JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

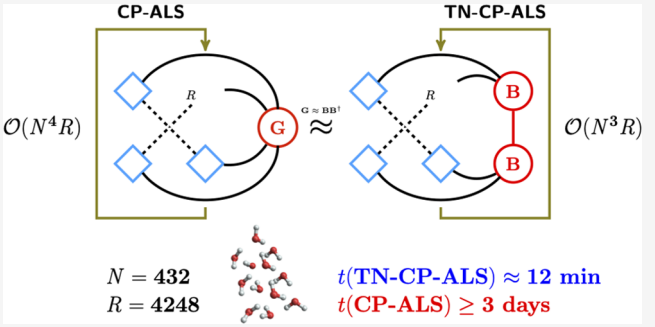
Efficient Construction of Canonical Polyadic Approximations of Tensor Networks

Karl Pierce and Edward F. Valeev*

Cite This: J. Chem. Theory Comput. 2023, 19, 71−81 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: We consider the problem of constructing a canonical polyadic (CP) decomposition for a tensor network, rather than a single tensor. We illustrate how it is possible to reduce the complexity of constructing an approximate CP representation of the network by leveraging its structure in the course of the CP factor optimization. The utility of this technique is demonstrated for the order-4 Coulomb interaction tensor approximated by two order-3 tensors via an approximate generalized square-root (SQ) factorization, such as density fitting or (pivoted) Cholesky. The complexity of constructing a four-way CP decomposition is reduced from $O(n^4R_{CP})$ (for the non-approximated Coulomb tensor) to $O(n^3R_{CP})$ (for the SQ-factorized Coulomb tensor), where n and R_{CP} are the basis and CP ranks, respectively. This reduces the cost of constructing the CP approximation of two-body interaction tensors of relevance to accurate many-body electronic structure by up to 2 orders of magnitude for systems with up to 36 atoms studied here. The full four-way CP approximation of the Coulomb interaction tensor is shown to be more accurate than the known approaches which utilize CP-factorizations of the SQ factors (which are also constructed with an $O(n^3R_{CP})$ cost), such as the algebraic pseudospectral and tensor hypercontraction approaches. The CP-decomposed SQ factors can also serve as a robust initial guess for the four-way CP factors.



1. INTRODUCTION

The canonical polyadic (CP) decomposition,[1,2] also known as CANDECOMP,[3] PARAFAC,[4] separated representation,[5] and other names,[6] is a popular data-sparse representation of higher-order tensors as a data analysis tool in many settings as well as a way to reduce the cost and/or complexity of the tensor algebra in some physical simulation applications. The CP decomposition of an order-n tensor can be viewed as a tensor network (TN) of n factor matrices, each describing a single mode of the target tensor. What makes the CP TN unique is that each factor matrix is connected directly to every other factor via a single hyperedge. By treating all modes of the target tensor on equal footing, the CP decomposition has a universal structure that simplifies its use compared to other TNs with more specialized topologies (such as matrix product states, etc.). Formally, the CP decomposition reduces the storage complexity of higher-order tensors from exponential to linear in the tensor order. However, finding the minimal extent of the CP hyperedge (CP rank) sufficient for exact CP representation of a given tensor or its approximate CP representation to a fixed given precision is hard.[7,8]

Whether CP decomposition and/or TNs with CP-like hyperedges can be gainfully used in quantum many-body simulation, despite the challenges of computing the CP decomposition, is an open question. Beylkin and Mohlenkamp motivated broad utility of polyadic separation of multidimensional functions and operators (i.e., an infinite-dimensional analogue of CP decomposition) and illustrated its effectiveness for model quantum simulation problems.[5] The only uses of the CP decomposition for global approximation of the electronic Hamiltonian and/or wave function have been demonstrated by Auer and co-workers;[9−11] the scope of these applications was limited to small systems by the high cost of the computing and updating the CP decomposition. More limited uses of the CP decomposition appeared in the context of compressed spectral element representation for constructing numerical (real-space) orbitals by Bischoff and Valeev[12] and for compressing real-space representations of various quantities in the electronic structure context by Chinnamsetty et al.[13] More common scenario in the quantum many-body simulation is the

Received: August 21, 2022
Published: December 9, 2022

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81

appearance of tensor networks featuring CP-like hyperedges, e.g., as a result of approximating an integral by a quadrature. Such examples include the approximation of the Coulomb operator in the pseudospectral (PS) family of methods,[14] the related tensor hypercontraction (THC) factorization of two-particle tensors[15] and its close cousin, interpolative separable density fitting,[16] robust CP-DF factorization marrying the PS and THC approaches,[17] and the CD-SVD compression of the Coulomb interaction tensor.[18] These decompositions have productive uses in quantum simulation contexts on both classical[17−21] and quantum hardware.[22] Finally, the use of Laplace transform in the electronic structure[23] also leads to networks with CP-like hyperedges.

Our recent experiences[17] suggest that more productive uses of the CP factorization can be found when focusing on approximating not individual tensors but on tensor networks. Thus, here we consider a problem of constructing a CP approximation for a tensor network. To the best of our knowledge, such class of problems has not been yet considered. It is clear that leveraging the structure of the tensor network in the course of constructing its CP approximation can lead to notable savings. Here we show that that is indeed the case in a practically important scenario; namely, we show how to construct a full (four-way) CP decomposition of a two-particle interaction represented in a molecular orbital basis with reduced complexity by leveraging the approximate generalized square root (SQ) factorization of such a tensor. This allows us to construct the full CP decomposition of two-particle interaction tensors for systems of unprecedented sizes. Such productive combination of CP and other tensor network approximations can be further improved by other CP acceleration strategies such as parallelization[24] and improved solver strategies.[25−27]

The rest of the manuscript is organized as follows. Section 2 introduces the mathematical formalism for SQ-accelerated CP4 decomposition and its application to the two-body interaction tensors in quantum electronic structure. Technical details of the computational experiments are described in Section 3. In Section 4, we compared the accuracy of the CP4 decomposition to the more common three-way CP decompositions as well as illustrated the efficiency of the accelerated CP4 solver relative to its standard (naive) counterpart. Our findings are summarized in Section 5.

2. FORMALISM

Our objective is to approximate a two-particle interaction represented in a generic basis $\{\phi_p\}$

$$g_{st}^{pq} = \iint \phi_s^*(\mathbf{r}_1)\phi_t^*(\mathbf{r}_2)g(\mathbf{r}_1, \mathbf{r}_2)\phi_p(\mathbf{r}_1)\phi_q(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \tag{1}$$

all numerical experiments in this work will use $g(\mathbf{r}_1,\mathbf{r}_2) \equiv |\mathbf{r}_1 - \mathbf{r}_2|^{-1}$ (Coulomb interaction), but other positive kernels can be used. A CP factorization of g is represented using a set of factor matrices $\mathbf{x} = \{\gamma,\rho,\beta,\kappa\}$

$$g_{st}^{pq} \overset{\text{CP4}}{\approx} \hat{g}_{st}^{pq} = \sum_r^{R_{\text{CP4}}} \lambda_r \gamma_r^p \rho_r^q \beta_s^r \kappa_t^r \tag{2}$$

moniker "CP4" will be used to help distinguish such four-way CP factorization from other factorizations involving CP decomposition.

In this work, we impose the CP factors to be column-wise unit-normalized, i.e., $\sum_p |\gamma_r^p|^2 = 1$, etc. In general, the CP "singular values" λ_r can be absorbed into the factors;[a] the weighted factors will be denoted by a caron, $\check{\gamma}_r^p \equiv \lambda_r \gamma_r^p$. Thus, eq 2 then can be written equivalently using weighted factors

$$\begin{aligned} g_{st}^{pq} &\overset{\text{CP4}}{\approx} \sum_r^{R_{\text{CP4}}} \check{\gamma}_r^p \rho_r^q \beta_s^r \kappa_t^r = \sum_r^{R_{\text{CP4}}} \gamma_r^p \check{\rho}_r^q \beta_s^r \kappa_t^r = \sum_r^{R_{\text{CP4}}} \gamma_r^p \rho_r^q \check{\beta}_s^r \kappa_t^r \\ &= \sum_r^{R_{\text{CP4}}} \gamma_r^p \rho_r^q \beta_s^r \check{\kappa}_t^r \end{aligned} \tag{3}$$

There exists no finite algorithm to determine the exact or approximate (sufficient for a given accuracy) CP rank of a general tensor.[7,8] Thus, in practice, the CP rank is determined heuristically. For a fixed CP rank, eq 2 is typically constructed by minimizing the norm of the error,

$$f(\mathbf{x}) \equiv \frac{1}{2} \sum_{pqst} |g_{st}^{pq} - \hat{g}_{st}^{pq}|^2 \tag{4}$$

