A NOVEL DETERMINISTIC METHOD FOR LARGE-SCALE BLIND SOURCE SEPARATION

Martijn Bousse[*] *Otto Debals*[*†] *Lieven De Lathauwer*[*†]

[*]Department of Electrical Engineering (ESAT), KU Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium. [†]Group Science, Engineering and Technology, KU Leuven Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium.

ABSTRACT

A novel deterministic method for blind source separation is presented. In contrast to common methods such as independent component analysis, only mild assumptions are imposed on the sources. On the contrary, the method exploits a hypothesized (approximate) intrinsic low-rank structure of the mixing vectors. This is a very natural assumption for problems with many sensors. As such, the blind source separation problem can be reformulated as the computation of a tensor decomposition by applying a low-rank approximation to the tensorized mixing vectors. This allows the introduction of blind source separation in certain big data applications, where other methods fall short.

Index Terms— blind source separation, big data, higher-order tensor, tensor decomposition, low-rank approximation

1. INTRODUCTION

The goal of blind source separation (BSS) is to reconstruct a collection of unobserved source signals based only on a collection of observed sensor signals. The latter are unknown linear instantaneous mixtures of the unknown sources. Applications can be found in telecommunications, signal processing and biomedical sciences [1–3].

There is no *a priori* information regarding the sources and the mixture. This "lack" of information is often solved by imposing assumptions on the sources. A common BSS method, called independent component analysis (ICA), assumes statistically independent sources [4]. Recently, a deterministic method based on block term decompositions, called block component analysis (BCA), was introduced in [5]. Elaborating on this idea, new methods were proposed that assume the sources can be written as exponential polynomials or rational functions, which allows broad deterministic modeling [6, 7].

Previous methods are not applicable to problems with a massive amount of sensors because they do not scale well. This is especially true for ICA, of which several variants use higher-order statistics (HOS) that suffer from the curse of dimensionality [8, 9]. However, the mixture has possibly some smooth structure when one considers many and closely located sensors. It becomes ever more relevant to exploit this structure for an increasing number of sensors. Hence, we present a new method that can cope with large-scale problems by using this underlying compactness. More specifically, we approximate the tensorized mixing vectors with higher-order low-rank terms. A comparable strategy has proven to be very successful in the field of tensor-based scientific computing [10]. As such, we will demonstrate that the BSS problem boils down to the computation of a tensor decomposition. This method uniquely determines both the sources and the mixing vectors of large-scale problems under mild conditions. Moreover, it imposes only mild assumptions on the sources and because of its deterministic nature it even works well if only a few samples are present.

1.1. Notation and basic definitions

Tensors, denoted by calligraphic characters (e.g., $\mathcal{A}$), are higher-order generalizations of vectors and matrices, denoted by bold lowercase (e.g., $\mathbf{a}$) and bold uppercase (e.g., $\mathbf{A}$) characters, respectively. The $(i_1, i_2, \ldots, i_N)$th entry of an Nth-order tensor $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ is denoted by $a_{i_1 i_2 \ldots i_N}$ with $\mathbb{K}$ meaning $\mathbb{R}$ or $\mathbb{C}$. The nth element in a sequence is indicated by a superscript between parentheses (e.g., $\{\mathbf{a}^{(n)}\}_{n=1}^{N}$).

A mode-n vector of a tensor is defined by fixing every index except the nth and is a natural extension of the rows and columns of a matrix. The mode-n unfolding of a tensor $\mathcal{A}$ is a matrix $\mathbf{A}_{(n)}$ with the mode-n vectors of $\mathcal{A}$ as its columns. The vectorization, denoted as $\text{vec}(\mathcal{A})$, of a tensor $\mathcal{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N}$ maps each element $a_{i_1, i_2 \cdots i_N}$ onto $\text{vec}(\mathcal{A})_j$ with $j = 1 + \sum_{k=1}^{N} (i_k - 1) J_k$ and $J_k = \prod_{m=1}^{k-1} I_m$. The

This research is funded by (1) a Ph.D. grant of the Agency for Innovation by Science and Technology (IWT), (2) Research Council KU Leuven: GOA/10/09 MaNet, CoE PFV/10/002 (OPTEC), (3) FWO: project: G.0830.14N, G.0881.14N, (4) iMinds Medical Information Technologies SBO 2015, (5) the Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, Dynamical systems, control and optimization, 2012-2017), (6) EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (no. 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

outer product $\mathbf{a} \in \mathbb{K}^I$ and $\mathbf{b} \in \mathbb{K}^J$ is defined element-wise as $(\mathbf{a} \circ \mathbf{b})_{ij} = a_i b_j$. The Kronecker product is defined as $\mathbf{a} \otimes \mathbf{b} = \begin{bmatrix} a_1\mathbf{b} & a_2\mathbf{b} & \cdots & a_I\mathbf{b} \end{bmatrix}^T$.

