Multi-dimensional model order estimation using LineAr Regression of Global Eigenvalues (LaRGE) with applications to EEG and MEG recordings

Alexey A. Korobkov[*], *Marina K. Diugurova*[*], *Jens Haueisen*[†], and *Martin Haardt*[‡]
[*]Institute of Radio Electronics and Telecommunications
Kazan National Research Technical University n.a. A.N Tupolev-KAI, Kazan, Russia, Email: korobkov@inbox.ru
[†]Institute of Biomedical Engineering and Informatics, Ilmenau University of Technology, Ilmenau, Germany
[‡]Communications Research Laboratory, Ilmenau University of Technology, Ilmenau, Germany
Email: martin.haardt@tu-ilmenau.de

***Abstract*—The efficient estimation of an approximate model order is very important for real applications with multi-dimensional data if the observed low rank data is corrupted by additive noise. In this paper, we present a novel robust method for model order estimation of multi-dimensional data based on the LineAr Regression of Global Eigenvalues (LaRGE). The LaRGE method uses the multi-linear singular values obtained from the HOSVD of the measurement tensor to construct global eigenvalues. In contrast to the Modified Exponential Test (EFT) that also exploits the approximate exponential profile of the noise eigenvalues, LaRGE does not require the calculation of the probability of false alarm. Therefore, it is well suited for the analysis of biomedical data. The excellent performance of the LaRGE method is illustrated via simulations and results obtained from EEG as well as MEG recordings.**

***Index Terms*—tensor, the rank of the tensor, global eigenvalue, eigenvalue**

I. Introduction

Multi-dimensional models are widespread in a variety of applications, for example, radar, sonar, channel modeling in wireless communications, image processing, the estimation of MIMO channels parameters, blind source separation and many more [1]. According to these models the measured signals or the data can be staked into multi-dimensional arrays or tensors. Moreover, in biomedical data processing multi-dimensional models have been widely used recently. For example, biomedical signals like Electroencephalograms (EEG), Magnetoencephalograms (MEG) or Electrocardiograms (ECG) are recorded from many sensors simultaneously. Therefore, it is natural to use multi-dimensional models or tensors for representing the such sets of signals.

Different types of tensor decompositions are used for the extraction of features from the data or to denoise measured and recorded signals. However, in biomedical signal processing the most frequently used decompositions of multi-dimensional data are an approximate low-rank Canonical Polyadic (CP) decomposition [2] also known as Parallel Factor (PARAFAC) analysis [3] or Canonical Decomposition (CANDECOMP) [4] and the truncated Multi-Linear Singular Value Decomposition (MLSVD) [5] also known as Higher Order Singular Value Decomposition (HOSVD).

According to the CP model, a tensor is decomposed into the minimum number R of rank one components. Hence, the proper choice of the model order affects the accuracy of the processing and subsequently also the interpretable results. Obviously, this problem is very important when measured and recorded biomedical data is processed and analyzed. However, in most cases the observed data is corrupted by noise. Therefore, the problem of estimating the order of an approximate low-rank model is a non-trivial task.

The first attempts to develop method for overcoming this problem were made in the beginning of 1970s. In [6] Akaike's information criterion (AIC) was proposed. This criterion takes the observed data structure into account. In [7] another criterion was proposed that penalizes the over parameterization more strongly than AIC. G.Schwarz proposed a Bayesian information criterion (BIC) for the estimation the order of linear models that describe independent and identically distributed observations [8]. M.Wax and T.Kailath implemented AIC and MDL for the detection of the number of signals in a multichannel time series [9]. However, these methods often fail when the number of the snapshots of the observed data is small.