The nonlinear optimization uses well-known heuristics [alternating least squares (ALS),[3,5,29] accelerated gradient,[9,30−33] nonlinear least squares,[34] etc.] whose cost-determining step is the evaluation of the objective function's gradient (or its components)

$$\begin{aligned} \frac{\partial f}{\partial \check{\gamma}_r^p} &= \sum_{qst} (\hat{g}_{pq}^{st} - g_{pq}^{st}) \rho_r^q \beta_s^r \kappa_t^r \\ &= \left(\sum_{r'} \check{\gamma}_p^{r'} (\mathbf{W}^{\{\gamma\}})_{r'}^{r} - \sum_{qst} g_{pq}^{st} \rho_r^q \beta_s^r \kappa_t^r \right) \end{aligned} \tag{5}$$

where

$$(\mathbf{W}^{\{\gamma\}})_{r'}^{r} \equiv \left(\sum_q \rho_q^{r'} \rho_r^q \right) \left(\sum_s \beta_{r'}^s \beta_s^r \right) \left(\sum_t \kappa_r^t \kappa_t^r \right) \tag{6}$$

The simplest way to update CP factors is via ALS, by (1) solving the linear least squares problem $\partial f/\partial \check{\gamma}_p^r = 0$ for $\check{\gamma}$[b] and (2) normalizing its columns to obtain updated γ and λ, and repeating for other factors until the desired ALS convergence criterion (ϵ_{ALS} has been achieved (here we use the criterion of ref 17)), and repeating as necessary. Computing the last term in eq 5, which is needed both in ALS and in the gradient-based approaches, is the cost-determining step, with a complexity of $O(N^4R)$, where N is proportional to the mode dimension and R is the CP rank (in practice $R \propto O(N)$). The high-order cost of the gradient evaluation is aggravated by the difficulty of the optimization problem (strong nonlinearity, sensitivity to the initial guess, and often poor conditioning; e.g., see ref 35). These factors explain why the CP4 decomposition has barely been explored in the context of quantum simulation, as mentioned in Section 1.

Popular CP-like approximations of g can be determined by starting from the generalized[c] square root (SQ) factorization

$$g_{st}^{pq} \overset{\text{SQ}}{\approx} \sum_X B_s^{pX} B_t^{qX} \tag{7}$$

The two most popular ways to determine the square root factor B are the density fitting[36−39] (DF) approximation, in which X represents a predetermined fitting ("auxiliary") basis that grows linearly with the system size, and the rank-revealing Cholesky decomposition (CD).[40−43] The error of eq 7 can be

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81

controlled in the case of DF using a more complete fitting basis, or in the case of CD by increasing the Cholesky rank.

Although the generalized square root factorization is extremely popular, due to its ability to reduce the storage complexity of g from N^4 to N^3, it rarely reduces the computational complexity of electronic structure methods by itself. However, by combining eq 7 with the CP decomposition, it is possible to reduce the computational complexity for several important algorithms, such as the exact exchange evaluation.[14,44] Traditionally, this is accomplished by CP factorizing the tensor B

$$B_s^{pX} \approx \hat{B}_s^{pX} = \sum_r^{R_{\mathrm{CP3}}} \lambda_r \gamma_r^p \beta_s^r \chi_r^X = \sum_r^{R_{\mathrm{CP3}}} \gamma_r^p \beta_s^r \check{\chi}_r^X \tag{8}$$

where R_{CP3} is the CP rank of this decomposition; we will refer to a three-way CP decomposition by moniker "CP3". Note that although eq 8 does not impose the hermiticity of B (with respect to swapping of p and s), in practice $\gamma_r^p = (\beta_p^r)^*$ can be imposed.[17,45] The CP3 approximation can then be introduced into eq 7

- once

$$g_{st}^{pq} \approx \sum_X \hat{B}_s^{pX} B_t^{qX} = \sum_r^{R_{\mathrm{CP3}}} \gamma_r^p \beta_s^r \left(\sum_X \check{\chi}_r^X B_t^{qX} \right) \equiv \sum_r^{R_{\mathrm{CP3}}} \gamma_r^p \beta_s^r \tilde{B}_{qr}^t \tag{9}$$

  to produce a factorization of the pseudospectral (PS) method;[14,46−54]

- twice,

$$g_{st}^{pq} \approx \sum_X \hat{B}_s^{pX} \hat{B}_t^{qX} = \sum_{r,r'}^{R_{\mathrm{CP3}}} \gamma_r^p \beta_s^r \gamma_{r'}^q \beta_t^{r'} \left(\sum_X \check{\chi}_r^X \check{\chi}_{r'}^X \right) \equiv \sum_{r,r'}^{R_{\mathrm{CP3}}} \gamma_r^p \beta_s^r \gamma_{r'}^q \beta_t^{r'} \tilde{\chi}_{r,r'} \tag{10}$$

  to produce the tensor hypercontraction format;[15,20,44,45,55−60] or

- robustly[17] using a combination of single and double substitutions

$$g_{st}^{pq} \approx \sum_X \hat{B}_s^{pX} B_t^{qX} + \sum_X B_s^{pX} \hat{B}_t^{qX} - \sum_X \hat{B}_s^{pX} \hat{B}_t^{qX} \tag{11}$$

The latter, denoted as rCP3-DF, has been shown recently[17] to be more accurate than either eqs 9 or 10. As we demonstrate next, the "square-root" factorization can be also useful as a precursor to the CP4 approximation of g. Specifically, its use allows one to reduce the complexity of constructing the CP4 approximation from $O(N^5)$ to $O(N^4)$.

As mentioned above, evaluation of the gradient of the CP objective function, eq 5, costs $O(N^4R)$. However, it is possible to reduce the complexity of the CP optimization problem by replacing tensor g with its SQ-factorization, eq 7

$$\frac{\partial f}{\partial \check{\gamma}_r^p} \overset{\mathrm{SQ}}{=} \sum_{qst} \left(\hat{g}_{pq}^{st} - \sum_X \sum_X B_{pX}^s B_{qX}^t \right) \rho_r^q \beta_s^r \kappa_t^r = \left(\sum_{r'} \check{\gamma}_p^{r'} (\mathbf{W}^{\{\gamma\}})_{r'}^r - \sum_X (B_{pX}^s \beta_s^r) \left(\sum_{qt} B_{qX}^t \rho_r^q \kappa_t^r \right) \right) \tag{12}$$

Using the SQ-network approximated CP4 loss function, we reduce the cost of the gradient of CP4 from $O(N^4R)$ to $O(N^2RX)$. Furthermore, we reduce the storage requirements of the approach from $O(N^4 \approx N^3R)$ to $O(N^3 \approx N^2R)$ Although others have introduced TNs into the CP loss function,[59] this work is first to construct factor matrices for only the external modes of a TN. The CP decomposition, in this way, provides a route to approximate the result of a TN without a priori construction of the tensor from the network. Clearly, such use of CP can be more generally useful than just for the case of the Coulomb integrals.

Now, formally, the CP4 ALS optimization has the same computation scaling as the CP3 ALS optimization, $O(N^4)$. However, it is known that the efficiency of the CP optimization process and the accuracy of the resulting factorization can strongly depend on an initial guess. In this work, we utilize an initial guess scheme that uses the CP3-SQ-factorization of g as well as vectors filled with quasi-random numbers taken from a uniform distribution on $[-1,1]$. To motivate our initialization strategy, we compare the CP4-approximated g to its THC-like CP3-SQ counterpart

$$g_{st}^{pq} \overset{\mathrm{eq}\ 2}{\approx} \sum_r^{R_{\mathrm{CP4}}} \gamma_r^p \beta_s^r \rho_r^q \kappa_t^r \lambda_r \tag{13}$$

$$g_{st}^{pq} \overset{\mathrm{eq}\ 10}{\approx} \sum_{r,r'}^{R_{\mathrm{CP3}}} \gamma_r^p \beta_s^r \gamma_r^q \beta_t^{r'} \tilde{\chi}_{r,r'} \tag{14}$$

