20) | 30.20 (30.20) | 30.20 [30.20] (30.20) |
| (2,2) | 38.71 | 35.32 | 30.20 | 32.72 (31.98) | 31.14 [36.09] (31.01) | 32.08 (31.32) | 32.24 (31.47) | 30.65 [35.58] (30.51) |
| (6,4) | 38.71 | 28.74 | 30.20 | 32.73 (31.98) | 31.14 [36.09] (31.01) | 33.00 (32.29) | 33.07 (32.36) | 31.48 [36.47] (31.39) |
| (2,4) | 38.71 | 46.02 | 30.20 | 28.93 (28.81) | 27.82 [31.89] (28.25) | 32.55 (31.84) | 32.66 (32.11) | 31.55 [35.20] (31.56) |
| (6,6) | 41.23 | 30.28 | 32.18 | 35.15 (34.45) | 34.00 [37.67] (33.79) | 35.04 (34.34) | 35.13 (34.43) | 33.98 [37.65] (33.77) |
| (8,7) | 41.23 | 29.29 | 32.17 | 35.17 (34.46) | 34.01 [37.68] (33.80) | 35.17 (34.46) | 35.17 (34.46) | 34.01 [37.68] (33.80) |

[a]Relative energies in kcal mol$^{-1}$. All calculations with the 6-311++G (2df,2pd) basis set. The (0,0) space denotes HF, MP2, HFLYP, HFCS reference energies. Active spaces: (2,2): minimal; (6,4): minimal + all occupied valence orbitals; (2,4): minimal + all virtual valence orbitals; (6,6): full valence; (8,7): full valence 1 s (C). Values in parentheses were obtained with the SPP functional (see text). CS: Colle–Salvetti functional; FCS and FOS: scalling factor *f* calculated for a closed- or open-shell situation, respectively; SPP: Stoll–Pavlidou–Preuss functional for elimination of equal-spin correlation; DS: Davidson–Staroverov correction for low-spin cases; level 1 and 2: Different Levels of distingusihing between core and active electron correlation.

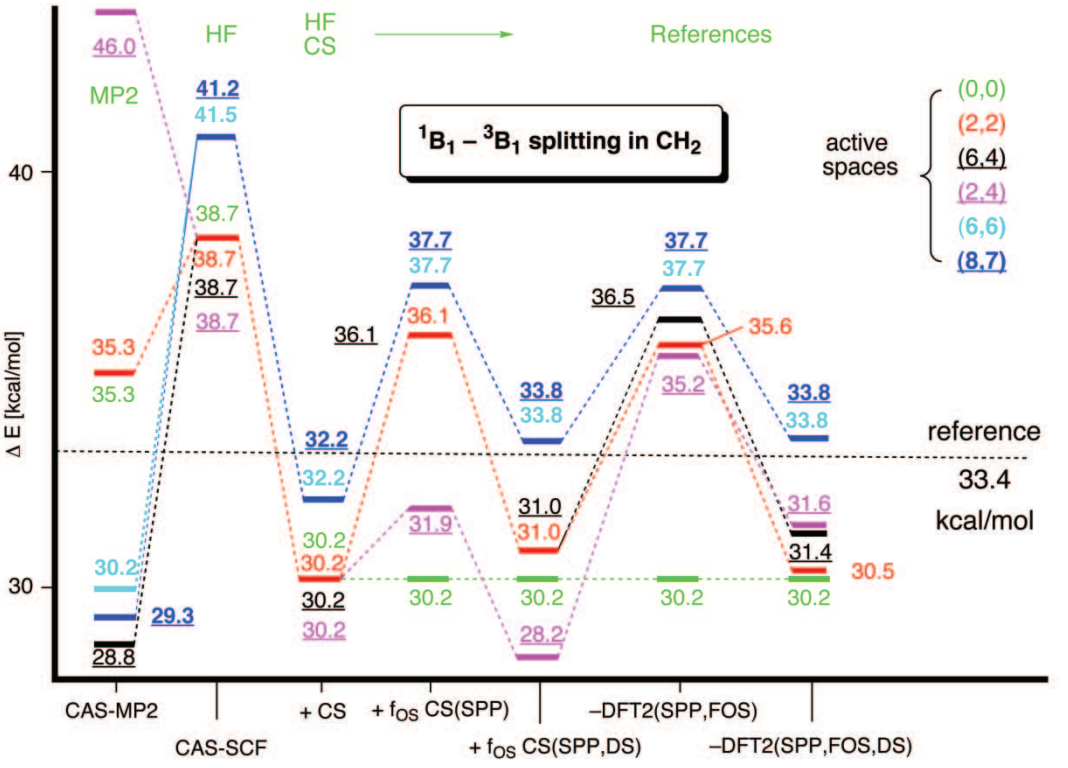


Figure 8. Changes in the excitation energy $E(^1A_1) - E(^3B_1)$ of $CH_2$ for an increasing active space calculated at different levels of theory with a 6-311 + + G(2df,2pd) basis set. CS: Colle–Salvetti functional; FCS and FOS: scaling factor $f$ calculated for a closed-shell and open-shell situation, respectively; SPP: Stoll–Pavlidou–Preuss functional for elimination of equal-spin correlation; DS: Davidson–Staroverov correction for low-spin cases; levels 1 and 2: different levels of distinguishing between core and active electron correlation.

In this work, we have solved the various problems by (a) improving the scaling by a new procedure based on a different scaling of active and inactive orbital effects, (b) correcting the scaling procedure by eliminating equal-spin correlation effects that are exaggerated when using the HEG as reference for the derivation of the scaling factor, (c) including spin polarization in the scaling procedure to distinguish between correlation effects in closed- and open-shell systems, and (d) using the Davidson–Staroverov correction to treat low-spin open-shell systems in a balanced way together with high-spin open-shell and closed-shell systems. The most advanced CAS-DFT approach obtained in this way is CAS-DFT2 (CS,SPP,FOS,DS), i.e. CAS-DFT2 carried out with the Colle–Salvetti functional using the Stoll–Pavlidou–Preuss functional for equal-spin correlation corrections, including spin-polarization in the scaling procedure, and correcting with the Davidson–Staroverov density for low-spin cases.

The new CAS-DFT methods have been tested for the three lowest states of $CH_2$ using six different active spaces of increasing sophistication. A sensitive test has been created by analysing the changes in the absolute energies for the three states (only the $^1A_1$ state is discussed here) for increasing active space. This test reflects any inconsistencies in the method used. Especially, it reveals whether in the case of a lack of non-dynamical correlation effects, the scaling procedure can reduce DFT correlation in the correct way independent of the state considered. Analysing in addition the calculated relative energies of the two excited states relative to that of the triplet ground state has provided a means to test whether certain shortcomings of the method used annihilate for relative energies.