To overcome this problem, the Exponential Fitting Test (EFT) based on the geometrical profile of noise-only eigenvalues has been developed and presented in [10]. The EFT allows the probability of false alarm to be controlled and predefined, which is a crucial point for systems such as RADARs. Moreover, in [11] the EFT was improved for multidimensional signals via the Modified Exponential Fitting Test (M-EFT) that outperforms all other schemes for the cases when the data is corrupted by Gaussian noise. Furthermore, the authors showed how the classical AIC and MDL methods can be improved. In [12] the authors presented R-dimensional (R-D) EFT that outperforms the other techniques for cases with Gaussian noise. For colored noise a new Closed-Form PARAFAC-based Model Order Selection (CFP-MOS) was proposed. Moreover, the AIC and MDL techniques were extended to the multi-dimensional cases. Several methods for estimating the order of CP models were presented in [13], [14], [15].

In this paper we present a novel robust method for model order estimation based on the the LineAr Regression of

Global Eigenvalues (LaRGE). In contrast to the EFT, it does not require an estimate of the probability of false alarm. In biomedical signal processing problems, an estimate of the model order is often based on the results of the visual inspection of the singular values or some assumptions about the structure of the data. Instead of this, the proposed LaRGE method helps to overcome this "inspection" problem and provides a systematic approach for model order estimation in practical applications.

In Section II the multi-dimensional data model is presented. Section III describes the proposed method. The simulation results that are presented in Section IV confirm the accuracy of the LaRGE method. Inspired by the good performance of the proposed method, we also implemented it for model order estimation of EEG and MEG recordings. The results are presented in Section V. In Section VI we conclude the paper.

II. Data model and notation

In this paper we use the following notation, a, $\mathbf{a}$, $\mathbf{A}$ and $\boldsymbol{\mathcal{A}}$ are used to denote a scalars, column vectors, matrices, and tensors, respectively. Moreover, $\mathbf{a}(i)$ defines the element (i) of a vector $\mathbf{a}$. The same applies to a matrix $\mathbf{A}$ (i,j) and a tensor $\boldsymbol{\mathcal{A}}$ (i,j,k). The tensor $\boldsymbol{\mathcal{I}}_{D,R}$ is an D-dimensional identity tensor of size $R\times R\times\ldots\times R$, which is equal to one if all D indices are equal and zero otherwise. The d-mode product between a D-way tensor of size M_d along mode $d = 1,2,\ldots,D$ represented as $\boldsymbol{\mathcal{A}}\in\mathbb{C}^{M_1\times M_2\times\cdots\times M_D}$ and a matrix $\boldsymbol{U}\in\mathbb{C}^{J\times M_d}$ is written as $\boldsymbol{\mathcal{A}}\times_d\boldsymbol{U}$. It is computed by multiplying all d-mode vectors of $\boldsymbol{\mathcal{A}}$ with $\boldsymbol{U}$, whereas the d-mode vectors of $\boldsymbol{\mathcal{A}}$ are obtained by varying the d-th index from 1 to M_D and keeping all other indices fixed. Aligning all d-mode vectors as the columns of a matrix yields the d-mode unfolding of $\boldsymbol{\mathcal{A}}$ which is denoted by $[\boldsymbol{\mathcal{A}}]_{(d)}\in\mathbb{C}^{M_d\times M_{d+1}\cdot\ldots\cdot M_D\cdot M_1\cdot\ldots\cdot M_{d-1}}$. The order of the columns is chosen according to [2].

The CP decomposition of a D-way noiseless tensor $\boldsymbol{\mathcal{X}}_0\in\mathbb{C}^{M_1\times M_2\times M_3\times\cdots\times M_D}$ is represented as

$$\boldsymbol{\mathcal{X}}_0=\boldsymbol{\mathcal{I}}_{D,R}\times_1\boldsymbol{F}_1\times_2\boldsymbol{F}_2\times_3\boldsymbol{F}_3\times\cdots\times_D\boldsymbol{F}_D, \tag{1}$$

where $\boldsymbol{F}_d\in\mathbb{C}^{M_d\times R}(d=1,2,3,\ldots D)$ are the factor matrices and R is the order of the CP model or the rank of the tensor $\boldsymbol{\mathcal{X}}_0$.

