Efficient computation of the PARAFAC2 decomposition

Yao Cheng and *Martin Haardt*

Communications Research Laboratory
Ilmenau University of Technology
P.O. Box 10 05 65, D-98684 Ilmenau, Germany

***Abstract*—The PARAFAC2 decomposition is regarded as a promising multi-linear signal processing tool in a variety of scientific fields. Designing efficient computation algorithms for PARAFAC2 is a long-standing topic. In this contribution, we propose to incorporate a dimension reduction in direct fitting-based schemes for the computation of PARAFAC2. Extensive simulations show that the proposed algorithms are able to achieve the same performance in terms of the residual as their counterpart schemes in the literature, while requiring a significantly reduced computation time.**

I. Introduction

The PARAFAC2 decomposition has been widely used in a great number of areas, including the food industry for the processing of gas chromatography-mass spectrometry (GC-MS) data [1]–[5], semi-blind channel estimation and symbol detection for MIMO systems [6], and the multi-way component analysis of biomedical signals, e.g., measured visual evoked potentials [7], somatosensory evoked magnetic fields (SEFs) and somatosensory evoked electrical potentials (SEPs) [8], [9], [10].

To compute PARAFAC2, the direct fitting (DF) algorithm [11] is usually employed. Due to its alternating least squares (ALS) nature, the DF scheme is prone to a low convergence especially when the collinearity in the data is strong. In the literature, line search methods have been proposed to improve the convergence behaviour of the DF algorithm [12], [13]. Moreover, a geometric search approach has been reported in [14], where its superiority over existing line search schemes has been shown. Recently, we have investigated the incorporation of semi-algebraic approaches in the DF algorithm and have thereby developed enhanced versions featuring improved efficiency and robustness against collinearity in the data [15].

In this contribution, we propose a further extension of the aforementioned DF schemes based on a new dimension reduction concept, targeting at an even higher efficiency. Extensive simulations have been conducted to assess the performance of the proposed approach. The obtained numerical results indicate that it can be applied to various versions of the DF algorithm, leading to the same performance and a much lower computational complexity in the meantime.

The remainder of the paper is organized as follows: the DF algorithm [11] for the computation of PARAFAC2 is briefly reviewed in Section II. Then Section III presents the efficiency enhancement thanks to the dimension reduction. Numerical results are shown in Section IV, before conclusions are drawn in Section V.

To facilitate the distinction between scalars, vectors, matrices, and tensors, the following notation is used throughout this work: scalars are represented by italic letters, vectors by lower-case bold-faced letters, matrices by upper-case bold-faced letters, and tensors as bold-faced calligraphic letters. We use the superscript $^{\mathrm{T}}$ for transpose and $\circ$ to denote the outer product operation. The i-th row and the j-th column of a matrix $\boldsymbol{A} \in \mathbb{C}^{I\times J}$ is symbolized by $\boldsymbol{A}(i,:) \in \mathbb{C}^{1\times J}$ and $\boldsymbol{A}(:,j) \in \mathbb{C}^{I}$, respectively, where $i = 1,\ldots,I$ and $j = 1,\ldots,J$.

An R-way tensor with size I_r along mode $r = 1,2,\ldots,R$ is represented as $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1\times I_2\times\ldots\times I_R}$. Its r-mode vectors are obtained by varying the r-th index from 1 to I_r and keeping all other indices fixed. Aligning all r-mode vectors as the columns of a matrix yields the r-mode unfolding of $\boldsymbol{\mathcal{A}}$ which is denoted by $[\boldsymbol{\mathcal{A}}]_{(r)} \in \mathbb{C}^{I_r\times I_{r+1}\cdot\ldots\cdot I_R\cdot I_1\cdot\ldots\cdot I_{r-1}}$. Here the reverse cyclical ordering of the columns [16] is used. The r-mode product between a tensor $\boldsymbol{\mathcal{A}}$ and a matrix $\boldsymbol{U}$ is written as $\boldsymbol{\mathcal{A}} \times_r \boldsymbol{U}$. It is computed by multiplying all r-mode vectors of $\boldsymbol{\mathcal{A}}$ with $\boldsymbol{U}$. For a three-dimensional tensor $\boldsymbol{\mathcal{B}} \in \mathbb{C}^{I_1\times I_2\times I_3}$, its k-th frontal slice (also called three mode slice) is represented by $[\boldsymbol{\mathcal{B}}]_{(:,:,k)} \in \mathbb{C}^{I_1\times I_2}$, where $k = 1,\ldots,I_3$. In addition, we denote the higher-order norm of a tensor $\boldsymbol{\mathcal{A}}$ by $\|\boldsymbol{\mathcal{A}}\|_{\mathrm{H}}$. It is defined as the square root of the sum of the squared magnitude of all elements in $\boldsymbol{\mathcal{A}}$. A $d \times d$ identity matrix is denoted by $\boldsymbol{I}_d$.

