# Fast linear scaling second-order Møller-Plesset perturbation theory (MP2) using local and density fitting approximations

Hans-Joachim Werner, Frederick R. Manby, and Peter J. Knowles

Citation: J. Chem. Phys. **118**, 8149 (2003); doi: 10.1063/1.1564816
View online: http://dx.doi.org/10.1063/1.1564816
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v118/i18
Published by the American Institute of Physics.

# Fast linear scaling second-order Møller-Plesset perturbation theory (MP2) using local and density fitting approximations

Hans-Joachim Werner[a)]
*Institut für Theoretische Chemie, Universität Stuttgart, Pfaffenwaldring 55, D-70569 Stuttgart, Germany*

Frederick R. Manby
*School of Chemistry, University of Bristol, Cantocks Close, Bristol BS8 1TS, United Kingdom*

Peter J. Knowles
*School of Chemical Sciences, University of Birmingham, Birmingham B15 2TT, United Kingdom*

(Received 23 December 2002; accepted 10 February 2003)

We apply density fitting approximations to generate the 2-electron integrals in local MP2 (LMP2) to produce a method denoted DF-LMP2. The method can equally be seen as a local version of the well-known RI-MP2 method, which in this work is referred to as DF-MP2. Local approximations reduce the asymptotic scaling of computational resources to $\mathcal{O}(N)$, and the most expensive step of DF-MP2 [the $\mathcal{O}(N^5)$ assembly) is rendered negligible in DF-LMP2. It is demonstrated that for large molecules DF-LMP2 is much faster (1–2 orders of magnitude) than either LMP2 or DF-MP2. The availablility of LMP2, DF-MP2 and DF-LMP2 has for the first time made it possible to assess the accuracy of local and density fitting approximations for extended molecules using cc-pVDZ and cc-pVTZ basis sets. The density fitting errors are found to be consistently small, but the errors arising from local approximations are somewhat larger than expected from calculations on smaller systems. It is proposed to apply local density fitting approximations also for the Fock matrix construction in Hartree-Fock calculations. Preliminary results demonstrate that this can lead to significant savings in the Hartree-Fock calculation. © *2003 American Institute of Physics.*
[DOI: 10.1063/1.1564816]

## I. INTRODUCTION

During the last few years much progress has been made in developing high-level local electron correlation methods with low-order scaling of the computational cost as a function of molecular size. Based on the local correlation approach originally proposed by Pulay,[1–4] linear scaling has been achieved for local second-order Møller-Plesset perturbation theory (LMP2),[5,6] local coupled-cluster with single and double excitations (LCCSD),[7–9] and even for the local perturbative or iterative treatment of triple excitations[10–12] in LCCSD(T) and related methods. Other recent variants of low-order scaling MP2 approaches are the Laplace-transform MP2 method of Ayala and Scuseria,[13] the AO-based LMP2 methods of Lee, Maslen, and Head-Gordon,[14] a local MP2 implementation of Saebø and Pulay,[15] and the pseudospectral LMP2 methods of Friesner *et al.*[16] and Carter *et al.*[17] These methods have extended the applicability of wave function based correlation methods to much larger systems, and LCCSD(T) calculations for molecules with more than 100 atoms are now feasible with basis sets of double-zeta plus polarization quality.

However, larger basis sets, at least of triple-zeta quality, are usually needed to obtain sufficiently accurate results. Unfortunately, the computational time as well as the disk-space requirement scale with the fourth power of the basis set size per atom. This is the same in conventional and local methods, and currently the limiting factor in accurate calculations for larger systems. When the correlation-consistent cc-pV$n$Z basis sets of Dunning[18] are used, the basis set size increases approximately as $2^n$, and thus an increase of the CPU time by typically one order of magnitude is expected when going to the next larger basis set. Another difficulty is caused by the fact that the linear scaling regime is reached only for rather extended systems, and therefore the savings achieved by linear scaling methods are less for compact three-dimensional molecules than for the extended systems, such as one-dimensional alkane or peptide chains, which are commonly used to demonstrate low-order scaling behavior.

Both in the LMP2 and LCCSD methods, the bottleneck in calculations with large basis sets is the calculation of the 4-index 2-electron integrals and their transformation from the atomic orbital (AO) into the local orbital basis. This is also the case for the Laplace-transform linear scaling MP2 method of Ayala and Scuseria.[13] It is mainly the integral evaluation and transformation which causes the $N_{\mathrm{AO}}^4$ dependence of the computational cost with basis set size. An alternative to the exact calculation of the 2-electron integrals and their subsequent transformation into the MO basis is their approximation by density fitting methods.[19–24] In this case the one-electron charge densities in the 2-electron integrals, which are binary products of orbitals, are approximated by linear expansions in an auxiliary basis set. This leads to an approximation of the 4-index 2-electron integrals in terms of 2- and 3-index 2-electron integrals. It should be mentioned that a different approach in a similar spirit is the pseudospec-

a)Electronic mail: werner@theochem.uni-stuttgart.de

 

tral approximation of the integrals.[16,17] This method does not use basis functions but a grid as an auxiliary basis.

Density fitting has a long history in *ab initio* correlation methods,[19,21,25] but is usually cited in terms of its early uses in the Coulomb problem in Hartree-Fock-Slater and Kohn-Sham theory.[20,22,23] By now, density fitting is a well established approach, and it has been used for Coulomb fitting in DFT and Hartree-Fock (HF),[26,27] for exchange fitting in HF,[28] as well as for approximating the 2-electron integrals in MP2[29] and CCSD(T).[30] The accuracy of the method has been carefully investigated, and it has been shown that with suitable fitting basis sets the errors are much smaller than other typical errors in the calculations, as for instance basis set errors.[31] Optimized fitting basis sets are available for Coulomb[32] and exchange[28] fitting, as well as for MP2.[31,33]

Density fitting was first applied to MP2 by Feyereisen *et al.*,[26,29] and subsequently implemented by several authors;[34–38] for a review see Ref. 39. Nowadays this method is widely used and known as RI-MP2. Density fitting mathematically resembles a resolution of the identity (RI) in the specific case where the fitting criterion and target integral type coincide. However, RIs in quantum mechanics usually involve a summation over states and an implied overlap metric, neither of which appear in density fitting. Furthermore RIs do not offer a framework in which to discuss fitting criteria, constraints or robust fitting. The method is better thought of in terms of minimizing the Coulomb energy of a fitting residual, as first described by Whitten[21] and introduced in density functional theory by Dunlap *et al.*[22,23] In this work we therefore use the term DF-MP2 as a synonym for RI-MP2, and hope that other authors will accept this as the standard name.

The advantages of density fitting methods are twofold: first, the $N_{\mathrm{AO}}^4$ dependence of the computational cost is reduced to $N_{\mathrm{AO}}^3$, which makes the method particularly useful for calculations with larger basis sets. Second, the 3-index integrals are much faster to transform than the 4-index integrals, and therefore the method has a low prefactor for medium sized molecules. However, in DF-MP2 the scaling of the cost with molecular size is still $\mathcal{O}(N^5)$, as in normal MP2. In fact, in DF-MP2 the $\mathcal{O}(N^5)$ scaling cannot be reduced by screening techniques, while in integral-direct MP2 a scaling of about $\mathcal{O}(N^3)$ can be achieved in practice.[40] Thus, the range of applicability of the DF-MP2 method is limited to small and medium sized molecules.

In the present work we will demonstrate that the $\mathcal{O}(N^5)$ bottleneck in DF-MP2 can be removed by introducing local approximations. These approximations involve the use of individual excitation subspaces (*domains*) for each electron pair, and the use of multipole expansions for generating the transformed 2-electron integrals for distant pairs.[41] In this way, the $\mathcal{O}(N^5)$ scaling can be reduced to $\mathcal{O}(N^2)$ without any further screening. By introducing different fitting bases for each electron pair, the scaling can be further reduced to $\mathcal{O}(N)$. For large molecules, this leads to a dramatic reduction of CPU-time, in particular when accurate basis sets are used.

In Sec. II we will outline the method. In Sec. III we will demonstrate the accuracy and efficiency of the method for a number of test cases. In Sec. IV we demonstrate that local density fitting can also be used to reduce the cost of HF calculations. Finally, in Sec. V we will discuss further possible improvements as well as the an extension of local density fitting to higher-level methods as LCCSD(T).

## II. THEORY

In the following section we briefly summarize the conventional DF-MP2 method using a canonical orbital basis, and discuss the scaling behavior of the individual computational steps with molecular size. This will be needed for the further discussion of the local method later on.

