Explicitly correlated local second-order perturbation theory with a frozen geminal correlation factor

Frederick R. Manby, Hans-Joachim Werner, Thomas B. Adler, and Andrew J. May

Additional information on J. Chem. Phys.
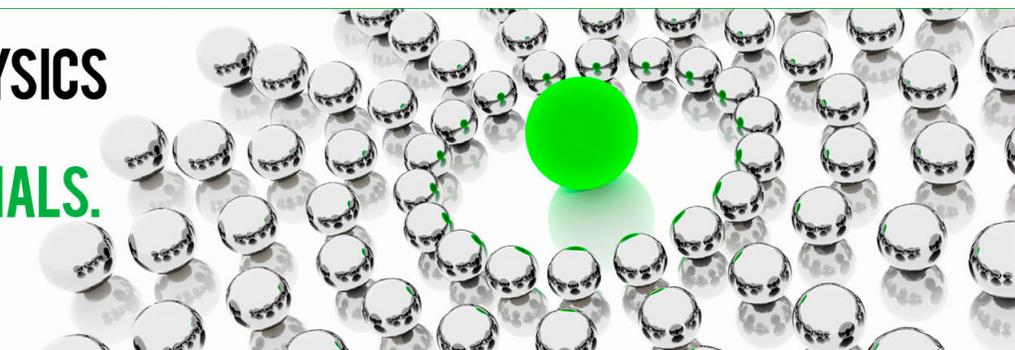
Explicitly correlated local second-order perturbation theory with a frozen geminal correlation factor

Frederick R. Manby[a),b)]
School of Chemistry, University of Bristol, Cantocks Close, Bristol BS8 1TS, United Kingdom

Hans-Joachim Werner[a),c)] and Thomas B. Adler
Institut für Theoretische Chemie, Universität Stuttgart, Pfaffenwaldring 55, D-70569 Stuttgart, Germany

Andrew J. May[d)]
School of Chemistry, University of Bristol, Cantocks Close, Bristol BS8 1TS, United Kingdom

The recently introduced MP2-R12/2^*A(loc) and LMP2-R12/2^*A(loc) methods are modified to use a short-range correlation factor expanded as a fixed linear combination of Gaussian geminals. Density fitting is used to reduce the effort for integral evaluation, and local approximations are introduced to improve the scaling of the computational resources with molecular size. The MP2-F12/2^*A(loc) correlation energies converge very rapidly with respect to the atomic orbital basis set size. Already with the aug-cc-pVTZ basis the correlation energies computed for a set of 21 small molecules are found to be within 0.5% of the MP2 basis set limit. Furthermore the short-range correlation factor leads to an improved convergence of the resolution of the identity, and eliminates problems with long-range errors in density fitting caused by the linear r_{12} factor. The DF-LMP2-F12/2^*A(loc) method is applied to compute second-order correlation energies for molecules with up to 49 atoms and more than 1600 basis functions. © *2006 American Institute of Physics.*
[DOI: 10.1063/1.2173247]

I. INTRODUCTION

Three major problems in *ab initio* quantum chemistry are steep scaling of the computational cost with system size, slow convergence of the correlation energy with the basis set size, and high cost if large basis sets are used. In our efforts to overcome these problems we have recently introduced a method[1] (hereafter **I**) which uses local approximations[2–4] to reduce the scaling with system size, density fitting to reduce the cost for integral evaluation,[5–9] and a linear r_{12} correlation factor as proposed by Klopper and Samson and Kutzelnigg and Klopper[10,11] to improve the convergence of the energy with basis set size. It was demonstrated that this method, denoted by DF-LMP2-R12/2^*A(loc), is applicable to molecules with 50 atoms or more, and yields rather accurate correlation energies at moderate cost. However, several problems were identified when applying this method to larger molecules. First, the basis set convergence was still not quite satisfactory: Using atomic orbital (AO) basis sets of triple-zeta quality, correlation energies of just quadruple-zeta quality were obtained. Secondly, the convergence of the results with respect to the resolution of the identity (RI) basis was poor, and even with uncontracted quintuple-zeta basis sets convergence was not reached. And third, it was shown that the linear increase of the r_{12} correlation factor with the interelectronic distance leads to significant errors in the density

fitting procedure, which do not systematically reduce with increasing distance or increasing basis sets.

Recently, Ten-no[12] proposed to replace the linear correlation factor r_{12} by a short-range correlation factor of the form $f(r_{12}) = \exp(-\beta r_{12})$. May and Manby and May *et al.*[9,13] expanded this function in a linear combination of Gaussians, and showed how to compute the resulting integrals using efficient density fitting techniques. Very recently, Tew and Klopper[14] applied a similar method without density fitting and also tested alternative short-range correlation factors. All these authors found that the MP2-F12 method yields greatly improved results as compared with the MP2-R12 method of Kutzelnigg and Klopper, and a recent, thorough analysis showed that the largest source of error in MP2-R12 theory arose from the poor choice of the correlation factor.[13] This is attributed to the fact that the linear increase of the r_{12} correlation factor at long range is unphysical.