II. PARAFAC2 and direct fitting algorithm

PARAFAC2 decomposes the tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{M_1\times M_2\times M_3}$ such that the k-th frontal slice of $\boldsymbol{\mathcal{X}}$ ($k = 1,\ldots,M_3$) is written as[1]

$$[\boldsymbol{\mathcal{X}}]_{(:,:,k)} = \boldsymbol{A}\cdot\mathrm{diag}\left\{\boldsymbol{C}(k,:)\right\}\cdot\boldsymbol{B}_k^{\mathrm{T}} + [\boldsymbol{\mathcal{N}}]_{(:,:,k)} \in \mathbb{R}^{M_1\times M_2},$$

where $\boldsymbol{A} \in \mathbb{R}^{M_1\times d}$, $\boldsymbol{C} \in \mathbb{R}^{M_3\times d}$, $\boldsymbol{B}_k \in \mathbb{R}^{M_2\times d}$ for $k = 1,\ldots,M_3$, and d represents the model order. The residuals are collected in $\boldsymbol{\mathcal{N}} \in \mathbb{R}^{M_1\times M_2\times M_3}$. We use $\mathrm{diag}\{\cdot\}$ to denote the operation of constructing a diagonal matrix whose diagonal

[1]PARAFAC2 can be applied to a set of M_3 matrices of size $M_1 \times M_{2_k}$ ($k = 1,\ldots,M_3$). For notation simplicity, we assume $M_{2_k} = M_2$ for $k = 1,\ldots,M_3$, while the proposed algorithms are also applicable to the general PARAFAC2 model.

elements are the entries of the input vector. The PARAFAC2 decomposition is illustrated in Fig. 1.

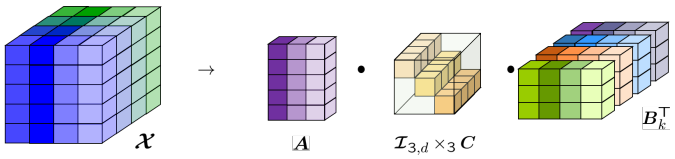

Fig. 1: Illustration of PARAFAC2

Note that the Harshman constraint [17] given by $\boldsymbol{B}_k^{\mathrm{T}} \cdot \boldsymbol{B}_k = \boldsymbol{H} \in \mathbb{R}^{d\times d}$ $(k = 1, \ldots, M_3)$ can be reformulated [11] such that $\boldsymbol{B}_k^{\mathrm{T}} = \boldsymbol{F}^{\mathrm{T}} \cdot \boldsymbol{V}_k$, where $\boldsymbol{F} \in \mathbb{R}^{d\times d}$, and $\boldsymbol{V}_k \in \mathbb{R}^{d\times M_2}$ for $k = 1, \ldots, M_3$, satisfying $\boldsymbol{V}_k \cdot \boldsymbol{V}_k^{\mathrm{T}} = \boldsymbol{I}_d$. Let us introduce a tensor $\widetilde{\boldsymbol{\mathcal{X}}}_0 \in \mathbb{R}^{M_1\times d\times M_3}$ where the k-th frontal slice of $\widetilde{\boldsymbol{\mathcal{X}}}_0$ is written as

$$\left[\widetilde{\boldsymbol{\mathcal{X}}}_0\right]_{(:,:,k)} = \boldsymbol{A} \cdot \mathrm{diag}\left\{\boldsymbol{C}(k,:)\right\} \cdot \boldsymbol{F}^{\mathrm{T}}, \tag{1}$$

indicating that $\boldsymbol{A}$, $\boldsymbol{C}$, and $\boldsymbol{F}$ are the factor matrices of the PARAFAC decomposition of $\widetilde{\boldsymbol{\mathcal{X}}}_0$. In a zero-residual case, the following equality holds

$$\left[\widetilde{\boldsymbol{\mathcal{X}}}_0\right]_{(:,:,k)} = [\boldsymbol{\mathcal{X}}]_{(:,:,k)} \cdot \boldsymbol{V}_k^{\mathrm{T}}. \tag{2}$$

These facts of PARAFAC2 have led to the DF algorithm [11] summarized as follows:

- **Step 1**: Initialize $\boldsymbol{A}$, $\boldsymbol{C}$, and $\boldsymbol{F}$.
- **Step 2**: Reconstruct $\widetilde{\boldsymbol{\mathcal{X}}}_0$ with $\boldsymbol{A}$, $\boldsymbol{C}$, and $\boldsymbol{F}$ according to (1), and update $\boldsymbol{V}_k$ $(k = 1, \ldots, M_3)$ with the generalized solution of the Orthogonal Procrustes Problem (OPP) [18] owing to (2).
- **Step 3**: Update $\widetilde{\boldsymbol{\mathcal{X}}}_0$ according to (2) with $\boldsymbol{V}_k$ $(k = 1, \ldots, M_3)$ computed in **Step 2**. Then compute its PARAFAC decomposition to update $\boldsymbol{A}$, $\boldsymbol{C}$, and $\boldsymbol{F}$.
- **Step 4**: Compute the reconstructed tensor $\hat{\boldsymbol{\mathcal{X}}}$ and then calculate the residual defined via $E_{\mathrm{R}} = \left(\left\|\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{X}}\right\|_{\mathrm{H}}^2\right) / \|\boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2$. Repeat **Step 2** – **Step 4** until the change of the residual quantified by $\Delta E_{\mathrm{R}} = \left(E_{\mathrm{R}}^{\mathrm{old}} - E_{\mathrm{R}}\right) / E_{\mathrm{R}}^{\mathrm{old}}$ is smaller than a predefined threshold, implying the convergence of the algorithm, where $E_{\mathrm{R}}^{\mathrm{old}}$ represents the residual in the previous iteration.