The results summarized in this work show that CAS-DFT, if properly designed, has the following advantages.

(1) Contrary to CAS-MP2 [6] (CASPT2 [5]), CAS-DFT is an economical method that is not more expensive than CAS-SCF. Once the scaling procedure and the corrections for open shell cases have been set up, no further time investments have to be made.
(2) Contrary to first expectations, there is no longer a need to work with a minimal active space [41, 42]. Improvements in the active space are reflected in the calculated relative CAS-DFT energies and

give in this way the method sufficient flexibility to adjust appropriately to a given chemical problem.

(3) Contrary to CAS-MP2 (CASPT2), CAS-DFT is more stable with regard to variations in the active space. This is of advantage when treating electronic systems for which non-dynamical correlation effects are difficult to predict, i.e. nearly all non-trivial systems as shown even in the case of methylene [91].

(4) The coupling of the DFT correlation functional via the CAS total density and on-top density turns out to be efficient even though *a priori* a coupling via the perturbation operator spanned in the active space and projected onto the orthogonal space seemed to be better.

(5) In the case of methylene, the accuracy of CAS-DFT2 meets that of CAS-MP2 (CASPT2) or even outdoes it at much lower cost. It appears realistic to generally reach CAS-MP2 (CASPT2) quality with the CAS-DFT2 ansatz.

(6) Questions concerning size-consistency (extensivity) apply to CAS-DFT in the same way as they apply to CAS-SCF: the choice of the active space decides on the size-consistency (extensivity) of the calculation. This is a large improvement compared to our first attempts and is mainly a result of using the DS correction.

(7) The use of exact (CASSCF) exchange and the CS correlation functional keeps the CAS-DFT method free from a direct SIE (i.e. a SIE caused by the functionals used). When the scaling factor is determined based on the SPP corrected RPA correlation energies, one avoids also the indirect SIE caused by the SIE in the scaling factor. The SPP corrected CAS-DFT methods are thus SIE-free.

Nevertheless, there is room for further improvements of CAS-DFT.

As reflected by the absolute energies there is still a trend of CAS-DFT to reduce correlation effects too strongly with an increasing number of active electrons. Several possibilities exist to tackle this problem and to try to improve CAS-DFT further.

(a) For the HEG, there are excitations from shortly below the Fermi edge to shortly above the Fermi edge possible, which mimic strong correlation effects and lead to artificially decreased scaling factors. This shortcoming may be eliminated in future work by introducing a system-specific 'HOMO–LUMO' gap that helps to exclude low-lying excitations. Savin and co-workers [92–93] have published investigations that might help to introduce a gap-correction into the scaling procedure.

(b) As shown in figure 1, the CS functional also introduces some back-scaling of correlation effects, which might have to be considered when setting up the scaling procedure. The results for $CH_2$ give at hand that this double back-scaling leads to inconsistencies between different active spaces. Future investigations are needed to clarify whether any back-scaling caused by the CS functional has to be considered in the scaling procedure. In this connection, we have also to consider recent criticism and improvement suggestions for the CS functional [94–96].

(c) The scaling is done on a basis of a LDA approach even though one uses nowadays GGA functionals to which also the CS functional belongs. There is a need for adjusting the scaling factor to modern GGA functionals in a more advanced fashion than has been done in our current approach.

(d) It will also be important to replace the assumption of a non-spin-polarized core in open-shell systems by a better description.

(e) Finally, it will be necessary to technically solve the fitting problems arising when trying to express the scaling factor as a function of four rather than just three parameters.

We have used CAS-DFT2(CS,SPP,FOS,DS) in a large set of calculations to determine dissociation energies, excitation energies and reaction energies. The accuracy of the results is equivalent to those obtained with strongly correlated WFT methods. This work will be published elsewhere.

Useful discussions with Michael Filatov are acknowledged. Calculations were done on the supercomputers of the Nationellt Superdatorcentrum (NSC), Linköping, Sweden. DC thanks the NSC for a generous allotment of computer time. JG thanks the Carl Tryggers Stiftelse for financial support.

## Appendix A: Truncated random phase approximation (RPA)

In this appendix we generalize the RPA calculation for the correlation energy of the homogeneous electron gas [97] to the case of truncated RPA.

The fluctuation-dissipation theorem [98] makes it possible to express the correlation energy per particle in terms of the dielectric polarizability of the HEG, provided the latter is available for all relative

electron–electron coupling constants λ from 0 (interaction-free electron gas) to 1 (electron gas with full electron–electron interaction). Let $\alpha_\lambda(Q,\omega)$ be the polarizability for the momentum transfer $Q$, the frequency $\omega$, and the coupling constant λ. Then with the definitions

$$\overline{\alpha}_\lambda(Q,U;r_s,\zeta)=\alpha_\lambda(Q,\mathrm{i}QU;r_s,\zeta), \tag{A 1 a}$$

$$\overline{\alpha}_c(Q,U;r_s,\zeta)=\int_0^1 \mathrm{d}\lambda[\overline{\alpha}_\lambda(Q,U;r_s,\zeta)-\overline{\alpha}_0(Q,U;r_s,\zeta)], \tag{A 1 b}$$

one has

$$\underbrace{\frac{E_c}{N}}_{=\varepsilon_c(r_s,\zeta)}=-\frac{r_s^3}{12\pi}\int \mathrm{d}^3Q\, v_Q\int_0^{+\infty}\mathrm{d}U\,\overline{\alpha}_c(Q,U;r_s,\zeta). \tag{A 2}$$

Here, $v_Q=4\pi/Q^2$ is the Fourier transform of the Coulomb potential. $r_s$ is the Wigner–Seitz radius, $r_s=(\alpha k_F)^{-1}$, where $k_F$ is the Fermi wave vector and $\alpha=(4/9\pi)^{1/3}$.