Like MP2-R12, the MP2-F12 method includes terms for which the computational effort scales steeply with molecular size. For instance, if the *Ansatz* **2**A′ is used,[10] the computational effort scales as $\mathcal{O}(\mathcal{N}^8)$, where $\mathcal{N}$ is a measure of the system size. In small molecules these terms have a small prefactor, but they become quickly dominant in large molecules. This high scaling was eliminated in our local variant, namely, the DF-LMP2-R12/2^*A(loc) method mentioned above. In the present work we generalize this method to use f_{12} correlation factors and investigate how this affects the accuracy and performance of the method.

In Sec. II the method will be briefly outlined. In Sec. III A we will test the accuracy of the method for a set of small molecules and chemical reactions.

[a)]Authors to whom correspondence should be addressed.

[b)]fred.manby@bris.ac.uk

[c)]Electronic mail: werner@theochem.uni-stuttgart.de

[d)]Present address: School of Chemistry, Cardiff University, Park Place, Cardiff, CF10 3AT, UK.

124, 094103-1

II. THEORY

A. The DF-MP2-F12/2*A(loc) method

The theory of the DF-MP2-R12/2*A(loc) method has been outlined in **I**. Here we will just summarize the basic working equations and point out the differences which arise when the linear r_{12} correlation factor is replaced by a general correlation factor $f_{12} \equiv f(r_{12})$. In the following $i, j, \ldots$ will denote occupied orbitals, $a, b, \ldots$ virtual orbitals, and $r, s, \ldots$ any orbitals. Summation over repeated dummy indices will be implied throughout the paper.

In the MP2-R12 and MP2-F12 methods the pair functions are written as

$$|u_{ij}\rangle = T_{ab}^{ij}|ab\rangle + \hat{Q}_{12} t_{kl}^{ij} f_{12}|kl\rangle, \tag{1}$$

where $T_{ab}^{ij} = T_{ba}^{ji}$ and $t_{kl}^{ij} = t_{lk}^{ji}$ are the amplitudes and $|rs\rangle$ denotes an orbital product $\phi_r(\mathbf{r}_1)\phi_s(\mathbf{r}_2)$. The first term in Eq. (1) is the conventional expansion in products of virtual orbitals, and the second term adds explicit correlation by introducing the correlation factor f_{12}. The summation over occupied orbitals kl ensures orbital invariance.[15,16] The projection operator $\hat{Q}_{12}$ is necessary to keep the first-order wave function orthogonal to the Hartree-Fock reference function. This operator can be chosen in various ways,[10] leading to different *Ansätze*. In the current work $\hat{Q}_{12}$ is defined as

$$\hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2)(1 - \hat{V}_1\hat{V}_2), \tag{2}$$

where $\hat{O}_i = |m(i)\rangle\langle m(i)|$ and $\hat{V}_i = |a(i)\rangle\langle a(i)|$ are one-electron projectors onto the occupied and virtual orbital spaces, respectively.

In the LMP2 method the occupied orbitals are localized, and instead of the virtual orbitals a and b nonorthogonal projected atomic orbitals (PAOs) are used. The summation in the first term of Eq. (1) can then be restricted to subspaces (domains) of PAOs. Furthermore, in the current implementation of the LMP2-R12/2*A(loc) and LMP2-F12/2*A(loc) methods the summation over kl in the second term is restricted to include only $kl = ij$ and $kl = ji$ (*diagonal* approximation). Furthermore, we will assume the generalized and extended Brillouin conditions (GBC and EBC, respectively) to be fulfilled, leading to the *Ansatz* 2*A(loc). These simplifications cause small errors,[13] but these do not affect the conclusions of the current work.

In this approximation, the contributions of the conventional and explicit terms in the first-order wave functions are uncoupled and can be determined independently. In the following, we will only consider the f_{12} corrections. As in **I** these will be added to the canonical, i.e., nonlocal, MP2 correlation energies, in order to clearly distinguish the errors caused by the domain approximation from those in the f_{12} treatment. This method is denoted MP2-F12/2*A(loc).