It is known that the PARAFAC decomposition can be computed via simultaneous matrix diagonalizations (SMDs), e.g., [19]. Given the truncated higher-order SVD (HOSVD) of $\widetilde{\boldsymbol{\mathcal{X}}}_0$ as

$$\widetilde{\boldsymbol{\mathcal{X}}}_0 = \boldsymbol{\mathcal{S}}^{[\mathrm{s}]} \times_1 \boldsymbol{U}_1^{[\mathrm{s}]} \times_2 \boldsymbol{U}_2^{[\mathrm{s}]} \times_3 \boldsymbol{U}_3^{[\mathrm{s}]}, \tag{3}$$

where $\boldsymbol{\mathcal{S}}^{[\mathrm{s}]} \in \mathbb{R}^{d\times d\times d}$, $\boldsymbol{U}_1^{[\mathrm{s}]} \in \mathbb{R}^{M_1\times d}$, $\boldsymbol{U}_2^{[\mathrm{s}]} \in \mathbb{R}^{d\times d}$, and $\boldsymbol{U}_3^{[\mathrm{s}]} \in \mathbb{R}^{M_3\times d}$, finding the factor matrices $\boldsymbol{A}$, $\boldsymbol{F}$, and $\boldsymbol{C}$ is translated into the estimation of three non-singular transform matrices $\boldsymbol{T}_1$, $\boldsymbol{T}_2$, and $\boldsymbol{T}_3 \in \mathbb{R}^{d\times d}$ [19]. Once these transform matrices are computed via SMDs, the factor matrices are obtained as [19]

$$\boldsymbol{A} = \boldsymbol{U}_1^{[\mathrm{s}]} \cdot \boldsymbol{T}_1 \tag{4}$$

$$\boldsymbol{F} = \boldsymbol{U}_2^{[\mathrm{s}]} \cdot \boldsymbol{T}_2 \tag{5}$$

$$\boldsymbol{C} = \boldsymbol{U}_3^{[\mathrm{s}]} \cdot \boldsymbol{T}_3. \tag{6}$$

If multiple SMDs are calculated to find these transform matrices, this approach leads to potential multiple solutions for $\boldsymbol{A}$, $\boldsymbol{F}$, and $\boldsymbol{C}$ that, in general, have different accuracies. Therefore, it is very challenging to incorporate this powerful scheme to the DF algorithm. In [15], the impact of these diverse PARAFAC factor estimates obtained via SMDs on the DF algorithm is investigated. Two schemes, namely "DF-SMD one batch" and "DF-SMD parallel", have been designed to exploit this diversity while taking the complexity aspect into account. For a special two-component case ($d = 2$), we have observed that, instead of SMDs, only the eigenvalue decompositions (EVDs) are needed. The resulting DF scheme is referred to as "DF-EVD". In Section IV, the aforementioned DF algorithms are used as benchmarks for the performance evaluation.

III. Enhanced version with a dimension reduction

Considering a zero-residual case, let $\boldsymbol{U}_1 \in \mathbb{R}^{M_1\times d}$ denote the one-mode left singular matrix and $\boldsymbol{U}_{2_k} \in \mathbb{R}^{M_2\times d}$ represent the right singular matrix of the k-th frontal slice of $\boldsymbol{\mathcal{X}}$. Next let us introduce an equivalent data tensor $\boldsymbol{\mathcal{X}}' \in \mathbb{R}^{d\times d\times M_3}$ such that its k-th frontal slice is written as

$$\left[\boldsymbol{\mathcal{X}}'\right]_{(:,:,k)} = \boldsymbol{U}_1^{\mathrm{T}} \cdot [\boldsymbol{\mathcal{X}}]_{(:,:,k)} \cdot \boldsymbol{U}_{2_k} \tag{7}$$

$$= \boldsymbol{U}_1^{\mathrm{T}} \cdot \boldsymbol{A} \cdot \mathrm{diag}\left\{\boldsymbol{C}(k,:)\right\} \cdot \boldsymbol{B}_k^{\mathrm{T}} \cdot \boldsymbol{U}_{2_k}. \tag{8}$$

