BLIND SYSTEM IDENTIFICATION AS A COMPRESSED SENSING PROBLEM

Frederik Van Eeghem[⋆†‡] *Lieven De Lathauwer*[⋆†‡]

[⋆]Group Science, Engineering and Technology, KU Leuven Kulak, E. Sabbelaan 53,
B-8500 Kortrijk, Belgium. [†]Department of Electrical Engineering (ESAT),
KU Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium.
[‡]iMinds Medical IT, KU Leuven. Kasteelpark Arenberg 10, 3001 Leuven, Belgium

ABSTRACT

If the inputs of a convolutive system are hard or expensive to measure, one can resort to blind system identification (BSI). BSI tries to identify a system using only output information. By computing a structured canonical polyadic decomposition of a cumulant tensor, the sought system coefficients can be extracted. However, these structured decompositions may be cumbersome to model. Here, we propose an alternative approach based on compressed sensing which captures the convolution structure in a known matrix. This allows us to obtain the system coefficients from an unstructured CPD of a smaller tensor that is implicitly available.

Index Terms— independent component analysis, canonical polyadic decomposition, tensor, compressed sensing, blind system identification

1. INTRODUCTION

As opposed to classical identification techniques, blind system identification (BSI) tries to estimate a system using output measurements only. This is especially useful when the inputs are hard or even impossible to measure, as is often the case in domains such as telecommunications, acoustics and biomedical data processing [1–3]. Blindly identifying a system is infeasible without making some assumptions, either on the input signals or on the system itself. Though there are many assumptions possible, we will assume independent inputs, which has proven to be a useful approximation in various applications [2]. The term BSI will thus be used here to denote blind identification of systems with independent inputs.

Frederik Van Eeghem is supported by an Aspirant Grant from the Research Foundation – Flanders (FWO). This research is funded by (1) Research Council KU Leuven: C1 project c16/15/059-nD and CoE PFV/10/002 (OPTEC), (2) F.W.O.: project G.0830.14N and G.0881.14N, (3) the Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO II, Dynamical systems, control and optimization, 2012-2017), (4) EU: The research leading to these results has received funding from the European Research Council under the European Unions 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.

The common independent inputs assumption bridges the gap between BSI and independent component analysis (ICA). In its most basic form, ICA attempts to retrieve the different statistically independent components from an instantaneous mixture. The convolutive extension of ICA is usually modeled as a linear finite impulse response (FIR) system with independent inputs. In this way ICA and BSI are linked, even though there is a slight difference in terminology as ICA focuses on retrieving the input signals whereas BSI concentrates on estimating the system coefficients. For overdetermined systems, both approaches are interchangeable. However, if the system is underdetermined, an estimate of the mixing coefficients may not suffice to uniquely extract the input signals.

Tensor-based methods for instantaneous ICA usually involve second- or higher-order statistics and have already been studied well [2, 4, 5]. For convolutive mixtures however, relatively few algebraic results are available [6–9]. The existing tensor-based approaches can be classified as either time or frequency domain methods. In time domain methods higher-order statistics are computed, which will contain a particular structure due to the convolutive nature of the mixture. However, this structure often remains unexploited or is cumbersome to model. Frequency domain methods try to deal with this structure by transforming the convolutive problem to a set of instantaneous mixtures in different frequency bands [3]. This simplifies the further analysis as conventional techniques can be used for each instantaneous mixture. However, this strategy suffers from several disadvantages. For instance, the different solutions to the instantaneous problems may be differently permuted, which has been coined the permutation ambiguity problem. Several attempts to deal with this have been proposed, for instance in [8, 10]. A more detailed discussion of the advantages and disadvantages of time domain and frequency domain methods can be found in [3].

In this paper, a time domain method will be presented, in which the structure will be exploited using a compressed sensing (CS) approach. Compressed sensing, also called compressive sampling, is able to reconstruct signals with far fewer measurements than traditional methods use [11]. To do this, the sought signal has to be compressible, which will be satisfied due to the low-rank canonical polyadic decomposition (CPD) structure present in our case.