The key differences between these expressions are: (1) factors γ and β in eqs 13 and 14 were determined to CP-approximate different tensors (g_{st}^{pq} and B_s^{pX}, respectively), and (2) eq 14 is explicitly particle-symmetric. Importantly, however, the CP3-SQ approximation (eq 14) can be rewritten in the CP4 form exactly by merging indices into a composite index $w \equiv \{r, r'\}$

$$g_{st}^{pq} \overset{\mathrm{eq}\ 14}{\approx} \sum_w^{R_{\mathrm{CP3}}^2} \gamma_w^p \beta_s^w \rho_w^q \kappa_t^w \lambda_w \tag{15}$$

where

$$\gamma_w^p \equiv \gamma_{\{r,r'\}}^p \equiv \gamma_r^p \otimes \mathbb{I}_{r'} \tag{16}$$

$$\beta_s^w \equiv \beta_s^{\{r,r'\}} \equiv \beta_s^r \otimes \mathbb{I}_{r'} \tag{17}$$

$$\rho_w^q \equiv \rho_{\{r,r'\}}^q \equiv \gamma_{r'}^q \otimes \mathbb{I}_r \tag{18}$$

$$\kappa_t^w \equiv \kappa_t^{\{r,r'\}} \equiv \beta_t^{r'} \otimes \mathbb{I}_r \tag{19}$$

$$\lambda_w \equiv \tilde{\chi}_{r,r'} \tag{20}$$

such that $\otimes$represents the outer product and $\mathbb{I}_r$ is an identity vector with dimension r. In practical applications of THC, the CP3 rank is $O(N)$; thus, it may appear that the CP4 rank required to represent it exactly is $O(N^2)$. However since

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81

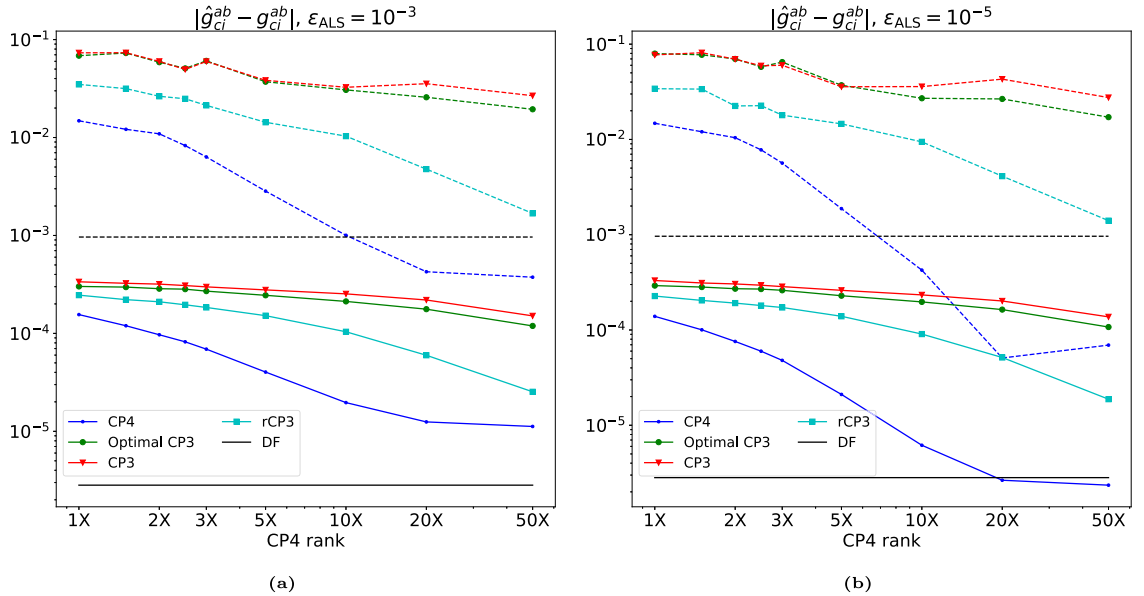


Figure 1. Absolute average element error for approximated g^{ab}_{ci} tensor using an ALS solver threshold[17] of (a) $\epsilon_{ALS} = 10^{-3}$ and (b) $\epsilon_{ALS} = 10^{-5}$. Solid lines represent mean error, and dashed lines represent max error. Note that the CP4 rank is in units of a factor times the DF fitting basis, a metric which grows linearly with rank. Data were collected using the S66/5 dataset with an aVDZ/aVDZ-RI basis.

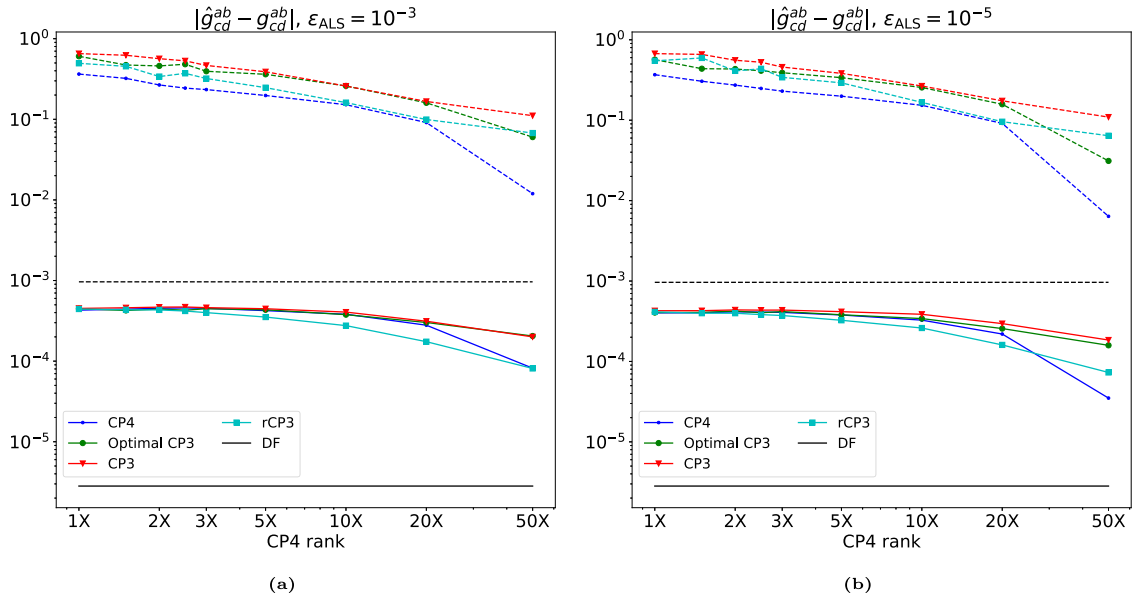


Figure 2. Absolute average element error for approximated g^{ab}_{cd} tensor using an ALS solver threshold[17] of (a) $\epsilon_{ALS} = 10^{-3}$ and (b) $\epsilon_{ALS} = 10^{-5}$. Solid lines represents mean error, and dashed lines represent max error. Note that the CP4 rank is in units of a factor times the DF fitting basis, a metric which grows linearly with rank. Data were collected using the S66/5 dataset with an aVDZ/aVDZ-RI basis.

FMM[61] and related approximations[62] can be used to construct representations of sufficiently well-behaved integral operators with $O(N \log N)$ complexity, this limits the asymptotic CP4 rank of g to $O(N \log N)$ as well. In the context of THC, this asymptotic limit means that the $\tilde{\chi}$ matrix can be made sparse for sufficiently large systems by an appropriate transformation. However, for small systems where the $\tilde{\chi}$ matrix is dense, direct construction of CP4 can be more compact than the THC representation, i.e., a CP4 approximation with an accuracy of ϵ_{THC} can be found via an optimization of a rank $R_{CP4} \ll R^2_{CP3}$ CP4 ALS decomposition (see Section 4 for details).

The formal connection between eqs 2 and 14 via eq 15 thus suggests the use of unit-normalized CP3 factors as the initial guess for the CP4 decomposition. Further, because the CP optimization recomputes λ every iteration, it is not necessary to keep elements of the diagonalized $\tilde{\chi}$ matrix (in eq 14). If the target CP4 rank is greater than the CP3 rank, the CP4 factors can be padded with vectors of quasi-random numbers taken from the uniform distribution on $[-1,1]$ and normalized to unity. This is sufficient to bootstrap the ALS CP4 solver since, again, the singular values are not needed to compute the gradient with respect to the weighted factors (see eq 5).

3. COMPUTATIONAL DETAILS

The CP3 and CP4 optimization computations have been computed using the Basic Tensor Algebra Subroutine (BTAS)

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81

software package.[63] All of these computations were run on the Virginia Tech Advanced Research Computing's (ARC) Cascades cluster which utilizes standard nodes that contain 2 Intel Xeon E5-2683 v4 CPUs. The CP approximations were computed using the standard alternating least squares (ALS) method.[5,64] In this work, as was done in our previous work,[17] initial CP3 factors were generated using quasi-random numbers taken from the uniform distribution on $[-1,1]$. As discussed earlier in this work, we utilized both optimized CP3 factors and quasi-random numbers as an initial guess strategy for the CP4 decomposition.