Defining

$$\boldsymbol{T}_1' = \boldsymbol{U}_1^{\mathrm{T}} \cdot \boldsymbol{A} \tag{9}$$

$$\boldsymbol{T}_{2_k}'^{\mathrm{T}} = \boldsymbol{B}_k^{\mathrm{T}} \cdot \boldsymbol{U}_{2_k} \tag{10}$$

yields

$$\left[\boldsymbol{\mathcal{X}}'\right]_{(:,:,k)} = \boldsymbol{T}_1' \cdot \mathrm{diag}\left\{\boldsymbol{C}(k,:)\right\} \cdot \boldsymbol{T}_{2_k}'^{\mathrm{T}}. \tag{11}$$

This observation implies that with $\boldsymbol{U}_1 \in \mathbb{R}^{M_1\times d}$ and $\boldsymbol{U}_{2_k} \in \mathbb{R}^{M_2\times d}$ computed at the beginning, we only have to calculate the PARAFAC2 decomposition of $\boldsymbol{\mathcal{X}}' \in \mathbb{R}^{d\times d\times M_3}$ instead of the original data tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{M_1\times M_2\times M_3}$. Consequently, the computational complexity is significantly reduced especially in large-dimension problems where $M_1, M_2 \gg d$, as later shown in Section IV. The resulting factor matrices of the decomposition are $\boldsymbol{T}_1' \in \mathbb{R}^{d\times d}$, $\boldsymbol{T}_{2_k}' \in \mathbb{R}^{d\times d}$, $\boldsymbol{C} \in \mathbb{R}^{M_3\times d}$. It is worth noting that a dimension reduction in the three-mode is not possible, as the size of the three-mode determines the number of frontal slices, i.e., the number of two-mode factor matrices.

Finally the one- and two-mode factor matrices are obtained as

$$\boldsymbol{A} = \boldsymbol{U}_1 \cdot \boldsymbol{T}_1', \tag{12}$$

$$\boldsymbol{B}_k = \boldsymbol{U}_{2_k} \cdot \boldsymbol{T}_{2_k}'. \tag{13}$$

The enhanced version of the DF algorithm benefiting from the dimension reduction is summarized as follows[2]:

[2]Here we take the DF algorithm with ALS for the PARAFAC decomposition [11] as an example, as the reader can conveniently refer to the corresponding part in Section II. A similar extension of the "DF-EVD", "DF-SMD one batch" and "DF-SMD parallel" schemes in [15] can be accordingly derived.

- **Step 1**: Compute $\boldsymbol{U}_1 \in \mathbb{R}^{M_1 \times d}$ and $\boldsymbol{U}_{2_k} \in \mathbb{R}^{M_2 \times d}$, and construct $\boldsymbol{\mathcal{X}}' \in \mathbb{R}^{d \times d \times M_3}$.
- **Step 2**: Initialize $\boldsymbol{T}_1'$, $\boldsymbol{C}$, and $\boldsymbol{F}'$ with $\boldsymbol{T}_{2,k}'^{\mathrm{T}} = \boldsymbol{F}'^{\mathrm{T}} \cdot \boldsymbol{V}_k'$. As an alternative to a random initialization, the columns of $\boldsymbol{T}_1'$ can be initialized as the first d left singular vectors of $\sum_{k=1}^{M_3} \left[\boldsymbol{\mathcal{X}}'\right]_{(:,:,k)} \cdot \left[\boldsymbol{\mathcal{X}}'\right]_{(:,:,k)}^{\mathrm{T}}$.
- **Step 3**: Reconstruct $\widetilde{\boldsymbol{\mathcal{X}}}_0'$ and update $\boldsymbol{V}_k'$ with the generalized solution of the Orthogonal Procrustes Problem (OPP) [18] owing to (2).
- **Step 4**: Update $\widetilde{\boldsymbol{\mathcal{X}}}_0'$ and then compute its PARAFAC decomposition to update $\boldsymbol{T}_1'$, $\boldsymbol{C}$, and $\boldsymbol{F}'$, and accordingly $\boldsymbol{A} = \boldsymbol{U}_1 \cdot \boldsymbol{T}_1'$ and $\boldsymbol{B}_k = \boldsymbol{U}_{2_k} \cdot \boldsymbol{V}_k'^{\mathrm{T}} \cdot \boldsymbol{F}'$.
- **Step 5**: Compute the reconstructed tensor $\hat{\boldsymbol{\mathcal{X}}}$ and then calculate the residual defined via $E_{\mathrm{R}} = \left(\left\|\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{X}}\right\|_{\mathrm{H}}^2\right) / \|\boldsymbol{\mathcal{X}}\|_{\mathrm{H}}^2$. Repeat **Step 3** – **Step 5** until the change of the residual quantified by $\Delta E_{\mathrm{R}} = \left(E_{\mathrm{R}}^{\mathrm{old}} - E_{\mathrm{R}}\right) / E_{\mathrm{R}}^{\mathrm{old}}$ is smaller than a predefined threshold, implying the convergence of the algorithm, where $E_{\mathrm{R}}^{\mathrm{old}}$ represents the residual in the previous iteration.