In practice, the recorded data is corrupted by noise. The tensor that is constructed from observations can be defined as

$$\boldsymbol{\mathcal{X}}=\boldsymbol{\mathcal{X}}_0+\boldsymbol{\mathcal{N}}, \tag{2}$$

where $\boldsymbol{\mathcal{N}}\in\mathbb{C}^{M_1\times M_2\times M_3\times\cdots\times M_D}$ is the additive noise tensor. Therefore, (1) can be rewritten as

$$\boldsymbol{\mathcal{X}}=\boldsymbol{\mathcal{I}}_{D,R}\times_1\boldsymbol{F}_1\times_2\boldsymbol{F}_2\times_3\boldsymbol{F}_3\times\cdots\times_D\boldsymbol{F}_D+\boldsymbol{\mathcal{N}}. \tag{3}$$

Obviously the rank of the tensor $\boldsymbol{\mathcal{X}}$ is not equal R. In general it is bigger. Therefore, after a CP decomposition of the tensor with observations $\boldsymbol{\mathcal{X}}$ we obtain the estimates $\hat{\boldsymbol{F}}_d\in\mathbb{C}^{M_d\times R}(d=1,2,3,\ldots D)$ of factor matrices $\boldsymbol{F}_d\in\mathbb{C}^{M_d\times R}(d=1,2,3,\ldots D)$

$$\boldsymbol{\mathcal{X}}\approx\boldsymbol{\mathcal{I}}_{D,R}\times_1\hat{\boldsymbol{F}}_1\times_2\hat{\boldsymbol{F}}_2\times_3\hat{\boldsymbol{F}}_3\times\cdots\times_D\hat{\boldsymbol{F}}_D. \tag{4}$$

On the other side, the HOSVD model of a tensor $\boldsymbol{\mathcal{X}}$ is defined by

$$\boldsymbol{\mathcal{X}}=\boldsymbol{\mathcal{S}}\times_1\boldsymbol{U}_1\times_2\boldsymbol{U}_2\times_3\boldsymbol{U}_3\times\cdots\times_D\boldsymbol{U}_D, \tag{5}$$

where $\boldsymbol{\mathcal{S}}\in\mathbb{C}^{M_1\times M_2\times M_3\times\cdots\times M_D}$ is the core tensor and $\boldsymbol{U}_r\in\mathbb{C}^{M_d\times M_d}$, $(d=1,2,3,\ldots D)$ are the unitary factor matrices.

As shown in [10], [11], [12] the d-mode singular values that are contained in the core tensor $\boldsymbol{\mathcal{S}}$ play a major role in the problem of multi-dimensional model order estimation. The d-mode singular values can be computed via the Singular Value Decomposition (SVD) of the d-mode unfolding of the tensor $\boldsymbol{\mathcal{X}}$ according to

$$[\boldsymbol{\mathcal{X}}]_{(d)}=\boldsymbol{U}_d\cdot\boldsymbol{\Sigma}_d\cdot\boldsymbol{V}_d^{\mathrm{H}}, \tag{6}$$

where $\boldsymbol{U}_d\in\mathbb{C}^{M_d\times M_d}$, $\boldsymbol{V}_d\in\mathbb{C}^{\tilde{M}_d\times\tilde{M}_d}$ are unitary matrices and $\boldsymbol{\Sigma}_d\in\mathbb{C}^{M_d\times\tilde{M}_d}$ is a diagonal matrix that has the d-mode singular values $\sigma_i^{(d)}$ on the main diagonal and $\tilde{M}_d=\frac{M}{M_d}$.