After introducing the notation, the CPD will be discussed. Next, the actual problem statement is presented, after which the developed method is introduced. Finally, some experiments are conducted.

Notations Scalars are denoted by lowercase letters (e.g., a), vectors by bold lowercase letters (e.g., $\mathbf{a}$), matrices by bold uppercase letters (e.g., $\mathbf{A}$), and tensors by uppercase calligraphic letters (e.g., $\mathcal{A}$). The outer product is denoted by $\otimes$. The Kronecker product is denoted by $\otimes$. The complex conjugate is given by a bar atop, e.g., $\bar{a}$ and the Moore-Penrose pseudoinverse is given by $\cdot^{\dagger}$. Estimates are written with a hat, e.g., $\widehat{\mathbf{P}}$. The mathematical expectation is denoted by $\mathbb{E}\{\cdot\}$ and the Frobenius norm is given by $||\cdot||_{\mathrm{F}}$.

2. CANONICAL POLYADIC DECOMPOSITION

The polyadic decomposition (PD) writes a tensor as a linear combination of rank-1 terms:

$$\begin{aligned}\mathcal{A} &= \sum_{r=1}^{R} \lambda_r \mathbf{u}_r^{(1)} \otimes \mathbf{u}_r^{(2)} \otimes \cdots \otimes \mathbf{u}_r^{(N)} \qquad (1)\\ &= \left[\!\left[\boldsymbol{\lambda}; \mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(N)} \right]\!\right],\end{aligned}$$

in which the factor matrix $\mathbf{U}^{(n)}$ has R columns $\mathbf{u}_r^{(n)}$ and $\boldsymbol{\lambda}$ is a vector containing the scaling coefficients λ_r. We say that an Nth order tensor has rank 1 if and only if it equals the outer product of N nonzero vectors. By extension, the rank of a tensor $\mathcal{A}$ is defined as the minimal number of rank-1 tensors yielding $\mathcal{A}$ in a linear combination. If R in equation (1) is minimal, and $\mathcal{A}$ thus has rank R, we call the decomposition canonical. Other terms sometimes used for this decomposition are CANDECOMP and PARAFAC [12, 13]. Contrary to a decomposition of a matrix in rank-1 terms, the CPD of higher-order tensors is essentially unique under fairly mild conditions [14–16].

3. PROBLEM STATEMENT

Consider a discrete linear time-invariant system with R inputs, M outputs, and with system coefficients $h_{mr}[l]$ with $l \in \{0, \ldots, L\}$ for the filter from input r to output m. The mth output signal of this system can now be written as

$$x_m[n] = \sum_{r=1}^{R} \sum_{l=0}^{L} h_{mr}[l] s_r[n-l] + v_m[n],$$

in which $s_r[n]$ represents the rth input and $v_m[n]$ denotes additive noise on the mth output channel. To blindly identify this system, we make following assumptions:

A) The inputs $s_r[n]$ are stationary, non-Gaussian, mutually independent and the samples within each input are independent and identically distributed (i.i.d.).

B) The additive noise is randomly sampled from a Gaussian distribution and is independent of the input signals.

4. CONVOLUTIVE ICA AS CS

We will turn to fourth-order cumulants to blindly retrieve the system coefficients, which is a common approach in tensor-based methods for instantaneous mixtures [2]. To take the time shifts due to the convolution into account, the spatio-temporal cumulant is computed here. This leads to a 7th-order tensor $\mathcal{C}^{\mathbf{x}} \in \mathbb{C}^{M \times M \times M \times M \times (2L+1) \times (2L+1) \times (2L+1)}$. If we use the shorthand x_{m_x,τ_y} for $x_{m_x}[t+\tau_y]$, its elements can be found by