Figure 2 illustrates the dimension reduction concept. Specifically it contributes to a reduced computational complexity in the following two procedures:

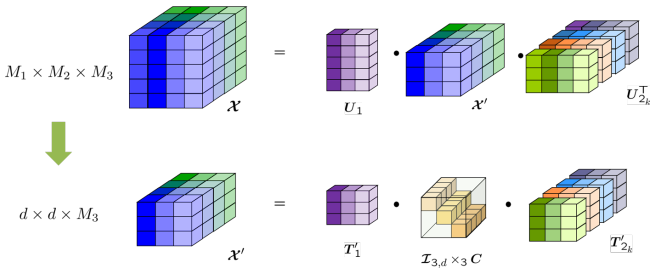

Fig. 2: Illustration of the dimension reduction

- The dimension of the OPP is reduced from $\{M_1 \times M_2, M_1 \times d\}$ to $\{d \times d, d \times d\}$.
- The dimension of the PARAFAC decomposition is reduced from $M_1 \times d \times M_3$ to $d \times d \times M_3$.

It is worth mentioning that a dimension reduction is inherent in SMD-based approaches owing to the truncated HOSVD. This fact should be taken into account when extending the "DF-EVD", "DF-SMD one batch" and "DF-SMD parallel" schemes in [15].

IV. Numerical Results

To evaluate the performance of the proposed algorithm, extensive numerical simulations have been carried out. The factor matrices are randomly generated according to the PARAFAC2 model, while $\boldsymbol{\mathcal{N}}$ contains zero-mean uncorrelated Gaussian entries with variance σ_n^2. Accordingly, we define $\mathrm{SNR} = 1/\sigma_n^2$. The threshold for convergence is set to 10^{-6}, and for all scenarios shown below, 2000 Monte-Carlo trials have been conducted. We use DF-ALS to refer to the original DF scheme in [11], whereas "DF-EVD", "DF-SMD one batch" and "DF-SMD parallel" represent the DF schemes proposed in [15] as mentioned in the previous sections. Having "RD-" as a prefix in the name of an algorithm indicates that the proposed dimension reduction concept is employed.

First, a two-component scenario is considered with $M_1 = 20$, $M_2 = 30$, $M_3 = 10$, and $d = 2$.

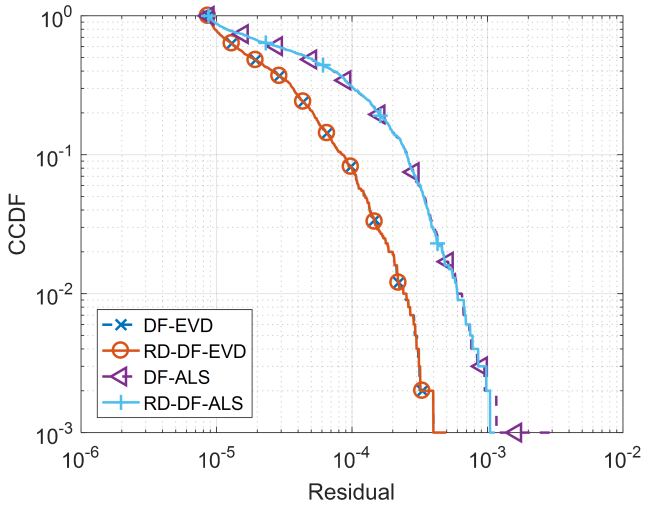

Fig. 3: CCDF of residuals for a scenario with $M_1 = 20$, $M_2 = 30$, $M_3 = 10$, $d = 2$, SNR = 50 dB

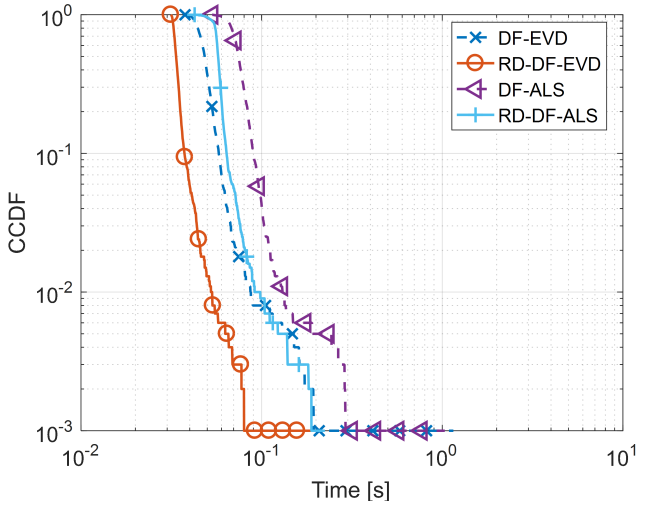

Fig. 4: CCDF of run time for a scenario with $M_1 = 20$, $M_2 = 30$, $M_3 = 10$, $d = 2$, SNR = 50 dB

A comparison of different versions of the DF algorithm in terms of the complementary cumulative distribution function (CCDF) of the residual is shown in Figure 3 whereas the CCDF of the corresponding run time is presented in Figure 4. The SNR is set to 50 dB. It can be observed in Figure 3 that for the two DF algorithms considered here, introducing the dimension reduction does not cause any performance

degradation. Figure 4 indicates that the proposed schemes require a significantly reduced run time.