Assessment of the CP factorizations utilized a small subset of five complexes from the S66 (S66/5) benchmark set of weakly-bound complexes.[65,d] The S66 geometries were taken from the Benchmark Energy and Geometry Database (BEGDB).[66] Additionally, the conformer of $(H_2O)_{12}$[67,68] is used to demonstrate the capacity of the reduced-scaling CP4 ALS optimization strategy. All of the above computations utilized the aug-cc-pVDZ[69,70] (aVDZ) orbital basis set (OBS). The two-electron interaction tensors were approximated using standard Coulomb-metric density fitting using the aug-cc-pVDZ-RI (abbreviated as aVDZ-RI) density fitting basis set (DFBS).[71] Assessment of the basis set variation in the performance of CP4 used the following additional OBS/DFBS pairs: the cc-pVDZ-F12 (abbreviated as DZ-F12)[72] OBS paired with the aVDZ-RI DFBS, the aug-cc-pVTZ (aVTZ) OBS[69,70] paired with the aug-cc-pVTZ-RI[71] (aVTZ-RI) DFBS, and the cc-pVTZ-F12[72] (TZ-F12) OBS paired with the aVTZ-RI DFBS.

Finally, to make comparisons between different systems, the CP ranks will be defined in units of size X of the DF fitting (auxiliary) basis, which grows proportionally to the number of atoms in the system.

4. RESULTS

Figures 1 and 2 illustrate the accuracy of the CP4 approximation compared to CP3-DF-based approximations. To compare tensor network approximations to each other, it is necessary to map all of them to a common network topology, i.e., the CP3-DF tensor networks are mapped their CP4 counterpart via eq 15. Since a CP3-DF network with rank R_{CP3} is represented exactly by a CP4 network with rank $R_{\mathrm{CP4}} = R_{\mathrm{CP3}}^2$ (see the discussion around eq 15), the CP4 network of rank R_{CP4} is compared to the CP3 counterparts of rank $R_{\mathrm{CP3}} = \sqrt{R_{\mathrm{CP4}}}$. Furthermore, since the robust CP3-DF approximant (eq 11) is a sum of three contributions, it makes sense to use even a lower CP3 rank ($\sim\sqrt{R_{\mathrm{CP4}}/3}$) compared to a CP4 network. For the sake of simplicity, we conservatively use the same CP3 rank for all CP3-based networks.

In Figures 1 and 2, the {dashed, solid} lines denoted "DF" show the {maximum, average} element-wise errors due to the DF approximation of the exact g tensors, while the other lines show the corresponding element-wise errors associated with the CP decomposition of the DF-approximated g tensors. The lines labeled CP3 correspond to the THC-like CP3-DF approximation of the g tensors (eq 14), and the lines labeled "rCP3" correspond to the robust CP3-DF approximation (eq 11). The lines labeled CP4 correspond to the CP4 decomposition using the DF-approximated g tensor in the CP4 ALS optimization. Finally, the lines labeled "Optimal CP3" correspond to the CP3-DF approximation (eq 14) in which the constituent factors were optimized to make stationary an augmented CP loss function defined for, e.g., g_{ci}^{ab} as

$$f(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{2}\left\| B_c^{aX} B_i^{bX} - \hat{B}_c^{aX} \hat{B}_i^{bX} \right\|^2 + \frac{1}{2}\left\| B_c^{aX} - \hat{B}_c^{aX} \right\|^2 + \frac{1}{2}\left\| B_i^{bX} - \hat{B}_i^{bX} \right\|^2 \tag{21}$$

where $\mathbf{x}_{\{1,2\}}$ are the CP3 parameters defining $\{\hat{B}_c^{aX}, \hat{B}_i^{bX}\}$. The augmented CP loss function is designed to simultaneouly solve three coupled least-squares problems, namely (1) finding the optimal CP3 factors for approximating the (DF-approximated) g_{ci}^{ab} using the CP3-DF network, (2,3) finding the optimal CP3 factors for approximating the $\{B_c^{aX}, B_i^{bX}\}$ DF factors, respectively. Coupling the three CP problems helps the ALS solver to avoid excessively large steps that may occur when solving problem 1 alone, which stems from a lack of convergence guarantees for the ALS algorithm in general.[28] Similar modifications of the objective functions to penalize large steps have been used by others.[22] The use of more robust direct minimization solvers for the CP and general tensor network optimization problems (e.g., nonlinear quasi-Newton solvers with step restriction) should alleviate the need for augmenting the CP3-DF objective function.

As Figures 1 and 2 clearly demonstrate, at equivalent ranks, the CP4 approximation is more accurate than the CP3-DF counterparts. CP4 even surpasses the accuracy of the robust CP3-DF variant, which we recently showed to be more accurate than the corresponding PS-like and THC-like CP3-DF counterparts. Namely, CP4 is always more accurate than rCP3-DF for the g_{ci}^{ab} tensor and surpasses the accuracy of the latter with the largest CP4 rank for the g_{cd}^{ab} tensor. However, note that at modest ranks, the accuracy of the CP4 approximation of g_{cd}^{ab} was found to be lower than that of the robust CP3 approximation; see the solid lines in Figure 2. As discussed above, this is likely due to the overly conservative choice of the relationship between the CP3 and CP4 ranks for the robust case (i.e., the rank for the rCP3 approximation should have been smaller by a factor of $\sqrt{3}$).

CP4 errors usually, but not always, decrease monotonically: for example, the average CP4 error for g_{ci}^{ab} (Figure 1a) does not decrease substantially when the CP rank is increased from $20X$ to $50X$. This can be rationalized if we recall that the error of an approximate CP4 decomposition is due to two factors: finiteness of the CP4 rank (hence controlled by the rank) and the suboptimality of the factors (controlled by the ALS convergence threshold ϵ_{ALS}). As the CP4 rank increases the factor optimization error becomes dominant. Indeed, as the ALS threshold decreases from 10^{-3} to 10^{-5}, this artifact disappears as well as the total errors for large ranks ($\geq 5X$) decrease significantly. This result is consistent with our previous finding for the CP3 optimization.[17] It is well understood that more robust CP solvers[25,73−75] can be employed to improve the convergence rate-related issues. Since our focus here is primarily on addressing the universal concerns (the cost of gradient evaluation, initial guess) that apply to all CP solvers, here, we only use ALS and postpone the use of improved CP solvers to later studies.

Another important phenomenon becomes apparent when we plot the distribution of tensor element magnitudes for the two types of tensors in the two test systems (Figure 3). One of

