Tensor decomposition in post-Hartree–Fock methods. II. CCD implementation

Udo Benedikt, Karl-Heinz Böhm, and Alexander A. Auer

Citation: The Journal of Chemical Physics **139**, 224101 (2013); doi: 10.1063/1.4833565
View online: http://dx.doi.org/10.1063/1.4833565
View Table of Contents: http://aip.scitation.org/toc/jcp/139/22
Published by the American Institute of Physics

Tensor decomposition in post-Hartree–Fock methods. II. CCD implementation

Udo Benedikt, Karl-Heinz Böhm, and Alexander A. Auer[a)]
Max-Planck-Institute for Chemical Energy Conversion, Stiftstraße 34 - 36, D-45470 Mülheim an der Ruhr, Germany

(Received 4 September 2013; accepted 29 October 2013; published online 9 December 2013)

In a previous publication, we have discussed the usage of tensor decomposition in the canonical polyadic (CP) tensor format for electronic structure methods. There, we focused on two-electron integrals and second order Møller-Plesset perturbation theory (MP2). In this work, we discuss the CP format for Coupled Cluster (CC) theory and present a pilot implementation for the Coupled Cluster Doubles method. We discuss the iterative solution of the CC amplitude equations using tensors in CP representation and present a tensor contraction scheme that minimizes the effort necessary for the rank reductions during the iterations. Furthermore, several details concerning the reduction of complexity of the algorithm, convergence of the CC iterations, truncation errors, and the choice of threshold for chemical accuracy are discussed. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4833565]

I. INTRODUCTION

Nowadays, highly accurate *ab initio* calculations find widespread application in many fields of modern theoretical chemistry. Especially Coupled Cluster (CC) type calculations are used routinely to predict molecular properties.[1] These methods allow highly accurate calculations but a routine application for large systems is not feasible due to the very steep scaling of computational costs with system size ("curse of dimensionality").[2]

Hence, many research groups try to devise more efficient schemes through more efficient algorithms, or by reducing the number of parameters needed for the description of the wavefunction. One ansatz is to use tensor decomposition/representation techniques to compress the amount of data. Methods like the Cholesky decomposition (CD),[3–11] density fitting techniques,[12–14] or the famous resolution of identity (RI)[15–18] can be used to represent four dimensional arrays like two-electron integrals or amplitudes with two and three dimensional quantities. These techniques are mostly used not only in density functional theory (DFT) but also in explicit correlated methods[19–25] to describe the additional integrals. For an overview of various approaches from a perspective of general tensor decomposition techniques, see Ref. 26.

Likewise, further possibilities to reduce the number of parameters in post-Hartree-Fock calculations exist: the usage of pair natural orbital (PNO)[27–31] in efficient local approximations,[32–36] recently Yang *et al.*[37–40] proposed the orbital specific virtual approximation (OSV) in the framework of the Pulay Saebø local approximation (non-orthogonal projected atomic orbitals PAO)[41–43] and a series of optimized virtual orbital space methods (OVOS)[44–48] minimizes the size of the virtual space.

Besides these physically motivated approximations based on the locality of electron-electron correlation other schemes derived from numerical approaches based on tensor representation schemes have been proposed. Hino *et al.* describe an ansatz in which the computational effort for the CCSD(T) method can be reduced by applying an approach related to Singular Value Decomposition (SVD)[49–51] to the t_3 amplitudes.[52,53] In the tensor hypercontraction density fitting (THC-DF) proposed by Hohenstein *et al.*,[54–56] three center integrals arising in the RI procedure are decomposed with the CANDECOMP/PARAFAC procedure.[50,57] This way, a factorized representation of the two-electron integrals in the canonical polyadic or canonical product tensor format (CP) is obtained and the representation can be used to reduce the scaling of second and third order Møller-Plesset perturbation,[54] CISD,[58] or methods from the CC hierarchy like CC2.[56,59]

In a recent series of papers, we investigated representations in the Matrix Product State (MPS) format[26] or the CP format[60,61] in order to reduce the steep scaling of post Hartree-Fock correlation methods.

For low rank representation in the CP format, the necessary schemes and a detailed investigation of the AO and MO two electron integrals as well as the approximation for wavefunction parameters based on second order Møller-Plesset perturbation theory (MP2) were carried out. The scaling of ranks with increasing system and basis set size were found to be approximately $\mathcal{O}(N)$ for the wavefunction parameters and $\mathcal{O}(N^2)$ for two-electron integrals in the MO basis. This leads to overall tensor contraction costs that scale approximately as $\mathcal{O}(N^4)$ to $\mathcal{O}(N^5)$ depending on the type of contraction.

In this proof of principle study, the concept of tensor decomposition in the CP format is extended to CC theory and a scheme for the iterative solution of amplitude equations using tensors in CP representation is presented. In Sec. II, a short

a)Electronic mail: alexander.auer@cec.mpg.de

introduction to the CP representation including examples for tensor manipulations in this format is given. Section III outlines a CC algorithm for the example of the Coupled Cluster Double (CCD) t_2 amplitude equation. Furthermore, a novel scheme to prevent the drastic increase of ranks for certain tensor contractions is presented. In Sec. IV, the results of benchmark calculations are presented. Here, the focus lies on the error propagation during the iterative solution and the possibility to minimize the rank reduction costs by initialization with tensors from the previous iteration. Finally, in Sec. V, conclusions are drawn and an outlook is given.

II. THEORY

A. Canonical polyadic tensor format

Many problems related to high-dimensional tensors can be resolved by applying low rank tensor representations. One possible representation format is the CP tensor format where a general d-dimensional tensor is factorized into a sum of representing vectors.[62,63] For example, a tensor $\mathbf{A} \in \mathbb{R}^n$ in 4-dimensions ($A(w,x,y,z)$ with $w,x,y,$ and $z \in \mathbb{N} \leq n$) can be written as

$$\mathbf{A} = \sum_{r=1}^{R} a_r^{(w)} \otimes a_r^{(x)} \otimes a_r^{(y)} \otimes a_r^{(z)}, \tag{1}$$

with the summation length R, called rank of the representation, and the four representing vectors $a_r^{(w)}$, $a_r^{(x)}$, $a_r^{(y)}$, and $a_r^{(z)}$ that are connected with the Kronecker product. It should be noted that throughout this paper we refer to a tensor or matrix by writing its name in bold font. This notation indicates the full tensor and not a special value for a given index tuple. Furthermore, the superscripts in parenthesis do not denote a special entry in the vector but are given to distinguish the representing vectors for the different dimensions, as the representing vectors for the different dimensions do not have to be the same.

In contrast to the full 4-dimensional tensor, for example, that requires to store n^4 entries, the memory requirement is reduced to $4 \cdot n \cdot R$. The complexity of algebraic operations can also be decreased to linear scaling with respect to the dimension, so that the curse of dimensionality can be overcome. Off cause, this ansatz shows its full potential only if the tensor can be represented by a low-rank approximation, where the rank is significantly smaller than n^3. The CP format can easily be extended to higher dimensions by extending the number of representing vectors.

In contrast to the 2-dimensional case there does not exist a unique procedure to obtain an optimal low-rank approximation for an arbitrary tensor in d dimensions. The basic scheme currently applied is to find a trivial decomposition[60,61] of this tensor with a high rank followed by a rank-reduction.[64–71] To obtain a low rank approximation, the large rank has to be reduced within a given accuracy ϵ,

$$\|\mathbf{A} - \tilde{\mathbf{A}}\| \leq \epsilon \quad \text{with} \quad \tilde{\mathbf{A}} = \sum_{r=1}^{\tilde{R}} \tilde{a}_r^{(w)} \otimes \tilde{a}_r^{(x)} \otimes \tilde{a}_r^{(y)} \otimes \tilde{a}_r^{(z)}, \tag{2}$$

by fitting a new tensor $\tilde{\mathbf{A}}$ with a lower rank $\tilde{R} \ll R$ in an iterative procedure that successively increases the new rank one by one till the desired accuracy is reached. For this procedure, several algorithms like an alternating least square scheme,[66,67] a modified Newton method,[68,70] or an accelerated gradient[60,61,71] algorithm, that is used in this work, exist.

B. Tensor operations in the CP format

Given two tensors $\mathbf{A}$ and $\mathbf{B}$ in the CP representation, all tensor operations required for post-HF methods can be carried out in this format. Below, some examples for the most important tensor manipulations are given.