Figure 5 and Figure 6 depict another two-component scenario with a much higher dimension, where $M_1 = 100, M_2 = 100, M_3 = 10$, and $d = 2$. In addition, the columns of the three-mode factor matrix $\boldsymbol{C}$ are strongly correlated with a correlation factor of 0.98. Similar observations are obtained

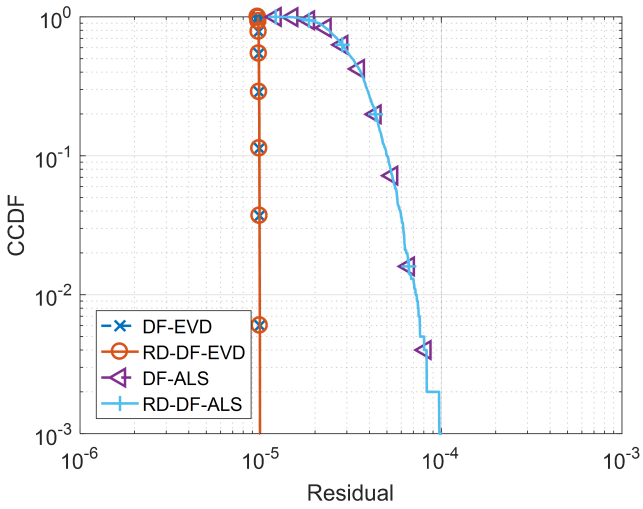

Fig. 5: CCDF of residuals for a scenario with $M_1 = 100, M_2 = 100, M_3 = 10$, $d = 2$, SNR = 50 dB, three-mode factor matrix $\boldsymbol{C}$ having correlated columns

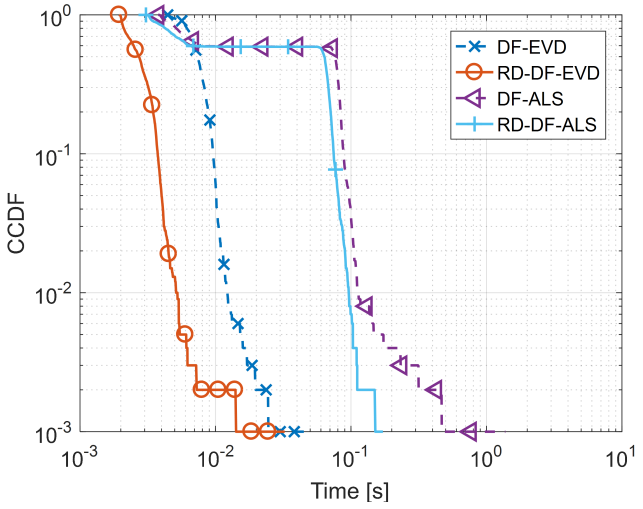

Fig. 6: CCDF of run time for a scenario with $M_1 = 100, M_2 = 100, M_3 = 10$, $d = 2$, SNR = 50 dB, three-mode factor matrix $\boldsymbol{C}$ having correlated columns with a correlation factor of 0.98

as those in Figure 3 and Figure 4. Furthermore, the superiority of the RD-DF-EVD and DF-EVD algorithms over their ALS-based counterparts is much more pronounced, as they are more immune to strong collinearity in the data.

Next we take a look at the impact of the size of the one-mode on the performance. Two three-component scenarios are

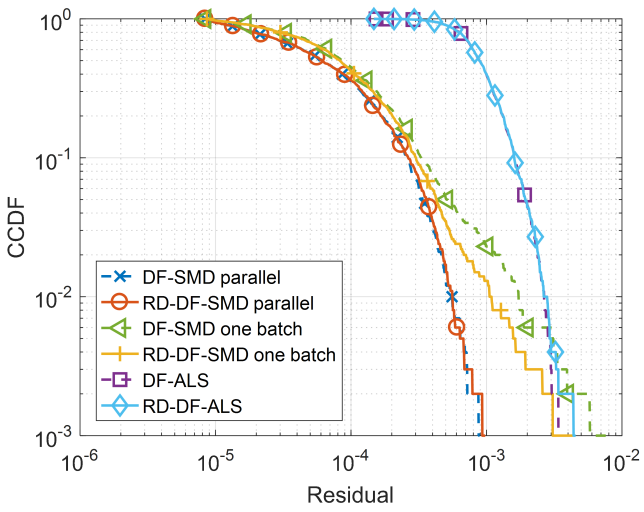

Fig. 7: CCDF of residuals for a scenario with $M_1 = 20, M_2 = 30, M_3 = 10, d = 3$, SNR = 40 dB