III. Multi-dimensional model order estimation

As is shown in (5), the d-mode singular values represent the internal structure of the data and should be exploited for model order estimation. However, tensors with real data may have different d-ranks in different modes. Moreover, when an approximate low-rank CP decomposition is computed, an estimate of the rank of the noiseless model that is equal for all d-mode unfoldings has to be found. Therefore, we can define *global eigenvalues* as proposed in [11] and [12] as the product of all d-mode singular values with the same indices as follows

$$\tilde{\lambda}_i^{(G)}=\prod_{d=1}^{D}\frac{M_d}{M}\left(\sigma_i^{(d)}\right)^2, i=1,2,3,\ldots,M^{(G)}, \tag{7}$$

where $M^{(G)}=\min(M_d), d=1,2,3,\ldots,D$. The global eigenvalues include the information about the signals and the noise. The product of the d-mode singular values increases the gap between signal global eigenvalues and noise global eigenvalues. Starting from the smallest global eigenvalue, the noise global eigenvalues follow a straight line on a logarithmic scale (due to their approximate exponential distribution)

$$\lambda_i^{(G)}=\lg\left(\tilde{\lambda}_i^{(G)}\right). \tag{8}$$

Once an actual eigenvalue is significantly larger than the one predicted by the straight line approximation, we have a significant signal component. The corresponding value of its index i is the estimated model order.

Consider the linear approximation or linear eigenvalue profile (LEP) $\hat{\lambda}_i^{(G)}$ of the actual global eigenvalues $\lambda_i^{(G)}$ based on least square optimization problem

$$\min\left(\sum_{i=M^{(G)}}^{M^{(G)}-k}\left(\hat{\lambda}_i^{(G)}-\lambda_i^{(G)}\right)^2\right), \tag{9}$$

where $\hat{\lambda}_i^{(G)}=a_1\cdot i+a_2, i=M^{(G)},M^{(G)}-1,M^{(G)}-2,\ldots,M^{(G)}-k$. We approximate the global noise eigenvalue via a straight line starting from the smallest global eigenvalues with indices $i=M^{(G)}$ and $i=M^{(G)}-1$ for $k=1$. Using this linear approximation the next largest noise eigenvalue can be

found. For each value of $k = 1, 2, ..., M^{(G)} - 1$ the variance of approximation error $\sigma^2_{M^{(G)}-k}$ is calculated

$$\sigma^2_{M^{(G)}-k} = \frac{1}{M^{(G)} - k} \sum_{i=M^{(G)}}^{M^{(G)}-k} \left(\Delta_i^{(G)} - m_{\Delta^{(G)}_{M^{(G)}-k}} \right)^2, \quad (10)$$

where $\Delta_i^{(G)}$ is the absolute prediction error defined as

$$\Delta_i^{(G)} = \lambda_i^{(G)} - \hat{\lambda}_i^{(G)} \quad (11)$$

and $m_{\Delta^{(G)}_{M^{(G)}-k}}$ is the expected value of $\Delta_i^{(G)}$ for $i = M^{(G)} - k, ...M^{(G)}$ that is defined as

$$m_{\Delta^{(G)}_{M^{(G)}-k}} = \frac{1}{M^{(G)}} \sum_{i=M^{(G)}}^{M^{(G)}-k} \Delta_i^{(G)} \quad (12)$$

The relative prediction error of $\lambda^{(G)}_{M^{(G)}-k-1}$ is defined as

$$\delta_{M^{(G)}-k-1} = \frac{\Delta^{(G)}_{M^{(G)}-k-1}}{\left|\hat{\lambda}^{(G)}_{M^{(G)}-k-1}\right|} = \frac{\lambda^{(G)}_{M^{(G)}-k-1} - \hat{\lambda}^{(G)}_{M^{(G)}-k-1}}{\left|\hat{\lambda}^{(G)}_{M^{(G)}-k-1}\right|}. \quad (13)$$

The model order or rank can be estimated by considering the Prediction Error to Standard Deviation Error Ratio (PESDR)

$$\text{PESDR}_i = \frac{\Delta_i}{\sigma_i}. \quad (14)$$

We define the normalized PESDR (nPESDR) according to

$$\text{PESDR}_{\text{NORM}_i} = \frac{\delta_i}{\sigma_i}, \quad (15)$$

where the δ_i is the relative prediction error that is defined in (13).