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81

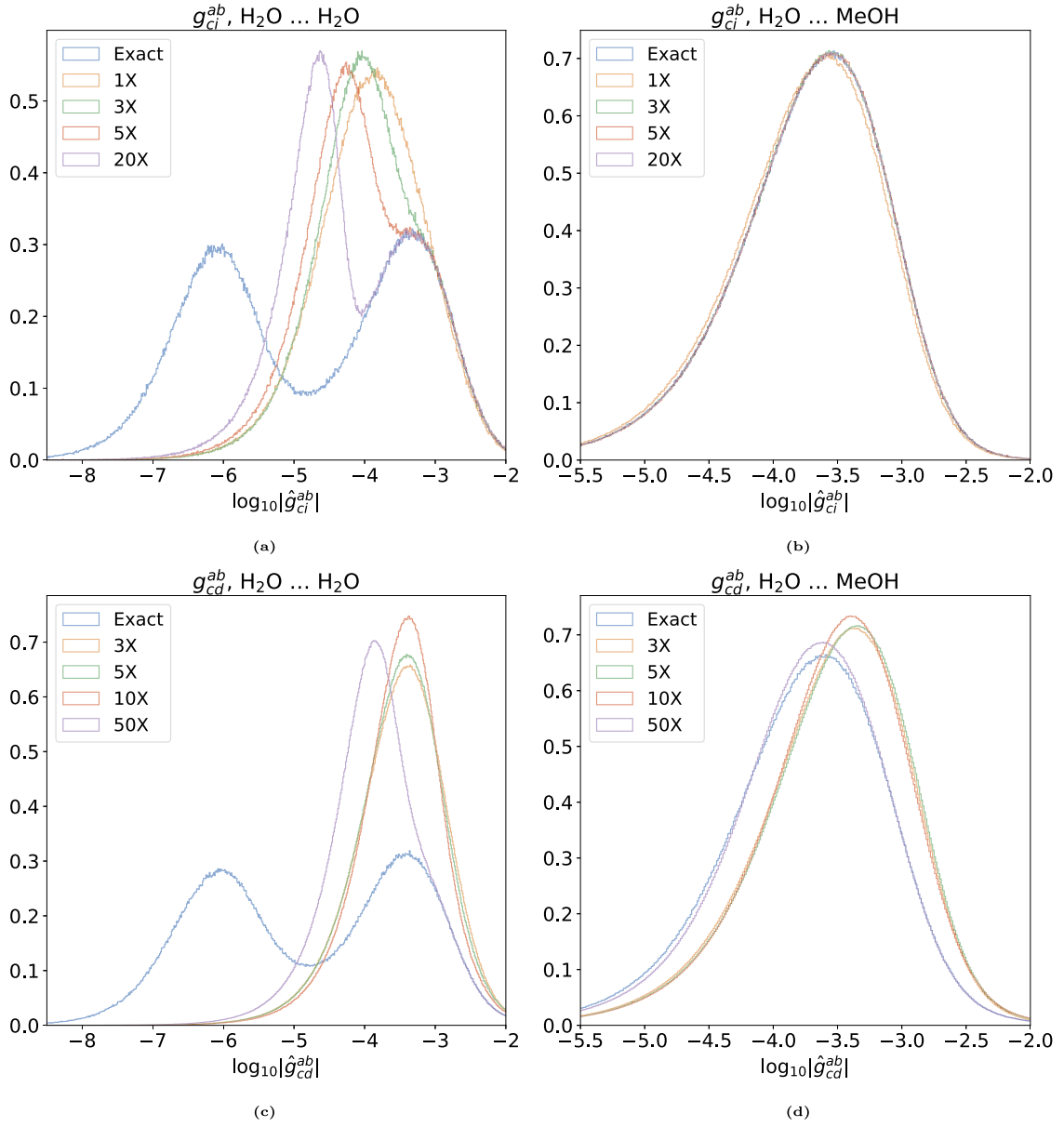


Figure 3. Distributions of the absolute element values for g^{ab}_{ci} (a, b) and g^{ab}_{cd} (c, d) tensors for representative systems. "Exact" refers to the DF-based tensor free of CP approximations, while "*nX*" corresponds to the counterpart approximated with CP4 decomposition of rank *nX*. All computations utilized an aVDZ/aVDZ-RI basis and $\epsilon_{ALS} = 10^{-3}$.

the two test systems (namely, the water dimer) has geometrical symmetry that makes many tensor elements nearly zero; note that our orbital solver does not exploit geometrical symmetry and therefore the zeroes are soft (i.e., the elements that are zero by symmetry are merely small). The presence of symmetry results in markedly bimodal distributions of the element magnitudes for both types of tensors for the water dimer (Figure 3a,c), with the "left" peak corresponding to the elements of the exact tensor that are nearly zero by symmetry. However, the distribution of the elements in the CP-approximated tensors is unimodal, except for the highest-rank approximants to the g^{ab}_{ci} tensor. As the CP4 rank increases, the largest tensor elements get approximated better and better, whereas the near-zero elements are not well approximated (in relative sense) even with the largest CP4 ranks employed here. The inability to preserve the blocked structure of the tensor is due to the global character of the CP4 factorization; to impose a blocked structure on the CP approximant would sacrifice the simplicity of the ansatz, hence was not done here.

Besides the block-sparse structure of the tensor approximate CP decomposition also fails to preserve permutational symmetries of the tensor. For example, we consider permutational symmetry of the (real) g^{ab}_{cd} tensor

$$g^{ab}_{cd} = g^{ad}_{cb} = g^{cb}_{ad} = g^{cd}_{ab} = g^{ba}_{dc} = g^{bc}_{da} = g^{da}_{bc} = g^{dc}_{ba} \tag{22}$$

The CP4 approximant $\hat{g}^{ab}_{cd}$ (eq 2) in general (i.e., for an arbitrary choice of CP4 parameters) lacks the symmetries of the exact tensor. Since it is possible in general to represent g^{ab}_{cd} exactly in the CP4 form, the CP4 factors must have a particular structure for the approximant to possess the exact symmetries of the tensor. Hence, we inspected the magnitude of the deviation of the CP4-approximated g^{ab}_{cd} tensors from the exact 8-fold symmetry in eq 22 as a function of the CP4 rank. Table 1 illustrates how the error of the standard non-symmetrized

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81

Table 1. Comparison of Standard (Equation 2) and Symmetrized (Equation 23) CP4 Approximants to DF-Approximated g_{cd}^{ab} Evaluated with the aVDZ/aVDZ-RI Basis Set Pair[a]

system	measure	R_{CP4}/X								
		1	1.5	2	2.5	3	5	10	20	50
$(H_2O)_2$	$\|g - \hat{g}\|$	17.0	16.1	15.3	14.6	13.9	11.6	8.01	3.15	0.854
	$\|g - \hat{g}_{sym}\|$	16.9	15.8	14.9	14.2	13.5	11.2	7.61	2.63	0.612
	$\|\hat{g} - \hat{g}_{sym}\|$	2.34	2.96	3.38	3.45	3.29	2.77	2.52	1.74	0.595
$H_2O\cdots$MeOH	$\|g - \hat{g}\|$	23.1	22.1	21.3	20.5	19.9	17.6	13.9	8.71	1.05
	$\|g - \hat{g}_{sym}\|$	22.9	21.8	20.9	20.1	19.4	17.1	13.3	8.01	0.763
	$\|\hat{g} - \hat{g}_{sym}\|$	3.31	3.60	4.01	4.13	4.12	3.84	3.88	3.41	0.718

[a]The corresponding error 2-norms are denoted by $\|g - \hat{g}\|$ and $\|g - \hat{g}_{sym}\|$, respectively. $\|\hat{g} - \hat{g}_{sym}\|$ denotes the 2-norm of the difference between the standard and symmetrized CP4 approximants.

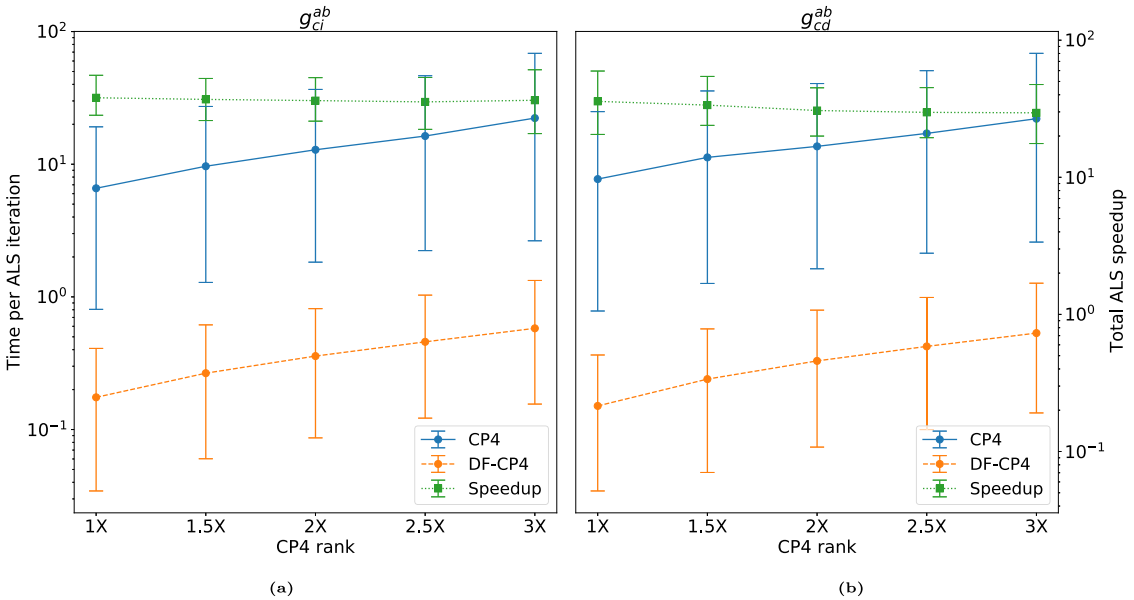


Figure 4. Cost per iteration of the exact and DF-approximated CP4 ALS decomposition and the total speedup associated with the DF approximation, i.e., $\frac{t_{exact}}{t_{DF}}$. The error bars represent the maximum and minimum per-iteration cost. Data were collected using the S66/5 dataset and the aVDZ/aVDZ-RI basis pair.

CP4 approximant compares to that of a (posteriori) symmetrized CP4 approximant

$$\hat{S}\hat{g}_{cd}^{ab} \equiv \frac{1}{8}(\hat{g}_{cd}^{ab} + \hat{g}_{cb}^{ad} + \hat{g}_{ad}^{cb} + \hat{g}_{ab}^{cd} + \hat{g}_{dc}^{ba} + \hat{g}_{da}^{bc} + \hat{g}_{bc}^{da} + \hat{g}_{ba}^{dc}) \tag{23}$$

There are several conclusions that can be drawn from these data. First, the violation of the exact symmetries is relatively small compared to the error due to the error of the CP4 approximant, ranging from ∼1% with $R_{CP4} = X$ to ∼30% with $R_{CP4} = 50X$. Second, enforcing the symmetry always reduces the error. Since for some applications it may be worthwhile to enforce the symmetry by construction, it would be potentially useful in the future to explore symmetry-adapted CP ansatz.