### A. Canonical DF-MP2

In MP2 theory, 2-electron integrals $K_{ab}^{ij}=(ai|bj)$ over two occupied orbitals $\phi_i$, $\phi_j$ and two virtual orbitals $\phi_a$, $\phi_b$ are needed, and are defined as the electrostatic repulsion between two orbital-product densities,

$$K_{ab}^{ij}=\int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{\rho_{ai}(\mathbf{r}_1)\rho_{bj}(\mathbf{r}_2)}{r_{12}}. \tag{1}$$

Here and in the following we assume that the orbitals are real. In the DF-MP2 method the one-electron densities $\rho_{ai}(\mathbf{r}_1)=\phi_a(\mathbf{r}_1)\phi_i(\mathbf{r}_1)$ are approximated as

$$\overline{\rho}_{ai}(\mathbf{r})=\sum_A^{N_{\mathrm{fit}}} d_A^{ai}\chi_A(\mathbf{r}), \tag{2}$$

where $\chi_A(\mathbf{r})$ are fitting basis functions (e.g., atom-centred Gaussian-type orbitals, GTOs). The expansion coefficients $d_A^{ai}$ can be obtained by minimizing the positive definite functional[21–23]

$$\Delta_{ai}=\int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{[\rho_{ai}(\mathbf{r}_1)-\overline{\rho}_{ai}(\mathbf{r}_1)][\rho_{ai}(\mathbf{r}_2)-\overline{\rho}_{ai}(\mathbf{r}_2)]}{r_{12}}. \tag{3}$$

This leads to

$$d_B^{ai}=\sum_A (ai|A)[\mathbf{J}^{-1}]_{AB}, \tag{4}$$

$$\overline{K}_{ab}^{ij}=\sum_B d_B^{ai}(B|bj)=\sum_{AB}(ai|A)[\mathbf{J}^{-1}]_{AB}(B|bj), \tag{5}$$

where

$$J_{AB}=\int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{\chi_A(\mathbf{r}_1)\chi_B(\mathbf{r}_2)}{r_{12}}, \tag{6}$$

$$(ai|A)=\int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{\phi_a(\mathbf{r}_1)\phi_i(\mathbf{r}_1)\chi_A(\mathbf{r}_2)}{r_{12}}. \tag{7}$$

As shown by Dunlap *et al.*[22,23] this form of fitting, which uses the weight operator $1/r_{12}$, minimizes the least squares error of the electric field. Other possibilities, such as minimizing the error in the density using the weight operator $\delta(r_{12})$, have also been proposed,[20] but have been found to be less accurate.[26] It has not been established whether the idea of minimizing the error in the Coulomb potentials by using the weight $-r_{12}$ (Ref. 42) offers any advantage.

If the MOs are expanded in a basis of GTOs $\{\chi_\mu\}$, the 3-index integrals in the MO basis are obtained by a two-step transformation of the 3-index integrals $(\mu\nu|A)$ in the AO basis

$$(\mu i|A)=\sum_\nu C_{\nu i}(\mu\nu|A), \tag{8}$$

$$(ai|A)=\sum_\mu C_{\mu a}(\mu i|A). \tag{9}$$

According to the Gaussian product theorem, the size of an integral $(\mu\nu|A)$ decreases exponentially with the square of the distance between the basis functions $\chi_\mu$ and $\chi_\nu$, and therefore the number of non-negligible integrals scales asymptotically as $\mathcal{O}(\mathcal{N}^2)$, where $\mathcal{N}$ is a measure of the molecular size. The number of occupied (correlated) orbitals, $N_{\mathrm{occ}}$ is proportional to the molecular size, and thus the first transformation step in Eq. (8) scales as $\mathcal{O}(\mathcal{N}^3)$. Since the canonical MOs are usually delocalized over the whole molecule, the half transformed integrals $(\mu i|A)$, are not sparse, and therefore the second half transformation scales as $\mathcal{O}(\mathcal{N}^4)$. The next step is the solution of the linear equations in (4). Since $\mathbf{J}^{-1}$ cannot be expected to be sparse, this also scales as $\mathcal{O}(\mathcal{N}^4)$. Finally, the integral assembly step in Eq. (5) requires $N_{\mathrm{occ}}(N_{\mathrm{occ}}+1)N_{\mathrm{fit}}N_{\mathrm{virt}}^2$ floating point operations, and thus scales as $\mathcal{O}(\mathcal{N}^5)$. In practice, this step by far dominates the total computational cost in calculations for large molecules. Due to the delocalized character of the canonical MOs, no sparsity can be exploited to reduce the cost. On the other hand, all the four steps described above require simple matrix multiplications, and can be performed very efficiently on modern computers. This leads to a low prefactor for the algorithm, and despite the $\mathcal{O}(\mathcal{N}^5)$ scaling quite large molecules (about 40 nonhydrogen atoms) can be handled.[28]

### B. Local density fitting MP2 (DF-LMP2)

In the LMP2 method the occupied space is spanned by localized molecular orbitals (LMOs), which can be obtained from the canonical orbitals by standard localization procedures as proposed by Boys[43] or Pipek and Mezey.[44] The virtual space is spanned by a basis of nonorthogonal projected atomic orbitals (PAOs), which are obtained from the AO basis functions by projecting out the occupied orbital space.[1] In the following, PAOs will be labeled $r,s$. Since these functions are inherently local, one can introduce two approximations: First, excitations from a pair of occupied LMOs can be restricted to subsets of PAOs that are spatially close to the two LMOs. The number of functions $N_{[ij]}$ in each of these subsets (*pair domains*) is independent of the molecular size, and the number of excitations for each electron pair reduces from $N_{\mathrm{virt}}^2$ to $N_{[ij]}^2$. Second, the integrals $(ri|sj)$ for distant orbitals $i$ and $j$ can be approximated by multipole expansions[41] or neglected. The remaining number of nondistant orbital pairs $(ij)$, and therefore the total number of excitations, scales linearly with molecular size. Since in LMP2 there is a one-to-one correspondence between the number of excitations and integrals, it is obvious that the number of integrals $(ri|sj)$ to be calculated also scales linearly. Here the range of $r,s$ is restricted to the pair domain $[ij]$. For further details we refer the reader to previous work.[5,7,8]

These approximations have a profound effect on the last, most expensive computational step in DF-MP2: since the number of pairs $(ij)$ scales linearly, and the number of $r,s$ is independent of the molecular size, the computational effort for the assembly step [Eq. (5)] is dramatically reduced from $\mathcal{O}(\mathcal{N}^5)$ in the canonical case to just $\mathcal{O}(\mathcal{N}^2)$. Moreover, since the PAOs $r$ must be within a finite range of the LMOs $i$ and $j$, one requires only those transformed 3-index integrals $(ri|A)$ with $r$ in the *united pair domain* of the associated orbital $i$. This domain comprises all PAOs that belong to any pair domain $[ij]$ in which orbital $i$ occurs. For large molecules, the size of the united pair domains also become independent of the molecular size, and therefore the number of integrals $(ri|A)$ scales only as $\mathcal{O}(\mathcal{N}^2)$. This reduces the computational effort for solving the linear equations (4) from $\mathcal{O}(\mathcal{N}^4)$ to $\mathcal{O}(\mathcal{N}^3)$. Finally, since the occupied orbitals $i$ are now local, the number of half transformed integrals $(\mu i|A)$ scales only as $\mathcal{O}(\mathcal{N}^2)$. Using prescreening techniques, the first and second half transformations [Eqs. (8) and (9)] should then scale only as $\mathcal{O}(\mathcal{N}^2)$. Thus, without further approximations in the fitting basis, the bottleneck for large molecules will be to perform the matrix multiplications in Eq. (4). The inversion of the matrix $\mathbf{J}$ scales also with $\mathcal{O}(\mathcal{N}^3)$, but this has a much lower prefactor.

As will be shown in the following, linear scaling can be achieved by using domains also for the fitting basis. In the present work we have implemented and tested two different approximations, which both lead to linear scaling for the all steps of the calculation.

In the first case, a different fitting basis is used for each electron pair, and the linear equations are solved for each pair individually. The fitting basis for a given pair comprises all fitting functions at the atoms belonging to the pair domain, and in addition includes the functions at all atoms within a given distance $R_d$ from any atom in the pair domain. For large molecules, the size of these pair fitting basis sets $[ij]_{\mathrm{fit}}$ is independent of the molecular size, and therefore the computational effort scales linearly, provided the number of electron pairs scales linearly (i.e., distant pairs are neglected or treated by multipole expansions). For a given orbital $i$, the integrals $(ri|A)$ are needed only for the *orbital fit domain* $[i]_{\mathrm{fit}}$, which is the union of all pair fit domains $[ij]_{\mathrm{fit}}$ containing $i$. This can be exploited in the transformation, skipping all fitting functions $A$ that are not needed for orbital $i$. For a given LMO $i$, the number of PAOs $r$ and of fitting functions $A$ is asymptotically independent of molecular size, and linear scaling is achieved for both CPU time and disk space. The prefactor and thus the total cost of this method depends on the size of the fitting domains. As will be shown in Sec. III, the errors are small if $R_d$ is chosen to be 3–5 bohr, and, despite the fact that very many systems of linear equations have to be solved, the method is competitive for large molecules.

In the second case the fit is not performed for each pair individually, but only once for each orbital $i$, using the orbital fit domains $[i]_{\mathrm{fit}}$ as a fitting basis. As discussed above,

the size of $[i]_{\text{fit}}$ is asymptotically independent of the molecular size, and therefore linear scaling is achieved for both the linear equations and integral assembly. The number of sets of linear equations to be solved is typically 15 times smaller than in the first method, but the number of fitting functions in the orbital fit domains, as well as the number of PAOs $r$, is larger, causing a higher cost for solving each set of equations. Since the second method includes more functions in the fit for each pair than the first, it is more accurate. In fact, it turns out that it is sufficient in this case to use the pair fit domains without extensions ($R_d=0$ bohr).