At the RPA level, $\overline{\alpha}_c$ is given by

$$\overline{\alpha}_c(Q,U;r_s,\zeta)=\frac{Q^2}{4\pi}\Phi\left[\frac{\pi}{Q^3}\sum_\sigma k_{F,\sigma}^2 I_0\left(\frac{Q}{k_{F,\sigma}},\frac{U}{k_{F,\sigma}}\right)\right], \tag{A 3}$$

where $k_{F,\sigma}=z_\sigma k_F$, $z_\sigma=(1+\sigma\zeta)^{1/3}$ and

$$\Phi(x)=x-\ln(1+x)\to\begin{cases}x^2, & \text{for } x\to 0,\\ x, & \text{for } x\to+\infty,\end{cases} \tag{A 4}$$

$$I_0(q,u)=\frac{1}{\pi}\int \mathrm{d}^3p\,\Theta(p-1)\Theta(1-|\mathbf{p}-\mathbf{q}|)\frac{\mathbf{p}\widehat{\mathbf{q}}-q/2}{(\mathbf{p}\widehat{\mathbf{q}}-q/2)^2+u^2}. \tag{A 5}$$

In equation (A 5), $I_0$ is defined in terms of the dimensionless variables $q=Q/k_F$ and $u=U/k_F$, likewise, a dimensionless integration variable $p$ is used. Furthermore, $\widehat{\mathbf{q}}=\mathbf{q}/q$.

Finally,

$$\varepsilon_c(r_s,\zeta)=-\frac{3}{4\pi\alpha^2 r_s^2}\int_0^\infty \mathrm{d}q\int_0^\infty \mathrm{d}u\,q^3\times\Phi\left[\frac{\alpha r_s}{\pi}\frac{1}{Q^3}\sum_\sigma z_\sigma^2 I_0\left(\frac{q}{z_\sigma},\frac{u}{z_\sigma}\right)\right]. \tag{A 6}$$

For $\zeta=0$, this expression is equivalent to equation (19) in [93].

Truncated RPA implies that only excitations from orbitals with a momentum $k\geq k_{core}=:\eta_{core}k_F$ into orbitals with a momentum $k\leq k_{ref}=:\eta_{ref}k_F$ are considered in the dielectric polarizability. This implies that the expression for $\overline{\alpha}$ from equation (A 3) has to be replaced by the more general expression

$$\overline{\alpha}_c(Q,U;r_s,\zeta,\eta_{ref},\eta_{core})=\frac{q^2}{4\pi}\Phi\left[\frac{\pi}{Q^3}\sum_\sigma k_{F,\sigma}^2 I\left(\frac{Q}{k_{F,\sigma}},\frac{U}{k_{F,\sigma}};\frac{\eta_{ref}}{z_\sigma},\frac{\eta_{core}}{z_\sigma}\right)\right] \tag{A 7}$$

with

$$I(q,u;\eta_{ref},\eta_{core})=\frac{1}{\pi}\int \mathrm{d}^3p\,\Theta(p-1)\Theta(1-|\mathbf{p}-\mathbf{q}|)\times\Theta(\eta_{ref}-p)\Theta(|\mathbf{p}-\mathbf{q}|-\eta_{core})\times\frac{\mathbf{p}\widehat{\mathbf{q}}-q/2}{(\mathbf{p}\widehat{\mathbf{q}}-q/2)^2+u^2}. \tag{A 8}$$

This gives finally

$$\varepsilon_c(r_s,\zeta,\eta_{ref},\eta_{core})=-\frac{3}{4\pi\alpha^2 r_s^2}\int_0^{+\infty}\mathrm{d}q\int_0^{+\infty}\mathrm{d}u\times\underbrace{q^3\Phi\left\{\frac{\alpha r_s}{\pi q^3}\sum_\sigma z_\sigma^2 I\left[\frac{q}{z_\sigma},\frac{u}{z_\sigma},\frac{\eta_{ref}}{z_\sigma},\frac{\eta_{core}}{z_\sigma}\right]\right\}}_{=F(q,u;r_s,\zeta,\eta_{ref},\eta_{core})} \tag{A 9}$$

For $\zeta=0$ and $\eta_{core}=0$, this is equivalent to Savin's [39] equation (A 3).

The integral $I$ can be treated analytically. It can be represented as follows:

$$I(q,u;\eta_{ref},\eta_{core})=\overline{I}(q,u;\eta_{ref},1)-\overline{I}(q,u;\eta_{ref},\eta_{core}). \tag{A 10}$$

$$\overline{I}(q,u;\eta_{ref},\eta_{core})=\begin{cases}0 & \text{for } q\leq 1-\eta_{core};\\ I_{up}(q,u;\eta_{ref},\eta_{core})-I_{low}(q,u;\eta_{ref},\eta_{core}) & \text{for } 1-\eta_{core}<q<\eta_{ref}+\eta_{core};\\ 0 & \text{for } q\geq\eta_{ref}+\eta_{core};\end{cases} \tag{A 11 a}$$

$$I_{up}(q,u;\eta_{ref},\eta_{core})=\begin{cases}I_+(q,u;\eta_{core}) & \text{for } q\leq\eta_{ref}-\eta_{core};\\ I_0(q,u;\eta_{ref},\eta_{core}) & \text{for } q>\eta_{ref}-\eta_{core};\end{cases} \tag{A 11 b}$$

$$I_{low}(\eta_{ref},\eta_{core})=\begin{cases}I_0(q,u;1,\eta_{core}) & \text{for } q<\eta_{core}+1;\\ I_-(q,u;\eta_{core}) & \text{for } q\geq\eta_{core}+1;\end{cases} \tag{A 11 c}$$

For $\eta_{\mathrm{core}} = 1$, the expressions in equations (A 11) can be simplified to

$$\overline{I}(q,u;\eta_{\mathrm{ref}},1) = \begin{cases} I_+(q,u;1) - I_-(q,u;1) & \text{for } q \le \eta_{\mathrm{ref}} - 1; \\ I_0(q,u;\eta_{\mathrm{ref}},1) - I_-(q,u;1) & \text{for } \eta_{\mathrm{ref}} - 1 < q < \eta_{\mathrm{ref}} + 1; \\ 0 & \text{for } q < \eta_{\mathrm{ref}} + 1. \end{cases} \tag{A 12}$$

The expressions $I_+$, $I_-$ and $I_0$ introduced in equations (A 11) and (A 12) are defined as

$$\begin{aligned} I_0(q,u;p,\eta_{\mathrm{core}}) = {} & \frac{1}{2}\left(p^2 - \frac{q^2}{4} + u^2\right) \ln\left[\left(p - \frac{q}{2}\right)^2 + u^2\right] \\ & + qu \arctan\frac{p - q/2}{u} \\ & - \frac{1}{2}\left(p^2 - \eta_{\mathrm{core}}^2\right) \ln\left[\left(\frac{p^2 - \eta_{\mathrm{core}}^2}{2q}\right)^2 + u^2\right] \\ & - 2qu \arctan\frac{p^2 - \eta_{\mathrm{core}}^2}{2qu} + \frac{1}{2}p^2 - \frac{1}{2}pq, \end{aligned} \tag{A 13 a}$$

$$\begin{aligned} I_\pm(q,u;\eta_{\mathrm{core}}) = {} & \frac{1}{2}\left(\eta_{\mathrm{core}}^2 - \frac{q^2}{4} + u^2\right) \ln\left[\left(\frac{q}{2} \pm \eta_{\mathrm{core}}\right)^2 + u^2\right] \\ & - qu \arctan\frac{\eta_{\mathrm{core}} \pm q/2}{u} \\ & \pm \frac{1}{2}\eta_{\mathrm{core}} q + \frac{1}{2}\eta_{\mathrm{core}}^2. \end{aligned} \tag{A 13 b}$$