1.2. Multilinear algebraic prerequisites

An Nth-order tensor has rank 1 if it can be written as the outer product of N non-zero vectors. The rank of a tensor is defined as the minimal number of rank-1 terms that generate the tensor as their sum. The multilinear rank of an Nth-order tensor is equal to the tuple of mode-n ranks, which are defined as the ranks of the mode-n unfoldings of the tensor.

Definition 1. A *polyadic decomposition* (PD) writes a tensor $\mathcal{A} \in \mathbb{K}^{I_1 \times I_2 \times \cdots \times I_N}$ as a sum of R rank-1 terms:

$$\mathcal{A} = \sum_{r=1}^{R} \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \cdots \circ \mathbf{u}_r^{(N)}. \tag{1}$$

The columns of the factor matrices $\mathbf{U}^{(n)} \in \mathbb{K}^{I_n \times R}$ are equal to the factor vectors $\mathbf{u}_r^{(n)}$, $r = 1, \ldots, R$. The PD is called *canonical* (CPD) when R is equal to the rank of $\mathcal{A}$.

The CPD is a powerful model for several applications within signal processing, biomedical applications, computer vision, machine learning and data mining [11, 12]. The decomposition is *essentially unique* (i.e., up to trivial permutations and scalings) under rather mild conditions [13–15].

A decomposition that writes a tensor as a sum of low multilinear rank terms and associated uniqueness results were introduced in [6, 16]. The block terms in this decomposition are more general than the simple rank-1 terms in the CPD. Hence, they allow the modeling of more complex phenomena, see e.g., [5, 17]. We only use the following specific case.

Definition 2. A *block term decomposition* (BTD) *in rank-*$(L, L, 1)$ *terms* writes a third-order tensor $\mathcal{X} \in \mathbb{K}^{I \times J \times K}$ as a sum of R multilinear rank-$(L, L, 1)$ terms:

$$\mathcal{X} = \sum_{r=1}^{R} (\mathbf{A}_r \mathbf{B}_r^T) \circ \mathbf{c}_r, \tag{2}$$

in which $\mathbf{A}_r \in \mathbb{K}^{I \times L}$ and $\mathbf{B}_r \in \mathbb{K}^{J \times L}$ are rank-L.

2. BLIND SOURCE SEPARATION BASED ON LOW-RANK TENSOR APPROXIMATIONS

Blind source separation (BSS) uses the following model [4]:

$$\mathbf{X} = \mathbf{M}\mathbf{S} + \mathbf{N}, \tag{3}$$

with $\mathbf{X} \in \mathbb{K}^{M \times K}$ and $\mathbf{S} \in \mathbb{K}^{R \times K}$ containing K samples of the M sensor signals and R source signals, respectively; $\mathbf{M} \in \mathbb{K}^{M \times R}$ is the mixing matrix and $\mathbf{N} \in \mathbb{K}^{M \times K}$ is the additive noise. The goal of BSS is to retrieve the unknown mixing vectors in $\mathbf{M}$ and/or the unknown sources in $\mathbf{S}$, given only the observed data $\mathbf{X}$. In the derivation of our method we ignore the noise $\mathbf{N}$ for notational simplicity; the influence will be investigated in Subsection 3.1 by means of simulations.

2.1. Kronecker product structure

The fact that many real-life signals admit a low-rank representation is at the heart of compressed sensing [18]; this notion is exploited in a novel way in this work. Particularly, the mixing vectors in (3) possibly have some intrinsic low-rank structure in real-life applications. If the number of sensors M increases, it becomes very interesting to exploit this structure in order to cope with large-scale problems.Let us illustrate how this can be achieved with Kronecker products. Assume we approximate a vector $\mathbf{m} \in \mathbb{K}^M$ with a Kronecker product of two smaller vectors, i.e., $\mathbf{m} = \mathbf{b} \otimes \mathbf{a}$, with $\mathbf{a} \in \mathbb{K}^I$, $\mathbf{b} \in \mathbb{K}^J$ and $M = IJ$. This is actually a rank-1 approximation; indeed, the Kronecker product is related to the outer product through a vectorization: $\mathbf{b} \otimes \mathbf{a} = \text{vec}(\mathbf{a} \circ \mathbf{b})$, i.e., a (second-order) rank-1 term. Note that the number of coefficients decreases from $M = IJ$ to $I + J$. Suppose $I = J$, then the number of coefficients decreases one order of magnitude.