In order to achieve linear scaling for the integral transformation, a further fitting domain has to be used: since in the integral assembly step [Eq. (5)] the orbital fit domain $A \in [i]_{\text{fit}}$ is used to multiply $d^A_{ri}$ with the integrals $(sj|A)$, the latter integrals for a fixed $j$ must be available for $A$ belonging to the union of all $[i]_{\text{fit}}$ of orbitals $i$ forming pairs with $j$. These domains are denoted *united orbital fit domains* $[j]^{\text{u}}_{\text{fit}}$. They are larger than the orbital fit domains $[j]_{\text{fit}}$, but their size is asymptotically still independent of the molecular size. We found that in large cases the computational effort for the second method is comparable to the first method for $R_d=3$ bohr, but is more accurate. In Sec. III we will demonstrate the accuracy and efficiency of these approximations.

The final problem is to achieve linear scaling also for the evaluation of the 3-index integrals $(\mu\nu|A)$. As already pointed out, the total number of significant integrals scales quadratically. Since the number of AOs $\mu$ contributing to an LMO $i$ is asymptotically independent of the molecular size, the first transformation step (8) scales quadratically if all fitting functions $A$ are used for each $i$. The same holds for the number of half-transformed integrals $(\mu i|A)$. However, as we have shown above, for each $i$ only a constant number of $A$ is needed. This can be exploited by first estimating the magnitude of the integral $(\mu\nu|A)$ by the Schwarz inequality

$$(\mu\nu|A) \leqslant (\mu\nu|\mu\nu)^{1/2}(A|A)^{1/2}. \tag{10}$$

On the basis of this value it is checked which orbitals $i$ will contribute to the transformation. For the list of surviving orbitals $i$ (independent of the molecular size), a lookup table is used to test if $A$ is needed. If not, the integral and the transformation is skipped. This makes it possible to achieve linear scaling also in the integral evaluation. For the sake of efficiency, the screening is performed using blocks of integrals, as described in the next section.

### C. Technical aspects

As pointed out above, it should be possible to reduce the scaling of the transformation steps in Eqs. (8) and (9) using prescreening techniques. The problem here is that without prescreening the matrix multiplications are very efficient, and screening of individual integrals would very much reduce the number of floating point operations per second (FLOPS) and thus strongly increase the prefactor. In order to keep the matrix multiplications, we have decided to split the integrals $(\mu\nu|A)$ for each fixed fitting index $A$ into blocks of AOs $\mu$, $\nu$. Typically, one block comprises all AOs at one atom. For very large basis sets these blocks may be further split. On the other hand, the AOs at neighboring atoms can be merged into one block if otherwise the blocks were too small. This often happens for the AOs at hydrogen atoms, which are then merged with the functions at the heavy atom to which the H is attached. For each AO basis block, the LMOs are sorted according to decreasing maximum absolute coefficients, and the maximum coefficients as well as the permutation list is stored in memory. Second, the integrals $(\mu\nu|A)$ are evaluated block by block, and for each block the largest element is also stored (for fixed A). In the first half transformation [Eq. (8)] only those LMOs are included for which the product of the largest coefficient in the block and the largest integral is larger than a threshold THRPROD. Typically, this threshold is chosen to be $10^{-7}$. Of course, prescreening is also used in the integral evaluation, in order to avoid the calculation of small integrals when $\mu$ and $\nu$ are far apart. This is controlled by a second threshold THRAO.

The nonzero half-transformed integrals are written to disk for blocks of fitting functions. A bucket sort is used to resort the integrals such that the fitting index $A$ becomes the fastest and the orbital index $i$ the slowest in the second half transformation. In the second half transformation only the nonzero blocks of half-transformed integrals are further transformed. The final integrals $(ir|A)$ are accumulated for all $r$ belonging to the united pair domain of orbital $i$ (see previous section). The transformed integrals are written to disk for batches of orbitals $i$ (unless all transformed integrals fit into the memory). The final fitting and assembly stage [Eq. (4) and Eq. (5)] is then driven by the orbital batches. In practice we found that the algorithm is CPU-bound and I/O is not a bottleneck.

## III. TEST APPLICATIONS

In this section we demonstrate the accuracy and efficiency of the DF-LMP2 method, which has been implemented as part of the MOLPRO suite of *ab initio* programs.[45]

Tables I and II show for a number of molecules the errors introduced by the local and fitting approximations, respectively, using the cc-pVTZ basis set[18] and the corresponding fitting basis optimized by Weigend *et al.*[31] In these calculations the full fitting basis has been used for all pairs. The orbitals were localized using the Pipek-Mezey procedure.[44] The excitation domains were determined using the automatic procedure of Boughton and Pulay[46] and a completeness criterion of 0.985 (corresponding of a least squares residual of 0.015, as defined in Ref. 46). The average error of the local approximation (in this case only the use of domains, since for molecules of this size there are no distant pairs) amounts to 1% relative to the nonlocal correlation energy for the same basis, while the fitting errors are typically two orders of magnitude smaller, and therefore negligible. Somewhat surprisingly, the error of the local approximation is largest for aliphatic molecules like pentane (error 1.5%), in which the orbitals can be well localized. Interestingly, the fitting errors are also largest for these cases. The reason for this behavior is still unclear. Possibly, this is due to the elimination of basis set superposition errors (BSSE) in the local calculations (see below). Table II also shows that the fitting errors are consis-

TABLE I. Comparison of conventional and local MP2 correlation energies. For each local method the second column contains the percentage of the canonical correlation energy.

| | | LMP2 | | | | | |
|---|---|---|---|---|---|---|---|
| Molecule | MP2[a] | Standard | | $R^s_{ext}=3$ bohr | | $R_{ext}=3$ bohr | |
| water | −0.261 836 | −0.260 358 | 99.44 | −0.261 836 | 100.00 | −0.261 836 | 100.00 |
| formaldehyde | −0.395 047 | −0.393 345 | 99.57 | −0.395 011 | 99.99 | −0.395 015 | 99.99 |
| methylamine | −0.404 625 | −0.400 147 | 98.89 | −0.403 952 | 99.83 | −0.404 380 | 99.94 |
| hydrazine | −0.442 778 | −0.438 684 | 99.08 | −0.442 172 | 99.86 | −0.442 590 | 99.96 |
| propane | −0.543 975 | −0.536 553 | 98.64 | −0.542 359 | 99.70 | −0.543 226 | 99.86 |
| oxirane | −0.575 545 | −0.570 608 | 99.14 | −0.575 141 | 99.93 | −0.575 364 | 99.97 |
| dimethylether | −0.598 373 | −0.592 191 | 98.97 | −0.597 098 | 99.79 | −0.597 845 | 99.91 |
| ethanol | −0.601 787 | −0.594 691 | 98.82 | −0.600 128 | 99.72 | −0.601 123 | 99.89 |
| butadiene | −0.649 923 | −0.642 571 | 98.87 | −0.648 916 | 99.85 | −0.649 291 | 99.90 |
| isobutene | −0.685 241 | −0.675 624 | 98.60 | −0.682 955 | 99.67 | −0.684 095 | 99.83 |
| thiophene | −0.812 760 | −0.804 784 | 99.02 | −0.811 501 | 99.85 | −0.812 032 | 99.91 |
| furan | −0.869 014 | −0.860 975 | 99.07 | −0.867 673 | 99.85 | −0.868 269 | 99.91 |
| imidazole | −0.881 400 | −0.872 479 | 98.99 | −0.879 789 | 99.82 | −0.880 571 | 99.91 |
| pentane | −0.892 456 | −0.878 700 | 98.46 | −0.888 488 | 99.56 | −0.890 658 | 99.80 |
| benzene | −0.949 757 | −0.938 641 | 98.83 | −0.947 850 | 99.80 | −0.948 551 | 99.87 |
| hexatriene | −0.966 507 | −0.954 822 | 98.79 | −0.964 459 | 99.79 | −0.965 281 | 99.87 |
| glycine | −1.014 178 | −1.004 380 | 99.03 | −1.011 357 | 99.72 | −1.012 976 | 99.88 |
| benzenethiol | −1.124 330 | −1.110 172 | 98.74 | −1.121 546 | 99.75 | −1.122 694 | 99.85 |
| alanine | −1.186 356 | −1.172 645 | 98.84 | −1.182 033 | 99.64 | −1.184 462 | 99.84 |
| oxalic acid | −1.241 657 | −1.232 669 | 99.28 | −1.239 405 | 99.82 | −1.240 494 | 99.91 |
| benzoquinone | −1.372 438 | −1.360 151 | 99.10 | −1.370 023 | 99.82 | −1.370 977 | 99.89 |
| maleic acid | −1.555 492 | −1.540 893 | 99.06 | −1.551 796 | 99.76 | −1.553 168 | 99.85 |
| average | | | 98.96 | | 99.80 | | 99.90 |

[a]Basis cc-pVTZ (see text). The geometries have been optimized at the B3LYP/cc-pVTZ(d/p) level.

tently smaller in the local than in the nonlocal calculations if the standard domains are used. If the domain sizes are extended (see below), the fitting errors approach the nonlocal ones.