The MP2-F12/2*A(loc) energy correction can be written as

$$E_{\text{F12}} = -\sum_{p=\pm 1} (2-p) \frac{[V_{ij}^{ijp}]^2}{B_{ij,ij}^{(p)}}, \tag{3}$$

with

$$V_{ij}^{ijp} = \frac{1}{2}(V_{ij}^{ij} + pV_{ji}^{ij}), \tag{4}$$

$$B_{ij,ij}^{(p)} = \frac{1}{2}(B_{ij,ij} + pB_{ij,ji}). \tag{5}$$

The general form of the quantities $B_{kl,mn}$ and V_{kl}^{ij} are

$$V_{kl}^{ij} = \langle ij|r_{12}^{-1}\hat{Q}_{12}f_{12}|kl\rangle, \tag{6}$$

$$B_{kl,mn} = (A_{kl,mn} + A_{mn,kl})/2, \tag{7}$$

where

$$A_{kl,mn} = \langle kl|f_{12}\hat{Q}_{12}[\hat{t}_1 + \hat{t}_2, f_{12}]|mn\rangle. \tag{8}$$

These involve three-electron integrals, which are approximated using the RI approximation.[11] This leads to

$$V_{kl}^{ij} = (FK)_{kl}^{ij} - F_{rs}^{ij}K_{rs}^{kl} - F_{ir'}^{ij}K_{ir'}^{kl} - F_{r'i}^{ij}K_{r'i}^{kl} + F_{ir}^{ij}K_{ir}^{kl} + F_{ri}^{ij}K_{ri}^{kl}, \tag{9}$$

$$A_{kl,mn} = (FU)_{mn}^{kl} - F_{rs}^{kl}U_{rs}^{mn} - F_{ir'}^{kl}U_{ir'}^{mn} - F_{r'i}^{kl}U_{r'i}^{mn} + F_{ir}^{kl}U_{ir}^{mn} + F_{ri}^{kl}U_{ri}^{mn}, \tag{10}$$

where the indices r' denote functions in the (orthonormalized) RI basis. The two-electron integrals are defined as

$$K_{pq}^{kl} = \langle kl|r_{12}^{-1}|pq\rangle, \tag{11}$$

$$F_{pq}^{kl} = \langle kl|f_{12}|pq\rangle, \tag{12}$$

$$U_{pq}^{kl} = \langle kl|[\hat{t}_1 + \hat{t}_2, f_{12}]|pq\rangle, \tag{13}$$

$$(FK)_{mn}^{ij} = \langle ij|f_{12}r_{12}^{-1}|mn\rangle, \tag{14}$$

$$(FU)_{mn}^{ij} = \langle ij|f_{12}[\hat{t}_1 + \hat{t}_2, f_{12}]|mn\rangle. \tag{15}$$

These equations reduce to those for MP2-R12/2*A(loc) by replacing f_{12} with r_{12}. In this case the last two integral types simplify to $(FK)_{mn}^{ij} = \delta_{im}\delta_{jn}$ and $(FU)_{mn}^{ij} = \delta_{im}\delta_{jn}$.

The four-index integrals are computed using *robust* fitting approximations,[5,9,17] e.g.,

$$F_{pq}^{kl} = \bar{J}_A^{pk}F_A^{ql} + F_B^{pk}\bar{J}_B^{ql} - \bar{J}_A^{pk}F_{AB}\bar{J}_B^{ql}, \tag{16}$$

$$U_{pq}^{kl} = \bar{Y}_A^{pk}F_A^{ql} + X_A^{pk}\bar{J}_A^{ql} + F_A^{pk}\bar{Y}_A^{ql} + \bar{J}_A^{pk}X_A^{ql} - \bar{Y}_A^{pk}F_{AB}\bar{J}_B^{ql} - \bar{J}_A^{pk}F_{AB}\bar{Y}_B^{ql}, \tag{17}$$

where A and B denote fitting basis functions and the two-index and three-index integrals are defined as

$$J_{AB} = (A|r_{12}^{-1}|B), \tag{18}$$

$$F_{AB} = (A|f_{12}|B), \tag{19}$$

$$J_A^{pk} = (A|r_{12}^{-1}|pk), \tag{20}$$

$$F_A^{pk} = (A|f_{12}|pk), \tag{21}$$

$$X_A^{pk} = (A|[\hat{t}_1, f_{12}]|pk), \qquad (22)$$

$$Y_A^{pk} = (A|[\hat{t}_1, r_{12}^{-1}]|pk), \qquad (23)$$

$$\bar{J}_A^{pk} = (J^{-1})_{AB} J_B^{pk}, \qquad (24)$$

$$\bar{Y}_A^{pk} = (J^{-1})_{AB} Y_B^{pk}. \qquad (25)$$

The evaluation of the integrals $(FK)_{mn}^{ij}$ and $(FU)_{mn}^{ij}$ is analogous to F_{pq}^{kl}. Analytical expressions for the two- and three-index integrals have been given in a previous work.[9]

Previous tests of various correlation factors have shown that very accurate results are obtained with a simple Slater function $e^{-\beta r}$. Even though the integrals involving this functions can in principle be evaluated analytically,[12] here we use a simpler approximation in which the Slater function is expanded in a linear combination of Gaussians,[9]

$$e^{-\beta r} \approx \sum_i^n c_i e^{-\alpha_i r^2}. \qquad (26)$$

In this work, we chose $\beta = 1.4 a_0^{-1}$ and fitted the Slater function in a basis of seven Gaussians. The exponents were taken to be an even tempered series with center $1 a_0^{-2}$ and a ratio of 3.0.