The integrations over $u$ and $q$ in equation (A 9) have to be done numerically. The $u$ integration is improper in any case, the $q$ integration is improper only for $\eta_{\mathrm{ref}} = +\infty$ and amounts to a finite integration over the interval $[0, \eta_{\mathrm{ref}} + z_{\max}]$, $z_{\max} = \max(z_\alpha, z_\beta)$ otherwise. Care needs to be taken to ensure a sufficient numerical accuracy. We use Euler integration with Romberg extrapolation for automatic step-width control. Also, a Romberg-type extrapolation with respect to the upper integration limit is used for the treatment of improper integrals, where use is made of the asymptotic behaviour

$$\begin{aligned} F(q,u;r_{\mathrm{s}},\zeta,+\infty,\eta_{\mathrm{core}}) &\propto q^{-4} \quad \text{for } q \to +\infty, \\ F(q,u;r_{\mathrm{s}},\zeta,\eta_{\mathrm{ref}},\eta_{\mathrm{core}}) &\propto u^{-4} \quad \text{for } u \to +\infty \end{aligned} \tag{A 14}$$

of the integrand. Given the efficiency of the Romberg extrapolation for the upper integration limit, the $q$ integration for large but finite $\eta_{\mathrm{ref}}$ is carried out as

$$\begin{aligned} \int_0^{\eta_{\mathrm{ref}}+z_{\max}} \mathrm{d}q\, F(q \ldots \eta_{\mathrm{ref}} \ldots) = {} & \left[\int_0^{+\infty} \mathrm{d}q - \int_{\eta_{\mathrm{ref}}-z_{\max}}^{+\infty} \mathrm{d}q\right] F(q \cdots + \infty \ldots) \\ & + \int_{\eta_{\mathrm{ref}}-z_{\max}}^{\eta_{\mathrm{ref}}+z_{\max}} \mathrm{d}q\, F(q \ldots \eta_{\mathrm{ref}} \ldots) \end{aligned} \tag{A 15}$$

(note $F(q,u;r_{\mathrm{s}},\zeta,\eta_{\mathrm{ref}},\eta_{\mathrm{core}}) = F(q,u;r_{\mathrm{s}},\zeta,+\infty,\eta_{\mathrm{core}})$ for $q \le \eta_{\mathrm{ref}} - z_{\max}$).

The expressions (A 13) for the integral $I$ may lead to small differences of large values for small or large values of the arguments. In particular, the limit $u \to \infty$, $q \ll u$ had to be taken into account explicitly:

$$\overline{I}(q,u;\eta_{\mathrm{ref}},\eta_{\mathrm{core}}) = \tilde{I}(p_{\mathrm{up}},\eta_{\mathrm{ref}},\eta_{\mathrm{core}}) - \tilde{I}(p_{\mathrm{low}},\eta_{\mathrm{ref}},\eta_{\mathrm{core}}); \tag{A 16 a}$$

$$\begin{aligned} \tilde{I}(p,\eta_{\mathrm{ref}},\eta_{\mathrm{core}}) = {} & -\frac{1}{24\,q^2}p^6 + \left(\frac{1}{4} + \frac{\eta_{\mathrm{core}}^2}{8q^2}\right)p^4 \\ & - \frac{q}{3}p^3 + \left(\frac{q^2}{8} - \frac{\eta_{\mathrm{core}}^4}{8q^2}\right)p^2, \end{aligned} \tag{A 16 b}$$

$$\to \frac{1}{2}(\eta_{\mathrm{core}}^2 - 1)q \text{ for } q \downarrow 0, \tag{A 16 c}$$

$$p_{\mathrm{up}} = \min(q+1, \eta_{\mathrm{ref}}), \tag{A 16 d}$$

$$p_{\mathrm{low}} = \max(q-1, 1). \tag{A 16 e}$$

The correlation energies per particle were calculated with a numerical accuracy of $10^{-6}$ for $r_{\mathrm{s}} < 0.1$ and $10^{-7}$ otherwise.

## Appendix B: The parametrization of the results

In the numerical calculations, the scaling factor $f$ is treated as a function of the parameter $r_{\mathrm{s}}$ and the factors $\eta_{\mathrm{ref}}$, $\eta_{\mathrm{core}}$ rather than the densities $\varrho$, $\varrho_{\mathrm{ref}}$ and $\varrho_{\mathrm{core}}$. In this appendix, we represent therefore $f$ in the form $f(r_s, \eta_{\mathrm{ref}}, \eta_{\mathrm{core}}, \zeta)$, where we use the shorthand notations $f_{\mathrm{ref}}(r_s, \eta_{\mathrm{ref}}) = f(r_s, \eta_{\mathrm{ref}}, \eta_{\mathrm{core}} = 0, \zeta = 0)$ and $f_{\mathrm{CS}}(\varrho, \eta_{\mathrm{ref}}, \eta_{\mathrm{core}}) = f(\varrho, \eta_{\mathrm{ref}}, \eta_{\mathrm{core}}, \zeta = 0)$.

The results obtained pointwise for $f(r_{\mathrm{s}}, \zeta, \eta_{\mathrm{ref}}, \eta_{\mathrm{core}}, \zeta)$ need to be fitted to an analytical form for use in CAS-DFT calculations. Miehlich and co-workers [40] use the ansatz

$$\tilde{f}_{\mathrm{ref}}(r_{\mathrm{s}}, \eta_{\mathrm{ref}}) = \left(\sum_{m=1}^{M}\sum_{n=1}^{N} b_{mn} t^{m-1} \eta_{\mathrm{ref}}^{n-1}\right)^{-1}, \tag{B 1 a}$$

$$t = \ln r_{\mathrm{s}} \tag{B 1 b}$$

(we use a tilde to mark fitted quantities in distinction from the original ones). We found that this approach has a number of limitations.