More generally, larger reductions can be obtained by considering a Kronecker product of N vectors, which is the vectorization of an outer product of N vectors, i.e., $\bigotimes_{n=1}^{N} \mathbf{u}^{(n)} = \text{vec}\left(\circ_{n=1}^{N} \mathbf{u}^{(N-n+1)}\right)$ with $\mathbf{u}^{(n)} \in \mathbb{K}^{I_n}$, $n = 1, \ldots, N$. On the other hand, a rank-1 approximation might not be sufficient; hence, we take a sum of P such Kronecker products resulting in a low-rank approximation. Consequently, we have:

$$\mathbf{m} = \sum_{p=1}^{P} \bigotimes_{n=1}^{N} \mathbf{u}_p^{(n)} = \text{vec}\left(\sum_{p=1}^{P} \circ_{n=1}^{N} \mathbf{u}_p^{(N-n+1)} \right), \tag{4}$$

in which $\mathbf{u}_p^{(n)} \in \mathbb{K}^{I_n}$ and $M = \prod_{n=1}^{N} I_n$. In other words, approximating a vector with a sum of Kronecker products amounts to a low-rank approximation of the 'folded' vector, i.e., a low-rank tensor approximation. Note that the number of coefficients decreases from $M = \prod_{n=1}^{N} I_n$ to $P \sum_{n=1}^{N} I_n$. If $I_n = I$ for $n = 1, \ldots, N$, then $M = I^N$ reduces to PNI, i.e., a reduction of $N - 1$ orders of magnitude; thus, a decrease in function of the number of Kronecker products. On the other hand, the number of coefficients grows proportionally with the number of rank-1 terms P in the sum.

2.2. Tensorization

We now demonstrate that, assuming Kronecker product structured mixing vectors, the BSS problem can be reformulated as the computation of a CPD of the tensorized observations. Let us illustrate this as follows: assume the mixing vectors have a simple Kronecker product structure, e.g., $\mathbf{m}_r = \mathbf{b}_r \otimes \mathbf{a}_r$ with $\mathbf{a}_r \in \mathbb{K}^I$, $\mathbf{b}_r \in \mathbb{K}^J$, and assume that $M = IJ$, then:

$$\mathbf{X} = \sum_{r=1}^{R} (\mathbf{b}_r \otimes \mathbf{a}_r) \circ \mathbf{s}_r. \tag{5}$$

Equation (5) can be tensorized by stacking each matricized column of the data matrix $\mathbf{X}$ in a third-order tensor $\mathcal{X} \in$

$\mathbb{K}^{I\times J\times K}$. Remember that a Kronecker product is a vectorized outer product; consequently, we have:

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{s}_r. \tag{6}$$

Equation (6) is a CPD as defined in (1). Hence, the above described tensorization strategy reformulates the BSS problem as the computation of a CPD of a third-order tensor of rank R. The number of coefficients reduces from $RM = RIJ$ to $R(I+J)$ (cf. above). The idea can be generalized to a sum of P Kronecker products of N vectors analogous as in (4):

$$\mathbf{X} = \sum_{r=1}^{R} \left(\sum_{p=1}^{P} \bigotimes_{n=1}^{N} \mathbf{u}_{pr}^{(n)} \right) \circ \mathbf{s}_r, \tag{7}$$

in which $\mathbf{u}_{pr}^{(n)} \in \mathbb{K}^{I_n}$, $n = 1, \ldots, N$ and $M = \prod_{n=1}^{N} I_n$, where we have chosen a fixed P for each source r. Applying the same tensorization as before amounts to:

$$\mathcal{X} = \sum_{r=1}^{R} \left(\sum_{p=1}^{P} \circ_{n=1}^{N} \mathbf{u}_{pr}^{(N-n+1)} \right) \circ \mathbf{s}_r. \tag{8}$$

Hence, the BSS problem boils down to the decomposition of an $(N+1)$th-order tensor in PR rank-1 terms. The number of coefficients decreases from RM to $R\sum_{n=1}^{N} I_n$ (cf. above). Equation (8) is a BTD in rank-$(P, P, 1)$ terms if $N = 2$.