The method was implemented as part of the MOLPRO package of *ab initio* programs.[18]

III. RESULTS AND DISCUSSION

A. Reaction energies

In order to investigate the performance of the MP2-F12/$\mathbf{2}^*$A(loc) method and to compare with the MP2-R12/$\mathbf{2}^*$A(loc) results we consider the same set of 21 molecules and 16 reactions as in **I**. The geometries were optimized at the MP2 level using the aug-cc-pVTZ basis set. For simplicity, we will in the following denote the cc-pVnZ and aug-cc-pVnZ bases[19,20] simply as VnZ and AVnZ, respectively.

As before, results will be compared for the AVTZ, AVQZ, and AV5Z ($n=3-5$) basis sets. In combination with each AO basis, various density fitting (DF) and RI bases were tested. For the density fitting, MP2 fitting basis sets of Weigend *et al.* and Hättig were used.[21,22] Density fitting was also employed in the Hartree-Fock calculations,[23] using the V5Z JK fitting basis set of Weigend.[24] The errors introduced by the DF approximations are negligible and not further considered in this paper. In the following, the MP2 and JK fitting bases will be denoted as (A)VnZ/MP2FIT and VnZ/JKFIT, respectively. The Hartree-Fock JK fitting bases have been used as RI basis sets as well (see below).

As pointed out before, the local F12/$\mathbf{2}^*$A(loc) correction is entirely decoupled from the conventional amplitudes and unaffected by the LMP2 approximations. In order to eliminate any errors arising from the domain approximation in the LMP2, the F12/$\mathbf{2}^*$A(loc) corrections are added to the canonical MP2 energies. Errors introduced by the local treatment of LMP2 in the pure orbital basis will not be considered in this paper. Furthermore, as in **I**, we will only consider

correlation energies or contributions of the correlation energy to energy differences, in order to remove HF errors as much as possible. As shown in **I** and below, the HF errors are not entirely negligible with the AVTZ orbital basis.

The computed correlation energies obtained with MP2-F12/$\mathbf{2}^*$A(loc) method for the 21 molecules are listed in Table I. The AVnZ AO basis sets have been used along with the associated AVnZ/MP2FIT and VnZ/JKFIT sets as DF and RI bases, respectively. For comparison, estimates for the complete basis set (CBS) limits are also given. The first set (CBS[45]) was obtained by extrapolating the DF-MP2 correlation energies for $n=4$ and $n=5$, assuming $E_n = E_{CBS} + An^{-3}$ with A being fitted. For most molecules also calculations with the $n=4$ (AVQZ) to $n=6$ (AV6Z) basis sets have been performed without density fitting. By comparing the CBS[45] results with and without density fitting the errors of the density fitting on the extrapolated results were found to be at most 0.03 mH. Most accurate results were obtained by using in each case the MP2 fitting basis sets which correspond to the AO basis. Using the AV5Z fitting basis for both the AVQZ and AV5Z calculations slightly deteriorated the agreement with the nonfitted results. Comparison of the MP2-F12/AV5Z with the MP2/CBS[45] and MP2/CBS[56] correlation energies in Table I shows that in most cases the deviations of the MP2-F12/AV5Z values from the MP2/CBS[56] ones are significantly smaller than from the MP2/CBS[45] ones, and the extrapolated values converge towards to MP2-F12/AV5Z ones (in particular, for the larger molecules). The average absolute deviations between MP2-F12/AV5Z and MP2/CBS[45] or MP2/CBS[56] amount to 0.24 or 0.08 kcal/mol, respectively. The corresponding maximum deviations are 0.50 and 0.18 kcal. We therefore believe that the MP2-F12/AV5Z correlation energies are most accurate and very close to the true CBS limit. In the following, these values will be taken as CBS reference data.

Figure 1 shows the absolute errors of the correlation energies of the test molecules relative to the basis set limit (in millihartree). Amazingly, already with the AVTZ basis the CBS limit is reached within 0.5% for all molecules; in most cases the F12/$\mathbf{2}^*$A(loc) correction slightly overshoots. This is almost certainly a result of the A approximation (neglect of commutators of the form $[\hat{K}_1, f_{12}]$) and can be corrected, if necessary, by incorporating these terms.[10] If the basis set is increased, the MP2-F12 energy approaches the CBS limit smoothly. For the AVTZ basis the absolute errors are all (except for the last two molecules) smaller than 2 millihartree, while for AVQZ they are all below 0.3 millihartree.

