Jussi Salmi, Andreas Richter, and Visa Koivunen. 2008. Sequential Unfolding SVD for low rank orthogonal tensor approximation. In: Michael B. Matthews (editor). Conference Record of the 42nd Asilomar Conference on Signals, Systems and Computers (ACSSC 2008). Pacific Grove, CA, USA. 26-29 October 2008, pages 1713-1717

Sequential Unfolding SVD for Low Rank Orthogonal Tensor Approximation

Jussi Salmi, Andreas Richter and Visa Koivunen
Department of Signal Processing and Acoustics/SMARAD CoE
Helsinki University of Technology
Espoo, Finland

***Abstract*— This paper contributes to the field of N-way ($N \geq 3$) tensor decompositions, which are increasingly popular in various signal processing applications. A novel PARATREE decomposition structure is introduced, accompanied with Sequential Unfolding SVD (SUSVD) algorithm. SUSVD applies a matrix SVD sequentially on the unfolded tensor, which is reshaped from the right hand basis vectors of the SVD of the previous mode. The consequent PARATREE model is related to the well known family of PARAFAC tensor decompositions, describing a tensor as a sum of rank-1 tensors. PARATREE is an efficient model to be used for orthogonal lower rank approximations, offering significant computational savings in algorithm implementations due to a hierarchical tree structure. The performance of the proposed algorithm is illustrated through an application of measurement noise suppression in wideband MIMO measurements.**

I. Introduction

A tensor is any N-dimensional collection of data (a second order tensor $N = 2$ is a matrix). In many applications, measurement data is composed of information in more than two dimensions. Recently, researchers in several application domains have contributed to generalizing well established matrix operations to their tensor counter parts. Unfortunately, these extensions from their matrix counterparts are not trivial. For instance, the Singular Value Decomposition (SVD) has proven to be a powerful tool for analyzing two way tensors (matrices). However, its generalization to higher order tensors is not straightforward. There are in practice two major classes of models for higher order tensor decomposition, namely Tucker [1] and PARAFAC [2], [3] (parallel factorization), which is also known as CANDECOMP (canonical decomposition [4]). A well written overview of the two approaches can be found in [5]. We use the same notational conventions in this paper, too.

PARAFAC tensor decompositions are based on multilinear analysis in the fields of psychometrics [1], sociology, chromatography and chemometrics [6]. PARAFAC has been recently applied in several signal processing applications, such as image recognition, acoustics, wireless channel estimation [7] and array signal processing [8]. Also Tucker-type (HOSVD) models have been proposed for multidimensional harmonic retrieval [9]. In this paper a novel tensor representation stemming from the PARAFAC model is introduced. The new model will be referred to as PARATREE. The key idea is to sequentially unfold the tensor and to apply the singular value decomposition (SVD) on the resulting matrix. This is repeated for the remaining tensor dimensions remaining in the right-hand singular vectors of the SVD, until a hierarchical tree structure for the factors in each dimension is obtained. This decomposition procedure will be referred to as Sequential Unfolding SVD (SUSVD).

In general, PARAFAC does not provide an orthogonal decomposition unless some additional constraints are imposed. The proposed SUSVD forms an inherently *orthogonal decomposition*, which allows efficient selection of the number of factors used in the resulting PARATREE model — independently for each branch of factors.

The formulation of a tensor decomposition as a sum of orthogonal rank-1 tensors is suitable for many applications (see e.g. [10]), where there is advantage of processing or reconstructing multidimensional data from smaller dimensional orthogonal elements. One important class of applications is the reduction of computational complexity in linear algebraic expressions, involving tensor valued measurement data impaired by colored noise. An example of such application was discussed in [11] involving low rank tensor approximation to reduce the complexity of finding a Fisher Information Matrix for a tensor valued data model.

In this paper the low rank PARATREE approximation is applied to suppress measurement noise in wideband MIMO channels, obtained by channel sounding measurements [12], [13]. Improving the SNR is crucial in order to apply noisy measured channel realizations in simulations (see e.g. [14]). An improved estimate of the channel is obtained by estimating the signal subspace from the measured tensor.