TABLE II. Fitting errors of DF-MP2 and DF-LMP2 calculations.[a]

| Molecule | DF-MP2[b] | DF-LMP2 | DF-LMP2[c] | DF-LMP2[d] |
|---|---|---|---|---|
| water | 26 | 17 | 26 | 26 |
| formaldehyde | 28 | 22 | 28 | 28 |
| methylamine | 67 | 50 | 67 | 65 |
| hydrazine | 70 | 53 | 70 | 68 |
| propane | 93 | 69 | 93 | 89 |
| oxirane | 44 | 29 | 43 | 42 |
| dimethylether | 66 | 50 | 67 | 64 |
| ethanol | 77 | 58 | 79 | 74 |
| butadiene | 102 | 72 | 98 | 96 |
| isobutene | 112 | 82 | 112 | 106 |
| thiophene | 100 | 67 | 96 | 93 |
| furan | 79 | 51 | 77 | 73 |
| imidazole | 98 | 70 | 96 | 91 |
| pentane | 152 | 114 | 153 | 145 |
| benzene | 128 | 86 | 122 | 118 |
| hexatriene | 148 | 103 | 142 | 138 |
| glycine | 91 | 65 | 92 | 86 |
| benzenethiol | 153 | 104 | 146 | 140 |
| alanine | 119 | 89 | 123 | 114 |
| oxalic acid | 54 | 34 | 54 | 49 |
| benzoquinone | 154 | 108 | 148 | 144 |
| maleic acid | 107 | 71 | 105 | 97 |

[a]Basis and geometries as in Table I.
[b]Energy differences relative to the corresponding calculations without density fitting in microhartree.
[c]Extended domains for strong pairs, $R^s_{ext}=3$ bohr, see text.
[d]Extended domains for all pairs, $R_{ext}=3$ bohr, see text.

Previous work has shown that the local approximation hardly affects properties like equilibrium structures[47] or harmonic frequencies.[48] However, due to the somewhat different size of the absolute errors in different molecules, energy differences as reaction enthalpies can be more strongly affected. Despite the fact that basis set errors are certainly much larger than the local errors (the cc-pVTZ basis accounts for only about 85% of the valence-shell correlation energy), it might be desirable to reduce the local errors and to obtain results more consistent with conventional calculations. This can be achieved by extending the domains. A simple possibility is to include all PAOs into an orbital domain which are associated with atoms within a certain radius of any atom in the original domains. In the present work we have used a radius of 3 bohr (3.5 bohr for molecules containing second-row atoms). This value is somewhat larger than the longest bond distances, and therefore the domains are extended by the PAOs at the atoms which are directly attached to the ones in the original domains. The results of calculations with such extended domains are also shown in Tables I and II. Two different cases have been tested: In the first case, only the domains of the *strong pairs* are extended ($R^s_{ext}=3$ bohr). Here, strong pairs are those in which the (nonextended) domains of the two LMOs share at least the PAOs at one atom. These pairs typically account for 90%–95% of the valence correlation energy. The domain extension reduces the average error of the local approximation to 0.2% (largest error 0.4% for pentane). In the second case ($R_{ext}=3$ bohr) the domains of *all* pairs are extended. This further reduces the average error to 0.1%. The comparison of the

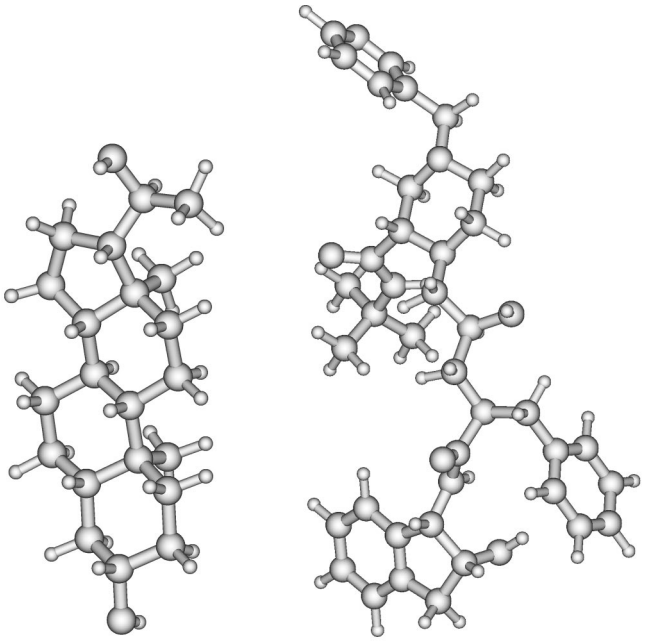


FIG. 1. Structures of pregnanediol and indinavir.

two cases is of interest since the domain sizes affect the integral transformation times and in particular the CPU time for solving the linear LMP2 equations. Calculations in which only the domains of the strong pairs are extended are significantly cheaper, and therefore this appears to be a cost effective choice. A relative error of 0.2% of the correlation energy appears acceptable since it is very much smaller than typical errors due to the basis set (about 15% for cc-pVTZ) or intrinsic errors of the MP2 approximation.

In order to demonstrate the efficiency of the DF-LMP2 method we have arbitrarily chosen two larger molecules, namely pregnanediol ($C_{21}H_{36}O_2$) and indinavir ($C_{36}H_{47}N_5O_4$), shown in Fig. 1. The structure of pregnanediol has been optimized at the B3LYP/cc-pVDZ level, and that of indinavir at the B3LYP/6-31G** level. Pregnanediol has no double bonds and it is therefore expected that the orbitals can be well localized. On the other hand, indinavir contains some benzene rings, and is therefore less well localizable. In all calculations the prescreening thresholds THRAO and THRPROD were chosen to be $10^{-7}$. It has been checked that this leads to energies within a microhartree of the values obtained with tighter thresholds of $10^{-9}$. The same thresholds were used in the LMP2 algorithm[5] and in the non-local DF-MP2 calculations. As usual, a completeness criterion of 0.98 was used in the Boughton-Pulay procedure[46] for selecting the domains for the cc-pVDZ basis, and 0.985 for the cc-pVTZ basis. These thresholds lead to almost identical domains for both basis sets. In all calculations distant pairs have been treated by multipole expansions,[41] using a distance criterion $R_{\mathrm{dist}}=8$ bohr. Very distant pairs ($R_{\mathrm{vdist}}=15$ bohr) have been neglected.

In Table III the CPU times to generate the integrals $K_{rs}^{ij}$ are compared for LMP2, DF-MP2, and DF-LMP2 calculations for pregnanediol ($C_{21}H_{36}O_2$) using cc-pVDZ and cc-pVTZ basis sets. We first consider the CPU times for generating the transformed integrals $K_{rs}^{ij}$. For a molecule of this medium size (59 atoms, including 36 hydrogens) the conventional DF-MP2 calculation is still 5.5 times faster than the local direct integral transformation as described in Ref. 5. The time is further reduced by a factor of 3–4 in the DF-LMP2 case. This is mainly due to dramatic savings in the assembly step [Eq. (5)], for which the CPU time is reduced by a factor of 150 from 1.5 hours (DF-MP2/cc-pVTZ) to 36 seconds (DF-LMP2/cc-pVTZ). Also the transformation of the 3-index integrals [Eqs. (8) and (9)] and the fitting [Eq. (4)] becomes faster, due to the better screening and the reduced number of transformed integrals in the local case. Furthermore, Table III demonstrates the increasing savings with increasing basis size: while the cost of the LMP2 calculation increases by about a factor of 14 when going from cc-pVDZ to cc-pVTZ, the cost of the DF-LMP2 calculations increases only by a factor of 7. This increase is smaller than for the nonlocal DF-MP2 (10). All these factors are smaller than the theoretical ones (32.0 and 13.5 for $N^4$ and $N^3$ scaling, respectively), which is partly due to the increased efficiency of the matrix multiplications with larger matrix sizes and partly to the fact that the size of the fitting basis sets increases somewhat more slowly than that of the orbital basis sets.

TABLE III. Analysis of CPU-times[a] for pregnanediol.

| | LMP2 | DF-MP2 | DF-LMP2[b] | DF-LMP2[b,c] |
|---|---|---|---|---|
| cc-pVDZ: | | | | |
| Integrals | 1256 | 111 | 111 | 111 |
| Transformation | 2893 | 110 | 89 | 90 |
| Solve | 0 | 61 | 49 | 50 |
| Assemble | 0 | 465 | 7 | 10 |
| Total $K_{rs}^{ij}$ | 4150 | 749 | 259 | 261 |
| Iteration | 133 | 0 | 132 | 174 |
| Total MP2 | 4313 | 749 | 422 | 494 |
| cc-pVTZ: | | | | |
| Integrals | 15 312 | 704 | 711 | 702 |
| Transformation | 42 886 | 895 | 660 | 665 |
| Solve | 0 | 441 | 301 | 301 |
| Assemble | 0 | 5399 | 36 | 54 |
| Total $K_{rs}^{ij}$ | 58 197 | 7440 | 1710 | 1724 |
| Iteration | 1009 | 0 | 1021 | 1290 |
| Total MP2 | 59 414 | 7447 | 2941 | 3413 |

[a]In seconds for Pentium4/2 GHz.
[b]Full fitting basis used for all pairs.
[c]Extended domains for strong pairs, $R_{\mathrm{ext}}^{s}=3$ bohr.