The proposed method simultaneously and uniquely determines the mixing vectors *and* the sources under the mild uniqueness conditions of the CPD or BTD. It is 1) applicable for large M, as is clear from the possible large reductions in the number of coefficients, and 2) is deterministic, meaning that it even works well if the number of samples K is small, which is a clear advantage compared to stochastic methods such as ICA. Moreover, 3) the method imposes only mild conditions on the sources in contrast to existing methods (e.g., linear independence instead of statistical independence).

3. RESULTS AND DISCUSSION

3.1. Simulations

We present two experiments in which we investigate 1) the influence of the noise level for several K with exact second-order rank-2 mixing vectors ($N = 2$ and $P = 2$) and 2) the influence of deviations from a second-order rank-1 structure ($N = 2$ and $P = 1$) for several SNRs. Both experiments have $M = 2500$ sensors, and $R = 2$ i.i.d. zero-mean unit-variance Gaussian random source signals of length K. Note that ICA techniques based on (full) HOS would be infeasible here as the number of entries in a Qth-order statistic equals $\mathcal{O}(M^Q)$. Moreover, ICA is not applicable here because of the specific sources used in this experimental setup [4]. We use a second-order rank-2 and rank-1 approximation with $I = J = 50$ for

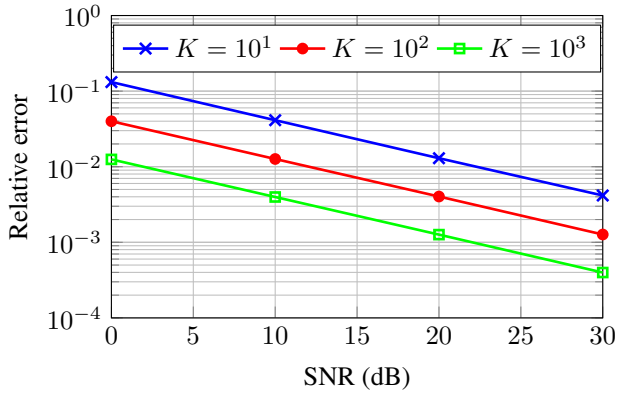

Fig. 1. Median across 100 experiments of the relative error on the mixing vectors for several K. The mixing vectors exactly satisfy the assumed second-order rank-1 structure.

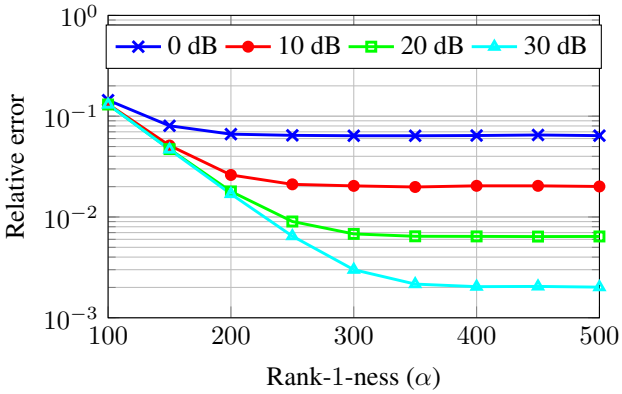

Fig. 2. Median across 100 experiments of the relative error on the mixing vectors for several SNRs. The rank-1-ness of the mixing vectors is controlled with α.

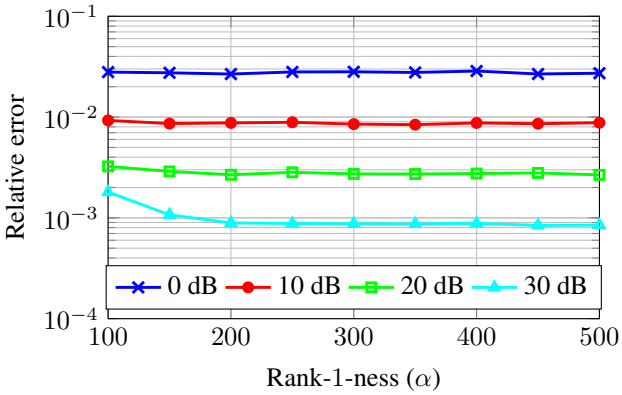

Fig. 3. Median across 100 experiments of the relative error on the sources for several SNRs. The rank-1-ness of the mixing vectors is controlled with α.

the first and second experiment respectively. Note that only 4% of the coefficients per rank-1 term have to be estimated, which is the maximal reduction for the given order. The mixing vectors and sources can only be found up to scaling and permutation, which are the standard indeterminacies in BSS. Hence, for each case we report the median across 100 experiments of the relative error, defined as the relative difference in Frobenius norm after optimal scaling and permutation, on the mixing vectors and/or the sources. The computations are carried out in the MATLAB package Tensorlab [19].