To summarize, the benefits of the proposed SUSVD/PARATREE approach include:

- Measurement noise suppression
- Compression of data (similar to low rank matrix approximation)
- Reduced data dimensions and consequently relaxed memory requirements
- Orthogonality, which facilitates scalable decomposition order selection
- Fast and adaptive computation.

II. PARATREE Model

The PARATREE model is related to the family of PARAFAC [2], [3] tensor decompositions. The basic idea in PARAFAC is, as with SVD for matrices, to express the tensor as a sum of rank-1 components. Each of these rank-1 components is an outer product of the N corresponding basis vectors. Hence, for a three dimensional ($N = 3$) tensor $\mathcal{X}\mathbb{C}^{M_1 \times M_2 \times M_3}$, a rank-$R$ PARAFAC decomposition is given by

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \mathbf{a}_r^{(3)}, \tag{1}$$

where $\circ$ denotes the outer product (see e.g. [5]). The tensor $\mathcal{X}$ can be expressed element-wise as

$$x_{m_1,m_2,m_3} = \sum_{r=1}^{R} a_{r,m_1}^{(1)} \cdot a_{r,m_2}^{(2)} \cdot a_{r,m_3}^{(3)}, \tag{2}$$

where m_i denotes the index of the tensor element in i^{th} dimension. The PARAFAC model is illustrated in Figure 1.

The PARATREE differs from PARAFAC by introducing hierarchy for the factors in different dimensions. PARATREE decomposition for a three dimensional ($N = 3$) tensor is defined as

$$\mathcal{X} = \sum_{r_1=1}^{R_1} \mathbf{a}_{r_1}^{(1)} \circ \sum_{r_2=1}^{R_{2|r_1}} \left\{ \mathbf{a}_{r_{2|r_1}}^{(2)} \circ \mathbf{a}_{r_{2|r_1}}^{(3)} \right\}, \tag{3}$$

where each of the basis vectors $\mathbf{a}_{r_1}^{(1)}$ may be common to several ($R_{2_{|r_1}}$) vectors in latter dimensions. The subscript $|r^{(1)}, \ldots, r^{(n-1)}$ indicates the dependency of the basis vectors on the indexes of the previous factors of that branch in the decomposition tree. The structure of the factors in the full PARATREE decomposition for a $N = 4$ tensor $\mathcal{X} \in \mathbb{C}^{2\times2\times2\times2}$ is illustrated in Fig. 2.

The main difference between PARATREE and HOSVD [15] is that PARATREE allows an independent set of basis vectors $\mathbf{a}_{r_n|r_1,\ldots,r_{n-1}}^{(n)}$ for each branch $(|r_1, \ldots, r_{n-1})$, whereas in HOSVD the number is limited to $R_n = M_n$.

III. Sequential Unfolding SVD (SUSVD) Algorithm

The SUSVD algorithm can be applied to estimate the PARATREE model for given tensor-valued data. The power of the algorithm lies in its recursive and inherently orthogonal structure. Furthermore, it allows adaptive selection of the order of the decomposition. SUSVD can be applied to any N-dimensional (real or complex) tensor, and for $N = 2$ it equals the matrix SVD.

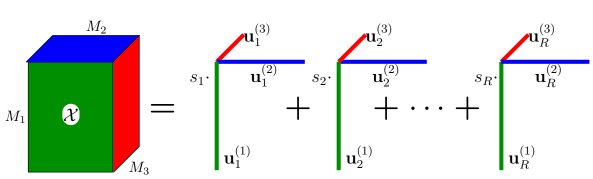


Fig. 1: Illustration of the PARAFAC decomposition — a sum of R rank-1 tensors. The relation to (1) is given by $\mathbf{a}_r^{(n)} = s_r \mathbf{u}_r^{(n)}$.