Figure 4 illustrates the wall time per ALS iteration with and without the DF approximation of the CP4 gradient (eq 12) as well as the corresponding speedup (i.e., reduction in computational time per ALS iteration), versus the CP4 rank for the molecules in the S66/5 dataset. With the aVDZ/aVDZ-RI basis set pair, the average observed reduction in per-iteration cost is ∼30. The speedup figure is clearly basis-dependent, as illustrated using four different OBSs for the water dimer in Figure 5; speedups as large as 100 have been observed.

Finally, we use the DF-accelerated CP4 ALS optimization to decompose the g_{ci}^{ab} tensor for a $(H_2O)_{12}$ cluster in the aVDZ/aVDZ-RI basis and report the results in Table 2. The size of this tensor is 31 GB, which is 2 orders of magnitude larger than the largest tensors that have previously been decomposed using CP4 in the chemistry context.[9] Not only are we able to construct a CP4 decomposition of this high-cost integral tensor, but we can systematically improve the accuracy by increasing rank while maintaining a relatively small number of ALS iterations and thus low computational cost.

5. SUMMARY AND PERSPECTIVE

We demonstrated how to efficiently compute full (four-way) CP decomposition of large two-body interaction tensors of relevance to electronic structure simulations by (1) reducing the cost of the CP gradient evaluation via a generalized square root (SQ) decomposition (such as density fitting or Cholesky decomposition) and (2) using three-way CP decomposition of the SQ factors to construct initial guess for the four-way CP decomposition. These improvements allow us to compute CP decompositions with greater speed and lower storage require-

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81

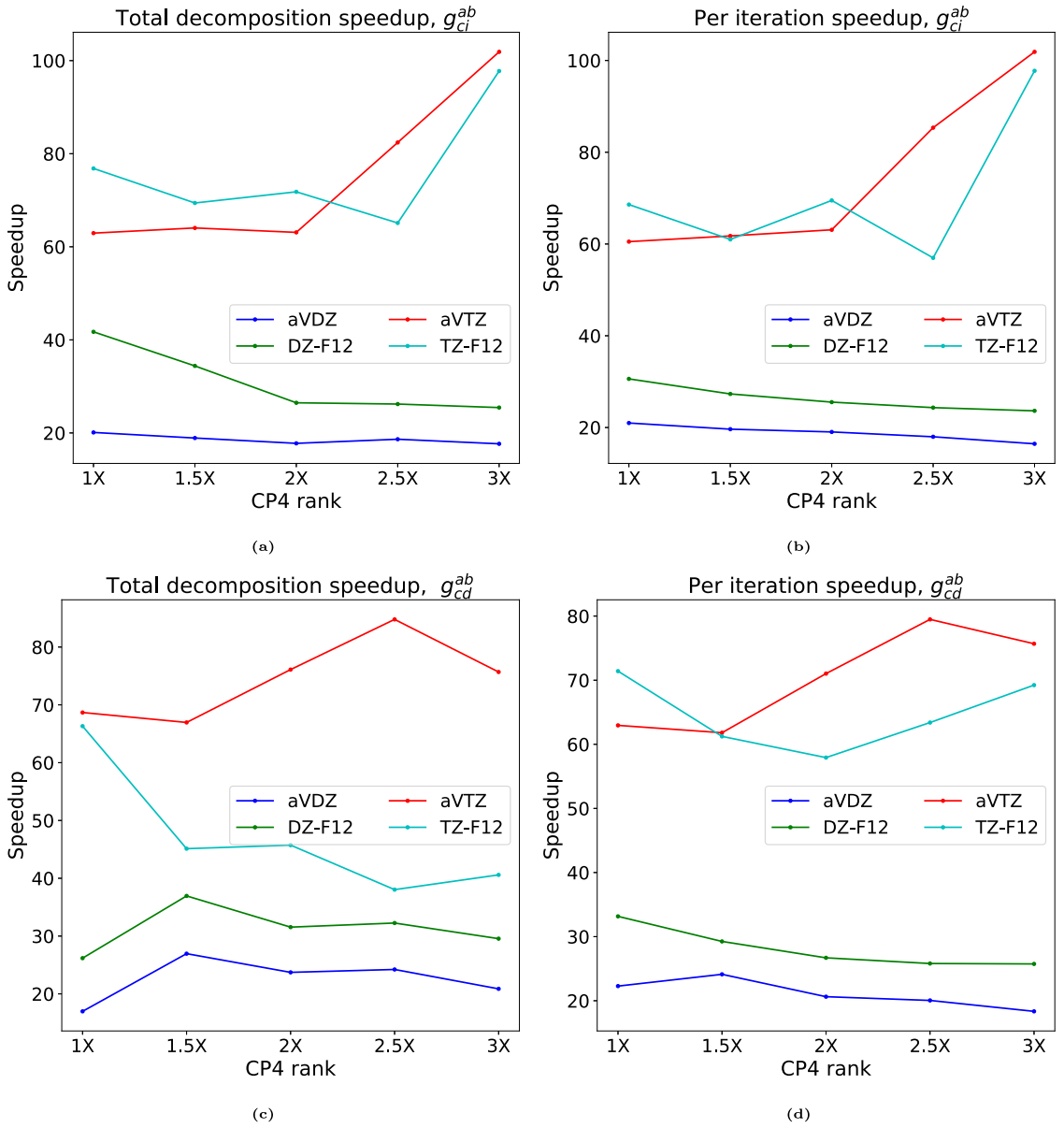


Figure 5. Total speedup of the CP4 ALS optimization of the (a) g_{ci}^{ab} tensor and (c) g_{cd}^{ab} tensor and the per-iteration speedup of the (b) g_{ci}^{ab} tensor and (d) g_{cd}^{ab} tensor using a $(H_2O)_2$ cluster from the S66 dataset and four different orbital and DF basis sets.

Table 2. Computation of the CP4 Decomposition of the g_{ci}^{ab} Tensor for a $(H_2O)_{12}$ Cluster Using DF-Approximated CP4 ALS Optimization Strategy and Using CP3 Factor Matrices as an Initial Guess Using $R_{CP3} = R_{CP4}$[a]

CP4 rank	1.0	1.5	2.0	2.5	3.0
CP3 iterations	10/8	10/8	9/9	9/10	8/11
CP4 iterations	21	21	22	24	24
per-iteration cost (s)	9.62	15.5	22.6	30.8	29.8
CP4 time (s)	202	325	498	739	716
max error	8.05×10^{-3}	7.72×10^{-3}	6.71×10^{-3}	6.17×10^{-3}	5.92×10^{-3}
mean error	2.48×10^{-5}	2.06×10^{-5}	1.82×10^{-5}	1.62×10^{-5}	1.49×10^{-5}

[a]The CP3 iteration's column shows the number of iterations required to compute the CP3 decomposition of B_c^{aX}/B_{iX}^c. The CP4 time reported does not include the time it takes to construct the CP3 initial guess because there is no one best, unique way to construct the guess. The max and mean errors are the element-wise error $|g_{ci}^{ab} - \hat{g}_{ci}^{ab}|$.

ments than previously possible. Four-way CP decomposition of Coulomb integral tensors with three and four virtual indices was demonstrated to be more compact and/or accurate than the corresponding combinations of three-way CP decompositions (similar to the pseudospectral and THC approximations) for several small clusters in the standard S66 benchmark set; practicality of a DF-accelerated four-way CP solver was demonstrated for the three-virtual-index integral tensor of a system with 36 atoms.

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81

The techniques described in this paper are more generally applicable than just for approximating the two-body interaction tensors in the electronic structure domain. The key lesson here is that it can be much more efficient to compute a CP approximation of a target tensor using a representative tensor network in place of the (reconstructed or exact) tensor itself. This suggests that the CP factorization, which offers the most complete decoupling of the tensor modes compared to other tensor network topology, may be more broadly deployable than currently thought, by combining it with other tensor network approximations. We will demonstrate productive uses of CP4 approximation in the context of correlated electronic structure models elsewhere soon.

AUTHOR INFORMATION

Corresponding Author

Edward F. Valeev − *Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States;* orcid.org/0000-0001-9923-6256; Email: efv@vt.edu

Author

Karl Pierce − *Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c00861

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy award DE-SC0022263 provided via the Scientific Discovery through Advanced Computing (SciDAC) program by the Offices of Advanced Scientific Computing Research (ASCR) and Basic Energy Sciences (BES). The authors also acknowledge Advanced Research Computing at Virginia Tech (www.arc.vt.edu) for providing computational resources and technical support that have contributed to the results reported within this paper.