Clearly, the rank of a tensor can be estimated by analyzing the behavior of the PESDR or nPESDR curves. If it is significantly larger than a predefined threshold ρ (we use $\rho = 1$ in this paper) we have found the smallest signal global eigenvalue, i.e., the approximate model order $\hat{R}$. In the following section, we present some results of the implementation of the proposed LaRGE method for synthetic data.

IV. Experimental results. Synthetic simulations

A data tensor for the CP model (1) with factor matrices $\boldsymbol{F}_d, d = 1, 2, 3$ is generated. The factor matrices contain elements drawn from a zero-mean Gaussian distribution with unit variance σ_s^2. To evaluate the robustness of the proposed LaRGE method, factor matrices with independent elements and with correlated elements are formed. The correlation of the columns of all factor matrices was defined by the vectors $\boldsymbol{r}_1 = [0\ 0\ 0]^{\text{T}}$ and $\boldsymbol{r}_2 = [0.6\ 0.7\ 0.3]^{\text{T}}$ for the uncorrelated and correlated cases of factor matrices, respectively [16]. According to the model (3) the elements of the noise tensor $\boldsymbol{\mathcal{N}}$ are drawn independently from a zero-mean Gaussian distribution with variance σ_n^2. The SNR is defined as

$$\text{SNR} = 10 \cdot \log_{10}\left(\frac{\sigma_s^2}{\sigma_n^2}\right), \text{ dB}. \quad (16)$$

The resulting tensor has the size $100 \times 350 \times 30$. In Fig. 1 the linear regression and normalized PESDR are depicted for the uncorrelated case and SNR=10 dB. We can clearly see that starting from the smallest global eigenvalues the value of the normalized PESDR exceeds the threshold $\rho = 1$ for the first time when the index of the global eigenvalue is equal to 5, which corresponds to the correct model order.

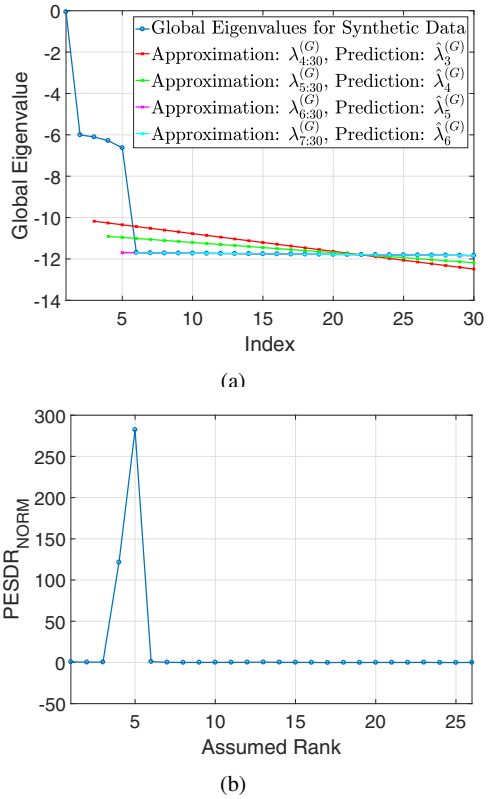


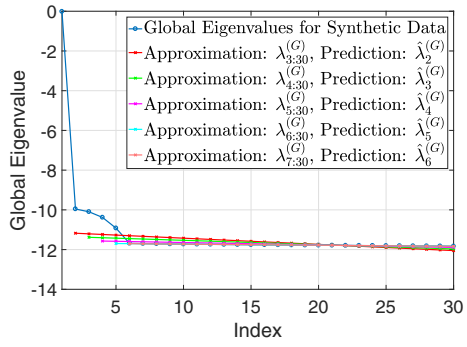
Fig. 1: Linear regression (a) and normalized PESDR (b) for uncorrelated case with SNR = 10 dB

Similarly, in Fig. 2 we show the linear regression of global eigenvalues and the computed nPESDR for the correlated case. Again, starting from the smallest global eigenvalues we can see that the first value of the nPESDR that exceeds the threshold $\rho = 1$ corresponds to the correct rank 5.