Table III also shows the total times for DF-MP2, DF-LMP2, and LMP2 calculations. The DF-LMP2 time relative to that of DF-MP2 is somewhat less favourable than for the integral generation alone, since in the LMP2 case the linear LMP2 equations have to be solved. This time is sensitive to the domain sizes, and significantly increases if the domains are extended. In addition, some time is needed for the generation of the transformation matrices to the pseudo-canonical basis for each pair, as needed for the update of the amplitudes[7] in the iterative scheme. This requires the diagonalization of a Fock matrix block for each pair and depends also on the domain sizes. Nevertheless, the total times are

TABLE IV. Analysis of CPU-times[a] for indinavir.

| | LMP2 | DF-MP2 | DF-LMP2[b] | DF-LMP2[b,c] |
|---|---|---|---|---|
| cc-pVDZ: | | | | |
| Integrals | 2869 | 588 | 537 | 540 |
| Transformation | 4199 | 550 | 129 | 127 |
| Solve | 0 | 497 | 84 | 84 |
| Assemble | 0 | 7314 | 7 | 9 |
| Total $K_{rs}^{ij}$ | 7068 | 8982 | 763 | 768 |
| Iteration | 478 | 0 | 479 | 690 |
| Total MP2 | 7587 | 8974 | 1287 | 1509 |
| cc-pVTZ: | | | | |
| Integrals | 25 540 | 2992 | 2816 | 2816 |
| Transformation | 56 620 | 4795 | 970 | 972 |
| Solve | 0 | 3364 | 362 | 362 |
| Assemble | 0 | 82 663 | 38 | 57 |
| Total $K_{rs}^{ij}$ | 82 160 | 93 900 | 4208 | 4220 |
| Iteration | 3772 | 0 | 3775 | 6666 |
| Total MP2 | 86 177 | 93 914 | 8247 | 11 221 |

[a]In seconds for HP ZX6000 Itanium2/900 MHz.
[b]Using linear scaling algorithm.
[c]Using extended domains for strong pairs, $R_{\text{ext}}^{s}=3$ bohr.

still a factor of 2–3 smaller than for DF-MP2, and is less than one hour on a low-cost PC for the cc-pVTZ basis set.

A second, larger, example is shown in Table IV for the indinavir molecule. The general pattern is similar to the previous case, but the savings become more pronounced. In this case the DF-MP2 is more expensive than the LMP2. The DF-LMP2 is six times faster than the LMP2 for the cc-pVDZ basis, and more than ten times faster for the cc-pVTZ basis. The savings are even larger relative to the DF-MP2. Again, a dramatic reduction of time is seen in the assembly step. For the cc-pVTZ basis the time is reduced from 23 hours to 38 seconds, i.e., by a factor of 2175. The smallest savings are achieved in the integral evaluation, and apart from the iterative solution of the LMP2 equations this dominates the computational effort. The CPU time of 2.3 hours for the DF-LMP2/cc-pVTZ calculation can be compared to the effort for the preceding the HF calculation. Using the orbitals of the cc-pVDZ calculation as a starting guess and an integral threshold of $10^{-11}$, this took about 40 hours (9 iterations) using two Athlon 1.2 GHz processors in parallel (the single processor speed is about 1.5 times lower than of the workstation used for the DF-LMP2 calculations). Thus, the DF-LMP2 calculation takes only about 4% of the HF CPU time, and can be considered as just a short post-processing after the HF calculation has been completed. Even though it might be possible to improve the efficiency of our HF program for extended systems, it is unlikely that the times would be reduced to the extent that DF-LMP2 would become the dominant computational step. The long HF times have prevented us from presenting even larger calculations, either with bigger basis sets or for even larger molecules, in this paper. A remedy of this problem could be to apply density fitting approximations also in the Hartree-Fock calculation. Work in this direction is in progress, and will be briefly discussed in the next section.

The correlation energies computed for pregnanediol and indinavir are summarized in Table V. In both cases the cc-pVDZ basis recovers about 82% of the DF-MP2 correlation energy obtained with the cc-pVTZ basis. By comparing the DF-MP2 and DF-LMP2 results it is for the first time possible to assess the accuracy of the local approximations for relatively large molecules. Unexpectedly, it is found that the relative errors are significantly bigger than for the set of test molecules shown in Table I. Particularly large errors are found for pregnanediol, a well localizable molecule without double bonds. With the cc-pVDZ basis, only 96.2% of the nonlocal correlation energy are recovered. With the larger cc-pVTZ basis, 97.8% are obtained, still 1.2% less than the average value in Table I. The errors are reduced by about a factor of 2 when the domains of the strong pairs are increased by a radius $R_{\text{ext}}^{s}=3$ bohr, but for pregnanediol the remaining error is still 1.1%, much larger than the average error of 0.2% in Table I. In view of the large effect of the basis set on the relative errors it is likely that a significant fraction of the difference between local and nonlocal correlation energies is due to basis set superposition effects, which are minimized in the local case.[7,49–51] Possibly, these effects are larger in saturated molecules in which the carbon atoms have a tetrahedral three-dimensional environment than in molecules with conjugated bonds, which contain more planar subunits. We have tested that neither the localization nor the distant pair approximations have a significant effect (Boys localization yields about 0.1% more correlation energy for pregnanediol than the Pipek-Mezey localization used for all calculations in this paper). A further systematic investigation of these effects is under way but beyond the scope of the present work.

By comparing the LMP2 and DF-LMP2 results it is also

TABLE V. Errors of local approximation for pregnanediol and indinavir as a function of the domain sizes.[a]

| | | cc-pVDZ | | | cc-pVTZ | | |
|---|---|---|---|---|---|---|---|
| | $R_{\text{ext}}^{s}$ | $E_{\text{corr}}$ | $\Delta E_{\text{corr}}$ | % | $E_{\text{corr}}$ | $\Delta E_{\text{corr}}$ | % |
| Pregnandiole: | | | | | | | |
| DF-LMP2 | 0 | −3.238 440 | 0.128 618 | 3.82 | −4.020 395 | 0.089 232 | 2.17 |
| DF-LMP2 | 3 | −3.305 368 | 0.061 690 | 1.84 | −4.064 465 | 0.045 161 | 1.10 |
| DF-MP2 | | −3.367 058 | 0.000 000 | 0.00 | −4.109 627 | 0.000 000 | 0.00 |
| Indinavir: | | | | | | | |
| DF-LMP2 | 0 | −6.244 577 | 0.187 209 | 2.91 | −7.731 696 | 0.139 197 | 1.77 |
| DF-LMP2 | 3 | −6.354 340 | 0.077 446 | 1.20 | −7.809 763 | 0.061 130 | 0.78 |
| DF-MP2 | | −6.431 786 | 0.000 000 | 0.00 | −7.870 893 | 0.000 000 | 0.00 |

[a]$R_{\text{ext}}^{s}$: Distance criterion for domain extensions in bohr.

TABLE VI. Effect of local approximations in the fitting basis on CPU-times and energies for indinavir.

| Fitting basis | $R_d$/bohr | CPU-time[a] | Error[b] |
|---|---|---|---|
| cc-pVDZ: | | | |
| Local pair | 3 | 732 | 686 |
| Local pair | 5 | 971 | 519 |
| United pair | 0 | 763 | 480 |
| Full | 0 | 1024 | 480 |
| cc-pVTZ: | | | |
| Local pair | 3 | 3878 | 1010 |
| Local pair | 5 | 4902 | 862 |
| United pair | 0 | 4208 | 832 |
| Full | 0 | 5853 | 832 |

[a]Time in seconds for integrals $K_{rs}^{ij}$ (HP ZX6000 Itanium2/900 MHz).
[b]Errors in microhartree relative to LMP2 ($R_s=0$ bohr); $E_{\mathrm{corr}}^{\mathrm{LMP2}}$(cc-pVDZ) $=-6.245\,057$ hartree; $E_{\mathrm{corr}}^{\mathrm{LMP2}}$(cc-pVTZ)$=-7.732\,528$ hartree.

possible to determine the fitting errors for large molecules.

Table VI summarizes some results for indinavir. Again somewhat unexpectedly, the fitting errors are larger with the cc-pVTZ basis than with the cc-pVDZ basis, but in both cases the error is below a millihartree and entirely negligible compared to the local or basis set errors. (Note that the difference of the correlation energies for the cc-pVDZ and cc-pVTZ bases amounts to 1.44 hartree, which is more than 1700 times as large as the fitting error). In the table the errors obtained with different approximations for the fitting basis are also compared. While the error introduced by orbital fit domains is sub-microhartree, somewhat larger fitting errors result if individual fit domains are used for each pair. In view of the fact that the use of pair fitting domains does not lead to additional savings we conclude that using orbital fit domains is most cost effective.