ADDITIONAL NOTES

[a]It should be noted that though λ_r's resemble singular values, the CP decomposition does not have the same properties as the singular values decomposition.[28]

[b]Since matrix **W** is Hermitian, fast (pivoted) Cholesky solver for the linear system can be used.

[c]A square root B of matrix/operator A satisfies $A = BB$. In some contexts, such as Kalman filtering, B such that $A = B^{\dagger}B$, for positive A, is also referred to as square root. To avoid confusion we will be refer to such B as generalized square root of A.

[d]1 water···water, 2 water···MeOH, 3 water···$MeNH_2$, 4 $MeNH_2$···MeOH, 5 ethyne···water (CH-O).

REFERENCES

(1) Hitchcock, F. L. The Expression of a Tensor or a Polyadic as a Sum of Products. *J. Math. Phys.* **1927**, *6*, 164−189.

(2) Kiers, H. A. L. Towards a Standardized Notation and Terminology in Multiway Analysis. *J. Chemometrics* **2000**, *14*, 105−122.

(3) Carroll, J. D.; Chang, J.-J. Analysis of Individual Differences in Multidimensional Scaling via an N-Way Generalization of "Eckart-Young" Decomposition. *Psychometrika* **1970**, *35*, 283−319.

(4) Harshman, R. A. In *Foundations of the PARAFAC Procedure: Models and Conditions for an "Explanatory" Multimodal Factor Analysis*, UCLA Working Papers in Phonetics, 1970; pp 1−84.

(5) Beylkin, G.; Mohlenkamp, M. J. Numerical Operator Calculus in Higher Dimensions. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10246−10251.

(6) Möcks, J. Topographic Components Model for Event-Related Potentials and Some Biophysical Considerations. *IEEE Trans. Biomed. Eng.* **1988**, *35*, 482−484.

(7) Håstad, J. Tensor rank is NP-complete. *J. Algorithms* **1990**, *11*, 644−654.

(8) Hillar, C. J.; Lim, L.-H. Most Tensor Problems Are NP-Hard. *J. ACM* **2013**, *60*, 1−39.

(9) Benedikt, U.; Auer, A. A.; Espig, M.; Hackbusch, W. Tensor Decomposition in Post-Hartree-Fock Methods. I. Two-electron Integrals and MP2. *J. Chem. Phys.* **2011**, *134*, No. 054118.

(10) Benedikt, U.; Bohm, K.-H.; Auer, A. A. Tensor decomposition in post-Hartree-Fock methods. II. CCD implementation. *J. Chem. Phys.* **2013**, *139*, No. 224101.

(11) Böhm, K.-H.; Auer, A. A.; Espig, M. Tensor representation techniques for full configuration interaction: A Fock space approach using the canonical product format. *J. Chem. Phys.* **2016**, *144*, No. 244102.

(12) Bischoff, F. A.; Valeev, E. F. Low-Order Tensor Approximations for Electronic Wave Functions: Hartree-Fock Method with Guaranteed Precision. *J. Chem. Phys.* **2011**, *134*, No. 104104.

(13) Chinnamsetty, S. R.; Espig, M.; Khoromskij, B. N.; Hackbusch, W.; Flad, H.-J. Tensor Product Approximation with Optimal Rank in Quantum Chemistry. *J. Chem. Phys.* **2007**, *127*, No. 084110.

(14) Friesner, R. A. Solution of Self-Consistent Field Electronic Structure Equations by a Pseudospectral Method. *Chem. Phys. Lett.* **1985**, *116*, 39−43.

(15) Hohenstein, E. G.; Parrish, R. M.; Martínez, T. J. Tensor Hypercontraction Density Fitting. I. Quartic Scaling Second- and Third-Order Møller-Plesset Perturbation Theory. *J. Chem. Phys.* **2012**, *137*, No. 044103.

(16) Lu, J.; Ying, L. Compression of the Electron Repulsion Integral Tensor in Tensor Hypercontraction Format with Cubic Scaling Cost. *J. Comput. Phys.* **2015**, *302*, 329−335.

(17) Pierce, K.; Rishi, V.; Valeev, E. F. Robust Approximation of Tensor Networks: Application to Grid-Free Tensor Factorization of the Coulomb Interaction. *J. Chem. Theory Comput.* **2021**, *17*, 2217−2230.

(18) Peng, B.; Kowalski, K. Highly Efficient and Scalable Compound Decomposition of Two-Electron Integral Tensor and Its Application in Coupled Cluster Calculations. *J. Chem. Theory Comput.* **2017**, *13*, 4179−4192.

(19) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, Approximate and Parallel Hartree-Fock and Hybrid DFT Calculations. A 'chain-of-Spheres' Algorithm for the Hartree-Fock Exchange. *Chem. Phys.* **2009**, *356*, 98−109.

(20) Parrish, R. M.; Sherrill, C. D.; Hohenstein, E. G.; Kokkila, S. I. L.; Martínez, T. J. Communication: Acceleration of Coupled Cluster Singles and Doubles via Orbital-Weighted Least-Squares Tensor Hypercontraction. *J. Chem. Phys.* **2014**, *140*, No. 181102.

(21) Hu, W.; Lin, L.; Yang, C. Interpolative Separable Density Fitting Decomposition for Accelerating Hybrid Density Functional Calculations with Applications to Defects in Silicon. *J. Chem. Theory Comput.* **2017**, *13*, 5420−5431.

(22) Goings, J. J.; White, A.; Lee, J.; Tautermann, C. S.; Degroote, M.; Gidney, C.; Shiozaki, T.; Babbush, R.; Rubin, N. C. Reliably Assessing the Electronic Structure of Cytochrome P450 on Today's Classical Computers and Tomorrow's Quantum Computers. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119*, No. e2203533119.

(23) Almlöf, J. Elimination of Energy Denominators in Møller−Plesset Perturbation Theory by a Laplace Transform Approach. *Chem. Phys. Lett.* **1991**, *181*, 319−320.