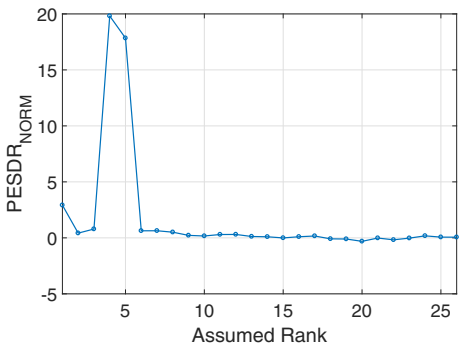
As expected, the assumption that the noise global eigenvalues have an approximate exponential distribution has been confirmed by the results of the simulations. As shown in Fig. 1, the linear regression provides a good fit for the noise global eigenvalues. Although the approximation error of the noise global eigenvalues is bigger in the correlated case (Fig. 2), the gap between them and the signal global eigenvalues is clearly seen.

V. Experimental results. Biomedical signal processing

The proposed LaRGE method presented in this paper is used to analyze measured EEG and MEG data. These data were recorded from twelve healthy volunteers at the Biomagnetic Center of the University Hospital in Jena, Germany, [17] for the investigation of the Photic Driving (PD) effect. This effect occurs when the brain is stimulated by Intermittent Photic Stimulation (IPS). The PD effect is widely used to assess

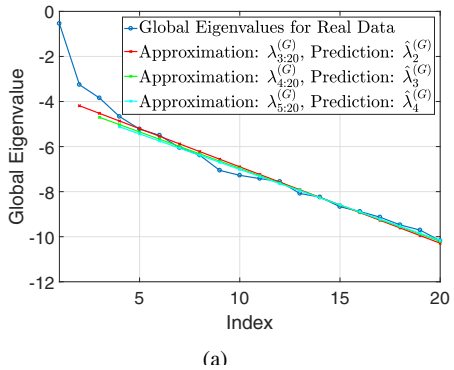


(a)

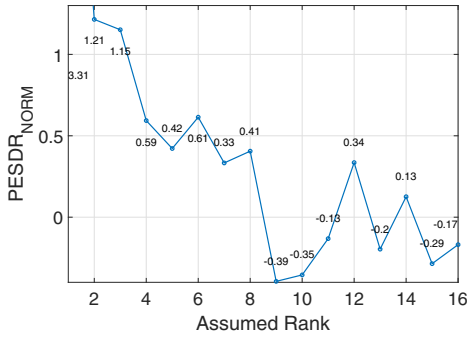


(b)

Fig. 2: Linear regression (a) and normalized PESDR (b) for correlated case with SNR = 10 dB



(a)



(b)

Fig. 3: Linear regression (a) and normalized PESDR (b) for EEG data, stimulation frequency $0.45 f_\alpha$, occipital zone, estimated model order $\hat{R} = 3$

the effects of medications and for the diagnosis of several neurophysiological diseases like Alzheimers, schizophrenia, and some forms of epilepsy.

During the mentioned experiment, the data have been recorded simultaneously from EEG and MEG channels. In the first step the individual alpha rhythm f_α is determined for each volunteer. To investigate the PD effect, twenty stimulation frequencies from $0.40 f_\alpha$ to $2.30 f_\alpha$ are used. Each stimulation of the particular frequency is presented to the volunteer 30 times during each train with a resting period of 4 seconds.