(1) It does not fulfill the boundary conditions for $f(r_{\mathrm{s}}, \eta_{\mathrm{ref}})$ automatically. For instance, the requirement $f(r_{\mathrm{s}}, 1) = 1$ leads to the constraints

$$\sum_{n=1}^{N} b_{mn} = \delta_{m,1}, \tag{B 2}$$

which have to be accounted for by $N-1$ Lagrangian multipliers.

(2) It yields deviations of up to 0.02 from the exact values for $f(r_s, \eta_{ref})$.
(3) Its accuracy cannot in practice be improved systematically by increasing $M$ and $N$. Increased $M$ and $N$ lead to an ill-conditioned equation system for the $b_{mn}$ and, eventually, may even result in an increased RMS deviation between exact and fitted values of $f$.

We therefore constructed a different model function that avoids problems (1) to (3). As a first step, we reparametrized $f_{ref}(r_s, \eta_{ref})$. The scaling factor has the properties

$$f_{ref}(r_s, +\infty) = 0, \tag{B 3 a}$$
$$f_{ref}(r_s, \eta_{ref}) \propto \eta_{ref}^{-3} \quad \text{for } \eta_{ref} \to +\infty, \tag{B 3 b}$$
$$f_{ref}(r_s, 1) = 1, \tag{B 3 c}$$
$$1 - f_{ref}(r_s, \eta_{ref}) = o(\eta_{ref} - 1) \quad \text{for } \eta_{ref} \downarrow 1. \tag{B 3 d}$$

In equation (B 3 $d$), $o(\cdots)$ is the Landau order symbol, indicating that the term on the left-hand side is of a lower order than $\eta_{ref} - 1$.

These requirements are fulfilled by the ansatz ($\kappa_{ref} = \eta_{ref}^{-1}$)

$$\tilde{f}_{ref}(r_s, \eta_{ref}) = \kappa_{ref}^3 + \sum_{m=1}^{M} \sum_{n=1}^{N_{ref}} d_{mn}^{ref} t^{m-1} K_n(\kappa_{ref}), \tag{B 4 a}$$

$$K_n(x) = \begin{cases} x^3(1-x)^{4/3} & \text{for } n = 1, \\ x^3(1-x)^{n-1} & \text{for } n > 1. \end{cases} \tag{B 4 b}$$

The power term with the exponent 4/3 has been found to reflect the behaviour close to $\eta_{ref} = 1$ efficiently.

This ansatz can be generalized straightforwardly to the case of $\eta_{core} > 0$. For the scaling factor $f_{CS}(r_s, \eta_{ref}, \eta_{core})$, we have the additional requirements

$$f_{CS}(r_s, \eta_{ref}, 0) = 0,\ f_{CS}(r_s, \eta_{ref}, \eta_{core}) \propto \eta_{core}^3 \quad \text{for } \eta_{core} \downarrow 0, \tag{B 5 a}$$

$$f_{CS}(r_s, \eta_{ref}, \eta_{core}) - f_{CS}(r_s, \eta_{ref}, 1) = o(1 - \eta_{core}) \quad \text{for } \eta_{core} \uparrow 1. \tag{B 5 b}$$

Attempts to build a model function for $f_{CS}$ in one step failed to provide the required accuracy: for small numbers of basis functions, the remaining error is too high, for larger numbers, one encounters stability problems. The parametrization was therefore done in three steps.

(1) The function $f_{core}(r_s, \eta_{core}) = f(r_s, 0, +\infty, \eta_{core})$ was fitted to a model function

$$\tilde{f}_{core}(r_s, \eta_{core}) = \eta_{core}^3 + \sum_{m=1}^{M} \sum_{n=1}^{N_{core}} d_{mn}^{core} t^{m-1} K_n(\eta_{core}). \tag{B 6}$$

(2) The model functions $\tilde{f}_{ref}(r_s, \eta_{ref})$ and $\tilde{f}_{core}(r_s, \eta_{core})$ are combined such that one gets a good approximation for $f(r_s, 0, \eta_{ref}, \eta_{core})$. The ansatz

$$\tilde{f}_1(r_s, \eta_{ref}, \eta_{core}) = \frac{\tilde{f}_{ref} + \tilde{f} - 2\tilde{f}_{ref}\tilde{f}_{core}}{1 - \tilde{f}_{ref}\tilde{f}_{core}} \tag{B 7}$$

reflects a number of features of $f_{CS}(r_s, \eta_{ref}, \eta_{core})$: (i) for $\kappa_{ref} \to 0$, it converges to $\tilde{f}_{core}(r_s, \eta_{core})$, for $\eta_{core} \to 0$, it converges to $\tilde{f}_{ref}(r_s, \eta_{ref})$; (ii) it is smaller or equal to either of $\tilde{f}_{core}(r_s, \eta_{core})$ and $\tilde{f}_{ref}(r_s, \eta_{ref})$.
(3) The difference $f_{rem} = f - \tilde{f}_1$ is fitted to a model function of the form (B 4 $b$), where we call the expansion coefficients for $d_{mnl}^{rem}$ instead of $d_{mnl}$.

Altogether, we get the model expression

$$f_{CS}(r_s, \eta_{ref}, \eta_{core}) = \tilde{f}_1(r_s, \eta_{ref}, \eta_{core}) + \sum_{m=1}^{M} \sum_{n=1}^{N_{ref}} \sum_{l=1}^{N_{core}} d_{mnl}^{rem} t^{m-1} K_n(\kappa_{ref}) K_l(\eta_{core}). \tag{B 8}$$

**The $\zeta$ dependence**

If $\zeta \neq 0$, the scaling factor is defined only for

$$\eta_{core} \leq (1-\zeta)^{1/3}, \tag{B 9 a}$$
$$\eta_{ref} \geq (1+\zeta)^{1/3}, \tag{B 9 b}$$

as the orbitals with $k < \eta_{core} k_F$ have to be doubly occupied and those with $k > \eta_{ref} k_F$ have to be empty. As a consequence, one should not use $\eta_{ref}$ and $\eta_{core}$ as independent variables in the parameter fit. A reasonable choice is

$$\tilde{\eta}_{ref} = \eta_{ref} + 1 - (1+\zeta)^{1/3}, \tag{B 10 a}$$
$$\tilde{\eta}_{core} = \eta_{core} + 1 - (1-\zeta)^{1/3}. \tag{B 10 b}$$