The results of the first experiment are computed with `btd_nls` of Tensorlab and shown in Figure 1: the relative error on the mixing vectors decreases as the SNR increases. Although the method is deterministic, and therefore requires only a few samples, it is beneficial to increase K under noisy conditions in order to obtain better accuracy. However, the number of samples can still be very low in contrast to ICA.

In the second experiment each mixing vector is defined as the vectorization of a random matrix with exponentially decaying singular values: $\mathbf{m}_r = \text{vec}(\mathbf{U}_r \text{diag}(\boldsymbol{\sigma})\mathbf{V}_r^T)$, in which $\boldsymbol{\sigma} = e^{-\alpha \mathbf{t}}$ with $\mathbf{t}$ a vector containing $\min(I, J)$ uniformly discretized samples on the interval $[0, 1]$, and $\mathbf{U}_r$ and $\mathbf{V}_r$ are random orthogonal matrices of compatible dimensions. The parameter α controls the exponential decay of the singular values; hence, increasing α results in more rank-1-like mixing vectors and vice versa. We take $K = 20$ and compute the results with `cpd` of Tensorlab. The relative error on the mixing vectors is shown for several SNRs in Figure 2: increasing α decreases the relative error; this is especially clear for high SNR. The decrease in relative error eventually stagnates because of the limited accuracy due to noise. Figure 3 shows the relative error on the source signals: the effect of α is less strong, probably because the rank-1 approximation already captures a lot of the information quite well. Also, the noise on the sources is possibly (largely) compensated because this is the shorter factor ($K < I, J$) in the decomposition.

3.2. Applications

Recent advances in biomedical sciences have led to applications with many sensors and/or high sensor density, e.g., wireless body area networks (WBANs). These are sensor systems, located in and/or around the body, that measure several different biomedical signals such as temperature, heartbeat and neuronal activity. They are currently investigated for electroencephalography (EEG) [20] and electrocorticography (ECoG) [21] with high spatial resolution as a possible means for long-term neuromonitoring. In such cases our method can provide an efficient separation of the different sources of neuronal activity. Moreover, the data rate of the wireless connection can be reduced because of the deterministic nature of the method. The problem becomes really large-scale when we consider neural dust; here, thousands of miniature sensors, called neural probes, are dispersed throughout the brain [22].

Our method can also be applied in cases where the mixing vectors have an exact low-rank structure. An example of this are uniform linear arrays (ULAs) with far-field sources that emit narrowband signals. Here, the mixing matrix $\mathbf{M}$ has Vandermonde structure because of the fact that (short) time delays amount to phase shifts of the signal [23]. Then, each mixing vector can be written as a Kronecker product of some smaller vectors, e.g., $[1,\ a,\ a^2,\ a^3]^T = \big([1,\ a^2] \otimes [1,\ a]\big)^T$, i.e., as a rank-1 term. If the signals propagate through P distinct paths [24], e.g., due to reflections or scattering caused by objects on the path, each mixing vector is low-rank. Note that this Vandermonde structure arises often in signal processing, see [25] and references therein. Assuming a rank-1 structure in such cases is relevant and provides an efficient and deterministic processing of the data. Finally, our method can also be applied when the sources are located in the near field. In this case the Vandermonde structure is only approximate and a low-rank approximation can be applied.

3.3. Future research

In this paper we have limited ourselves to sensor data that are linear instantaneous mixtures of the sources. However, the strategy outlined above can be extended to blind system identification (BSI), also known as equalization, in which the outputs are linear convolutive mixtures of the inputs [26]. More specifically, the same strategy can be applied to the identification of finite impulse response (FIR) systems of the form: $\mathbf{x}[k] = \sum_{r=1}^{R} \sum_{l=0}^{L-1} \mathbf{m}_r^{(l)} s_r[k-l]$, with memory L; note that this reduces to BSS if $L = 1$. This problem will be addressed in a follow-up paper with more simulations and applications.