In order to demonstrate the scaling of the DF-LMP2 method as a function of the molecular size we have used the same linear polyglycine chains $(\mathrm{Gly})_n$ as in our previous work.[5,9] In these very extended one-dimensional model systems screening of the 2-electron integrals is most effective, and they therefore represent an optimum case for the linear scaling LMP2 method of Ref. 5. Despite the fact that these model systems are quite unrealistic, they are useful for testing the asymptotic scaling of local methods. The upper panel of Fig. 2 shows the timings of the individual steps of the density fitting calculation as a function of the chain length $n$, up to $n=20$. The cc-pVDZ basis set has been used. It is found that linear scaling is achieved very early for the transformation, fitting, and assembly steps. The scaling of the fitting and assembly steps depends solely on the local approximations (treatment of distant pairs by multipole expansions, use of domains for the excitation and fitting spaces), and not on any integral screening. Therefore, the scaling behavior is expected to be rather insensitive to the molecular structure and the basis set. The integral evaluation and the transformation steps depend on integral screening. As seen in the figure, the integration time strongly dominates the total time, and linear scaling is achieved later than for the other steps. This is due to the fact that the united orbital fitting domains reach their maximum sizes later than the pair domains themselves (see Fig. 3). Furthermore, the integral estimates obtained by the Schwarz inequality [Eq. (10)] are not very accurate and in general much larger than the exact integrals. We found that a threshold of $10^{-4}$ can be used in the screening procedure without affecting the accuracy by more than a microhartree (this also holds for pregnanediol and indinavir, both for the cc-pVDZ and cc-pVTZ basis sets). Possibly, some further improvement of the efficiency could be achieved if better integral estimates were available.

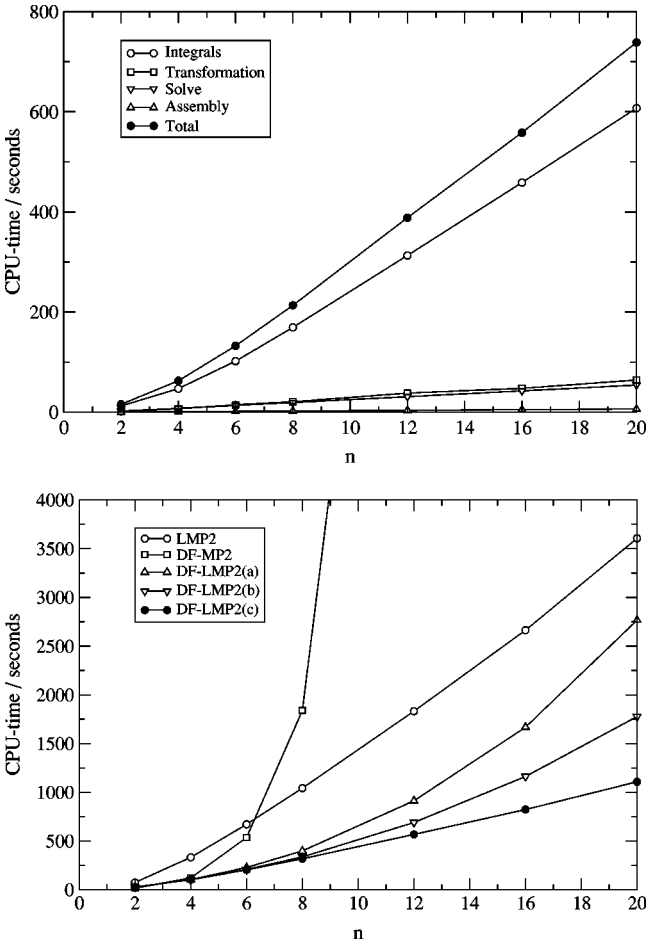


FIG. 2. CPU times (HP ZX6000, 900 MHz) of DF-LMP2 calculations for glycine polypeptide chains $(\mathrm{Gly})_n$ as function of chain length $n$. Upper panel: CPU times for integration, transformation, fit, and integral assembly in DF-LMP2 calculations [$\mathcal{O}(\mathcal{N})$ algorithm, see text]. The total times refer to the generation of the transformed integrals $K_{rs}^{ij}$. Lower panel: Comparison of total CPU times for DF-MP2, LMP2, and DF-LMP2 calculations. For the DF-LMP2 timings for the full fitting basis [case (a), $\mathcal{O}(\mathcal{N}^3)$] and for the local orbital fitting basis are shown. In the latter case, timings without Schwarz-screening [case (b), $\mathcal{O}(\mathcal{N}^2)$] and with full screening [case (c), $\mathcal{O}(\mathcal{N})$] are compared.

In the lower panel of Fig. 2 the total times of LMP2, DF-MP2, and DF-LMP2 calculations are compared as a function of the chain length $n$. In the LMP2 and DF-LMP2 cases these times include the generation of the pair domains and the solution of the linear LMP2 equations, as well as all other overheads. Due to the $\mathcal{O}(\mathcal{N}^5)$ scaling the DF-MP2 becomes very expensive for $n>8$. For the DF-LMP2 three cases are compared: (a) the full fitting basis is used for all pairs. As discussed in Sec. II B this leads to $\mathcal{O}(\mathcal{N}^3)$ scaling. In case (b) the fit is performed for each LMO in the basis of

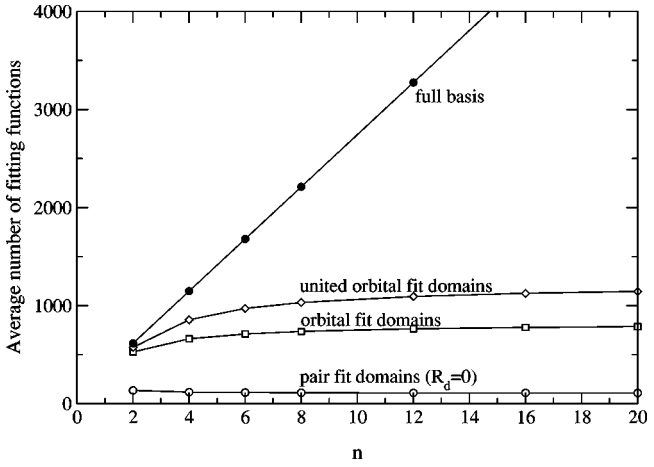


FIG. 3. Average sizes of fitting domains for glycine polypeptide chains $(\mathrm{Gly})_n$ (see text).

orbital fit domains. This leads to linear scaling for the transformation, fitting and assembly steps, but since no Schwarz screening has been used in the integration the scaling of the integration, and therefore of the overall cost, is quadratic. Finally, in case (c) the screening has also been used in the integration, leading to overall linear scaling (apart from very small nonlinear contributions, which arise, e.g., from the generation of the PAOs and the calculation of the 2-index integrals, but have a negligible prefactor). This latter case corresponds to the upper panel of the figure. For the present case, the DF-LMP2 method (case c) is about 3 times faster than the LMP2 without density fitting. As already pointed out, these calculations represent the optimum case for the LMP2, and the savings by the DF-LMP2 will be much larger in calculations for more compact systems, as already demonstrated for pregnanediol and indinavir. This is because the DF-LMP2 method is far less dependent on integral screening than the LMP2 method. The savings will also increase with increasing basis set.

Finally, Fig. 4 shows the timings (upper panel) and fitting errors (lower panel) for $(\mathrm{Gly})_n$ using different approximations for the fitting basis. The results obtained with the orbital fitting basis are within a microhartree of the ones for the full fitting basis. If individual pair fitting bases are used, the errors appear smaller, but this is due to a fortuitous error cancellation. In particular, for $R_d=3$ bohr the errors are very small, but they approach the one for the full fitting basis if $R_d$ is increased. Note that in the figure the negative of the errors is shown, i.e., the DF-LMP2 values are lower than the LMP2 ones. This is opposite to what was found for most molecules, including indinavir and pregnanediol. The upper panel of the figure shows that the CPU times are approximately the same if orbital fit domains are used as for pair fit domains with $R_d=3$ bohr, while the calculations for $R_d=5$ bohr are more expensive. Since the calculations with orbital fit domains are most accurate, this appears to be the best method, consistent with what was found above for indinavir.