We start with a fit for the case $\eta_{core} = 0$. A suitable model function is ($\tilde{\kappa}_{ref} = \tilde{\eta}_{ref}^{-1}$)

$$\tilde{f}(r_s, \eta_{ref}, 0, \zeta) = \tilde{\kappa}_{ref}^3 + \sum_{m=1}^{M} \sum_{n=1}^{N_{ref}} \sum_{l=1}^{N_\zeta} b_{mnl} x^{m-1} K_{nl}(\tilde{\kappa}_{ref}, \zeta), \tag{B 11 a}$$

$$K_{1,1}(\kappa, \zeta) = \kappa^3(1-\kappa)^{4/3}, \tag{B 11 b}$$
$$K_{2,1}(\kappa, \zeta) = \kappa^3(1-\kappa), \tag{B 11 c}$$
$$K_{2,2}(\kappa, \zeta) = \kappa^3\zeta, \tag{B 11 d}$$
$$K_{n+1,l}(\kappa, \zeta) = (1-\kappa)K_{n,l}(\kappa, \zeta) \quad \text{for } n \geq 2, \tag{B 11 e}$$
$$K_{n,l+1}(\kappa, \zeta) = \zeta K_{n,l}(\kappa, \zeta) \quad \text{for } n = 1 \text{ or } l \geq 2. \tag{B 11 f}$$

The choice for $K_{2,2}(\kappa, \zeta)$ ensures that the scaling factor no longer is kept fixed to 1 for $\tilde{\eta}_{\text{ref}} = 1$ as soon as $\zeta \neq 0$.

The parametrization of the results was done with a Fortran 77 program using the LAPACK subroutine library [99].

So far, no practicable generalization of equation (B 11) to the case of $\eta_{\text{core}} > 0$ has been found. The quasi-linear regression becomes instable (i.e. the resulting equation systems become ill-conditioned) for four independent variables and the corresponding number of expansion coefficients. A nonlinear ansatz will probably be necessary in this case.

## References


[1] P. Hohenberg, W. Kohn. *Phys. Rev.*, **136,** B864 (1964).
[2] W. Kohn, L.J. Sham. *Phys. Rev.*, **140,** A1133 (1965).
[3] R.G. Parr, W. Yang. *International Series of Monographs on Chemistry 16: Density-Functional Theory of Atoms and Molecules*, New York: Oxford University Press (1989).
[4] N.H. March. *Electron Density Theory of Atoms and Molecules*, New York: Academic Press (1992); W. Koch, M.C. Holthausen. *A Chemist's Guide to Density Functional Theory*, New York: Wiley (2000); D.P. Chong, (Ed.) *Recent Advances in Computational Chemistry*, Vol. 1, *Recent Advances in Density Functional Methods, Part II*, Singapore: World Scientific (1995); J.M. Seminario, P. Politzer (Eds) *Theoretical and Computational Chemistry*, Vol. 2, *Modern Density Functional Theory—A Tool for Chemistry*, Amsterdam: Elsevier (1996).
[5] K. Andersson, B.O. Roos. *Modern Electron Structure Theory, Advanced Series in Physical Chemistry*, Vol. 2, Part I, D.R. Yarkony (Ed.), p. 55, Singapore: World Scientific (1995).
[6] J.J.W. McDouall, K. Peasley, M.A. Robb. *Chem. Phys. Lett.*, **148,** 183 (1988).
[7] H.-J. Werner. Ab Initio *Methods in Quantum Chemistry, Part II*, *Advances in Chemical Physics*, Vol. 69, K.P. Lawley (Ed.), p. 1, Chichester: Wiley-Interscience (1988).
[8] P.G. Szalay. *Modern Ideas in Coupled-Cluster Methods*, R.J. Bartlett (Ed.), p. 81, Singapore: World Scientific (1997).
[9] J. Slater, J.B. Mann, T.M. Wilson, J.H. Wood. *Phys. Rev.*, **184,** 672 (1969).
[10] B.I. Dunlap. *Adv. Chem. Phys.*, **69,** 287 (1987).
[11] S.G. Wang, W.H.E. Schwarz. *J. Chem. Phys.*, **105,** 464 (1996).
[12] J.D. Goddard, G. Orlova. *J. Chem. Phys.*, **111,** 7705 (1999).
[13] M. Levy. *Phys. Rev.* A, **26,** 1200 (1982).
[14] E.H. Lieb. *Int. J. Quantum Chem.*, **24,** 243 (1983).
[15] J.T. Chayes, L. Chayes, M.B. Ruskai. *J. Stat. Phys.*, **38,** 497 (1985).
[16] H. Englisch, R. Englisch. *Phys. Status Solidi* (b), **123,** 711 (1984); *Phys. Status Solidi* (b), **124,** 373 (1984).
[17] E.K.U. Gross, L.N. Oliveira, W. Kohn. *Phys. Rev.* A, **37,** 2805 (1988).
[18] P.R.T. Schipper, O.V. Gritsenko, E.J. Baerends. *Theor. Chem. Acc.*, **99,** 329 (1998); *J. Chem. Phys.*, **111,** 4056 (1999).
[19] C.A. Ullrich, W. Kohn. *Phys. Rev. Lett.*, **87,** 093001 (2001).
[20] A. Savin, F. Colonna, R. Pollet. *Int. J. Quantum Chem.*, **93,** 166 (2003).
[21] R.C. Morrison. *J. Chem. Phys.*, **117,** 10506 (2002).
[22] M. Filatov, S. Shaik. *Chem. Phys. Lett.*, **304,** 429 (1999); *Chem. Phys. Lett.*, **332,** 409 (2000).
[23] M. Filatov, S. Shaik. *J. Phys. Chem.* A, **103,** 8885 (1999); M. Filatov, S. Shaik, M. Woeller, S. Grimme, S. Peyerimhoff. *Chem. Phys. Lett.*, **316,** 135 (2000); M. Filatov, S. Shaik. *J. Phys. Chem.* A, **104,** 6628 (2000).
[24] O.V. Gritsenko, P.R.T. Schipper, E.J. Baerends. *Chem. Phys.*, **107,** 5007 (1997); P.R.T. Schipper, O.V. Gritsenko, E.J. Baerends. *J. Chem. Phys.*, **111,** 4056 (1999).
[25] A. Savin, H.-J. Flad. *Int. J. Quantum Chem.*, **56,** 327 (1995).
[26] A. Savin. *Recent Advances in Density Functional Methods*, D.P. Chong. (Ed.), p. 129, Singapore: World Scientific (1995).
[27] A. Savin. *Recent Developments and Applications of Modern Density Functional Theory*, J.M. Seminario (Ed.), p. 327, Amsterdam: Elsevier (1996).
[28] T. Leininger, H. Stoll, H.-J. Werner, A. Savin. *J. Chem. Phys. Lett.*, **275,** 151 (1997).
[29] R. Pollet, A. Savin, T. Leininger, H. Stoll. *J. Chem. Phys.*, **116,** 1250 (2002).
[30] V.A. Rassolov. *J. Chem. Phys.*, **110,** 3672 (1999).
[31] I. Panas. *Chem. Phys. Lett.*, **245,** 171 (1995); *Molec. Phys.*, **89,** 239 (1996); I. Panas, A. Snis. *Theor. Chem. Acc.*, **97,** 232 (1997).
[32] R. Colle, O. Salvetti. *Theor. Chim. Acta*, **53,** 55 (1979).
[33] R. Colle, O. Salvetti. *Theor. Chim. Acta*, **37,** 329 (1975).
[34] G.C. Lie, E. Clementi. *J. Chem. Phys.*, **60,** 1275 (1974); *J. Chem. Phys.*, **60,** 1288 (1974).
[35] F. Moscardó, E. San-Fabián. *Phys. Rev.* A, **44,** 1549 (1991).
[36] E. Kraka. *Chem. Phys.*, **161,** 149 (1992).
[37] E. Kraka, D. Cremer, S. Nordholm. *Molecules in Natural Science and Biomedicine*, Z.B. Maksic, M. Eckert-Maksic (Eds), p. 351, Chichester: Ellis Horwood (1992).
[38] N.O.J. Malcolm, J.J.W. McDouall. *Chem. Phys. Lett.*, **282,** 122 (1998); *J. Phys. Chem.*, **100,** 10131 (1996); J.J.W. McDouall. *Molec. Phys.*, **101,** 361 (2003).
[39] A. Savin. *Int. J. Quantum Chem.* S, **22,** 59 (1988).
[40] B. Miehlich, A. Savin, H. Stoll. *Molec. Phys.*, **91,** 527 (1997).
[41] J. Gräfenstein, D. Cremer. *Chem. Phys. Lett.*, **316,** 569 (2000).
[42] J. Gräfenstein, D. Cremer. *Phys. Chem. Chem. Phys.*, **2,** 2091 (2000).
[43] S. Grimme, M. Waletzke. *J. Chem. Phys.*, **111,** 5645 (1999).
[44] R. Takeda, S. Yamanaka, K. Yamaguchi. *Chem. Phys. Lett.*, **366,** 321 (2002); *Int. J. Quantum Chem.*, **93,** 317 (2003).
[45] S. Gusarov, P.Å. Malmqvist, R. Lindh, B.O. Roos. *Theor. Chem. Acc.*, **112,** 84 (2004).
[46] J. Cioslowski. *Phys. Rev.* A, **43,** 1223 (1991).
[47] E. Valderrama, E.V. Ludena, J. Hinze. *J. Chem. Phys.*, **106,** 9227 (1997).
[48] D.K. Mok, R. Neumann, N. Handy. *J. Phys. Chem.*, **100,** 6225 (1996).
[49] C.H. Choi, M. Kertesz, A. Karpfen. *J. Chem. Phys.*, **107,** 6712 (1997).