The addition of two tensors can be done by simply adding the representing vectors of $\mathbf{B}$ to the set of already existing representing vectors of $\mathbf{A}$,

$$\begin{aligned}\mathbf{C} = \mathbf{A} + \mathbf{B} &= \sum_{r_A=1}^{R_A} a_{r_A}^{(w)} \otimes a_{r_A}^{(x)} \otimes a_{r_A}^{(y)} \otimes a_{r_A}^{(z)} \\ &\quad + \sum_{r_B=1}^{R_B} b_{r_B}^{(w)} \otimes b_{r_B}^{(x)} \otimes b_{r_B}^{(y)} \otimes b_{r_B}^{(z)} \\ &= \sum_{r_C=1}^{R_A+R_B} c_{r_C}^{(w)} \otimes c_{r_C}^{(x)} \otimes c_{r_C}^{(y)} \otimes c_{r_C}^{(z)}.\end{aligned} \tag{3}$$

This way, the rank is increased and the new tensor rank r_C is the sum of the CP ranks of the two other tensors.

The point-wise product of two tensors

$$C(w,x,y,z) = A(w,x,y,z) \cdot B(w,x,y,z) \tag{4}$$

in CP format is done by the Hadamard product of the representing vectors for each dimension

$$\begin{aligned}\mathbf{C} = \mathbf{A} \odot \mathbf{B} &= \sum_{r_A=1}^{R_A} a_{r_A}^{(w)} \otimes a_{r_A}^{(x)} \otimes a_{r_A}^{(y)} \otimes a_{r_A}^{(z)} \\ &\quad \odot \sum_{r_B=1}^{R_B} b_{r_B}^{(w)} \otimes b_{r_B}^{(x)} \otimes b_{r_B}^{(y)} \otimes b_{r_B}^{(z)} \\ &= \sum_{r_A=1}^{R_A} \sum_{r_B=1}^{R_B} \left(a_{r_A}^{(w)} \odot b_{r_B}^{(w)}\right) \otimes \left(a_{r_A}^{(x)} \odot b_{r_B}^{(x)}\right) \\ &\quad \otimes \left(a_{r_A}^{(y)} \odot b_{r_B}^{(y)}\right) \otimes \left(a_{r_A}^{(z)} \odot b_{r_B}^{(z)}\right) \\ &= \sum_{r_C=1}^{R_A \cdot R_B} c_{r_C}^{(w)} \otimes c_{r_C}^{(x)} \otimes c_{r_C}^{(y)} \otimes c_{r_C}^{(z)},\end{aligned} \tag{5}$$

leading to an increased rank of the product tensor, which is the product of the CP ranks of the two tensors. This representation of the product tensor can be compressed by application of the rank reduction procedure so that $\mathbf{C}$ can be represented with a low-rank approximation.

General tensor contractions that yield a resulting tensor

$$C(w,x,y,z) = \sum_{s,t} A(w,x,s,t) \cdot B(s,t,y,z) \tag{6}$$

can be carried out in the CP format as scalar product over the corresponding representing vectors that have to be summed up (s, t), while all remaining vectors are simply copied,

$$\begin{aligned}\mathbf{C} &= \sum_{s,t} \mathbf{A} \cdot \mathbf{B} = \sum_{s,t} \sum_{r_A=1}^{R_A} a_{r_A}^{(w)} \otimes a_{r_A}^{(x)} \otimes a_{r_A}^{(s)} \otimes a_{r_A}^{(t)} \\ &\quad \cdot \sum_{r_B=1}^{R_B} b_{r_B}^{(s)} \otimes b_{r_B}^{(t)} \otimes b_{r_B}^{(y)} \otimes b_{r_B}^{(z)} \\ &= \sum_{r_A=1}^{R_A} \sum_{r_B=1}^{R_B} \left(\sum_s a_{r_A}^{(s)} \cdot b_{r_B}^{(s)} \right) \left(\sum_t a_{r_A}^{(t)} \cdot b_{r_B}^{(t)} \right) a_{r_A}^{(w)} \\ &\quad \otimes a_{r_A}^{(x)} \otimes b_{r_B}^{(y)} \otimes b_{r_B}^{(z)} \\ &= \sum_{r_A=1}^{R_A} \sum_{r_B=1}^{R_B} \left\langle a_{r_A}^{(s)}, b_{r_B}^{(s)} \right\rangle_s \left\langle a_{r_A}^{(t)}, b_{r_B}^{(t)} \right\rangle_t a_{r_A}^{(w)} \otimes a_{r_A}^{(x)} \otimes b_{r_B}^{(y)} \otimes b_{r_B}^{(z)}. \end{aligned} \tag{7}$$

In order to obtain the product tensor in a CP representation, the two ranks are combined to a new rank $r_C = r_A \cdot r_B$ and the entries in the scalar products for each value of r_A and r_B are used to scale the remaining representing vectors. This way, the rank increases to $R_A \cdot R_B$ and the representation can be recompressed by application of the reduction algorithm. Tensor contractions with summation over all indices include the evaluation of all scalar products for each dimension and the final value is obtained by simple summation over all entries in this scalar product.

Comparing the computational costs for the contraction shown in Eq. (7), one can see that the index based representation scales as $\mathcal{O}(N^6)$ assuming the same lengths N in all dimensions. The same contraction in CP format scales as $\mathcal{O}(2 \cdot N \cdot R_A \cdot R_B)$, leading to a formal N^3 scaling assuming a linear scaling of ranks with N, or N^5 scaling for a quadratic dependence of ranks on N. This illustrates the crucial dependence of the resulting scheme on the size of the corresponding representation ranks. Furthermore, the reduction procedure is a very important aspect of all operations. If successive tensor contractions are required, for example in an iterative procedure, ranks might increase and re-compression of the resulting representation is necessary in order to prevent an increase of computational costs in further manipulations.

In summary, any tensor that results from a tensor manipulation of tensors in the CP format can again be obtained in the CP format. However, its rank might be increased upon the manipulation so that the application of the reduction procedure is necessary in order to maintain a low rank representation of the target tensor.

III. CCD ALGORITHM WITH TENSORS IN LOW-RANK CP REPRESENTATION

As an instructive example for using tensors in CP format in CC calculations, the CCDs method is chosen. The residual r_{ij}^{ab} in the index based format is calculated according to

$$\begin{aligned} r_{ij}^{ab} &= \langle ab\|ij\rangle + \frac{1}{2}\sum_{ef} t_{ij}^{ef} \langle ab\|ef\rangle + \frac{1}{2}\sum_{mn} t_{mn}^{ab} \langle mn\|ij\rangle \\ &\quad + P(ij)P(ab)\sum_{me} t_{im}^{ae} \langle mb\|ej\rangle \\ &\quad - P(ij)\sum_{m} f_j^m t_{im}^{ab} + P(ab)\sum_{e} f_e^b t_{ij}^{ae} \\ &\quad + \frac{1}{4}\sum_{mnef} t_{ij}^{ef} t_{mn}^{ab} \langle mn\|ef\rangle \\ &\quad - P(ab)\frac{1}{2}\sum_{mnef} t_{ij}^{ae} t_{mn}^{bf} \langle mn\|ef\rangle \\ &\quad - P(ij)\frac{1}{2}\sum_{mnef} t_{im}^{ab} t_{jn}^{ef} \langle mn\|ef\rangle \\ &\quad + P(ij)P(ab)\frac{1}{2}\sum_{mnef} t_{im}^{ae} t_{nj}^{fb} \langle mn\|ef\rangle , \end{aligned} \tag{8}$$

where $P(pq)$ is a permutation operator and f_q^p denotes one electron contributions from the Fock matrix. Note that in efficient production level implementations the computational complexity of these tensor contractions is minimized by factorizing the equations.[72] However, for simplicity the unfactorized equations are discussed for a detailed description of the equations in the CP format.

After each iteration, an amplitude update is performed with the residual tensor as $t_{ij}^{ab} = r_{ij}^{ab} \cdot D_{ij}^{ab}$, where $D_{ij}^{ab} = \frac{1}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$ contains the orbital energies as obtained from Hartree-Fock. In the first (or more convenient zeroth) iteration t_{ij}^{ab} is set to zero, so that only the first term in Eq. (8) contributes to the residuum as $r_{ij}^{ab(0)} = \langle ab\|ij\rangle$. The amplitudes can then be obtained after denominator weighting and the first iteration starts with the initial values of the amplitudes and a new residual tensor r_{ij}^{ab} can be calculated, followed by the amplitude update. With the correlation energy

$$E_{\mathrm{CCD}} = \frac{1}{4}\sum_{efmn} \langle ef\|mn\rangle\, t_{mn}^{ef}, \tag{9}$$

Eq. (8) is iterated to convergence.