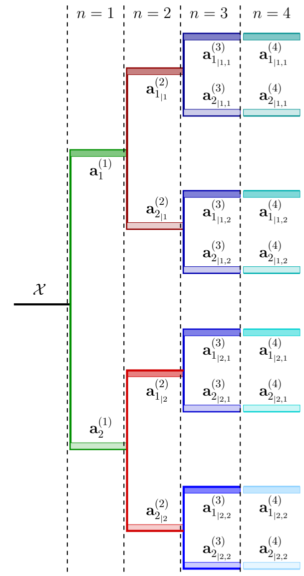


Fig. 2: An example of a full PARATREE for a four-way, $(2 \times 2 \times 2 \times 2)$ tensor. Different colors represent different modes $n = \{1, 2, 3, 4\} \equiv \{green, red, blue, cyan\}$, and the tone within each mode represents decreasing magnitude.

A. Full SUSVD

The description of a complete SUSVD for a N-way tensor $\mathcal{X}^{M_1 \times \cdots \times M_N}$ $(M_1 \geq \cdots \geq M_N)$ is provided in Table I, and visualized for a $2 \times 2 \times 2$ three-way tensor in Fig. 3a.

The core idea of the algorithm is to apply the matrix SVD on the 1-mode matrix unfolding[1] of the tensor to form the factors of the first mode. Then each of the conjugated right-hand singular vectors $\left(\mathbf{v}_{r^{(1)}}^{(1)}\right)^*$ are reshaped into tensors, and the procedure is repeated on the 1-mode unfoldings of these tensor. This is repeated to construct the PARATREE model, until there is only the last mode contained in the right-hand singular vectors. Note that for a full SUSVD (all possible factors included) described in Table I, the number of factors in each mode is the same for all branches and is given by

$$R_n = \max(\min(M_n, \prod_{j=n+1}^{N-1} M_j), 1). \tag{4}$$

Hence, the total number of orthogonal components in the decomposition is given by

$$R = \prod_{n=1}^{N-1} R_n. \tag{5}$$

The $2 \times 2 \times 2$ tensor in Fig. 3a can be reconstructed with the PARATREE model as

$$\mathcal{X} = \sum_{r_1=1}^{R_1} \sigma_{r_1} \cdot \mathbf{u}_{r_1}^{(1)} \circ \sum_{r_2=1}^{R_2} \sigma_{r_{2|r_1}}^{(2)} \cdot \mathbf{u}_{r_{2|r_1}}^{(2)} \circ \left(\mathbf{v}_{r_{2|r_1}}^{(2)}\right)^*. \tag{6}$$

[1]The 1-mode matrix unfolding sets the indexes of the first mode as the row indexes, and stacks the rest of the modes in the columns.

TABLE I: Description of the SUSVD algorithm for a complete decomposition. The output variables $\{S\}$, $\{U\}$, and $\{V\}$ denote abstract tree structures to store the elements of the decomposition.

$[\{S\},\{U\},\{V\}] = \mathrm{SUSVD}\,(\mathcal{X})$

- Set $\mathcal{T}_0^{(1)} = \mathcal{X}$.
- Set $R^{(0)} = 1$
- For each $n = \{1, \ldots, N-1\}$:
  - For each $r^{(n-1)} = \{1, \ldots, R^{(n-1)}\}$

    1) Unfold $\mathcal{T}^{(n)}_{|r^{(n-1)}}$ into a matrix $\mathbf{T}^{(n)}_{|r^{(n-1)}} \in \mathbb{C}^{M_n \times \prod_{q=n+1}^{N} M_q}$ on the first dimension.

    2) Compute the SVD $\mathbf{T}^{(n)}_{|r^{(n-1)}} = \mathbf{U}^{(n)}_{|r^{(n-1)}} \mathbf{\Sigma}^{(n)}_{|r^{(n-1)}} \left(\mathbf{V}^{(n)}_{|r^{(n-1)}}\right)^{\mathrm{H}}$.