It is important also to consider the errors of the Hartree-Fock energies, which are shown in Fig. 1 as well. Using the AVTZ basis the absolute HF errors are typically an order of magnitude larger than the errors of the MP2-F12 correlation energies obtained with the same basis. Even with the AVQZ basis the HF errors are still larger. Of course, it is well known that these errors mainly arise from the poor description of the core region and are therefore very systematic. But as shown in Fig. 1 of **I**, the HF errors of the reaction energies are not negligible if the AVTZ basis is used.

We now turn to the correlation contributions to the reac-

TABLE I. MP2-F12/**2**[*]A(loc) correlation energies ($-E_{corr}$ in millihartree) for AVnZ orbital basis sets. The corresponding AVnZ/MP2FIT and VnZ/JKFIT sets have been used as DF and RI bases, respectively. The complete basis set limit (CBS) has been estimated by extrapolating the canonical MP2 correlation energies (see text).

		MP2-F12/**2**[*] A(loc)			MP2	
No.	Molecule	AVTZ	AVQZ	AV5Z	CBS[45][a]	CBS[56][b]
1	H_2	34.26	34.21	34.21	34.24	34.24
2	CH_4	219.64	219.01	218.87	218.87	218.96
3	NH_3	264.94	264.36	264.28	264.39	264.29
4	H_2O	300.90	300.40	300.41	300.57	300.49
5	C_2H_2	347.18	346.25	346.00	345.60	345.82
6	C_2H_4	373.66	372.67	372.48	372.23	372.44
7	C_2H_6	410.85	409.64	409.34	409.23	409.44
8	CO	405.99	405.11	405.05	404.76	404.93
9	H_2CO	449.70	448.75	448.64	448.44	448.63
10	CH_3OH	486.08	484.99	484.86	484.79	484.95
11	H_2O_2	570.66	569.64	569.67	569.72	569.74
12	H_2CCO	610.53	609.17	608.98	608.48	608.82
13	C_2H_4O	650.18	648.67	648.43	648.02	648.41
14	CH_3CHO	641.80	640.29	640.03	639.64	640.01
15	C_2H_5OH	679.19	677.53	677.23	677.02	677.33
16	HNCO	657.56	656.18	655.99	655.59	655.82
17	$HCONH_2$	686.36	684.87	684.66	684.39	684.59
18	CO_2	689.32	687.99	687.90	687.43	687.77
19	HCOOH	720.33	718.90	718.79	718.51	718.77
20	NH_2CONH_2	919.74	917.70	917.41	917.07	
21	$HCOOCH_3$	909.78	907.74	907.50	907.00	

[a]DF-MP2/CBS[45] values obtained with AVQZ and AV5Z AO and DF bases.
[b]MP2/CBS[56] values obtained with AV5Z and AV6Z AO bases.

tion energies. These can be obtained from the data in Table I and therefore results will only be presented graphically. To facilitate comparison with the MP2-R12 results in **I**, the same scale (i.e., ±1 kcal/mol) is used as in the figures of our previous work. For convenience, the reactions are listed in Table II.

Figure 2 demonstrates the effect of the RI basis for the AVTZ orbital basis. In the upper panel uncontracted AVnZ orbital basis sets were employed as RI basis, while in the

lower panel the VnZ/JKFIT basis sets of Weigend[24] were used. It can be seen that the latter lead to much better convergence. This finding can be rationalized by the fact that the RI approximation may be viewed as the fitting of products of three occupied orbitals.[5] Apparently, the VnZ/JKFIT basis sets, which were optimized for exchange fitting in Hartree-Fock, are very well suited for this purpose. All other MP2-F12/**2**[*]A(loc) calculations in the current paper have therefore been performed using such RI basis sets.
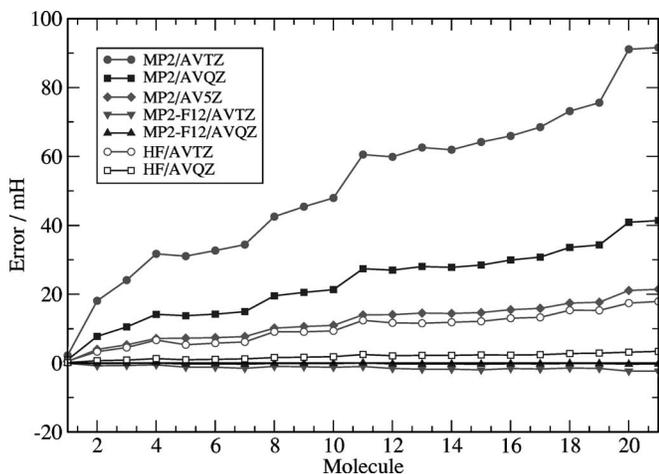
FIG. 1. Basis set truncation errors in Hartree-Fock total energies (open symbols, relative to HF/AV5Z) and in MP2 correlation energies (solid symbols, relative to MP2-F12/AV5Z). The molecules are in order of increasing total energy.