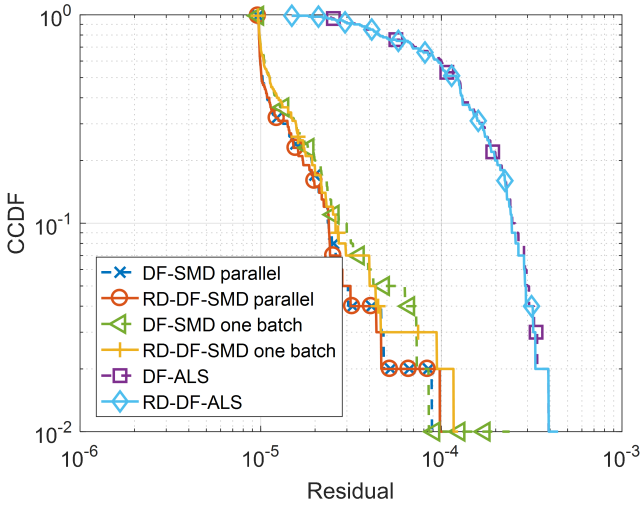

Fig. 8: CCDF of residuals for a scenario with $M_1 = 200, M_2 = 30, M_3 = 10$, $d = 3$, SNR = 50 dB

considered, i.e., for Figure 7 $M_1 = 20, M_2 = 30, M_3 = 10$, and $d = 3$, whereas $M_1 = 200, M_2 = 30, M_3 = 10$, and $d = 3$ for Figure 8. In the former, incorporating the dimension reduction gives rise to nearly 50 % reduction of the mean run time. As the size of the one-mode is increased to 200 in the second scenario, this figure approaches 80 %. Figure 7 and Figure 8 further corroborate the advantages of the proposed schemes by showing that there is no performance degradation at all.

Finally, we show a correlated three-component scenario, where $M_1 = 20, M_2 = 30, M_3 = 10$, and $d = 3$. The three-mode factor matrix $\boldsymbol{C}$ features correlated columns with a correlation factor of 0.98. The resulting Figure 9 and Figure 10 imply that the proposed extension is advantageous in terms of computation time.

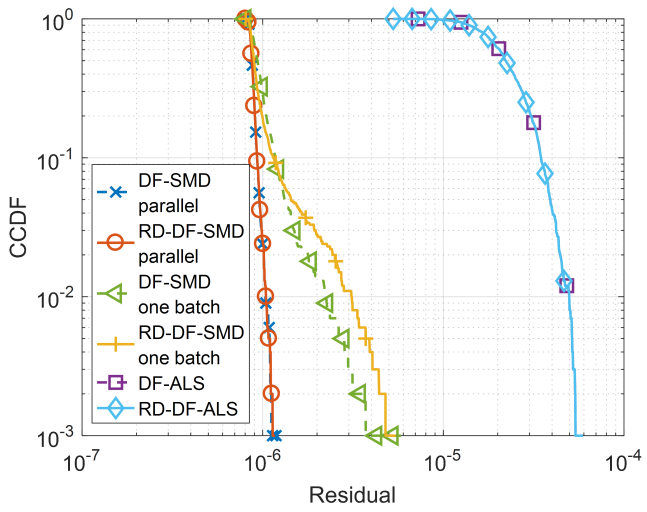

Fig. 9: CCDF of residuals for a scenario with $M_1 = 200, M_2 = 30, M_3 = 10$, $d = 3$, SNR = 60 dB, three-mode factor matrix $\boldsymbol{C}$ having correlated columns

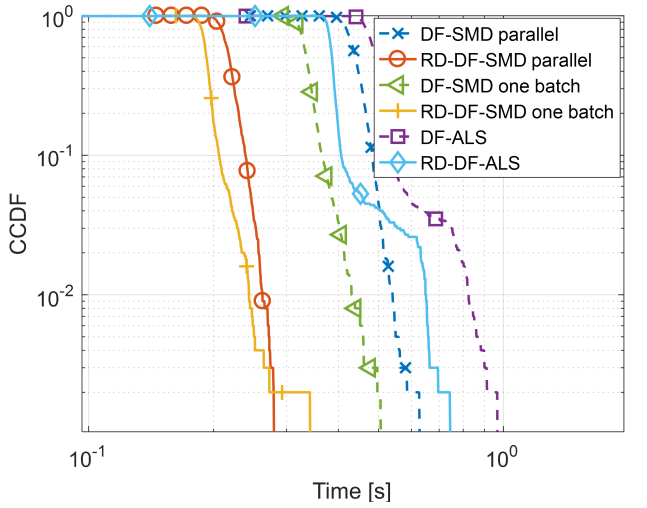

Fig. 10: CCDF of run time for a scenario with $M_1 = 200, M_2 = 30, M_3 = 10$, $d = 3$, SNR = 60 dB, three-mode factor matrix $\boldsymbol{C}$ having correlated columns with a correlation factor of 0.98

V. Conclusion

We have proposed an enhanced version of direct fitting-based schemes for the computation of PARAFAC2 based on a new dimension reduction concept aiming at a higher efficiency. The proposed extension can be applied to different versions of state-of-the-art computation algorithms of PARAFAC2. Numerical simulations show that the proposed algorithms are able to achieve the same performance in terms of the residual as their counterpart schemes in the literature. At the same time, they feature a lower computational complexity, require a significantly reduced computation time, and thus are very promising for applications with large-dimensional data.