To use tensors in the CP format in the evaluation of the t_2 amplitudes, it is necessary to provide all quantities like different types of MO integrals and Fock matrix elements in the CP format. The MO integrals can either be obtained by transformation of reduced AO integrals and a further reduction in the MO basis or by trivial decomposition of the transformed MO integrals. The Fock matrix elements can simply be cast into the CP format using the formalism of the trivial decomposition or by a Singular Value Decomposition. Both methods lead to a representation in the CP format with a maximal rank of n (number of basis functions),[73] so that a reduction of the ranks is not necessary.[60,61]

In order to keep the notation transparent, the following form of the integrals and Fock matrix elements (with the corresponding ranks) will be used for the derivation of the

algorithm using tensors in CP format:

$$\mathbf{v}_{ij}^{ab} \approx \sum_{r_1=1}^{R_1} v_{r_1}^{(a)} \otimes v_{r_1}^{(b)} \otimes v_{r_1}^{(i)} \otimes v_{r_1}^{(j)},$$

$$\mathbf{v}_{cd}^{ab} \approx \sum_{r_2=1}^{R_2} v_{r_2}^{(a)} \otimes v_{r_2}^{(b)} \otimes v_{r_2}^{(c)} \otimes v_{r_2}^{(d)},$$

$$\mathbf{v}_{kl}^{ij} \approx \sum_{r_3=1}^{R_3} v_{r_3}^{(i)} \otimes v_{r_3}^{(j)} \otimes v_{r_3}^{(k)} \otimes v_{r_3}^{(l)},$$

$$\mathbf{v}_{bj}^{ia} \approx \sum_{r_4=1}^{R_4} v_{r_4}^{(i)} \otimes v_{r_4}^{(a)} \otimes v_{r_4}^{(b)} \otimes v_{r_4}^{(j)},$$

$$\mathbf{f}_{j}^{i} \approx \sum_{u_1=1}^{U_1} f_{u_1}^{(i)} \otimes f_{u_1}^{(j)},$$

$$\mathbf{f}_{b}^{a} \approx \sum_{u_2=1}^{U_2} f_{u_2}^{(a)} \otimes f_{u_2}^{(b)}.$$

It should again be noted, that the representing vectors $v_{r_1}^{(a)}$ in $\mathbf{v}_{ij}^{ab}$ and $v_{r_2}^{(a)}$ in $\mathbf{v}_{cd}^{ab}$ and the corresponding ranks R_1 and R_2 can different from each other if the rank of these integrals is reduced in the MO basis.

As there are ten different contributions to the residual tensor (cf. Eq. (8)) that have to be expressed using tensors in the CP format and that might have different ranks, one needs to distinguish the residual contributions arising from the different tensor contractions. Therefore, the expansion lengths are labeled with an additional index indicating the term from which it is constructed. For example, the residual contribution from the first term in Eq. (8) is denoted as

$$[\mathbf{r}_1]_{ij}^{ab} = \sum_{l_1=1}^{L_1} \tau_{l_1}^{(a)} \otimes \tau_{l_1}^{(b)} \otimes \tau_{l_1}^{(i)} \otimes \tau_{l_1}^{(j)}, \tag{10}$$

using τ instead of r as label for the representing vectors to avoid confusion with the summation index over the rank of the MO integrals. The full residual tensor in the CP format will be denoted using the rank L. It is obtained from the sum of ten different contributions arising from the different tensor contractions in Eq. (8) according to Eq. (3),

$$\mathbf{r}_{ij}^{ab} = \sum_{x=1}^{10} [\mathbf{r}_x]_{ij}^{ab}. \tag{11}$$

Starting from a low-rank representation of the $\mathbf{v}_{ij}^{ab}$ integral

$$\mathbf{v}_{ij}^{ab} \approx \sum_{r_1=1}^{R_1} v_{r_1}^{(a)} \otimes v_{r_1}^{(b)} \otimes v_{r_1}^{(i)} \otimes v_{r_1}^{(j)}, \tag{12}$$

one immediately obtains a low-rank approximation of the residual increment for the zeroth iteration and in this case also for the full residual tensor as a simple copy of the integral tensor

$$\mathbf{r}_{ij}^{ab^{(0)}} = [\mathbf{r}_1]_{ij}^{ab} = \mathbf{v}_{ij}^{ab}. \tag{13}$$

If the amplitude update is to be carried out in the CP format, an appropriate representation of the denominator has to be chosen. For this purpose, a procedure called approximation with exponential sums[74] is used. Note that the rank of the approximate energy denominator is a constant and does not scale with the system size.[74] The actual expansion length depends on the desired accuracy and currently is set to 42, leading to a numerical error of 7.869×10^{-12} in an interval from zero up to 1×10^{11}. In order to perform the tensor contractions in the following iterations, the rank of the amplitudes has to be reduced by application of the rank reduction algorithm.

Once all quantities for an iterative solution of Eq. (8) are available in low-rank CP representation, the tensor contractions to build up the residual can be carried out in CP format. As already mentioned in Sec. II, most tensor contractions increase the rank of the representation of the target tensor. As a consequence, a rank reduction step is necessary after every contraction in order to maintain low ranks for the residual tensor. However, in some cases this is not necessary if the structure of the tensor contraction is exploited, for other cases we propose special algorithms to minimize the increase of the rank. This is illustrated in the following using the second and fifth term in Eq. (8) as examples:

The second term in Eq. (8) (also referred to as the "particle particle ladder" term) is a contraction of the amplitudes with the four virtual MO integrals,

$$r_{ij}^{ab} \hookleftarrow \frac{1}{2} \sum_{ef} t_{ij}^{ef} \langle ab \| ef \rangle. \tag{14}$$

To calculate the residual increment, the equation can be written as

$$\begin{aligned}[\mathbf{r}_2]_{ij}^{ab} &= \frac{1}{2} \sum_{ef} \mathbf{t}_{ij}^{ef} \cdot \mathbf{v}_{ef}^{ab} \\ &= \frac{1}{2} \sum_{ef} \left(\sum_{q=1}^{Q} t_q^{(e)} \otimes t_q^{(f)} \otimes t_q^{(i)} \otimes t_q^{(j)} \right) \\ &\quad \cdot \left(\sum_{r_2=1}^{R} v_{r_2}^{(a)} \otimes v_{r_2}^{(b)} \otimes v_{r_2}^{(e)} \otimes v_{r_2}^{(f)} \right). \end{aligned} \tag{15}$$

Finally, the two summations can be consolidated into one summation over l_2 which is obtained as a product of q and r_2:

$$\begin{aligned}[\mathbf{r}_2]_{ij}^{ab} = \frac{1}{2} \sum_{q=1}^{Q} \sum_{r_2=1}^{R_2} & \underbrace{\langle t_q^{(e)}, v_{r_2}^{(e)} \rangle \langle t_q^{(f)}, v_{r_2}^{(f)} \rangle v_{r_2}^{(a)}}_{:=\tau_{l_2}^{(a)}} \\ & \otimes \underbrace{v_{r_2}^{(b)}}_{:=\tau_{l_2}^{(b)}} \otimes \underbrace{t_q^{(i)}}_{:=\tau_{l_2}^{(i)}} \otimes \underbrace{t_q^{(j)}}_{:=\tau_{l_2}^{(j)}}. \end{aligned} \tag{16}$$

The corresponding matrix elements of the coefficient matrix can be multiplied into the representation system for the first dimension and the remaining vectors are simply copied. The prefactor of $\frac{1}{2}$ can also be multiplied into the representing system so that the CP format is recovered, but with an increased

rank of $L_2 = Q \cdot R_2$:

$$[\mathbf{r_2}]_{ij}^{ab} = \sum_{l_2=1}^{L_2=Q\cdot R_2} \tau_{l_2}^{(a)} \otimes \tau_{l_2}^{(b)} \otimes \tau_{l_2}^{(i)} \otimes \tau_{l_2}^{(j)}. \tag{17}$$

The resulting rank then has to be reduced using the reduction algorithm and a rank similar to the t_2 amplitudes can be expected for the low-rank approximation of the increment to the residual.