The recorded signals from 128 EEG and 102 MEG channels are filtered and the Fourier Transform is computed. The obtained frequency distributions for each channel and train are stacked into 3D 'ChFTr' tensors with dimensions channel × frequency × train. For further processing only reliable channels are used. Reliable channels are the channels that do not contain artifacts.

Moreover, to increase the accuracy of the processing, all channels can be divided into groups. The PD effect has highest amplitudes in the occipital zone. Therefore, in our investigation we divide the EEG and MEG channels into two groups that correspond to occipital and non-occipital zones of the head. The recorded signals from EEG and MEG sensors have different values (micro Volt and femto Tesla). Therefore, before the model order estimation and the CP decomposition, the tensors are normalized to unit norm according to

$$\boldsymbol{\mathcal{X}}_{\text{NORM,EEG}} = \frac{\boldsymbol{\mathcal{X}}_{\text{EEG}}}{\|\boldsymbol{\mathcal{X}}_{\text{EEG}}\|_{\text{H}}} \text{ and } \boldsymbol{\mathcal{X}}_{\text{NORM,MEG}} = \frac{\boldsymbol{\mathcal{X}}_{\text{MEG}}}{\|\boldsymbol{\mathcal{X}}_{\text{MEG}}\|_{\text{H}}}. \quad (17)$$

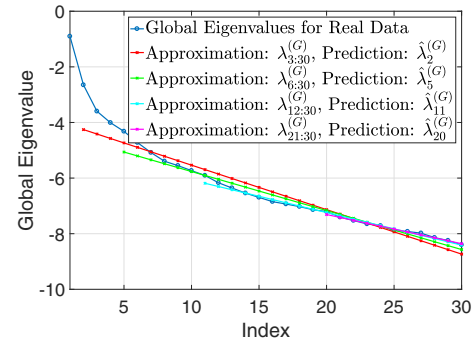
The LaRGE method is used for the model order estimation of EEG and MEG data that was recorded from volunteer 1. For the cases that are depicted in the Figs. 3- 5, the model order can be estimated as follows: EEG data, occipital zone the rank is three (Fig. 3), non occipital zone the rank is eleven (Fig. 4); MEG data, occipital zone the rank is one (Fig. 5).

Moreover, the proposed LaRGE method provides reliable model order estimates if the size of the multi-dimensional data is limited as, for example, in Fig. 5.

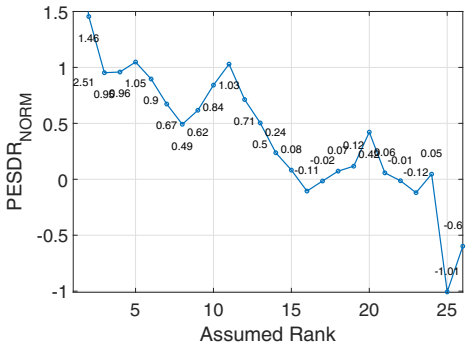
After the model order estimation, we compute the approximate low-rank CP decomposition via SECSI [16]. The resulting dominant components in all three dimensions confirm that the estimated model order is reasonable.

VI. Conclusion

In this paper we have presented the novel LaRGE method of model order estimation based on the global eigenvalues. Using the HOSVD of a measurement tensor, the global eigenvalues can be computed. Starting from the smallest global eigenvalues, the normalized PESDR is calculated using linear regression. If it is significantly larger than a predefined threshold ρ (we use $\rho = 1$ in this paper) we have found the smallest signal global eigenvalue, i.e., the approximate model

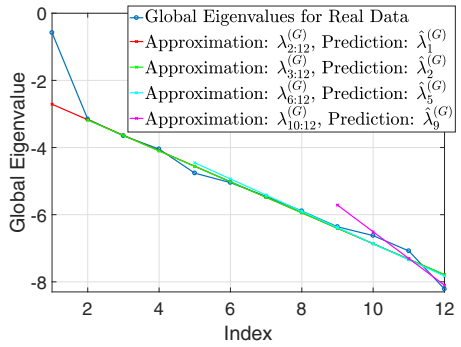


(a)