    3) Pick $R^{(n)} = \min\left(M_n, \prod_{q=n+1}^{N} M_q\right))$ largest singular values $\sigma^{(n)}_{r^{(n)}_{|r^{(n-1)}}}$, and for each $r^{(n)} \in \{1, \ldots, R^{(n)}\}$

       a) Store $\sigma^{(n)}_{r^{(n)}_{|r^{(n-1)}}}$ in $\{S\}$ and $\mathbf{u}^{(n)}_{r^{(n)}_{|r^{(n-1)}}}$ in $\{U\}$, and

       b) if $n < N-1$,
       - Reshape $\left(\mathbf{v}^{(n)}_{r^{(n)}_{|r^{(n-1)}}}\right)^*$ into a tensor $\mathcal{T}^{(n+1)}_{|r^{(n)}} \in \mathbb{C}^{M_{n+1} \times \cdots \times M_N}$.

       or else,
       - Store the vectors $\left(\mathbf{v}^{(N-1)}_{r^{(n)}_{|r^{(n-1)}}}\right)^*$ in $\{V\}$.

The full ($R_1 = 2$, $R_2 = 2$) reconstruction is illustrated in Fig. 3b. The relation of the values in (6) to the ones in the 3D-PARATREE formulation (3) is given by

$$\begin{aligned} \mathbf{a}^{(1)}_{r_1} &= \sigma^{(1)}_{r_1} \mathbf{u}^{(1)}_{r_1} \\ \mathbf{a}^{(2)}_{r_{2|r_1}} &= \sigma^{(2)}_{r_{2|r_1}} \mathbf{u}^{(2)}_{r_{2|r_1}} \\ \mathbf{a}^{(3)}_{r_{2|r_1}} &= \left(\mathbf{v}^{(2)}_{r_{2|r_1}}\right)^*. \end{aligned}$$

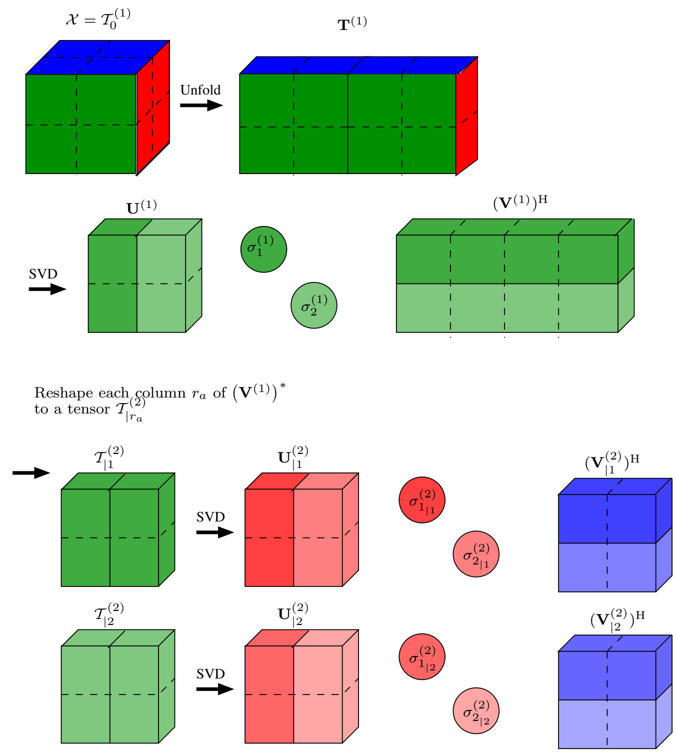
B. Low Rank SUSVD Approximation

The PARATREE model build with SUSVD can be deflated to form a reduced rank approximation of a tensor. This can be done either offline after bulding the full SUSVD, or online during the computation of the decomposition. Here the discussion is limited to the offline approach.