While in this scheme the rank increases quadratically for this type of contraction, an alternative scheme can be devised in which the rank only grows linearly upon contraction. For this purpose, one starts from Eq. (16) but rather than consolidating the rank indices, the explicit summation over r_2 is carried out, yielding an intermediate matrix $\mathbf{B}^{ab}$ for each value of q,[75]

$$[\mathbf{r_2}]_{ij}^{ab} = \frac{1}{2} \sum_{q=1}^{Q} \left(\mathbf{B}_q^{ab}\right) \otimes t_q^{(i)} \otimes t_q^{(j)}. \tag{18}$$

This matrix can be decomposed by SVD leading to a representation similar to the CP format

$$\mathbf{B}^{ab} = \sum_{x=1}^{virt} \left(\mathbf{U}_x^a \Sigma_x\right) \mathbf{V}_x^b. \tag{19}$$

Substituting the matrix in Eq. (18) leads to

$$[\mathbf{r_2}]_{ij}^{ab} = \frac{1}{2} \sum_{q=1}^{Q} \left(\sum_{x=1}^{virt} \tilde{u}_x^{(a)} \otimes v_x^{(b)} \right)_q \otimes t_q^{(i)} \otimes t_q^{(j)}, \tag{20}$$

where $\tilde{u}_x^{(a)}$ are rows from the left eigenvectors multiplied by the singular values and $v_x^{(b)}$ are rows from the transposed right eigenvectors for each value of q. Now the two ranks can again be consolidated so that a representation of the target tensor in CP format is achieved:

$$[\mathbf{r_2}]_{ij}^{ab} = \frac{1}{2} \sum_{\tilde{q}=1}^{Q\cdot virt} \tilde{u}_{\tilde{q}}^{(a)} \otimes v_{\tilde{q}}^{(b)} \otimes t_{\tilde{q}}^{(i)} \otimes t_{\tilde{q}}^{(j)}. \tag{21}$$

The rank of the target tensor $\tilde{q}$ is still increased but only by the factor of the number of virtual orbitals, $L_2 = Q \cdot virt$. This leads to large savings in computational effort as the number of virtuals is much lower than the expected ranks for the integrals or amplitudes. Note that this scheme is denoted as "SVD of reassembled intermediate matrices" (SVD-RIM) contraction scheme in the following. Furthermore, it is also possible to truncate each summation by neglecting eigenvalues that are below a given threshold.

As a further example, the fifth term in Eq. (8) is chosen. Here, Fock matrix elements are contracted with t_2 amplitudes and no two-electron integral is involved. For clarity, the permutation operator is skipped in the following, but has to be applied for constructing the final residual increment before adding it to the set of already obtained representing vectors,

$$[\mathbf{r_5}]_{ij}^{ab} = P(ij)[\mathbf{z_5}]_{ij}^{ab} \quad \text{with} \quad [\mathbf{z_5}]_{ij}^{ab} = \sum_m \mathbf{f}_j^m \mathbf{t}_{im}^{ab}. \tag{22}$$

The tensor contraction can be done in the same way as for the previous term by performing internal contractions as scalar products of the corresponding representing vectors,

$$\begin{aligned}
[\mathbf{z_5}]_{ij}^{ab} &= \sum_m \mathbf{f}_j^m \mathbf{t}_{im}^{ab} \\
&= \sum_m \left(\sum_{u_1=1}^{U_1} f_{u_1}^{(m)} \otimes f_{u_1}^{(j)} \right) \left(\sum_{q=1}^{Q} t_q^{(a)} \otimes t_q^{(b)} \otimes t_q^{(i)} \otimes t_q^{(m)} \right) \\
&= \sum_{u_1=1}^{U_1} \sum_{q=1}^{Q} \underbrace{\left\langle f_{u_1}^{(m)}, t_q^{(m)} \right\rangle}_{\mathbf{M}_{u_1 q}} t_q^{(a)} \otimes t_q^{(b)} \otimes t_q^{(i)} \otimes f_{u_1}^{(j)}.
\end{aligned} \tag{23}$$

As only the last representing vector and all scalar products depend on the first rank, this summation can be seen as a kind of transformation from the expansion length u_1 to q with the help of all scalar products. This becomes more clear if all scalar products in Eq. (23) are written as a matrix $\mathbf{M}_{u_1 q}$ and the summation over u is put in front of the last representing vector,

$$\begin{aligned}
[\mathbf{z_5}]_{ij}^{ab} &= \sum_{q=1}^{Q} \underbrace{t_q^{(a)}}_{:=z_{w_5}^{(a)}} \otimes \underbrace{t_q^{(b)}}_{:=z_{w_5}^{(b)}} \otimes \underbrace{t_q^{(i)}}_{:=z_{w_5}^{(i)}} \otimes \underbrace{\left(\sum_{u_1=1}^{U_1} \mathbf{M}_{u_1 q} f_{u_1}^{(j)} \right)}_{:=z_{w_5}^{(j)}} \\
&= \sum_{w_5=1}^{W_5=Q} z_{w_5}^{(a)} \otimes z_{w_5}^{(b)} \otimes z_{w_5}^{(i)} \otimes z_{w_5}^{(j)}.
\end{aligned} \tag{24}$$

This way, the last representing vector can be calculated through a matrix-matrix multiplication of the scalar product matrix $\mathbf{M}$ and all representing vectors written as a matrix. Accordingly, this is an example for a term, in which the rank does not change and the residual increment is obtained in the CP format with the same rank as the corresponding amplitude tensor. Note that this scheme is denoted as "rank transformed" (RT) contraction scheme in the following.

The residual increment is obtained by permuting the third and fourth dimension in the incremental tensor

$$\begin{aligned}
[\mathbf{r_5}]_{ij}^{ab} &= P(ij)[\mathbf{z_5}]_{ij}^{ab} = [\mathbf{z_5}]_{ij}^{ab} - [\mathbf{z_5}]_{ji}^{ab} \\
&= \sum_{w_5=1}^{W_5=Q} z_{w_5}^{(a)} \otimes z_{w_5}^{(b)} \otimes z_{w_5}^{(i)} \otimes z_{w_5}^{(j)} \\
&\quad - \sum_{w_5=1}^{W_5=Q} z_{w_5}^{(a)} \otimes z_{w_5}^{(b)} \otimes z_{w_5}^{(j)} \otimes z_{w_5}^{(i)} \\
&= \sum_{l_5=1}^{L_5=2\cdot W_5=2\cdot Q} \tau_{l_5}^{(a)} \otimes \tau_{l_5}^{(b)} \otimes \tau_{l_5}^{(i)} \otimes \tau_{l_5}^{(j)}.
\end{aligned} \tag{25}$$

This way, the rank of the residual increment is again increased and can be reduced by the application of the reduction algorithm.

After all contractions have been carried out, the final residual tensor can be set up by adding all representing vectors of the incremental tensors. The rank of the new residual tensor is the sum of the ranks of all increments and can be reduced using the rank reduction procedure. The amplitude update then yields new amplitudes, for which the rank can again be reduced to values which will be similar to the rank

of the previous iteration. Then the algorithm proceeds until convergence.

The evaluation of the correlation energy can be done according to Eq. (9) using tensors in CP format as well as a sum of all inner products of the amplitudes and the MO integrals,

$$\begin{aligned} E_{\mathrm{LCCD}} &= \frac{1}{4}\sum_{efmn}\left(\sum_{q=1}^{Q} t_q^{(e)}\otimes t_q^{(f)}\otimes t_q^{(m)}\otimes t_q^{(n)}\right) \\ &\quad\cdot\left(\sum_{r_1=1}^{R_1} v_{r_1}^{(e)}\otimes v_{r_1}^{(f)}\otimes v_{r_1}^{(m)}\otimes v_{r_1}^{(n)}\right) \\ &= \frac{1}{4}\sum_{q=1}^{Q}\sum_{r_1=1}^{R_1}\left\langle t_q^{(e)}, v_{r_1}^{(e)}\right\rangle\otimes\left\langle t_q^{(f)}, v_{r_1}^{(f)}\right\rangle \\ &\quad\otimes\left\langle t_q^{(m)}, v_{r_1}^{(m)}\right\rangle\otimes\left\langle t_q^{(n)}, v_{r_1}^{(n)}\right\rangle. \end{aligned} \tag{26}$$

Next, the sum over the two expansion lengths can be carried out as the Kronecker products can be substituted by normal products between the different scalars. Thus, only scalar products between the different representing vectors of $\mathbf{t}_{ij}^{ab}$ and $\mathbf{v}_{ij}^{ab}$ have to be evaluated and summed up.