(24) Karlsson, L.; Kressner, D.; Uschmajew, A. Parallel Algorithms for Tensor Completion in the CP Format. *Parallel Comput.* **2016**, *57*, 222−234.
(25) Acar, E.; Dunlavy, D. M.; Kolda, T. G. A Scalable Optimization Approach for Fitting Canonical Tensor Decompositions. *J. Chemometrics* **2011**, *25*, 67−86.
(26) Battaglino, C.; Ballard, G.; Kolda, T. G. A Practical Randomized CP Tensor Decomposition. *SIAM J. Matrix Anal. Appl.* **2018**, *39*, 876−901.
(27) Ma, L.; Solomonik, E. Accelerating Alternating Least Squares for Tensor Decomposition by Pairwise Perturbation. *Numer. Linear Algebra Appl.* **2022**, *29*, No. 2773.
(28) Kolda, T. G.; Bader, B. W. Tensor Decompositions and Applications. *SIAM Rev.* **2009**, *51*, 455−500.
(29) Harshman, R. In *Foundations of the PARAFAC Procedure: Models and Conditions for an Explanatory Multimodal Factor Analysis*, UCLA Working Papers in Phonetics, 1970; pp 1−84.
(30) Acar, E.; Dunlavy, D. M.; Kolda, T. G. A scalable optimization approach for fitting canonical tensor decompositions. *J. Chemometrics* **2011**, *25*, 67−86.
(31) Phan, A.-H.; Tichavský, P.; Cichocki, A. Low Complexity Damped Gauss-Newton Algorithms for CANDECOMP/PARAFAC. *SIAM J. Matrix Anal. Appl.* **2013**, *34*, 126−147.
(32) Tichavsky, P.; Phan, A. H.; Cichocki, A. In *A Further Improvement of a Fast Damped Gauss-Newton Algorithm for Candecomp-Parafac Tensor Decomposition*, 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 2013; pp 5964−5968.
(33) Espig, M.; Hackbusch, W.; Rohwedder, T.; Schneider, R. Variational calculus with sums of elementary tensors of fixed rank. *Numer. Math.* **2012**, *122*, 469−488.
(34) Sorber, L.; Van Barel, M.; De Lathauwer, L. Optimization-Based Algorithms for Tensor Decompositions: Canonical Polyadic Decomposition, Decomposition in Rank-$(L_r,L_r,1)$ Terms, and a New Generalization. *SIAM J. Optim.* **2013**, *23*, 695−720.
(35) Navasca, C.; De Lathauwer, L.; Kindermann, S. In *Swamp Reducing Technique For Tensor Decomposition*, 2008 16th European Signal Processing Conference, IEEE, 2008; pp 1−5.
(36) Whitten, J. L. Coulombic Potential Energy Integrals and Approximations. *J. Chem. Phys.* **1973**, *58*, 4496−4501.
(37) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On First-Row Diatomic Molecules and Local Density Models. *J. Chem. Phys.* **1979**, *71*, 4993.
(38) Vahtras, O.; Almlof, J.; Feyereisen, M. W. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514−518.
(39) Jung, Y. Chemical Theory and Computation Special Feature: Auxiliary Basis Expansions for Large-Scale Electronic Structure Calculations. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6692−6697.
(40) Beebe, N. H. F.; Linderberg, J. Simplifications in the Generation and Transformation of Two-Electron Integrals in Molecular Calculations. *Int. J. Quantum Chem.* **1977**, *12*, 683−705.
(41) Löwdin, P.-O. Studies in Perturbation Theory. IX. Connection between Various Approaches in the Recent Development—Evaluation of Upper Bounds to Energy Eigenvalues in Schrödinger's Perturbation Theory. *J. Math. Phys.* **1965**, *6*, 1341−1353.
(42) Löwdin, P.-O. Some properties of inner projections. *Int. J. Quantum Chem.* **2009**, *5*, 231−237.
(43) Folkestad, S. D.; Kjønstad, E. F.; Koch, H. An efficient algorithm for Cholesky decomposition of electron repulsion integrals. *J. Chem. Phys.* **2019**, *150*, No. 194112.
(44) Lee, J.; Lin, L.; Head-Gordon, M. Systematically Improvable Tensor Hypercontraction: Interpolative Separable Density-Fitting for Molecules Applied to Exact Exchange, Second- and Third-Order Møller-Plesset Perturbation Theory. *J. Chem. Theory Comput.* **2020**, *16*, 243−263.
(45) Hummel, F.; Tsatsoulis, T.; Grüneis, A. Low Rank Factorization of the Coulomb Integrals for Periodic Coupled Cluster Theory. *J. Chem. Phys.* **2017**, *146*, No. 124105.
(46) Friesner, R. A. Solution of the Hartree-Fock equations by a pseudospectral method: Application to diatomic molecules. *J. Chem. Phys.* **1986**, *85*, 1462−1468.
(47) Langlois, J.; Muller, R. P.; Coley, T. R.; Goddard, W. A.; Ringnalda, M. N.; Won, Y.; Friesner, R. A. Pseudospectral generalized valence-bond calculations: Application to methylene, ethylene, and silylene. *J. Chem. Phys.* **1990**, *92*, 7488−7497.
(48) Ringnalda, M. N.; Belhadj, M.; Friesner, R. A. Pseudospectral Hartree−Fock Theory: Applications and Algorithmic Improvements. *J. Chem. Phys.* **1990**, *93*, 3397−3407.
(49) Friesner, R. A. New Methods For Electronic Structure Calculations on Large Molecules. *Annu. Rev. Phys. Chem.* **1991**, *42*, 341−367.
(50) Martinez, T. J.; Carter, E. A. Pseudospectral multireference single and double excitation configuration interaction. *J. Chem. Phys.* **1995**, *102*, 7564−7572.
(51) Martinez, T. J.; Mehta, A.; Carter, E. A. Pseudospectral Full Configuration Interaction. *J. Chem. Phys.* **1992**, *97*, 1876−1880.
(52) Martinez, T. J.; Carter, E. A. Pseudospectral Møller−Plesset perturbation theory through third order. *J. Chem. Phys.* **1994**, *100*, 3631−3638.
(53) Ko, C.; Malick, D. K.; Braden, D. A.; Friesner, R. A.; Martínez, T. J. Pseudospectral time-dependent density functional theory. *J. Chem. Phys.* **2008**, *128*, No. 104103.
(54) Martinez, T. J.; Carter, E. A. Pseudospectral double excitation configuration interaction. *J. Chem. Phys.* **1993**, *98*, 7081−7085.
(55) Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Martínez, T. J. Communication: Tensor Hypercontraction. III. Least-squares Tensor Hypercontraction for the Determination of Correlated Wavefunctions. *J. Chem. Phys.* **2012**, *137*, No. 221101.
(56) Parrish, R. M.; Hohenstein, E. G.; Martínez, T. J.; Sherrill, C. D. Tensor Hypercontraction. II. Least-squares Renormalization. *J. Chem. Phys.* **2012**, *137*, No. 224106.
(57) Hohenstein, E. G.; Kokkila, S. I. L.; Parrish, R. M.; Martinez, T. J. Quartic scaling second-order approximate coupled cluster singles and doubles via tensor hypercontraction: THC-CC2. *J. Chem. Phys.* **2013**, *138*, No. 124111.
(58) Shenvi, N.; Van Aggelen, H.; Yang, Y.; Yang, W.; Schwerdtfeger, C.; Mazziotti, D. Low rank factorization of the Coulomb integrals for periodic coupled cluster theory. *J. Chem. Phys.* **2013**, *139*, No. 54110.
(59) Schutski, R.; Zhao, J.; Henderson, T. M.; Scuseria, G. E. Tensor-structured coupled cluster theory. *J. Chem. Phys.* **2017**, *147*, No. 184113.
(60) Parrish, R. M.; Zhao, Y.; Hohenstein, E. G.; Martinez, T. J. Rank reduced coupled cluster theory. I. Ground state energies and wavefunctions. *J. Chem. Phys.* **2019**, *150*, No. 164118.
(61) Greengard, L.; Rokhlin, V. A Fast Algorithm for Particle Simulations. *J. Comput. Phys.* **1987**, *73*, 325−348.
(62) Beylkin, G.; Cheruvu, V.; Perez, F. Fast Adaptive Algorithms in the Non-Standard Form for Multidimensional Problems. *Appl. Comput. Harmon. Anal.* **2008**, *24*, 354−377.
(63) BTAS Library, 2021. https://github.com/BTAS/BTAS (accessed Apr 9, 2019).
(64) Kroonenberg, P. M.; de Leeuw, J. Principal Component Analysis of Three-Mode Data by Means of Alternating Least Squares Algorithms. *Psychometrika* **1980**, *45*, 69−97.
(65) Řezáč, J.; Riley, K. E.; Hobza, P. S66: A Well-Balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427−2438.
(66) Řezáč, J.; Jurecka, P.; Riley, K. E.; Cerny, J.; Valdes, H.; Pluhackova, K.; Berka, K.; Rezac, T.; Pitonak, M.; Vondrasek, J.; Hobz, P. Quantum chemical benchmark energy and geometry database for molecular clusters and complex molecular systems (www.begdb.com): A users manual and examples. *Collect. Czech. Chem. Commun.* **2008**, *73*, 1261−1270.
(67) Jo/rgensen, P.; Simons, J. Ab Initio Analytical Molecular Gradients and Hessians. *J. Chem. Phys.* **1983**, *79*, 334−357.

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81

(68) Wales, D. J.; Hodges, M. P. Global minima of water clusters (H2O)n, n≤21, described by an empirical potential. *Chem. Phys. Lett.* **1998**, *286*, 65−72.
(69) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(70) Kendall, R. A.; Früchtl, H. A. The Impact of the Resolution of the Identity Approximate Integral Method on Modern Ab Initio Algorithm Development. *Theor. Chim. Acta* **1997**, *97*, 158−163.
(71) Weigend, F.; Köhn, A.; Hättig, C. Efficient Use of the Correlation Consistent Basis Sets in Resolution of the Identity MP2 Calculations. *J. Chem. Phys.* **2002**, *116*, 3175.
(72) Peterson, K. A.; Adler, T. B.; Werner, H.-J. Systematically Convergent Basis Sets for Explicitly Correlated Wavefunctions: The Atoms H, He, B-Ne, and Al-Ar. *J. Chem. Phys.* **2008**, *128*, No. 084102.
(73) Espig, M.; Hackbusch, W. A Regularized Newton Method for the Efficient Approximation of Tensors Represented in the Canonical Tensor Format. *Numer. Math.* **2012**, *122*, 489−525.
(74) Kazeev, V. A.; Tyrtyshnikov, E. E. Structure of the Hessian Matrix and an Economical Implementation of Newton's Method in the Problem of Canonical Approximation of Tensors. *Comput. Math. Math. Phys.* **2010**, *50*, 927−945.
(75) Phan, A.-H.; Tichavský, P.; Cichocki, A. Low Complexity Damped Gauss−Newton Algorithms for CANDECOMP/PARAFAC. *SIAM J. Matrix Anal. Appl.* **2013**, *34*, 126−147.

https://doi.org/10.1021/acs.jctc.2c00861
J. Chem. Theory Comput. 2023, 19, 71−81