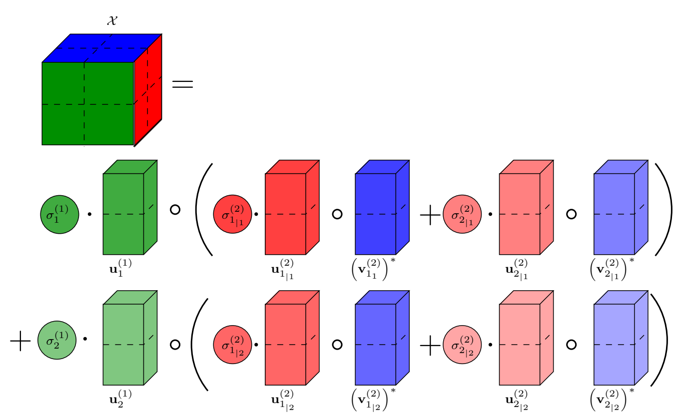
The approximation error is expressed as the normalized error, defined as

$$\epsilon_r = \frac{||\mathcal{X} - \mathcal{X}_A||_F}{||\mathcal{X}||_F}, \tag{7}$$

where $\mathcal{X}$ denotes the original tensor and $\mathcal{X}_A$ its approximation. Due to orthogonality of the decomposition [16], the approximation error can be equivalently expressed in terms of the sum of the product of the singular values of the factors. These



(a) The SUSVD computation. The tensor is first unfolded to a matrix $\mathbf{T}^{(1)}_0$. After applying SVD on this matrix, each of the right-hand singular vectors are reshaped (into $\mathbf{T}^{(n)}_{|r^{(n-1)}} \in \mathbb{C}^{M_n \times \prod_{k=n+1}^{N} M_k}$) and another SVD is applied on them. The procedure is repeated for each "branch" and "sub-branch", until no additional dimensions remain in the right hand basis vectors, i.e. $\left(\mathbf{v}^{(N-1)}_{i_j}\right)^*$ has only M_N elements.



(b) A PARATREE tensor is reconstructed as a sum of outer products of weighted (by $\sigma^{(n)}_{r_n|r_1 \cdots r_{n-1}}$) unitary basis vectors $\mathbf{u}^{(n)}_{r_{n|r_1 \cdots r_{n-1}}}$ and $\left(\mathbf{v}^{(N-1)}_{r_{N-1|r_1 \cdots r_{N-2}}}\right)^*$. The tree structure allows common basis vectors in the previous dimensions (main branch).

Fig. 3: SUSVD computation (Fig. 3a) and PARATREE reconstruction (Fig. 3b) for an arbitrary $2 \times 2 \times 2$ tensor . Different colors refer to different dimensions of the tensor. A circled σ denotes a singular value, dashed blocks are elements of the tensors, and solid lines are used to separate the column vectors.

can be interpreted as the magnitudes of the single rank-1 components in the PARATREE, given by

$$\tilde{\sigma}_{|r^{(1)},\ldots,r^{(N-1)}} = \sigma^{(1)}_{r^{(1)}} \cdot \sigma^{(2)}_{r^{(2)}_{|r^{(1)}}} \cdot \cdots \cdot \sigma^{(n-1)}_{r^{(n-1)}_{|r^{(1)},\ldots,r^{(n-2)}}}. \quad (8)$$

By stacking all the R (4)-(5) magnitude values (8) in descending order to a vector $\tilde{\boldsymbol{\sigma}} \in \mathbb{R}^{R\times 1}$, the normalized approximation error (7) can be expressed as

$$\epsilon_r = \frac{||\mathcal{X} - \mathcal{X}_A||_F}{||\mathcal{X}||_F} = \sqrt{1 - \frac{\sum_{i=1}^{R_A} \tilde{\sigma}_i^2}{\sum_{i=1}^{R} \tilde{\sigma}_i^2}}, \quad (9)$$

or equivalently

$$\epsilon_r^2 = 1 - \frac{||\tilde{\boldsymbol{\sigma}}_A||_F^2}{||\tilde{\boldsymbol{\sigma}}||_F^2}. \quad (10)$$

The described approach allows to define the achieved relative approximation error (7) precisely. It should be noted that applying a similar target error requirement for a general PARAFAC model would require a trial and error approach for finding a proper rank (as well as for determining the tensor-rank in general). Also the convergence of the alternating least squares (ALS [17]) algorithms used for PARAFAC is very slow for high dimensional or ill-conditioned problems [18], [19].