4. CONCLUSION

We have presented a new deterministic method for blind source separation (BSS) of large-scale problems under mild uniqueness conditions. The method can cope with many sensors by exploiting a hypothesized low-rank structure of the mixture. This is done by approximating the mixing vectors by sums of Kronecker products or, in other words, approximating the tensorized mixing vectors by sums of higher-order rank-1 terms. This strategy allowed us to reformulate the BSS problem as the computation of a tensor decomposition. Moreover, the method is deterministic, meaning that it does not need a large number of samples. Finally, in contrast to existing methods, the method imposes only mild assumptions on the sources. The generalization of this strategy to BSI, including FIR systems, will be addressed in a future paper.

REFERENCES

[1] A. Cichocki and S. Amari, *Adaptive blind signal and image processing: Learning algorithms and applications*, vol. 1, John Wiley Chichester, 2002.

[2] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, no. 4–5, pp. 411–430, June 2000.

[3] C.J. James and C.W. Hesse, "Independent component analysis for biomedical signals," *Physiological Measurement*, vol. 26, no. 1, pp. R15–R39, 2005.

[4] P. Comon and C. Jutten, *Handbook of blind source separation, independent component analysis and applications*, Academic press, 2009.

[5] L. De Lathauwer, "Block component analysis, a new concept for blind source separation," in *Latent Variable Analysis and Signal Separation*, vol. 7191 of *Lecture Notes in Computer Science*, pp. 1–8. Springer Berlin/Heidelberg, 2012.

[6] L. De Lathauwer, "Blind separation of exponential polynomials and the decomposition of a tensor in rank-$(l_r, l_r, 1)$ terms," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 4, pp. 1451–1474, Dec. 2011.

[7] O. Debals, M. Van Barel, and L. De Lathauwer, "Blind signal separation of rational functions using löwner-based tensorization," *Technical Report 14–167, ESAT-STADIUS, KU Leuven, Leuven, Belgium*, 2014.

[8] D.L. Donoho, "Aide-memoire. high-dimensional data analysis: The curses and blessings of dimensionality," in *AMS Conference on Math Challenges of the 21st Century (Los Angeles, USA*, Aug. 2000.

[9] N. Vervliet, O. Debals, L. Sorber, and L. De Lathauwer, "Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 71–79, Sept. 2014.

[10] L. Grasedyck, D. Kressner, and C. Tobler, "A literature survey of low-rank tensor approximation techniques," *GAMM-Mitteilungen*, vol. 36, no. 1, pp. 53–78, Feb. 2013.

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

[12] A. Cichocki, C. Mandic, A. Phan, C. Caifa, G. Zhou, Q. Zhao, and L. De Lathauwer, "Tensor decompositions for signal processing applications. from two-way to multiway component analysis," *IEEE Signal Processing Magazine*, Mar. 2015, (to appear).

[13] J.B. Kruskal, "Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra and its Applications*, vol. 18, no. 2, pp. 95–138, 1977.

[14] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — part i: Basic results and uniqueness of one factor matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 855–875, July-Sept. 2013.

[15] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — part ii: Uniqueness of the overal decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 876–903, July-Sept. 2013.

[16] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms — part II: Definitions and uniqueness," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1033–1066, Sept. 2008.

[17] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms — part III: Alternating least squares algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1067–1083, Sept. 2008.

[18] I. Oseledets and E. Tyrtyshnikov, "TT-cross approximation for multidimensional arrays," *Linear Algebra and its Applications*, vol. 432, no. 1, pp. 70–88, Jan. 2010.

[19] L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab v2.0," Jan. 2014, Available online at http://www.tensorlab.net/.

[20] A. Bertrand, "Distributed signal processing for wireless eeg sensor networks," *Technical Report 14-125, ESAT-STADIUS, KU Leuven, Leuven, Belgium*, 2014.

[21] B. Rubehn, C. Bosman, R. Oostenveld, P. Fries, and T. Stiegliz, "A mems-based flexible multichannel ecog-electrode array," *Journal of Neural Engineering*, vol. 6, no. 3, pp. 1–10, 2009.

[22] D. Seo, J.M. Carmena, J.M. Rabaey, E. Alon, and M.M. Maharbiz, "Neural dust: An ultrasonic, low power solution for chronic brain-machine interfaces," *arXiv preprint arXiv:1307.2196*, Jul. 2013.

[23] H. Krim and M. Viberg, "Two decades of array signal processing: The parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[24] W.C. Jakes, *Microwave mobile communications*, New York: Wiley, 1974.

[25] M. Sørensen and L. De Lathauwer, "Blind signal separation via tensor decomposition with vandermonde factor: Canonical polyadic decomposition," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5507–5519, Nov. 2013.

[26] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1310–1322, Aug. 1997.