[50] E. Valderrama, J.M. Mercero, J.M. Ugalde. *J. Phys.* B, **34,** 275 (2001); E. Valderrama, J.M. Ugalde. *Int. J. Quantum Chem.*, **86,** 40 (2002).
[51] J. Gräfenstein, D. Cremer, E. Kraka. *Chem. Phys. Lett.*, **288,** 593 (1998).
[52] F.W. Bobrowicz, W.A. Goddard III. *Methods of Electronic Structure Theory, Modern Theoretical Chemistry*, Vol. 3, H.F. Schaefer III (Ed.), p. 79, New York: Plenum Press (1977).
[53] A.D. Becke. *Advanced Series in Physical Chemistry*, Vol. 2, *Modern Electronic Structure Theory*, Part II, D.R. Yarkony (Ed.), p. 1022, Singapore: World Scientific (1995); *J. Chem. Phys.*, **104,** 1040 (1996).
[54] P.R.T. Schipper, O.V. Gritsenko, E.J. Baerends. *Phys. Rev.* A, **57,** 1729 (1998).
[55] O.V. Gritsenko, B. Ensing, P.R.T. Schipper, E.J. Baerends. *J. Phys. Chem.* A, **104,** 8558 (2000).
[56] Y. He, J. Gräfenstein, E. Kraka, D. Cremer. *Molec. Phys.*, **98,** 1639 (2000).
[57] N.C. Handy, A.J. Cohen. *Molec. Phys.*, **99,** 403 (2001).
[58] V. Polo, E. Kraka, D. Cremer. *Molec. Phys.*, **100,** 1771 (2002).
[59] V. Polo, E. Kraka, D. Cremer. *Theor. Chem. Acc.*, **107,** 291 (2002).
[60] D. Cremer. *Molec. Phys.*, **99,** 1899 (2001).
[61] V. Polo, J. Gräfenstein, E. Kraka, D. Cremer. *Chem. Phys. Lett.*, **352,** 469 (2002).
[62] V. Polo, J. Gräfenstein, E. Kraka, D. Cremer. *Theor. Chem. Acc.*, **109,** 22 (2003).
[63] D. Cremer, M. Filatov, V. Polo, E. Kraka, S. Shaik. *Int. J. Mol. Science*, **3,** 604 (2002).
[64] J. Gräfenstein, E. Kraka, M. Filatov, D. Cremer. *Int. J. Mol. Science*, **3,** 360 (2002).
[65] J. Gräfenstein, A.M. Hjerpe, E. Kraka, D. Cremer. *J. Phys. Chem.* A, **104,** 1748 (2000).
[66] For related work, see T. Ziegler, A. Rauk, E.J. Baerends. *Theor. Chim. Acta*, **43,** 261 (1997); L. Noodleman, D.A. Case. *Adv. Inorg. Chem.*, **38,** 423 (1992); T. Lovell, J.E. McGrady, R. Stranger, S. McGregor. *Inorg. Chem.*, **35,** 3079 (1996); L. Noodleman, D. Post, E.J. Baerends. *Chem. Phys.*, **64,** 159 (1982); C.J. Cramer, F.J. Dulles, D.J. Giesen, J. Almlöf. *Chem. Phys. Lett.*, **245,** 165 (1995); R. Caballol, O. Castell, F. Illas, P.R. Moreira, J.P. Malrieu. *J. Phys. Chem.* A, **101,** 7860 (1997); F. Illas, R.L. Martin. *J. Chem. Phys.*, **108,** 2519 (1998); F. Illas, I.D.R. Moreira, C. de Graaf, V. Barone. *Theor. Chem. Acc.*, **104,** 265 (2000).
[67] H. Fukutome. *Int. J. Quantum Chem.*, **20,** 955 (1981); K. Yamaguchi. *J. Molec. Struct.*, **103,** 101 (1983).
[68] I. Frank, J. Hutter, D. Marx, M. Parrinello. *J. Chem. Phys.*, **108,** 4060 (1998).
[69] M. Filatov, S. Shaik. *Chem. Phys. Lett.*, **288,** 689 (1998); *J. Chem. Phys.*, **110,** 116 (1999).
[70] E.R. Davidson. *Chem. Phys. Lett.*, **21,** 565 (1973).
[71] J.S. Andrews, C.W. Murray, N.C. Handy. *Chem. Phys. Lett.*, **201,** 458 (1993).
[72] P. Borowski, K.D. Jordan, J. Nichols, P. Nachtigall. *Theoret. Chem. Acc.*, **99,** 135 (1998).
[73] Y.G. Khait, M. Hoffmann. *J. Chem. Phys.*, **120,** 5005 (2004).
[74] P. Gombas. *Pseudopotentials*, New York: Springer (1967); *Die statistische Theorie des Atoms und ihre Anwendungen*, Wien: Springer-Verlag (1967).
[75] H. Stoll. *Chem. Phys. Lett.*, **376,** 141 (2003).
[76] M.J. Frisch, J.A. Pople, J.S. Binkley. *J. Chem. Phys.*, **80,** 3265 (1984); P.C. Hariharan, J.A. Pople. *Theoret. Chim. Acta*, **28,** 213 (1973).
[77] E. Kraka, J. Gräfenstein, M. Filatov, Y. He, J. Gauss, A. Wu, V. Polo, L. Olsson, Z. Konkoli, Z. He, D. Cremer. *COLOGNE 2004*, Göteborg: Göteborg University (2004).
[78] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery Jr, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople. *Gaussian 98*, Revision A.5, Pittsburgh PA: Gaussian, Inc. (1998).
[79] K. Ruedenberg, L.M. Cheung, S.T. Elbert. *Int. J. Quantum Chem.*, **16,** 1069 (1997).
[80] B.O. Roos. *Int. J. Quantum Chem. Symp.*, **14,** 175 (1980); Ab Initio *Methods in Quantum Chemistry*, Part II, *Advances in Chemical Physics*, Vol. 69, K.P. Lawley (Ed.), p. 399, Chichester: Wiley-Interscience (1980).
[81] D. Cremer, Z. He. *J. Phys. Chem.*, **100,** 6173 (1996).
[82] J.P. Perdew, A. Savin, K. Burke. *Phys. Rev.* A, **51,** 4531 (1995).
[83] C. Lee, W. Yang, R.P. Parr. *Phys. Rev.* B, **37,** 785 (1988).
[84] H. Stoll, C. Pavlidou, H. Preuss. *Theor. Chim. Acta*, **49,** 143 (1978).
[85] J.F. Dobson. *J. Phys.: Condens. Matter*, **4,** 7877 (1992).
[86] A.R.W. McKellar, P.R. Bunker, T.J. Sears, K.M. Evenson, R.J. Saykally, S.R. Langhoff. *J. Chem. Phys.* **79,** 5251 (1983).
[87] V.N. Staroverov, E.R. Davidson. *Chem. Phys. Lett.*, **330,** 161 (2000); *Chem. Phys. Lett.*, **340,** 142 (2001).
[88] A.E. Clark, E.R. Davidson, J.M. Zaleski. *J. Am. Chem. Soc.*, **123,** 2650 (2001).
[89] E.R. Davidson, V.N. Staroverov. *J. Am. Chem. Soc.*, **122,** 186.
[90] C.W. Bauschlicher. *Chem. Phys. Lett.*, **74,** 273 (1980).
[91] K. Andersson, B. Roos. *Int. J. Quantum Chem.* **45**, 591 (1993), find a questionable CAS-PT2/ANO value of $13.4\,\mathrm{kcal\,mol^{-1}}$ for the excitation energy $E(^1A_1) - E(^3B_1)$ of $CH_2$. In view of the CAS-SCF value they obtain ($9.9\,\mathrm{kcal\,mol^{-1}}$) and our CAS-MP2 values, their result has to be reconsidered; K. Nakayama, K. Hirao, R. Lindh. *Chem. Phys. Lett.*, **300,** 303 (1999), suggest to make CAS-MP2 (CASPT2) more stable by improving the virtual orbitals.
[92] C. Gutle, A. Savin, J.B. Krieger, J.O. Chen. *Int. J. Quantum Chem.*, **75,** 885 (1999).
[93] J. Rey, A. Savin. *Int. J. Quantum Chem.*, **69,** 581 (1998).
[94] J.C. Sancho-García, F. Moscardó. *J. Chem. Phys.*, **118,** 1054 (2002).
[95] R. Singh, L. Lassa, V. Sahni. *Phys. Rev.* A, **60,** 4135 (1999); S. Caratzoulas, P.J. Knowles. *Molec. Phys.*, **98,** 1811 (2000); L. Pastor-Abia, A. Pérez-Jiméńez,