IV. Application Example: Noise Suppression for Measured Wideband MIMO Channels

The SUSVD may be used for estimating the signal subspace of the wideband MIMO radio channel. A 3-way tensor model for an instantaneous measured MIMO channel [13], [20] transfer function is given

$$\mathcal{H} = \mathcal{H}_S(\boldsymbol{\theta}_S) + \mathcal{H}_D(\boldsymbol{\theta}_D) + \mathcal{H}_N(\sigma^2) \in \mathbb{C}^{M_f \times M_T \times M_R}, \quad (11)$$

where M_f, M_T and M_R denote the number of frequency samples, transmit antennas and receive antennas, respectively. The tensor $\mathcal{H}_S$ denotes the superposition of dominant propagation paths (deterministic plane waves) parameterized by $\boldsymbol{\theta}_S$, $\mathcal{H}_D$ models diffuse scattering (a colored complex Normal distributed noise process) parameterized by $\boldsymbol{\theta}_D$, and $\mathcal{H}_N$ is measurement noise modeled as a circular symmetric white Normal distributed process with variance σ^2. Detailed model description may be found in [20] and references therein.

The nominal signal-to-noise ratio (SNR) of the measurement model is defined as

$$S_{dB}(\mathcal{H}) = 10 \cdot \log_{10}\left(\frac{||\mathcal{H}_S + \mathcal{H}_D||_F^2}{||\mathcal{H}_N||_F^2}\right), \quad (12)$$

which can be approximated from measurement data as

$$S_{dB}(\mathcal{H}) \approx 10 \cdot \log_{10}\left(\frac{||\mathcal{H}||_F^2 - ||\mathcal{H}_N||_F^2}{||\mathcal{H}_N||_F^2}\right). \quad (13)$$

These quantities are assumed to be known, which is a valid assumption in channel sounding measurements . The suppression of the measurement noise is achieved by the following procedure:

1) Compute the SUSVD of $\mathcal{H}$, as described in Table I.
2) Define a threshold ϵ_r (7) for selecting the factors for the approximation. Here

$$\epsilon_r = \sqrt{\frac{||\mathcal{H}_N||_F^2}{||\mathcal{H}||_F^2}} \quad (14)$$

is chosen, i.e., only the factors whose cumulative energy exceeds the total noise energy are included in the decomposition.
3) Approximate $\mathcal{H}$ by $\hat{\mathcal{H}}_A$, where $\hat{\mathcal{H}}_A$ is obtained by sorting the values (8) and choosing the corresponding branches $\tilde{\boldsymbol{\sigma}}_A$ whose magnitude exceeds ϵ_r.

This method effectively suppresses the measurement noise as is shown in Section V.

V. Results

To test the performance of the proposed method, realizations of the wideband MIMO channel (11) were generated. The parameters $\boldsymbol{\theta}_S$ and $\boldsymbol{\theta}_D$ (11) were estimated (see [20]) from actual channel sounding measurements [12]. The SNR (12) was controlled by adjusting the level of receiver noise variance σ^2.

Fig. 4 shows Power-Delay Profiles (PDPs) of simulated data, transformed into delay domain and averaged over $M_T \cdot M_R$ antenna pairs. Fig. 4a shows a realization with $S_{dB} = 0$. One can observe that the proposed low rank approximation suppresses the noise level by $G(S_{dB} = 0) \approx 15$ dB. For the same data set, but $S_{dB} = 10$ in Fig. 4a, the suppression is about $G(S_{dB} = 10) \approx 12$ dB. This performance difference results from the definition of the signal/noise subspace threshold criteria (14). Higher initial SNR yields higher rank approximation, and the obtainable SNR improvement is upper bounded by the (inverse of the compression) ratio $10\log_{10}(R/R_A)$. Table II summarizes these numbers for the examples in Fig. 4 ($R = 5790$).