Acknowledgment

The authors gratefully acknowledge the financial support by the Carl-Zeiss-Foundation (http://carl-zeiss-stiftung.de/).

References

[1] L. G. Johnsen, P. B. Skou, B. Khakimov, and R. Bro, "Gas chromatography – mass spectrometry data processing made easy," *Journal of Chromatography A*, vol. 1503, pp. 57 – 64, 2017.

[2] B. Khakimov, R. J. Mongi, K. M. Sørensen, B. K. Ndabikunze, B. E. Chove, and S. B. Engelsen, "A comprehensive and comparative GCMS metabolomics study of non-volatiles in Tanzanian grown mango, pineapple, jackfruit, baobab and tamarind fruits," *Food Chem.*, vol. 213, pp. 691 – 699, 2016.

[3] M. Vosough and A. Salemi, "Exploiting second-order advantage using PARAFAC2 for fast HPLC-DAD quantification of mixture of aflatoxins in pistachio nuts," *Food Chem.*, vol. 127, pp. 827 – 833, 2011.

[4] A. Herrero, C. Reguera, M. C. Ortiz, and L. A. Sarabia, "Determination of dichlobenil and its major metabolite (BAM) in onions by PTV-GC-MS using PARAFAC2 and experimental design methodology," *Chemom. Intell. Lab. Syst.*, vol. 133, pp. 92 – 108, 2014.

[5] I. García, M. C. Ortiz, L. Sarabia, and J. M. Aldama, "Validation of an analytical method to determine sulfamides in kidney by HPLC-DAD and PARAFAC2 with first-order derivative chromatograms," *Anal. Chim. Acta.*, vol. 587, pp. 222 – 234, 2007.

[6] L. Khamidullina, Y. Cheng, and M. Haardt, "Constrained Tensor Decompositions for Semi-blind MIMO Detection," in *Proc. 23rd International ITG Workshop on Smart Antennas (WSA)*, Apr. 2019.

[7] M. Weis, D. Jannek, T. Guenther, P. Husar, F. Roemer, and M. Haardt, "Temporally resolved multi-way component analysis of dynamic sources in event-related EEG data using PARAFAC2," in *Proc. 18th European Signal Processing Conference (EUSIPCO 2010)*, Aug. 2010.

[8] J. Haueisen, B. Schack, T. Meier, G. Curio, and Y. Okada, "Multiplicity in the high-frequency signals during the short-latency somatosensory evoked cortical activity in humans," *Clinical Neurophysiology*, vol. 112, pp. 1316 – 1325, 2001.

[9] J. Haueisen, L. Leistritz, T. Süsse, G. Curio, and H. Witte, "Identifying mutual information transfer in the brain with differential-algebraic modeling: Evidence for fast oscillatory coupling between cortical somatosensory areas 3b and 1," *NeuroImage*, vol. 37, pp. 130 – 136, 2007.

[10] Y. Cheng, M. Haardt, T. Götz, and J. Haueisen, "Using PARAFAC2 for multi-way component analysis of somatosensory evoked magnetic fields and somatosensory evoked electrical potentials," in *Proc. 10th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2018)*, July 2018.

[11] H. A. L. Kiers, J. M. F. Ten Berge, and R. Bro, "PARAFAC2 Part I. A direct fitting algorithm for the PARAFAC2 model," *J. Chemometrics*, vol. 13, pp. 275 – 294, 1999.

[12] R.A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis," UCLA Work. Pap. Phonetics., 1970.

[13] J. Kiefer, "Sequential Minimax Search for a Maximum," *Proc. Am. Math. Soc.*, vol. 4, pp. 502 – 506, 1953.

[14] K. Tian, L. Wu, S. Min, and R. Bro, "Geometric search: A new approach for fitting PARAFAC2 models on GC-MS data," *Talanta*, vol. 185, pp. 378 – 386, 2018.

[15] Y. Cheng and M. Haardt, "Enhanced direct fitting algorithms for PARAFAC2 with algebraic ingredients," *IEEE Signal Processing Letters*, vol. 26, no. 4, pp. 533 – 537, Feb. 2019.

[16] L. de Lathauwer, B. de Moor, and J. Vanderwalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253 – 1278, 2000.

[17] R. A. Harshman, "PARAFAC2: Mathematical and technical notes," *UCLA Working Papers in Phonetics*, vol. 22, pp. 30 – 44, 1972.

[18] P. H. Schoenemann, "A generalized solution of the orthogonal Procrustes problem," *Psychometrika*, vol. 31, no. 1, 1966.

[19] F. Roemer and M. Haardt, "A semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI)," *Elsevier Signal Processing*, vol. 93, pp. 2722 – 2738, Sept. 2013.