$$\begin{aligned}&c^{\mathbf{x}}_{m_1,m_2,m_3,m_4}(t_1,t_2,t_3)\\ &\quad = \mathrm{Cum}\left[x_{m_1}, \bar{x}_{m_2,\tau_1}, \bar{x}_{m_3,\tau_2}, x_{m_4,\tau_3}\right]\\ &\quad = \mathbb{E}\left\{x_{m_1} \bar{x}_{m_2,\tau_1} \bar{x}_{m_3,\tau_2} x_{m_4,\tau_3}\right\}\\ &\qquad - \mathbb{E}\left\{x_{m_1} \bar{x}_{m_2,\tau_1}\right\} \mathbb{E}\left\{\bar{x}_{m_3,\tau_2} x_{m_4,\tau_3}\right\}\\ &\qquad - \mathbb{E}\left\{x_{m_1} \bar{x}_{m_3,\tau_2}\right\} \mathbb{E}\left\{\bar{x}_{m_2,\tau_1} x_{m_4,\tau_3}\right\}\\ &\qquad - \mathbb{E}\left\{x_{m_1} x_{m_4,\tau_3}\right\} \mathbb{E}\left\{\bar{x}_{m_2,\tau_1} \bar{x}_{m_3,\tau_2}\right\},\end{aligned}$$

in which $m_1, m_2, m_3, m_4 \in \{1, \ldots, M\}$ and $\tau_1, \tau_2, \tau_3 \in \{-L, \ldots, L\}$. In general, this cumulant tensor can be complex. Here, we will only consider real signals for simplicity, which will lead to a real cumulant tensor. In a next step, the tensor is permuted and reshaped as in [6] so that a fourth-order tensor $\mathcal{C}^{\mathbf{x},4} \in \mathbb{R}^{M \times M(2L+1) \times M(2L+1) \times M(2L+1)}$ is obtained. Mathematically, its entries $c^{\mathbf{x},4}_{i_1,i_2,i_3,i_4}$ can be found as

$$c^{\mathbf{x},4}_{m_1,i_2,i_3,i_4} = c^{\mathbf{x}}_{m_1,m_2,m_3,m_4}(t_1,t_2,t_3),$$

with $i_p = (t_{p-1} + L)M + m_p$ for $p \in \{2, \ldots, 4\}$. Because of the assumptions on the inputs and the additive noise, the tensor $\mathcal{C}^{\mathbf{x},4}$ admits a CPD consisting of $R(L+1)$ terms [6]. More specifically, it can be shown that

$$\mathcal{C}^{\mathbf{x},4} = \left[\!\left[\boldsymbol{\gamma}; \widetilde{\mathbf{H}}, \mathbf{H}, \mathbf{H}, \mathbf{H} \right]\!\right], \qquad (2)$$

in which $\boldsymbol{\gamma} = \mathbf{1}_{L+1} \otimes [\gamma_1; \ldots; \gamma_R]$, with γ_r the fourth-order cumulant of $s_r[t]$. To illustrate the structure within the factor matrices, first define the matrices

$$\mathbf{H}(l) = \begin{bmatrix} h_{11}(l) & h_{12}(l) & \cdots & h_{1R}(l) \\ \vdots & \vdots & & \vdots \\ h_{M1}(l) & h_{M2}(l) & \cdots & h_{MR}(l) \end{bmatrix},$$

for $l \in \{0, \ldots, L\}$. The factor matrix $\widetilde{\mathbf{H}}$ can then be constructed as

$$\widetilde{\mathbf{H}} = \begin{bmatrix} \mathbf{H}(0) & \mathbf{H}(1) & \cdots & \mathbf{H}(L) \end{bmatrix} \in \mathbb{C}^{M \times R(L+1)}.$$

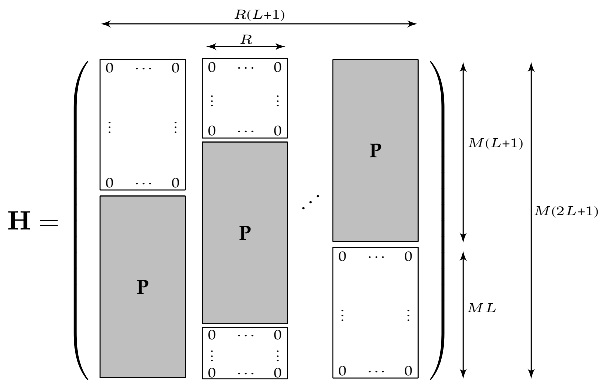

Fig. 1. Block-Hankel structure of the factor matrix $\mathbf{H}$.