TABLE II. Test reactions.

No.	Reaction
1	$C_2H_2 + H_2 \rightarrow C_2H_4$
2	$C_2H_4 + H_2 \rightarrow C_2H_6$
3	$C_2H_6 + H_2 \rightarrow 2CH_4$
4	$CO + H_2 \rightarrow H_2CO$
5	$H_2CO + H_2 \rightarrow CH_3OH$
6	$H_2O_2 + H_2 \rightarrow H_2O$
7	$C_2H_2 + H_2O \rightarrow CH_3CHO$
8	$C_2H_4 + H_2O \rightarrow C_2H_5OH$
9	$CH_3CHO + H_2 \rightarrow C_2H_5OH$
10	$CO + NH_3 \rightarrow HCONH_2$
11	$CO + H_2O \rightarrow CO_2 + H_2$
12	$HNCO + NH_3 \rightarrow NH_2CONH_2$
13	$CO + CH_3OH \rightarrow HCOOCH_3$
14	$HCOOH + NH_3 \rightarrow HCONH_2 + H_2O$
15	$CO + H_2O_2 \rightarrow CO_2 + H_2O$
16	$H_2CCO + H_2CO \rightarrow C_2H_4O + CO$

(a)

(b)
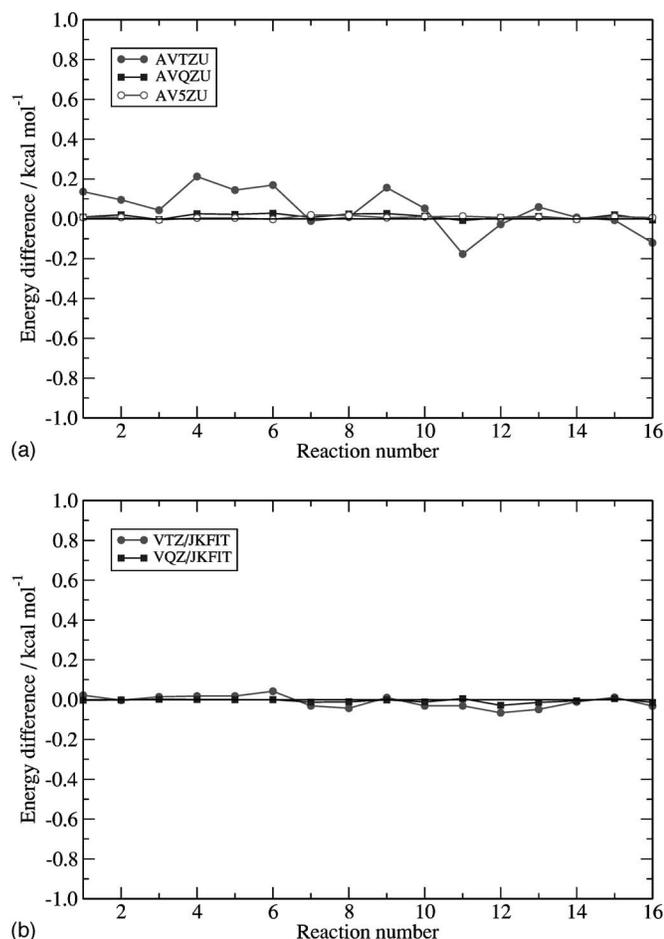
FIG. 2. Errors of the MP2-F12/2^{*}A(loc) correlation contributions to the reaction energies (in kcal/mol) for different RI basis sets (see text) using the AVTZ AO and DF basis sets. The reference values are obtained with the V5Z/JKFIT RI basis. Upper panel: Using *uncontracted* AVnZ AO basis sets (AVnZU). Lower panel: Using VnZ/JKFIT basis sets

Surprisingly, however, the JKFIT basis sets lead to huge errors when used as RI basis sets for MP2-R12 calculations. These errors are reduced to an acceptable amount when one shell diffuse basis functions is added. Since the MP2-F12 is much superior to MP2-R12, we will not consider the RI basis set issue for the MP2-R12 method any further.

Comparison with Fig. 4 of **I** shows that the dependence of the reaction energies on the choice of the RI basis is much weaker with f_{12} than with r_{12} correlation factors. With the MP2-F12/2^{*}A(loc) method the errors with the AVTZ orbital basis and VTZ/JKFIT RI basis are already much smaller than the ones obtained with the LMP2-R12 method, the AVQZ orbital basis, and uncontracted AV5Z for the RI basis. This result is of great importance for practical applications.

Figure 3 shows the error of the DF-MP2-F12/2^{*}A(loc) reaction energies for different basis sets. The errors obtained with AVTZ basis are less than 0.2 kcal/mol in all cases. This error is in many cases less than the HF error, and within the limits imposed by the 2^{*}A(loc) *Ansatz*. Possibly, even more accurate results will be obtained when the more expensive F12/2B method is used and the diagonal approximation is lifted. This will be investigated in future work.