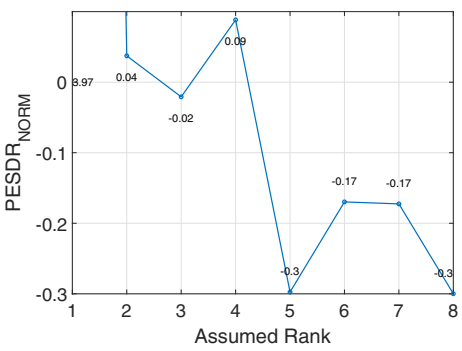


(b)

Fig. 4: Linear regression (a) and normalized PESDR (b) for EEG data, stimulation frequency $0.5f_\alpha$, not occipital zone, estimated model order $\hat{R}$ = 11



(a)



(b)

Fig. 5: Linear regression (a) and normalized PESDR (b) for MEG data, stimulation frequency $0.95f_\alpha$, occipital zone, estimated model order $\hat{R}$ = 1

order $\hat{R}$. To confirm its excellent performance, we have tested the proposed LaRGE method on synthetic data and measured EEG as well as MEG data.

In contrast to the EFT that is based on an estimate of the probability of false alarm and needs parameter tuning [11], [12] the proposed LaRGE method has less computational complexity. By replacing the global eigenvalues by the eigenvalues, a simplified version of LaRGE can also be used in the 2-D matrix case.

REFERENCES

[1] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, March 2015.

[2] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, September 2009.

[3] 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, pp. 1 – 84, 1970.

[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] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Analysis Applications*, vol. 21, pp. 1253–1278, 2000.

[6] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, p. 716–723, December 1974.

[7] J. Rissanen, "Modeling by shortest data description," *Automatic*, vol. 14, no. 5, p. 465–471, 1978.

[8] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, p. 461–464, 1978.

[9] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, April 1985.

[10] A. Quinlan, J.-P. Barbot, P. Larzabal, and M. Haardt, "Model order selection for short data: An exponential fitting test (EFT)," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, pp. 1–11, 2006.

[11] J. P. C. L. da Costa, M. Haardt, F. Roemer, and G. Del Galdo, "Enhanced model order estimation using higher-order arrays," in *Proc. 2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*, Nov 2007, pp. 412–416.

[12] J. P. C. L. da Costa, F. Roemer, M. Haardt, and R. T. de Sousa Jr, "Multi-dimensional model order selection," *EURASIP Journal on Advances in Signal Processing*, 2011.

[13] K. Liu, J. P. C. L. da Costa, H.-C. So, L. Huang, and J. Ye, "Detection of number of components in CANDECOMP/PARAFAC models via minimum description length," *Digit. Signal Process.*, vol. 51, pp. 110–123, 2016.

[14] T. Yokota, N. Lee, and A. Cichocki, "Robust multilinear tensor rank estimation using higher order singular value decomposition and information criteria," *IEEE Transactions on Signal Processing*, vol. 65, no. 5, pp. 1196–1206, March 2017.

[15] S. Pouryazdian, S. Beheshti, and S. Krishnan, "CANDECOMP/PARAFAC model order selection based on reconstruction error in the presence of Kronecker structured colored noise," *Digit. Signal Process.*, vol. 48, pp. 12–26, 2016.

[16] F. Roemer and M. Haardt, "A semi-algebraic framework for approximate cp decompositions via simultaneous matrix diagonalizations (SECSI)," *Signal Processing*, vol. 93, no. 9, pp. 2722 – 2738, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168413000704

[17] K. Schwab, C. Ligges, T. Jungmann, B. Hilgenfeld, J. Haueisen, and H. Witte, "Alpha entrainment in human electroencephalogram and magnetoencephalogram recordings," *Neuroreport*, vol. 17, pp. 1829 – 1833, 2006.