The other factor matrix $\mathbf{H}$ has the circular block-Hankel structure shown in Figure 1. Each block column is given by

$$\mathbf{H}^{(l)} = \begin{bmatrix} \mathbf{0}_{M(L-l),R} \\ \mathbf{P} \\ \mathbf{0}_{Ml,R} \end{bmatrix} \in \mathbb{C}^{M(2L+1)\times R},$$

with $l \in \{0,\ldots,L\}$. The matrix $\mathbf{P} \in \mathbb{C}^{M(L+1)\times R}$ contains all coefficients of the convolutive system and is given by

$$\mathbf{P} = \begin{bmatrix} \mathbf{H}(0) \\ \mathbf{H}(1) \\ \vdots \\ \mathbf{H}(L) \end{bmatrix} = \begin{bmatrix} h_{11}(0) & h_{12}(0) & \cdots & h_{1R}(0) \\ \vdots & \vdots & & \vdots \\ h_{M1}(0) & h_{M2}(0) & \cdots & h_{MR}(0) \\ h_{11}(1) & h_{12}(1) & \cdots & h_{1R}(1) \\ \vdots & \vdots & & \vdots \\ h_{M1}(L) & h_{M2}(L) & \cdots & h_{MR}(L) \end{bmatrix}.$$

Note that $\mathbf{P}$ is closely related to the factor matrix $\widetilde{\mathbf{H}}$.

Up to this point, the approach was similar to existing methods for instantaneous instantaneous mixtures, in which the cumulant tensor also admits a low-rank CPD [2]. However, in the convolutive case, far more structure is present in the CPD from equation (2). To effectively compute this decomposition, we would like to exploit the available structure. Though this is possible in some toolboxes for tensor computations, the implementation can be cumbersome. Here, we propose to view the structured CPD as un unstructured CS problem. If we knew the entries of a tensor $\mathcal{Y} = [\![\boldsymbol{\delta}; \mathbf{P}, \mathbf{P}, \mathbf{P}, \mathbf{P}]\!] \in \mathbb{R}^{M(L+1)\times M(L+1)\times M(L+1)\times M(L+1)}$, with $\boldsymbol{\delta} = [\gamma_1; \ldots; \gamma_R]$, the system coefficients could be easily obtained through an unstructured CPD. Moreover, this would be a natural generalization of the instantaneous ICA case, in which each factor matrix also is an unstructured collection of mixing coefficients.

We will show that the tensor $\mathcal{Y}$ is indeed available, but only implicitly. More specifically, its vectorized form is related to the vectorized form of $\mathcal{C}^{\mathbf{x},4}$ by multiplication with a sparse matrix:

$$\mathbf{B}_{\text{full}} \text{Vec}\,(\mathcal{Y}) = \text{Vec}\,\left(\mathcal{C}^{\mathbf{x},4}\right). \tag{3}$$

The definition of $\mathbf{B}_{\text{full}}$ requires some explanation. A conceptually simple (but not the most efficient) way to construct $\mathbf{B}_{\text{full}}$ is to start by constructing a related matrix. Consider the matrix $\mathbf{P}_{\text{pad}}$, which is a zero-padded version of $\mathbf{P}$:

$$\mathbf{P}_{\text{pad}} = \begin{bmatrix} \mathbf{0}_{ML,R} \\ \mathbf{P} \\ \mathbf{0}_{ML,R} \end{bmatrix}.$$