IV. BENCHMARKS

If a novel approximation is introduced in the framework of a quantum chemical method, it has to fulfill certain criteria to be applicable in practice. While this work is basically a proof-of-principle study and computational efficiency is not the main focus, several other vital questions arise, if the CP format is used throughout in the framework of a CC method: How do the ranks of t_2 evolve during different steps in the iterations? How does the error in the correlation energy correlate with the reduction threshold? Do numerical errors due to amplitude representation build up during the iterations and is the overall iteration scheme in the CP format stable? As the rank reduction algorithm is likely to be the most time consuming part, how to use it economically during the iterations? In this section, we try to address these questions using small numerical examples.

As an example of how the ranks change during the contractions and subsequent rank reductions in the iterations, the CCD method for water in a 6-31G basis set is chosen (R_{OH} = 95.720, θ_{HOH} = 104.52).[76] For this case, two-electron integrals in the MO basis have been cast into the CP format by trivial decomposition followed by rank reduction.[60,61] The resulting ranks for an accuracy of $\varepsilon = 10^{-4}$ are given in Table I.

Starting with the zeroth iteration, the rank of the first residual tensor is equal to that of the amplitude $L = R_1 = 87$. Performing the amplitude update is done with a fairly large expansion of the denominator (rank $S = 42$) in order to achieve high accuracy.[77] Therefore, the rank of the amplitudes grows to $Q = 3654$ and can be reduced by a sliced reduction (slice size is chosen as 200) followed by a full reduction with the same accuracy parameter ($\varepsilon = 10^{-4}$) to a rank of $Q = 84$. This is similar to the rank of the corresponding residual tensor or the $\mathbf{v}_{ij}^{ab}$ integral, respectively.

TABLE I. Rank of MO integrals for H_2O using the 6-31G basis set. The MO integral tensors are cast into the CP format by trivial decomposition. Rank reduction is done with a threshold of 10^{-4}.

Tensor	CP representation	Rank (R_μ) after reduction
$\mathbf{v}_{ij}^{ab}$	$\sum_{r_1}^{R_1} v_{r_1}^{(a)}\otimes v_{r_1}^{(b)}\otimes v_{r_1}^{(i)}\otimes v_{r_1}^{(j)}$	87
$\mathbf{v}_{cd}^{ab}$	$\sum_{r_2}^{R_2} v_{r_2}^{(a)}\otimes v_{r_2}^{(b)}\otimes v_{r_2}^{(c)}\otimes v_{r_2}^{(d)}$	107
$\mathbf{v}_{kl}^{ij}$	$\sum_{r_3}^{R_3} v_{r_3}^{(i)}\otimes v_{r_3}^{(j)}\otimes v_{r_3}^{(k)}\otimes v_{r_3}^{(l)}$	34
$\mathbf{v}_{bj}^{ia}$	$\sum_{r_4}^{R_4} v_{r_4}^{(i)}\otimes v_{r_4}^{(a)}\otimes v_{r_4}^{(b)}\otimes v_{r_4}^{(j)}$	59

The first iteration starts with a rank for t_2 of $Q = 84$. During the contractions, a large increase of ranks can be prevented by using the SVD-RIM or the RT scheme as described in Sec. III. A summary of the resulting ranks for the linear terms is given in Table II.

Here it should be noted, that the full increased ranks are reduced in a sliced reduction "on the fly." There, only a few of the scalar products are evaluated and multiplied with one representing vector while the remaining representing vectors are copied until the new rank reaches a certain size. The rank reduction is then applied only to this small slice of the full tensor and the reduced vectors are saved for a final full reduction. This way, the initial ranks for the rank reduction procedure are kept small and the reduction is much faster than the full reduction of the complete residual increment.

From Table II, it can be seen that the reduced ranks of the residual increments (note that for simplicity, permutations have been omitted) are almost the same for all given terms and approximately of the same size as in the zeroth iteration. To obtain the residual tensor in CP format, all permutations have to be applied which increases the rank of the tensor. Finally, all increments, also including higher order t_2 contractions, have to be summed up, again increasing the rank of the residual. In the example given here, this leads to a rank of L = 885, which can be reduced using the reduction procedure and a final low-rank representation of the residual tensor with a rank of L = 142 is obtained. Note that this rank is only

TABLE II. Evolution of ranks during the first CCD iteration for H_2O using the 6-31G basis set and $\varepsilon = 10^{-4}$. All ranks from the full procedure are reduced using a sliced reduction with a slice size of 200 followed by a full reduction. The ranks obtained with the SVD-RIM (first three columns) and RT (last two columns) procedure are obtained with a threshold of 10^{-6} for the SVD and reduced further with a full reduction to an accuracy of 10^{-4}. The permutations are skipped here and done with the reduced tensors.

Term	$\frac{1}{2}\mathbf{t}_{ij}^{ef}\mathbf{v}_{ef}^{ab}$	$\frac{1}{2}\mathbf{t}_{mn}^{ab}\mathbf{v}_{ij}^{mn}$	$\mathbf{t}_{im}^{ae}\mathbf{v}_{ej}^{mb}$	$\mathbf{f}_j^m\mathbf{t}_{im}^{ab}$	$\mathbf{f}_e^b\mathbf{t}_{ij}^{ae}$
Rank t_2	84	84	84	84	84
Rank v,f	107	34	59	5	8
Rank full residual	8998	2857	4956	420	672
Rank after reduction	84	86	84	84	84
Rank using SVD-RIM and RT	584	393	420	84	84
Rank after reduction	84	86	84	84	84

TABLE III. Error in CCD correlation energy using tensors in CP format for H_2O using 6-31G basis set throughout the first 10 iterations.

		CP $\varepsilon = 10^{-2}$		CP $\varepsilon = 10^{-3}$		CP $\varepsilon = 10^{-4}$	
Iterations	E_{LCCD}	E_{LCCD}	ΔE_{LCCD}	E_{LCCD}	ΔE_{LCCD}	E_{LCCD}	ΔE_{LCCD}
1	−0.13107	−0.13104	4.0×10^{-5}	−0.13107	4.0×10^{-6}	−0.13107	3.3×10^{-7}
2	−0.13342	−0.13325	1.7×10^{-4}	−0.13342	3.9×10^{-6}	−0.13342	1.1×10^{-6}
3	−0.13417	−0.13403	1.4×10^{-4}	−0.13417	2.5×10^{-6}	−0.13417	6.3×10^{-7}
4	−0.13446	−0.13428	1.7×10^{-4}	−0.13445	2.7×10^{-6}	−0.13446	3.2×10^{-7}
5	−0.13456	−0.13445	1.2×10^{-4}	−0.13456	2.7×10^{-6}	−0.13456	7.2×10^{-7}
6	−0.13461	−0.13447	1.4×10^{-4}	−0.13460	4.2×10^{-6}	−0.13461	6.5×10^{-7}
7	−0.13463	−0.13450	1.3×10^{-4}	−0.13462	3.2×10^{-6}	−0.13463	3.9×10^{-7}
8	−0.13463	−0.13447	1.6×10^{-4}	−0.13463	4.4×10^{-6}	−0.13463	3.8×10^{-7}
9	−0.13464	−0.13450	1.4×10^{-4}	−0.13464	1.2×10^{-6}	−0.13464	5.5×10^{-7}
10	−0.13464	−0.13454	9.5×10^{-5}	−0.13464	1.0×10^{-6}	−0.13464	2.5×10^{-7}

slightly larger than the rank in the zeroth iteration. The amplitude update is done with the new residual tensor leading to an increased rank of the amplitudes of $Q = 5964$, that can again be reduced to $Q = 87$. Thus, the next iteration starts with almost the same rank for the t_2 amplitudes and yields similar ranks for the residual increments after reduction. Overall, one CCD iteration consists of 10 tensor contractions, 11 rank reductions, 6 permutations, 1 summation of increments, 1 Hadamard product for the amplitude update, and 1 evaluation of all inner products for the energy.