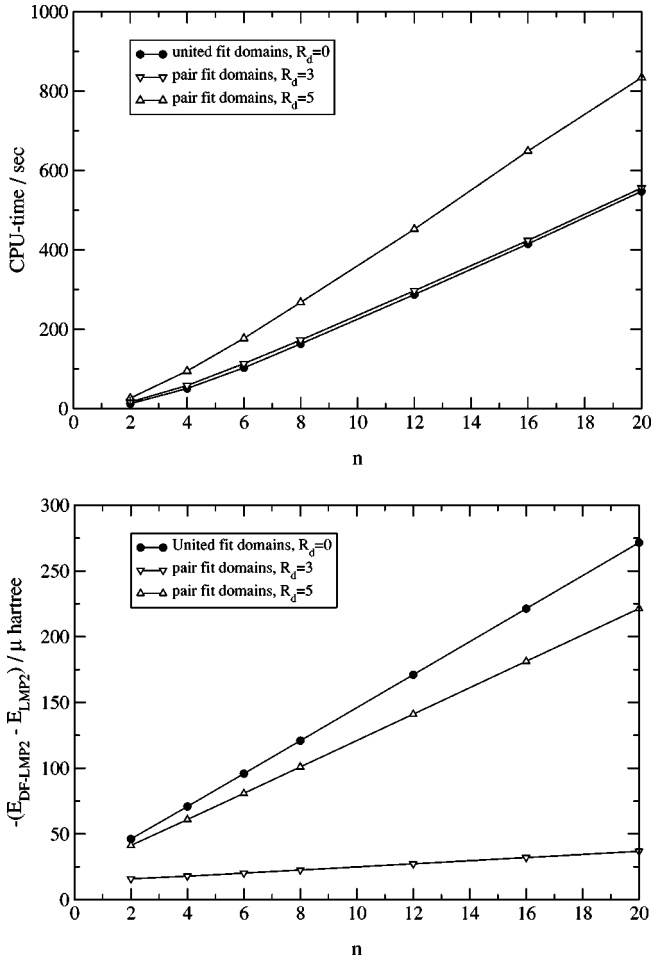


FIG. 4. Comparison of DF-LMP2 calculations for glycine polypeptide chains $(\mathrm{Gly})_n$ for different fitting approximations. Upper panel: CPU times (Pentium 4/2GHz) for generating the transformed integrals $K_{rs}^{ij}$. Lower panel: Fitting errors. The errors in the case with orbital fit domains are within a microhartree of those with the full fitting basis for all pairs.

## IV. LOCAL DENSITY FITTING IN HARTREE-FOCK

As pointed out in the previous section, DF-LMP2 calculations are much faster than the preceeding Hartree-Fock. In view of this fact one of the most important problems to solve next is to reduce the HF time. Density fitting can be readily applied to the Coulomb part of the Fock matrix, but previous attempts to fit also the exchange part lead to strict $\mathcal{O}(\mathcal{N}^4)$ scaling. Despite the fact that the prefactor is low and significant time can be saved for medium size molecules,[28] this is not a solution for large molecules. However, the current work offers a way forward: If the orbitals are localized in each iteration, similar techniques as described in this work for LMP2 can be applied to compute the exchange integrals $K_{\mu\nu}^{ii}$ in HF, and low-order—asymptotically linear—scaling should become possible. We have recently implemented such a local DF-HF method and first results are presented in the following. The Coulomb and exchange contributions to the Fock matrix are constructed from the half-transformed integrals $(\mu i|A)$ [cf. Eq. (8)] as

$$J_{\mu\nu}=\sum_A d_A(A|\mu\nu), \tag{11}$$

$$K_{\mu\nu}=\sum_i \sum_{A\in[i]_{\mathrm{fit}}} (\mu i|A) d_A^{\nu i}, \tag{12}$$

TABLE VII. DF-HF energies as a function of the domain sizes[a] in the HF orbital optimization for indinavir.[b]

| $R_{opt}/R_{energ}$ | $N_{fit}^{av}$ | $E_{HF}$[c] | Error[d] | CPU(tot)[e] | CPU(fit)[f] |
|---|---|---|---|---|---|
| 5/5 | 698 | −1962.987 417 | 10 599 | 70 | 13 |
| 5/10 | 1776 | −1962.996 701 | 1315 | 97 | 40 |
| 5/12 | 2269 | −1962.996 996 | 1020 | 117 | 60 |
| 5/14 | 2796 | −1962.997 103 | 913 | 134 | 77 |
| 5/full | 4965 | −1962.997 182 | 834 | 250 | 184 |
| full/full | 4965 | −1962.997 184 | 832 | 250 | 184 |
| 5/exact | | −1962.998 014 | 2 | 594 | |
| exact | | −1962.998 016 | 0 | 594 | |

[a]$R_{opt}$ bohr is the domain size $R_d$ (see text) in the orbital optimization, $R_{energ}$ the one for computing the final energy. *full* means the full fitting basis, and *exact* the HF energy without fitting. $N_{fit}^{av}$ is the average number of fitting functions per orbital in the energy calculation.
[b]The cc-pVTZ orbital basis has been used along with the corresponding JK-fitting basis of Weigend (Ref. 28).
[c]Local fitting for exchange only.
[d]Error in microhartree relative to the exact calculation.
[e]CPU time for Fock-matrix evaluation in minutes on Athlon 2200+.
[f]CPU time for transformation and fit in minutes on Athlon 2200+. The difference of the last two columns is the time for integral evaluation.

where

$$d_A^{\nu i} = \sum_{B \in [i]_{\mathrm{fit}}} (\nu i|B)[1/\mathbf{J}^i]_{AB}, \quad A \in [i]_{\mathrm{fit}}, \tag{13}$$

$$d_A = \sum_{\nu i} C_{\nu i} d_A^{\nu i} \tag{14}$$

and where $\mathbf{J}^i$ is a diagonal block of $\mathbf{J}$ involving only the fitting functions in $[i]_{\mathrm{fit}}$. Without local approximations, this corresponds to the method recently described by Weigend,[28] and the exchange part scales as $\mathcal{O}(\mathcal{N}^4)$. However, if localized orbitals are used, the fitting basis for a given orbital $i$ can be restricted to an orbital fitting domain $[i]_{\mathrm{fit}}$. In our implementation, this domain includes all functions at atoms which have a Löwdin charge $I_N^i = \Sigma_{\mu \in N}[\mathbf{S}^{1/2}\mathbf{C}]_{\mu i}^2$ of at least 0.05, plus all functions at atoms within a range $R_d$ of any of the primarily selected atoms. The size of these domains is asymptotically independent of the molecular size, and therefore the scaling of Eqs. (12) and (13) is reduced from $\mathcal{O}(\mathcal{N}^4)$ to $\mathcal{O}(\mathcal{N}^3)$ and $\mathcal{O}(\mathcal{N}^2)$, respectively. Furthermore, the integrals $(\mu i|A)$ will decrease quickly with increasing distance between $\mu$ and $i$. Asymptotically, the number of significant integrals will therefore scale linearly with molecular size, and if this is taken into account linear scaling can be achieved for the exchange part. The construction of the Coulomb matrix [Eq. (11)] scales quadratically. The prefactor depends sensitively on the size of the fitting domains. As will be demonstrated below, sufficiently accurate orbitals are obtained if a domain extension $R_d = 5$ bohr is used in the HF iterations. In order to compute an accurate energy, larger domains are needed, typical values of $R_d$ are 10–12 bohr. The Coulomb energy is more sensitive to the domain size than the exchange energy, and therefore the Coulomb part can optionally be computed in the usual way without local approximations. A detailed investigation of the acuracy and efficiency of this method will be presented in a forthcoming paper.[52]

Some preliminary results obtained with the new local DF-HF method are presented in Table VII for the indinavir molecule. The table demonstrate the dependence of the HF energy on the fitting domain size. Local approximations are applied only to the exchange part (local approximations in the construction of the Coulomb matrix work equally well for the orbital optimization, but lead to larger errors in the energies[52]). The orbitals were optimized using relatively small fitting domains, i.e., $R_d = 5$ bohr, and the final energy was then evaluated using larger domains. Since the (negative) exchange energies are underestimated with small domains, the total HF energies are too high in all cases and converge from above to the energy obtained with the full fitting basis. Despite quite large errors in the HF energies obtained with $R_d = 5$ bohr, the orbitals optimized with this small fitting basis yield accurate energies with larger fitting bases. The table shows that the energies computed with these orbitals and the full fitting basis agree within 2 microhartree with the ones obtained in a HF calculation with the full fitting basis. Using domain extensions $R_d$ of 10–12 bohr, the errors become significantly smaller than the errors caused by the nonlocal fitting itself.

The local approximations lead to significant savings. The construction of the Fock matrix for indinavir with the cc-pVTZ basis set and $R_d = 5$ bohr takes about 70 minutes on an Athlon 2200+ (1.8 GHz) machine. Of this time, 57 minutes are spent in the integral evaluation, and only 13 minutes are needed for the linear algebra [Eqs. (8), (12), and (13)]. The localization takes neglible time. Timings for larger fitting domains are also shown in Table VII. The nonlocal calculation of the Fock matrix with the full fitting basis takes 250 minutes. Of this time, 184 minutes are spent in the linear algebra, a factor of 14 more than in the local case. The exact integral-direct evaluation of the Fock matrix without density fitting takes between 594 minutes with the full density in the first iteration and 205 minutes with the difference density in the last iteration (using the SEWARD program of R. Lindh *et al.*[53] and a screening threshold of $10^{-11}$). The average time per iteration is about 2/3 of the initial time.

The DF-HF method can be applied to quite large molecules and basis sets. For instance, the local Fock matrix construction for indinavir with the full cc-pVQZ basis (3885 basis functions) takes 346 minutes, and for $(Gly)_{20}$ with cc-

pVTZ basis (3298 basis functions) 131 minutes ($R_d = 5$ bohr). For comparison, in the latter case the exact calculation of the Fock matrix takes 674 minutes. In our present DF-HF implementation 80–90% the CPU-time is spent in the evaluation of the 3-index integrals. We believe that our current integral code is far from being optimal, and that it will probably be possible to reduce the integral times at least by a factor of 2–3. This means that the local density fitting HF has the potential to become an order of magnitude faster than the conventional direct HF. The savings increase with increasing basis set size. Further reduction of the HF time can be achieved by using smaller basis sets in the HF calculation than in the subsequent correlation treatment, as first proposed by Jurgens-Lutovsky and Almlöf.[54] For instance, omitting in the cc-pVQZ basis set the highest angular momentum functions (both in the orbital and the fitting bases) reduces the time for the Fock matrix evaluation to from 346 to 182 minutes. We found that such approximations have only a negligible effect on the LMP2 total energies,[52] provided single excitations are taken into account in the LMP2.