Using this matrix, the tensor $\mathcal{Q} = [\![\boldsymbol{\delta}; \mathbf{P}, \mathbf{P}_{\text{pad}}, \mathbf{P}_{\text{pad}}, \mathbf{P}_{\text{pad}}]\!]$ can be constructed. Note that this tensor contains all elements of $\mathcal{Y}$, supplemented with zero entries. Let us first look at the matrix $\mathbf{B}^{\text{tmp}}$ of size $M^4(2L+1)^3 \times M^4(L+1)(3L+1)^3$ in the relation

$$\mathbf{B}^{\text{tmp}} \text{Vec}\,(\mathcal{Q}) = \text{Vec}\,\left(\mathcal{C}^{\mathbf{x},4}\right).$$

This matrix $\mathbf{B}^{\text{tmp}}$ is much easier to construct than $\mathbf{B}_{\text{full}}$. Let us consider its tensorized version $\mathcal{B}^{\text{tmp}}$, which is an 8th-order tensor of dimensions $M \times M(2L+1) \times M(2L+1) \times M(2L+1) \times M(L+1) \times M(3L+1) \times M(3L+1) \times M(3L+1)$. Its entries $b^{\text{tmp}}_{i_1,i_2,i_3,i_4,i_1+\ell M,i_2+\ell M,i_3+\ell M,i_4+\ell M}$ are equal to 1 for $\ell \in \{0,\ldots,L\}$, $i_1 \in \{1,\ldots,M\}$ and $i_2, i_3, i_4 \in \{1,\ldots,M(2L+1)\}$, and are 0 otherwise. The sparse and binary structure of this possibly large tensor can be exploited. By reshaping to a matrix again, $\mathbf{B}^{\text{tmp}}$ is found. The next step is the extraction of $\mathbf{B}_{\text{full}}$. As previously noticed, Vec $(\mathcal{Q})$ contains all elements of Vec $(\mathcal{Y})$, supplemented with zero entries. It thus suffices to drop the columns of $\mathbf{B}^{\text{tmp}}$ corresponding to these extra zero elements to obtain $\mathbf{B}_{\text{full}}$, which is straightforward to implement.

The dimensions of the sparse matrix can be reduced even further, since not all entries of $\mathcal{C}^{\mathbf{x},4}$ are needed in equation (3). This also allows us to reduce the cumulant estimation cost. Let us write Vec $\left(\mathcal{C}^{\mathbf{x},4}\right)_{\text{red.}}$ for the vector containing only the relevant entries of $\mathcal{C}^{\mathbf{x},4}$ to obtain

$$\mathbf{B} \text{Vec}\,(\mathcal{Y}) = \text{Vec}\,\left(\mathcal{C}^{\mathbf{x},4}\right)_{\text{red.}}. \tag{4}$$

The matrix $\mathbf{B}$ is obtained from $\mathbf{B}_{\text{full}}$ by dropping the zero rows of the latter. The structure of $\mathbf{B}$ for a system with $M = 3$, $L = 1$ and $R = 2$ is illustrated in Figure 2. Note that the structure from the original CPD in equation (2) is now ported to the known matrix $\mathbf{B}$. This matrix always has $(ML)^4$ fewer rows than columns, which can be verified by checking the number of entries of $\mathcal{Y}$ and Vec $\left(\mathcal{C}^{\mathbf{x},4}\right)_{\text{red.}}$ in Table 1. This implies that equation (4) is underdetermined, which is where CS comes in. To obtain Vec $(\mathcal{Y})$ from Vec $\left(\mathcal{C}^{\mathbf{x},4}\right)_{\text{red.}}$, the (vectorized) tensor $\mathcal{Y}$ has to be sufficiently compressible in some way. This is obviously the case if R is not excessively large since it can be written as a CPD consisting of R terms. In this way, the $M^4(L+1)^4$ entries of $\mathcal{Y}$ can be described by $4RM(L+1)$ parameters (ignoring scaling indeterminacies), and even by $RM(L+1)$ parameters if symmetry is taken into account.