As a small numerical error is introduced every time the rank reduction procedure is carried out, this error might build up and could even influence the overall convergence of the CC algorithm. To investigate whether this is the case in practice, the error of the CP-based algorithm compared to the conventional CCD correlation energy (calculated with NWChem, no convergence acceleration used; modified `tce` module)[78–80] are given in Table III and Fig. 1 for each CC iteration. (As a second example, methane has been chosen, see Figs. 3 and 4.) From this data, it can be seen that the errors in the calculated correlation energy are approximately 2 to 3 orders of magnitude lower than the given reduction threshold. For the iterative CC scheme, mH accuracy in every iteration is obtained for thresholds smaller than 10^{-3} which is in agreement with what has been found for MP2.[60]

From Figs. 1 and 3, it can be seen that the convergence using tensors in the CP format is not as smooth as the standard CCD algorithm when larger thresholds (10^{-2}) are used. Nevertheless, it is still possible to converge the correlation energy to values that are still 2 to 3 orders of magnitude lower than the given reduction threshold (cf. Figs. 2 and 4). For thresholds smaller than 10^{-3}, the convergence is much smoother and overall the algorithm appears very robust regarding the possible build-up of errors due to successive approximation of the amplitudes in each iteration.

To give an impression on the computational effort of the tensor contractions and the rank reduction compared to the standard calculation, timings for a single CCD iteration on a single core of a workstation are listed in Table IV. For the CP format, the full timing is given as contributions from the reduction and the pure tensor contraction. This shows that using tensors represented in the CP format, the rank reduction becomes the most dominant part, while the actual time for the pure tensor contraction is barely visible at all. The first implementation using a full rank reduction for tensors with increased ranks (after contraction) is very slow compared to the standard calculation. There, the increased initial ranks lead to higher CPU times as the rank reduction algorithm scales

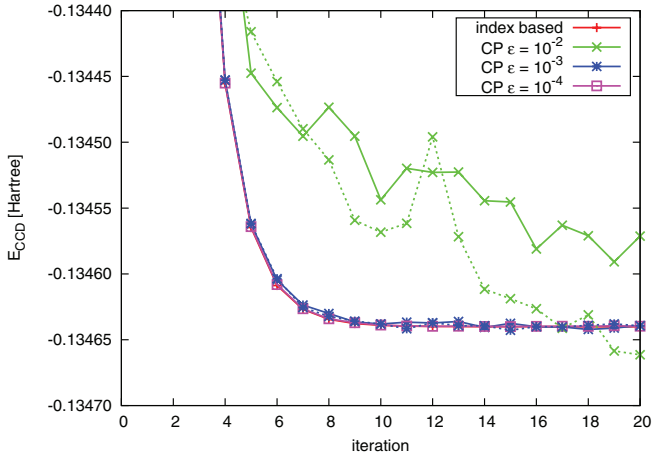

FIG. 1. CCD correlation energy using CP tensors for H_2O in 6-31G basis set. Solid lines show results using reductions from scratch, while dotted lines denote reductions with initial guess from the previous iterations. Note that the results for $\varepsilon = 10^{-4}$ is on top of the results of the conventional method on this scale.

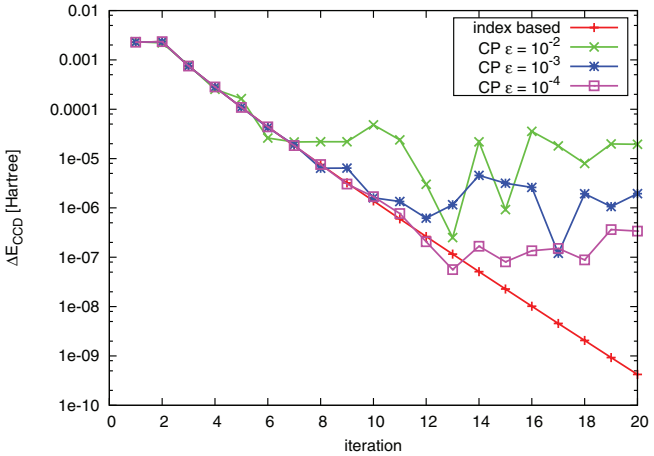

FIG. 2. Convergence of CCD correlation energy using CP tensors for H_2O in 6-31G basis set.

TABLE IV. CPU time per iteration on a single workstation core.[a] For the algorithm using tensors in CP format a threshold of $\varepsilon = 10^{-2}$ is applied.

Algorithm	Time per iteration [s]
Conventional	<1
CP, full reduction	677 (reduction: 676; contraction: <1)
CP, sliced reduction (250)	118 (reduction: 117; contraction: <1)
CP, SVD-RIM/RT, full reduction	35 (reduction: 34; contraction: <1)

[a]TCE implementation of NWChem on a workstation with exclusive usage of the system and the following specifications: 2× Quad-Core AMD Opteron™ Processor 2354 @ 2200 MHz, 32 GB RAM, 1 TB scratch disk space, openSUSE 10.3 (X86-64) with Linux Kernel 2.6.22.17.

with the initial rank, so that a sliced reduction with fixed initial rank can be chosen and the whole rank reduction procedure can be done in parallel. After this sliced reduction, the merged rank is still larger than one would expect from the full reduction and a further full reduction of the pre-reduced tensor can be done. This leads to similar ranks as obtained from the single full reduction and also the accuracy in the energy is comparable. The computational costs, however, is lowered significantly. Nevertheless, the algorithm using sliced reductions is more than 100 times slower than the standard calculation which takes in the order of a second for this small example. This is entirely due to the high prefactor of the rank reduction algorithm, which shows that beyond this proof of principle study, much more work has be invested in devising more efficient reduction schemes. The best scheme yet is to avoid the huge increase of ranks upon tensor contraction by applying the SVD-RIM and RT schemes.

As the rank reduction procedure becomes the time determining step, we propose another scheme to reduce the computational effort in the following. In fact, the iterative procedure of solving the amplitude equations allows for a more efficient reduction of increased CP ranks during the iterations. As for example, the entries in the amplitudes change only little during the iterations, the tensor representation from the last iteration can be used as an initial guess for the rank reduction. The algorithm then does not start with a rank one approximation but uses all representing vectors from the previous iteration as

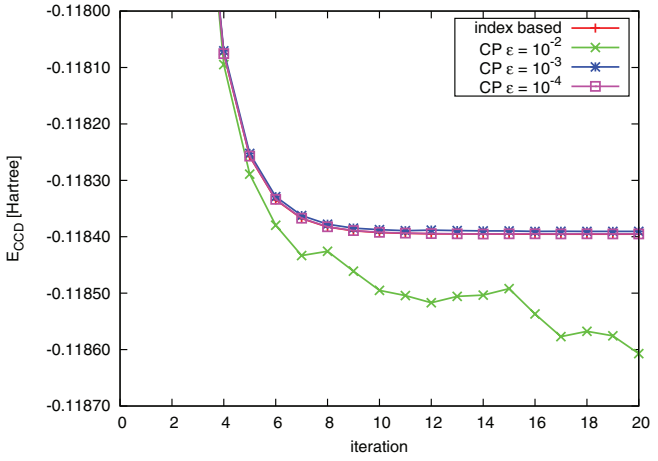

FIG. 3. CCD correlation energy using CP tensors for CH_4 in 6-31G basis set and a reductions from scratch. Note that the results for $\varepsilon = 10^{-4}$ is on top of the results of NWChem on this scale.

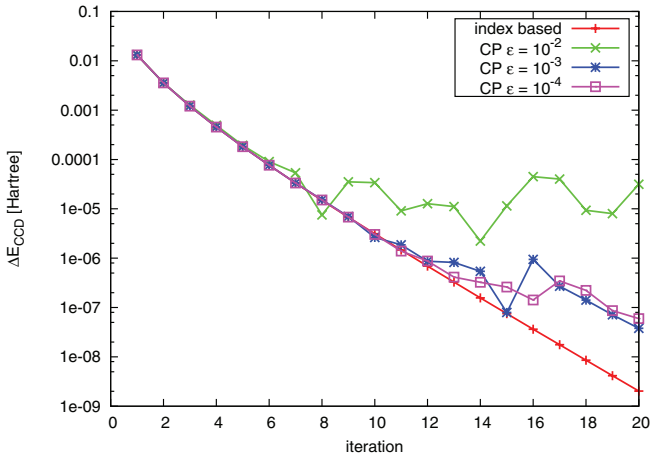

FIG. 4. Convergence of CCD correlation energy using CP tensors for CH_4 in 6-31G basis set.

an initial guess. By adjusting all representing vectors, the error in the ℓ^2-norm can be decreased significantly so that only a few new representing vectors are needed in order to reach the given error in the ℓ^2-norm. Thus, the rank is only increased slightly but the computational time for the rank reduction can be lowered significantly, as a large part of the iterative rank reduction procedure is skipped. A comparison between the evolution of ranks using reduction with initial guess and reductions starting from scratch is given in Table V for H_2O using the 6-31G basis set. From Fig. 1, it can also be seen that this hardly has any influence on the results, except for, as expected, fairly large thresholds.