## V. CONCLUSIONS

Density fitting approximations have been applied to the local MP2 method. Using the well known local approximations for the wave functions, and in addition different local fit domains for each orbital or orbital pair, linear scaling can be achieved. For large molecules, the fitting approximation reduces the CPU time by one order of magnitude or more. The largest calculation presented in this paper included 2008 basis functions and 5055 fitting functions, and took only 2.3 hours of CPU-time on a modern workstation (the times on a low-cost PC are approximately the same). This is only a very small fraction of the time needed for the preceding HF calculation (about 4% with our program), and thus a DF-LMP2 calculation can be considered as a short post-processing. This problem can be reduced by applying local density fitting also in the Fock matrix construction of the HF calculation. First results obtained with local density fitting in HF have been presented which show that the HF timings can be reduced by one order of magnitude. Full details of this method will be presented in a separate publication.[52]

Density fitting also strongly reduces the time to compute LMP2 gradients. A preliminary implementation has already been completed and will be described in a forthcoming publication.[55] Furthermore, the transformation of the 2-electron integrals into the LMO/PAO basis is even more a bottleneck in local coupled cluster methods as LCCSD[7–9] or LCCSD(T).[10–12] It appears readily possible to apply similar methods as described in the present work to all types of integrals, and first attempts in this direction[56] are very promising: the savings achieved by density fitting for integrals over four PAOs, as needed in all local correlation treatments beyond LMP2, are even larger than for LMP2. The implementation of a LCCSD(T) program in which all transformed integrals are generated by density fitting techniques is in progress.

As demonstrated by the test calculations in this paper, the evaluation of the 3-index integrals is the most expensive step in DF-HF and DF-LMP2 calculations. In our current implementation the integrals are evaluated once in the DF-LMP2 case and twice in the DF-HF per iteration. The half transformed integrals are stored on disk. Even though the number of significant integrals scales linearly, the disk storage and the I/O could become a bottleneck in calculations for large molecules with very accurate basis sets. Alternatively, one could recompute the integrals for batches of transformed integrals, and avoid the storage of the half transformed integrals. However, this would quite strongly increase the CPU time and deteriorate the scaling with molecular size. A remedy to this problem might be the use of Poisson fitting bases, an idea first discussed by Mintmire and Dunlap,[57] and recently turned into a practical approach by two of us.[58,59] In this case, most of the necessary 3-index 2-electron integrals reduce to simple 1-electron overlap integrals, which are much faster to evaluate than the 2-electron integrals. Furthermore, their number scales linearly with molecular size, without any local approximations. The application of such fitting basis sets to DF-HF and DF-LMP2 will be described in a forthcoming publication.[60]

## ACKNOWLEDGMENTS

This work has been supported by the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie, and the EU RTN Network *THEONET II* (HPRN-CT 1999 00005). One of the authors (F.R.M.) is grateful to The Royal Society for funding.

[1] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[2] S. Saebø and P. Pulay, Chem. Phys. Lett. **113**, 13 (1985).
[3] P. Pulay and S. Saebø, Theor. Chim. Acta **69**, 357 (1986).
[4] S. Saebø and P. Pulay, J. Chem. Phys. **86**, 914 (1987).
[5] M. Schütz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).
[6] G. Hetzer, M. Schütz, H. Stoll, and H.-J. Werner, J. Chem. Phys. **113**, 9443 (2000).
[7] C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).
[8] M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
[9] M. Schütz, Phys. Chem. Chem. Phys. **4**, 3941 (2002).
[10] M. Schütz and H.-J. Werner, Chem. Phys. Lett. **318**, 370 (2000).
[11] M. Schütz, J. Chem. Phys. **113**, 9986 (2000).
[12] M. Schütz, J. Chem. Phys. **116**, 8772 (2002).
[13] P. Y. Ayala and G. E. Scuseria, J. Chem. Phys. **110**, 3660 (1999).
[14] M. S. Lee, P. E. Maslen, and M. Head-Gordon, J. Chem. Phys. **112**, 3592 (2000).
[15] S. Saebø and P. Pulay, J. Chem. Phys. **115**, 3975 (2001).
[16] R. A. Friesner, R. B. Murphy, M. D. Beachy, M. N. Ringnalda, W. T. Pollard, B. D. Dunietz, and Y. Cao, J. Phys. Chem. A **103**, 1913 (1999).
[17] D. Walter, K. N. A. B. Szilva, and E. A. Carter, J. Chem. Phys. **117**, 1982 (2002).
[18] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[19] D. M. Schrader and S. Prager, J. Chem. Phys. **37**, 3927 (1962).
[20] E. J. Baerends, D. E. Ellis, and P. Ros, Chem. Phys. **2**, 41 (1973).
[21] J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).
[22] B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, Int. J. Quantum Chem., Symp. **11**, 81 (1977).
[23] B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys. **71**, 4993 (1979).
[24] B. I. Dunlap, Phys. Chem. Chem. Phys. **2**, 2113 (2000).
[25] S. F. Boys and I. Shavitt, University of Wisconsin, Report WIS-AF-13 (1969).
[26] O. Vahtras, J. Almlöf, and M. W. Feyereisen, Chem. Phys. Lett. **213**, 514 (1993).
[27] S. Ten-no and S. Iwata, Chem. Phys. Lett. **240**, 578 (1995).
[28] F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).
[29] M. W. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).

[30] A. P. Rendell and T. J. Lee, J. Chem. Phys. **101**, 400 (1994).
[31] F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[32] K. Eichkorn, F. Häser, O. Treutler, and R. Ahlrichs, Theor. Chim. Acta **97**, 119 (1997).
[33] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143 (1998).
[34] D. E. Bernholdt and R. J. Harrison, Chem. Phys. Lett. **250**, 477 (1996).
[35] D. Feller, E. Aprà, J. A. Nichols, and D. E. Bernholdt, J. Chem. Phys. **105**, 1940 (1996).
[36] D. E. Bernholdt and R. J. Harrison, J. Chem. Phys. **109**, 1593 (1998).
[37] F. Weigend and M. Häser, Theor. Chim. Acta **97**, 331 (1997).
[38] D. E. Bernholdt, Parallel Comput. **26**, 945 (2000).
[39] R. A. Kendall and H. A. Früchtl, Theor. Chem. Acc. **97**, 158 (1997).
[40] M. Schütz, R. Lindh, and H.-J. Werner, Mol. Phys. **96**, 719 (1999).
[41] G. Hetzer, P. Pulay, and H.-J. Werner, Chem. Phys. Lett. **290**, 143 (1998).
[42] P. M. W. Gill, B. G. Johnson, J. A. Pople, and S. W. Taylor, J. Chem. Phys. **96**, 7178 (1992).
[43] S. F. Boys, in *Quantum Theory of Atoms, Molecules, and the Solid State*, edited by P. O. Löwdin (Academic, New York, 1966), p. 253.
[44] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).
[45] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, *et al.* (see http://www.molpro.net), MOLPRO, *version 2002.5, a package of ab initio programs* (2003).
[46] J. W. Boughton and P. Pulay, J. Comput. Chem. **14**, 736 (1993).
[47] A. ElAzhary, G. Rauhut, P. Pulay, and H.-J. Werner, J. Chem. Phys. **108**, 5185 (1998).
[48] G. Rauhut, A. E. Azhary, F. Eckert, U. Schumann, and H.-J. Werner, Spectrochim. Acta, Part A **55**, 647 (1999).
[49] S. Saebø, W. Tong, and P. Pulay, J. Chem. Phys. **98**, 2170 (1993).
[50] M. Schütz, G. Rauhut, and H.-J. Werner, J. Phys. Chem. A **102**, 5997 (1998).
[51] N. Runeberg, M. Schütz, and H.-J. Werner, J. Chem. Phys. **110**, 7210 (1999).
[52] H.-J. Werner, F. R. Manby, and P. J. Knowles (unpublished).
[53] R. Lindh, U. Ryu, and B. Liu, J. Chem. Phys. **95**, 5889 (1991).
[54] R. Jurgens-Lutovsky and J. Almlöf, Chem. Phys. Lett. **178**, 451 (1991).
[55] M. Schütz, H.-J. Werner, F. R. Manby, and P. J. Knowles (unpublished).
[56] M. Schütz and F. R. Manby (unpublished).
[57] J. W. Mintmire and B. I. Dunlap, Phys. Rev. A **25**, 88 (1982).
[58] F. R. Manby and P. J. Knowles, Phys. Rev. Lett. **87**, 163001 (2001).
[59] F. R. Manby, P. J. Knowles, and A. W. Lloyd, J. Chem. Phys. **115**, 9144 (2001).
[60] R. Polly, F. R. Manby, H.-J. Werner, and P. J. Knowles (unpublished).