Solving equation (4) whilst imposing a vectorized CPD structure on Vec $(\mathcal{Y})$ can be done using optimization algorithms. From the decomposition of $\mathcal{Y}$, the matrix $\mathbf{P}$ and thus

Table 1. Number of entries in the full and reduced (vectorized) cumulant tensors and the desired tensor $\mathcal{Y}$ for convolutive systems with M outputs, R inputs and a maximum delay of L.

Tensor	Number of elements
$\mathcal{C}^{\mathbf{x},4}$	$M^4(2L+1)^3$
$\text{Vec}\left(\mathcal{C}^{\mathbf{x},4}\right)_{\text{red.}}$	$M^4(2L+1)(2L^2+2L+1)$
$\mathcal{Y}$	$M^4(L+1)^4$

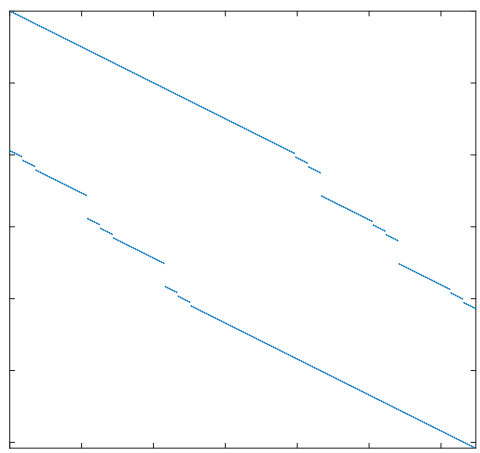

Fig. 2. Structure of sensing matrix $\mathbf{B}$ for a system with $M = 3$, $L = 1$ and $R = 2$. Each dot represents a value of 1 at that location in the matrix. All other values are zero.

the system coefficients can be extracted up to permutation and scaling of the columns of $\mathbf{P}$.

As noticed before, an interesting feature of this approach is that the structure due to the convolution has been ported from the factor matrix $\mathbf{H}$ in decomposition (2) to the known coefficient matrix $\mathbf{B}$ in the CS equation (4), with $\text{Vec}\left(\mathcal{C}^{\mathbf{x},4}\right)_{\text{red.}}$ having fewer entries than $\mathcal{C}^{\mathbf{x},4}$ in (2).

5. EXPERIMENTS

In our experiment, the obtained performance is expressed as the relative error norm (REN) of the matrix $\mathbf{P}$ after optimal scaling and permutation, mathematically described by

$$\text{REN}\left(\widehat{\mathbf{P}}\right) = 20\log_{10}\left(\frac{\left|\left|\mathbf{P} - \widehat{\mathbf{P}}\Delta_{\text{opt}}\Pi_{\text{opt}}\right|\right|_{\text{F}}}{||\mathbf{P}||_{\text{F}}}\right) \quad \text{(dB)},$$

in which Δ_{opt} and Π_{opt} represent the optimal scaling and permutation matrices respectively. The fully structured tensor decomposition was computed using the structured data fusion framework of Tensorlab [17–19].

The considered system has $M = 3$ outputs, $R = 2$ inputs and a maximum delay of $L = 1$. The measured outputs consist of 10^4 samples. The inputs are real signals randomly sampled from a uniform distribution on $[-0.5, 0.5]$. The system coefficients are randomly sampled from a standard normal distribution. The obtained average accuracy over 100 Monte Carlo runs is shown for various signal-to-noise ratio (SNR) values in Figure 3, for both a fully structured decomposition and the CS approach. For each experiment, the optimization algorithm was initialized 5 times with randomly chosen system coefficients from a standard normal distribution. The minimum of these 5 results was selected. The figure shows that the CS approach and a fully structured CPD yield comparable accuracy.