The reduced ranks of the amplitudes are almost constant during the iterations in all cases, so the increased ranks of the residual and its increments can always efficiently be reduced. Using the initial guess from previous iteration increases the ranks slightly compared to the reduction from scratch. Nevertheless, the ranks approach a constant value after the first few iterations and it can be concluded that the reduction scheme is capable of including small changes in the tensors by adjusting the representation vectors in the initial guess. Thus, the rank reduction during the iterative solution of the amplitude

TABLE V. Reduced ranks of t_2 amplitudes in CP format for H_2O using 6-31G basis set during iterative solution of amplitude equations. Shown are two different reduction procedures: starting each reduction from scratch or reductions with initial guess from previous iteration.

	Reduction from scratch			Reduction with guess		
Iterations	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$
1	48	76	87	50	82	101
2	48	76	88	52	85	108
3	48	76	87	53	85	110
4	48	75	88	53	85	111
5	48	76	88	52	85	111
6	48	77	89	54	85	111
7	48	76	88	53	85	111
8	48	77	87	53	85	111
9	48	76	88	53	85	111
10	48	77	88	53	85	111

TABLE VI. Scaling for the rank for the CCD T_2 amplitudes for a LiH chain in the STO-3G basis and H_2O with different basis sets. Scaling has been estimated by a fit of $a\,x^b$.

	Rank (T_2)		
System	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$
H_2O /STO-3G	9	14	15
H_2O /6-31G	56	85	111
H_2O /cc-pVDZ	154	226	305
H_2O /tzp	184	298	420
	$N^{1.4}$	$N^{1.5}$	$N^{1.6}$
LiH /STO-3G	8	9	9
$(LiH)_2$	25	34	50
$(LiH)_3$	40	67	120
$(LiH)_4$	59	106	191
$(LiH)_4$	82	154	292
	$N^{1.4}$	$N^{1.7}$	$N^{1.9}$

equation can be done efficiently using an initial guess from the previous iteration. First tests for small molecules (water and methane) show that the time for the rank reduction can be reduced at least by a factor of two. For larger molecules even larger factors can be expected, as a large part of the fitting procedure (constructing new representing vectors and adjust them one by one) can be avoided. While we only use this approach for the residual tensor here, it can also be applied to intermediate quantities during the tensor contractions. There, the low-rank approximation from the previous iteration could be used to speed up rank reduction of the next iteration. As storage in the CP format is not expensive (only short representing vectors are stored) all intermediate quantities can be stored and reused in the next iteration.

As mentioned before, a crucial quality of the tensor representation is the scaling of the ranks with system and basis set size. While in a previous study it has been found that the rank for the two electron integrals roughly scales as N^2, the scaling of the rank of the amplitudes has so far only been estimated by the zeroth order amplitudes as obtained from MP2. There, it is found that the rank scales between $N^{1.4}$ and $N^{2.1}$ for thresholds between 10^{-2} and 10^{-4}. In order to assess whether this is also true for iterative amplitudes, a test has been carried out using the algorithm discussed above for a LiH chain of increasing size and H_2O with different basis sets. The results are presented in Table VI.

In agreement with previous investigations it is found that the scaling of the T_2-rank with system size is in the order of N to N^2 depending on the threshold chosen. This also holds for the scaling with basis set size, which is found to be somewhat lower than the scaling with system size.

V. CONCLUSIONS

In this work, we present a CC algorithm in which the iterative procedure for solving the amplitude equations is carried out using tensors represented in the CP format. During successive tensor operations, the ranks increase and have to be reduced using a reduction algorithm. In general, tensor contraction costs scale with the ranks of the two corresponding tensors that are contracted and the contraction can be evaluated as simple scalar products between the representing vectors. Therefore, any (!) contraction between amplitudes (rank Q) and the two-electron integrals (rank R_μ) scales approximately with $\mathcal{O}(N \cdot Q \cdot R_\mu)$, where N is a measure of system size (orbital index). Taking into account the scaling of the rank of the amplitudes and integrals (see also Ref. 60), according to our results, the overall scaling of the tensor contraction is approximately $\mathcal{O}(N^4)$ to $\mathcal{O}(N^5)$. However, additionally to the computational costs for the tensor contraction the costs for the rank reduction has to be considered for the overall computational cost. The rank reduction procedure scales also approximately as $\mathcal{O}(N^5)$ if a sliced reduction is used, but the current scheme has a large prefactor. Thus, the time determining step is the rank reduction procedure itself, that has to be used during the iterative procedure multiple times in order to maintain a low-rank representation of the tensors. Herein, we also propose schemes to minimize the increase of ranks during contraction and show that rank reduction costs can be further reduced by utilizing quantities of previous iterations to initialize the reduction algorithm.

Numerical tests prove the robustness of the applied approximation. The reduced ranks stay almost constant during the iterative procedure and the overall error in the calculated correlation energy is approximately two orders of magnitude lower than the desired deviations in the ℓ^2-norm of the corresponding tensors. No significant error propagation due to the approximation of amplitudes in every iteration is observed. The CC iterations can be converged to changes in energy that are two orders of magnitude lower than the given threshold for the reduction procedure.

As, in its current status, the tensor representation scheme we present is characterized not only by low computational scaling but also by a very large prefactor, our current efforts are focused at devising a Full CI scheme in the CP format with rigorous control of accuracy and sub-exponential overall scaling.

ACKNOWLEDGMENTS

The authors would like to acknowledge Professor W. Hackbusch and Dr. M. Espig from the Max Planck Institute for Mathematics in the Sciences, Leipzig, for a lively and fruitful collaboration and Professor R. Schneider (TU Berlin) for various discussions on tensor decomposition related topics.

[1] R. J. Bartlett, "Coupled cluster theory: An overview of recent developments," in *Modern Electronic Structure Theory* (World Scientific, Singapore, 1995), Vol. II, pp. 1047–1131.
[2] R. Bellman, *Adaptive Control Processes: A Quided Tour* (Princeton University Press, Princeton, NJ, 1961).
[3] S. Wilson, Comput. Phys. Commun. **58**, 71 (1990).
[4] I. Røeggen and E. Wisløff-Nilssen, Chem. Phys. Lett. **132**, 154 (1986).
[5] R. Lindh, U. Ryu, and B. Liu, J. Chem. Phys. **95**, 5889 (1991).
[6] L. Boman, H. Koch, and A. S. de Merás, J. Chem. Phys. **129**, 134107 (2008).
[7] H. Koch, A. S. de Merás, and T. B. Pedersen, J. Chem. Phys. **118**, 9481 (2003).
[8] F. Aquilante, T. B. Pedersen, and R. Lindh, J. Chem. Phys. **126**, 194106 (2007).
[9] I. Røeggen and T. Johansen, J. Chem. Phys. **128**, 194107 (2008).