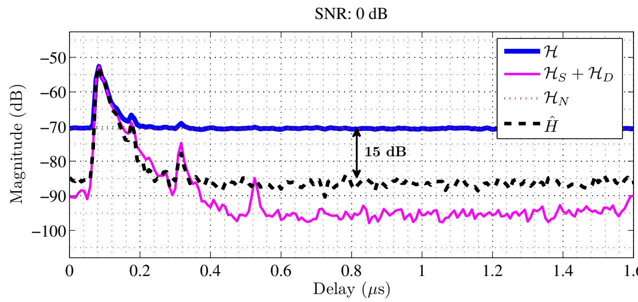
The performance of the algorithm using measured data [12] directly is shown in Fig. 5. The SNR in the beginning of the measurement was very low, and filtering provides about 15 dB improvement. In the end of the route the SNR is far better, and the improvement from filtering is not as significant.

VI. Conclusions

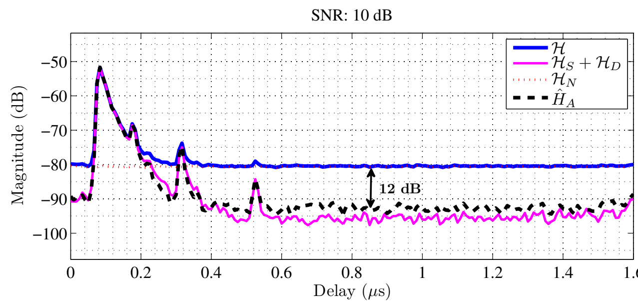
In this paper, we introduce a novel PARATREE tensor model and SUSVD algorithm. Together these form an efficient approach for applying an orthogonal tensor decomposition with a sum of rank-1 tensors. The proposed approach provides fast and reliable tensor compression, and it is applicable in

TABLE II: Rank of the approximation (total number of components $R = 5790$), compression ratio, and SNR improvement for the examples in Fig. 4.

	$S_{dB} = 0$	$S_{dB} = 10$
Rank R_A	99	249
$10\log_{10}(R/R_A)$	18 dB	14 dB
G_{dB}	15 dB	12 dB

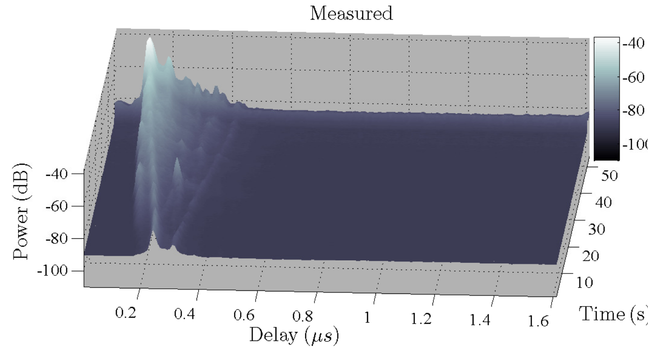


(a) Simulation with $S_{dB} = 0$.

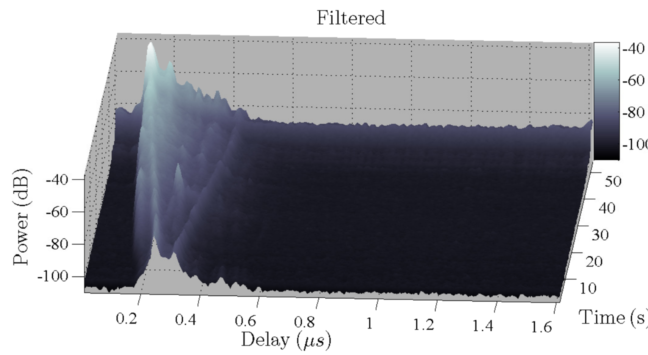


(b) Simulation with $S_{dB} = 10$.

Fig. 4: Example of the noise suppression performance with different SNR conditions.



(a) Before filtering.



(b) After filtering.