J.M. Peŕez-Jordá, J.C. Sancho-García, E. San-Fabián, F. Moscardó. *Theor. Chem. Acc.*, **111,** 1 (2004); L.P. Abia, J.M. Pérez-Jordá, E. San-Fabián. *J. Molec. Struct. (Theochem)*, **528,** 59 (2000); E. San-Fabián, L. Pastor-Abia. *Int. J. Quantum Chem.*, **91,** 451 (2003).

[96] F. Moscardó, A.J. Pérez-Jiménez. *Int. J. Quantum Chem.*, **61,** 313 (1997); T. Tsuneda, K. Hirao. *Chem. Phys. Lett.*, **268,** 510 (1997); T. Tsuneda, T. Suzumura, K. Hirao. *J. Chem. Phys.*, **110,** 10664 (1999); S. Caratzoulas. *Phys. Rev.* A, **63,** 062506 (2001); Y. Imamura, G.E. Scuseria, R.M. Martin. *J. Chem. Phys.*, **116,** 6458 (2002).

[97] M. Gell-Mann, K.A. Brueckner. *Phys. Rev.*, **106,** 364 (1957).

[98] H.B. Callen, T.A. Welton. *Phys. Rev.*, **83,** 34 (1951).

[99] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney and D. Sorensen. *LAPACK Users' Guide*, 3rd Edn, Society for Industrial and Applied Mathematics, Philadelphia, PA (1999).