[10]F. Aquilante, L. Gagliardi, T. B. Pedersen, and R. Lindh, J. Chem. Phys. **130**, 154107 (2009).
[11]F. Weigend, M. Kattannek, and R. Ahlrichs, J. Chem. Phys. **130**, 164106 (2009).
[12]F. R. Manby, J. Chem. Phys. **119**, 4607 (2003).
[13]H. J. Werner, F. R. Manby, and P. J. Knowles, J. Chem. Phys. **118**, 8149 (2003).
[14]M. Schütz and F. R. Manby, Phys. Chem. Chem. Phys. **5**, 3349 (2003).
[15]R. Kendall and H. Früchtl, Theor. Chem. Acc. **97**, 158 (1997).
[16]O. Vahtras, J. Almlöf, and M. Feyereisen, Chem. Phys. Lett. **213**, 514 (1993).
[17]M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).
[18]R. Ahlrichs, Phys. Chem. Chem. Phys. **6**, 5119 (2004).
[19]W. Kutzelnigg, Theor. Chim. Acta. **68**, 445 (1985).
[20]E. F. Valeev and H. F. Schaefer, J. Chem. Phys. **113**, 3990 (2000).
[21]J. Rychlewski, *Explicitly Correlated Wave Functions in Chemistry and Physics: Theory and Applications* (Kluwer Academic Publishers, Dordrecht, Boston, 2003).
[22]S. Ten-No, Chem. Phys. Lett. **398**, 56 (2004).
[23]W. Klopper, F. R. Manby, S. Ten-No, and E. F. Valeev, Int. Rev. Phys. Chem. **25**, 427 (2006).
[24]H. J. Werner, T. B. Adler, G. Knizia, and F. R. Manby, "Coupled cluster theory: An overview of recent developments," in *Recent Progress in Coupled Cluster Methods: Theory and Applications* (Springer, London, 2010), pp. 573–620.
[25]H. J. Werner, G. Knizia, and F. R. Manby, Mol. Phys. **109**, 407 (2011).
[26]U. Benedikt, H. Auer, M. Espig, W. Hackbusch, and A. A. Auer, Mol. Phys. **111**, 2398 (2013).
[27]W. Kutzelnigg, Theor. Chim. Acta **1**, 327 (1963).
[28]R. Ahlrichs, W. Kutzelnigg, and W. A. Bingel, Theor. Chim. Acta **5**, 289 (1966).
[29]R. K. Nesbet, Phys. Rev. **175**, 2 (1968).
[30]W. Meyer, Int. J. Quantum Chem. **5**, 341 (1971).
[31]W. Meyer, J. Chem. Phys. **58**, 1017 (1973).
[32]F. Neese, A. Hansen, and D. G. Liakos, J. Chem. Phys. **131**, 064103 (2009).
[33]F. Neese, F. Wennmohs, and A. Hansen, J. Chem. Phys. **130**, 114108 (2009).
[34]F. Neese, A. Hansen, F. Wennmohs, and S. Grimme, Acc. Chem. Res. **42**, 641 (2009).
[35]A. Anoop, W. Thiel, and F. Neese, J. Chem. Theory Comput. **6**, 3137 (2010).
[36]C. Riplinger and F. Neese, J. Chem. Phys. **138**, 034106 (2013).
[37]J. Yang, Y. Kurashige, F. R. Manby, and G. K. L. Chan, J. Chem. Phys. **134**, 044123 (2011).
[38]J. Yang, G. K.-L. Chan, F. R. Manby, M. Schütz, and H.-J. Werner, J. Chem. Phys. **136**, 144105 (2012).
[39]Y. Kurashige, J. Yang, G. K. L. Chan, and F. R. Manby, J. Chem. Phys. **136**, 124106 (2012).
[40]M. Schütz, J. Yang, G. K.-L. Chan, F. R. Manby, and H.-J. Werner, J. Chem. Phys. **138**, 054109 (2013).
[41]P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[42]P. Pulay and S. Saebo, Theor. Chim. Acta **69**, 357 (1986).
[43]S. Saebø and P. Pulay, J. Chem. Phys. **86**, 914 (1987).
[44]L. Adamowicz and R. J. Bartlett, J. Chem. Phys. **86**, 6314 (1987).
[45]P. Neogrády, M. Pitoňák, and M. Urban, Mol. Phys. **103**, 2141 (2005).
[46]M. Pitoňák, P. Neogrády, V. Kellö, and M. Urban, Mol. Phys. **104**, 2277 (2006).
[47]M. Pitoňák, F. Holka, P. Neogrády, and M. Urban, J. Mol. Struct.: THEOCHEM **768**, 79 (2006).
[48]L. Adamowicz, Mol. Phys. **108**, 3105 (2010).
[49]C. Eckart and G. Young, Psychometrika **1**, 211 (1936).
[50]J. D. Carroll and J. J. Chang, Psychometrika **35**, 283 (1970).
[51]G. Golub, *Matrix Computations* (Johns Hopkins University Press, Baltimore, 1996).
[52]T. Kinoshita, O. Hino, and R. J. Bartlett, J. Chem. Phys. **119**, 7756 (2003).
[53]O. Hino, T. Kinoshita, and R. J. Bartlett, J. Chem. Phys. **121**, 1206 (2004).
[54]E. G. Hohenstein, R. M. Parrish, and T. J. Martinez, J. Chem. Phys. **137**, 044103 (2012).
[55]R. M. Parrish, E. G. Hohenstein, T. J. Martinez, and C. D. Sherrill, J. Chem. Phys. **137**, 224106 (2012).
[56]E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, and T. J. Martinez, J. Chem. Phys. **137**, 221101 (2012).
[57]R. A. Harshman and M. E. Lundy, Comput. Stat. Data Anal. **18**, 39 (1994).
[58]N. Shenvi, H. Van Aggelen, Y. Yang, W. T. Yang, C. Schwerdtfeger, and D. Mazziotti, J. Chem. Phys. **139**, 054110 (2013).
[59]E. G. Hohenstein, S. I. L. Kokkila, R. M. Parrish, and T. J. Martinez, J. Chem. Phys. **138**, 124111 (2013).
[60]U. Benedikt, A. A. Auer, M. Espig, and W. Hackbusch, J. Chem. Phys. **134**, 054118 (2011).
[61]U. Benedikt, "Low-rank tensor approximation in post Hartree-Fock methods," Ph.D. thesis, TU Chemnitz, 2013.
[62]F. Hitchcock, J. Math. Phys. **6**, 164 (1927).
[63]T. G. Kolda and B. W. Bader, SIAM Rev. **51**, 455 (2009).
[64]C. F. Van Loan and N. Pitsianis, "Approximation with Kronecker products," in *Linear Algebra for Large Scale and Real-Time Applications*, NATO Advanced Science Institutes Series E: Applied Sciences Vol. 232 (Kluwer, Dordrecht, 1993), pp. 293–314.
[65]T. Zhang and G. H. Golub, SIAM J. Matrix Anal. Appl. **23**, 534 (2001).
[66]G. Beylkin and M. J. Mohlenkamp, P. Natl. Acad. Sci. U.S.A. **99**, 10246 (2002).
[67]G. Beylkin and M. J. Mohlenkamp, SIAM J. Sci. Comput. **26**, 2133 (2005).
[68]M. Espig, "Effiziente Bestapproximation mittels Summen von Elementartensoren in hohen Dimensionen," Ph.D. thesis, Universität Leipzig, 2008.
[69]M. Espig, L. Grasedyck, and W. Hackbusch, Constructive Approx. **30**, 557 (2009).
[70]M. Espig and W. Hackbusch, Numer. Math. **122**, 489 (2012).
[71]M. Espig, W. Hackbusch, T. Rohwedder, and R. Schneider, Numer. Math. **122**, 469 (2012).
[72]J. F. Stanton, J. Gauss, J. D. Watts, and R. J. Bartlett, J. Chem. Phys. **94**, 4334 (1991).
[73]In principle, the rank depends on the type of Fock matrix elements that are needed and is equal to the number of occupied orbitals for f^i_j or to the number of virtual orbitals for f^a_b, respectively.
[74]D. Braess and W. Hackbusch, IMA J. Numer. Anal. **25**, 685 (2005).
[75]Note that this intermediate format has some resemblance to the THC format proposed by Hohenstein *et al*.[54]
[76]A. R. Hoy, I. M. Mills, and G. Strey, Mol. Phys. **24**, 1265 (1972).
[77]Note that for practical reasons the rank of the denominator can be decreased to values around 15 still maintaining an accuracy of the numerical integration of the Laplace transformed denominator of 6.311×10^{-7} in an interval from zero to 2×10^6.[74] However, the larger value of $S = 42$ is used throughout this work.
[78]S. Hirata, J. Phys. Chem. A **107**, 9887 (2003).
[79]A. A. Auer, G. Baumgartner, D. E. Bernholdt, A. Bibireata, V. Choppella, D. Cociorva, X. Y. Gao, R. Harrison, S. Krishnamoorthy, S. Krishnan, C. C. Lam, Q. D. Lu, M. Nooijen, R. Pitzer, J. Ramanujam, P. Sadayappan, and A. Sibiryakov, Mol. Phys. **104**, 211 (2006).
[80]M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, and W. de Jong, Comput. Phys. Commun. **181**, 1477 (2010).