Fig. 5: Power-Delay Profile over measurement time before and after filtering. The PDPs are averaged over all Tx-Rx channels.

many signal processing applications. An application example was provided for a low rank approximation, providing measurement noise suppression for wideband MIMO channel sounding measurement data.

References

[1] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 36, pp. 279–311, 1966.

[2] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multi-modal factor analysis," *UCLA working papers in phonetics*, vol. 16, 1970.

[3] J. B. Kruskal, "Three-way arrays: rank uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, pp. 95–138, 1977.

[4] J. D. Carroll and J. J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition," *Psychometrika*, vol. 35, pp. 283–319, 1970.

[5] T. G. Kolda, "Multilinear operators for higher-order decompositions," Sandia National Laboratories, Albuquerque, NM and Livermore, CA, Technical Report SAND2006-2081, April 2006.

[6] N. K. M. Faber, R. Bro, and P. K. Hopke, "Recent developments in CANDECOMP/PARAFAC algorithms: a critical review," *Chemometrics and Intelligent Laboratory Systems*, vol. 65, no. 1, pp. 119–137, January 2003.

[7] A. L. F. de Almeida, G. Favier, J. C. M. Mota, and R. L. de Lacerda, "Estimation of frequency-selective block-fading MIMO channels using PARAFAC modeling and alternating least squares," in *Proc. The 40th Asilomar Conference on Signals, Systems and Computers*, Oct.-Nov. 2006, pp. 1630–1634.

[8] N. Sidiropoulos, R. Bro, and G. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Transactions on Signal Processing*, vol. 48, no. 8, pp. 2377–2388, Aug 2000.

[9] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3198–3213, July 2008.

[10] D. Muti and S. Bourennane, "Survey on tensor signal algebraic filtering," *Signal Processing*, vol. 87, no. 2, pp. 237–249, 2007.

[11] J. Salmi, A. Richter, and V. Koivunen, "Tracking of MIMO propagation parameters under spatio-temporal scattering model," in *Proc. The 41st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2007, pp. 666–670.

[12] P. Almers, K. Haneda, J. Koivunen, V.-M. Kolmonen, A. F. Molisch, A. Richter, J. Salmi, F. Tufvesson, and P. Vainikainen, "A dynamic multi-link MIMO measurement system for 5.3 GHz," in *Proc. URSI General Assembly*, Chicago, USA, Aug 7-16 2008.

[13] A. Richter, J. Salmi, and V. Koivunen, "Tensor decomposition of MIMO channel sounding measurements and its applications," in *Proc. URSI General Assembly*, Chicago, USA, Aug 7-16 2008.

[14] P. Hammarberg, P. S. Rossi, F. Tufvesson, O. Edfors, V.-M. Kolmonen, P. Almers, R. Mueller, and A. Molisch, "On the performance of iterative receivers for interfering MIMO-OFDM systems in measured channels," in *Proc. The 42th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Oct. 2008.

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

[16] J. Salmi, A. Richter, and V. Koivunen, "Sequential unfolding SVD for tensors with applications in array signal processing," *Submitted to IEEE Trans. on Signal Processing*, 2008.

[17] P. Kroonenberg and J. Leeuw, "Principal component analysis of three-mode data by means of alternating least squares algorithms," *Psychometrika*, vol. 45, no. 1, pp. 69–97, March 1980.

[18] G. Tomasi and R. Bro, "A comparison of algorithms for fitting the PARAFAC model," *Computational Statistics & Data Analysis*, vol. 50, no. 7, pp. 1700–1734, April 2006.

[19] P. K. Hopke, P. Paatero, H. Jia, R. T. Ross, and R. A. Harshman, "Three-way (PARAFAC) factor analysis: examination and comparison of alternative computational methods as applied to ill-conditioned data," *Chemometrics and Intelligent Laboratory Systems*, vol. 43, pp. 25–42, September 1998.

[20] J. Salmi, A. Richter, and V. Koivunen, "Detection and tracking of MIMO propagation path parameters using state-space approach," *Accepted for publication in IEEE Trans. on Signal Processing*, 2008.