In summary, we have shown in this section that with the
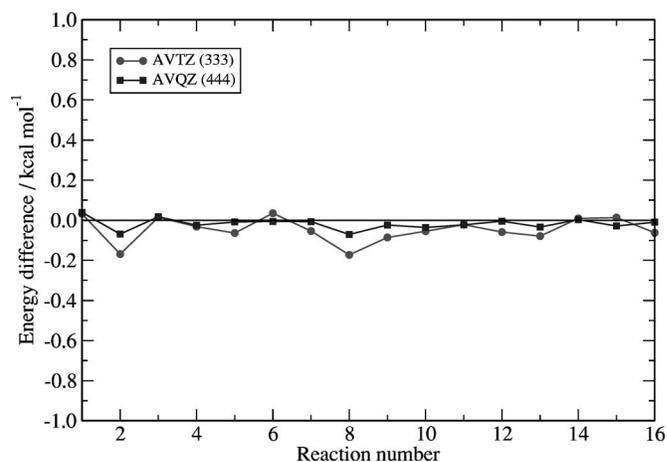
FIG. 3. Errors of the MP2-F12/2^{*}A(loc) correlation contributions to the reaction energies relative to the MP2-F12/2^{*}A(loc) (AV5Z) results (in kcal/mol) for different AO basis sets (see text). The VnZ/JKFIT sets have been used as RI bases.

MP2-F12/2^{*}A(loc) method excellent basis set convergence is achieved. Furthermore, the numerical errors due to the RI approximations converge rapidly to zero when JKFIT basis sets are used.

B. Weak pair approximations

In local calculations the pair correlation energies decrease quickly with the distance between the correlated local molecular orbitals (LMOs) i and j. Atoms are assigned to the LMOs using the Boughton-Pulay procedure,[25] and in this work a completeness parameter of 0.98 was used. For instance, in the glycine tetrapeptide (Gly)$_4$ in an extended conformation, 93.3% of the correlation energy is due to *strong pairs*, defined by the condition that the LMOs i and j are separated by no bonds; if all pairs separated by one or two bonds are added (close pairs), 98.3% or 99.6%, respectively, of the canonical correlation energy is recovered. As shown in Fig. 4, the computed f_{12} corrections show a very similar
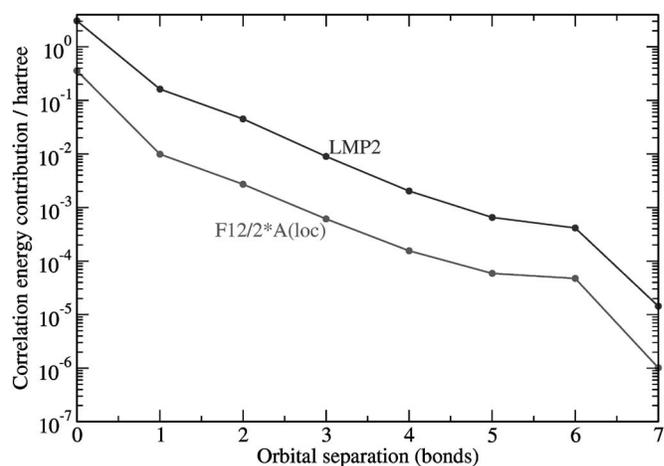
FIG. 4. Dependence of the correlation energy contributions of (Gly)$_4$ on the bond separation of the correlated orbital pairs (see text). There are only nine pairs separated by seven bonds, which explains the drop from 6 and 7. The AVTZ, VTZ/MP2FIT, and VTZ/JKFIT AO, DF, and RI basis sets, respectively, have been used.

behavior. In this case strong pairs contribute 96.4%, and if close pairs are included 99.0% (one bond) or 99.8% (two bonds) of the total f_{12} correction is recovered. As seen in Fig. 4, the remaining weak pair contributions are consistently about one order of magnitude smaller than the LMP2 contributions. We note that the unphysical increase of the contributions at long range reported in **I** for the DF-LMP2-R12/**2***A(loc) entirely disappears when the short range f_{12} correlation factor is used, as already anticipated in I.

Due to the rapid decrease of the f_{12} corrections with distance, it will in many applications be sufficient to compute them only for strong or strong and close pairs. Then the number of pairs included in the f_{12} calculations scales only linearly with molecular size, a prerequisite for obtaining low-order scaling of the computational cost. This affects directly the assembly stage of the density fitting procedure. For instance, if in the present example only strong pairs are included, the assembly time [Eqs. (16) and (17)] is reduced by almost a factor of 10 from 258 to 28 min (P4/2.8 GHz, AO basis: AVTZ, 1104 functions; DF basis: VTZ/MP2FIT, 1797 functions; RI basis: VTZ/JKFIT, 1763 functions). Similarly, the time needed for performing the RI summations [Eqs. (9) and (10)] is reduced from 32 to 4 min (this includes the transformation of the four-index integrals to an orthonormal basis, see **I**). These savings are even larger than implied by the reduction of pairs (182 strong pairs, total 922 pairs). This is due to the fact that the diagonal pairs ($i = j$), which constitute a larger fraction in the strong pair list, require less effort than the off-diagonal pairs.

C. Conclusions

In this work we have used a correlation factor of the form $f_{12} = \exp(-\beta r_{12})$ (Refs. 9 and 12) in a local orbital basis and combined with density fitting approximations.[5] The local orbital basis allows us to use a particularly simple and efficient *Ansatz*, in which the coupling of the f_{12} corrections of different pairs is neglected [diagonal approximation, here denoted as F12/**2***A(loc)]. Consistent with previous results with more elaborate *Ansätze* and canonical orbitals,[9,13,14] much more accurate results are obtained than with the linear r_{12} correlation factor. Using the AVTZ orbital basis and VTZ DF and RI basis sets the MP2-F12/**2***A(loc) method presented here consistently yields 0.3%–0.5% larger correlation energies than the MP2 basis set limit (taken as the MP2-F12/AV5Z result). At the same level, the largest error of reaction energies for 16 reactions involving small molecules amounts to only 0.17 kcal/mol, compared with the MP2/CBS limit. Also the convergence of the f_{12} corrections with the size of the RI basis set is much improved, and it has been found that the Hartree-Fock JK fitting basis sets of Weigend[24] lead to rapid convergence. Despite various shortcomings in the efficiency of our current implementation, the computation times for HF+MP2-F12/**2***A(loc) with AVnZ basis are already shorter than of corresponding HF+MP2 calculations with

AV(n+1)Z. The F12/AVnZ results are clearly better than those of conventional calculations with AV(n+2)Z basis. Even without linear scaling our method can be applied to molecules with 50 atoms or more, and it can be assumed that also for such large molecules the computed correlation energies are within 1% of the basis set limit. We believe that these results are of great importance for future applications of electron correlation methods in many fields.

More work is under way to further improve the method. This involves relaxing the diagonal approximation, which may work less well in cases with more delocalized orbitals; using further local approximations in the DF and RI expansions to achieve linear scaling of the computational resources with molecular size, avoiding the EBC approximation leading to the LMP2-F12/**2**A family of methods, avoiding the neglect of exchange terms leading to LMP2-F12/**2**B, treatment of open-shell cases, and of course extension to higher-level electron correlation methods.

[1] H.-J. Werner and F. R. Manby, J. Chem. Phys. **124**, 054114 (2006).
[2] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[3] C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).
[4] M. Schütz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).
[5] F. R. Manby, J. Chem. Phys. **119**, 4607 (2003).
[6] S. F. Boys and I. Shavitt, University of Wisconsin Report No. WIS-AF-13, 1959 (unpublished).
[7] J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).
[8] E. J. Baerends, D. E. Ellis, and P. Ros, Chem. Phys. **2**, 41 (1973).
[9] A. J. May and F. R. Manby, J. Chem. Phys. **121**, 4479 (2004).
[10] W. Kutzelnigg and W. Klopper, J. Chem. Phys. **94**, 1985 (1991).
[11] W. Klopper and C. C. M. Samson, J. Chem. Phys. **116**, 6397 (2002).
[12] S. Ten-no, Chem. Phys. Lett. **398**, 56 (2004).
[13] A. J. May, E. Valeev, R. Polly, and F. R. Manby, Phys. Chem. Chem. Phys. **7**, 2710 (2005).
[14] D. Tew and W. Klopper, J. Chem. Phys. **123**, 074101 (2005).
[15] W. Klopper, Chem. Phys. Lett. **186**, 583 (1991).
[16] W. Klopper and J. Almlöf, J. Chem. Phys. **99**, 5167 (1993).
[17] B. I. Dunlap, Phys. Chem. Chem. Phys. **2**, 2113 (2000).
[18] H.-J. Werner, P. J. Knowles, R. Lindh *et al.*, MOLPRO, version 2002.10, a package of *ab initio* programs (University of Cardiff, Cardiff, UK, 2005), see http://www.molpro.net
[19] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[20] A. Kendall, J. T. H. Dunning, and R. Harrison, J. Chem. Phys. **96**, 6769 (1992).
[21] F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[22] C. Hättig, Phys. Chem. Chem. Phys. **7**, 59 (2005).
[23] R. Polly, H.-J. Werner, F. R. Manby, and P. J. Knowles, Mol. Phys. **102**, 2311 (2004).
[24] F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).
[25] J. W. Boughton and P. Pulay, J. Comput. Chem. **14**, 736 (1993).