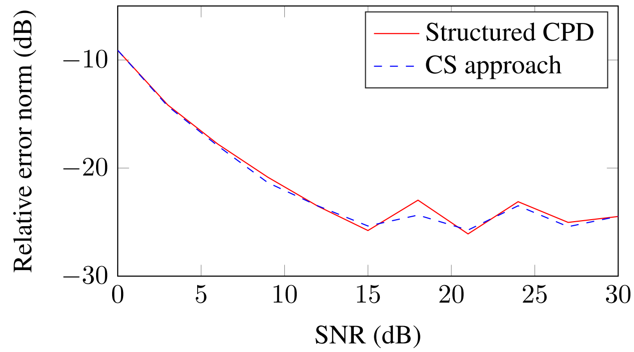

Fig. 3. Accuracy of both the fully structured CPD and the CS approach.

Computational requirements will be discussed in a full paper. For now, we argue that the computation time of the CS approach will not largely exceed that of a full structured decomposition, and may even improve on it. This is because the solution to (4) has lower rank and contains less structure than decomposition (2). Moreover, the sensing matrix $\mathbf{B}$ can be efficiently implemented since it is both sparse and binary.

6. CONCLUSION

Using a CS approach, the structure of the tensor decomposition arising in blind system identification can be ported to a known matrix. The remaining system of equations has no further structure apart from the low-rank CPD assumption, which simplifies the modeling of the problem. A numerical experiment has illustrated that this approach attains similar accuracy as the fully structured decomposition.

7. REFERENCES

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

[2] P. Comon and C. Jutten, Eds., *Handbook of blind source*

separation, independent component analysis and applications, Academic Press, 2010.

[3] M. Pedersen, J. Larsen, U. Kjems, and L. Parra, "A survey of convolutive blind source separation methods," *Multichannel Speech Processing Handbook*, pp. 1065–1084, 2007.

[4] L. De Lathauwer, J. Castaing, and J.-F. Cardoso, "Fourth-Order Cumulant-Based Blind Identification of Underdetermined Mixtures," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2965–2973, 2007.

[5] L. De Lathauwer and J. Castaing, "Blind Identification of Underdetermined Mixtures by Simultaneous Matrix Diagonalization," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1096–1105, 2008.

[6] C. E. R. Fernandes, G. Favier, and J. C. Mota, "PARAFAC-based blind identification of convolutive MIMO linear systems," in *System Identification*, 2009, vol. 15, pp. 1704–1709.

[7] H. Bousbia-Salah, A. Belouchrani, and K. Abed-Meraim, "Blind separation of convolutive mixtures using joint block diagonalization," in *Sixth International Symposium on Signal Processing and its Applications*. IEEE, 2001, vol. 1, pp. 13–16.

[8] Y. Yu and A.P. Petropulu, "PARAFAC-based blind estimation of possibly underdetermined convolutive MIMO systems," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 111–124, Jan. 2008.

[9] D. Nion, K. N. Mokios, N. D. Sidiropoulos, and A. Potamianos, "Batch and adaptive parafac-based blind separation of convolutive speech mixtures," *IEEE Trans. Audio, Speech, Language Process.*, vol. 18, no. 6, pp. 1193–1207, 2010.

[10] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," *IEEE Trans. Speech and Audio Process.*, vol. 12, no. 5, pp. 530–538, 2004.

[11] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, 2008.

[12] T. Kolda and B. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[13] A. Cichocki, C. Mandic, A. H. 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*, vol. 32, no. 2, pp. 145–163, Mar. 2015.

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

[15] N. D. Sidiropoulos and R. Bro, "On the Uniqueness of Multilinear Decomposition of N-way Arrays," *J. Chemometrics*, vol. 14, pp. 229–239, 2000.

[16] I. Domanov and L. De Lathauwer, "On the uniqueness of the canonical polyadic decomposition of third-order tensors — Part II: Overall uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 876–903, 2013.

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

[18] L. Sorber, M. Van Barel, and L. De Lathauwer, "Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank-$(L_r, L_r, 1)$ terms, and a new generalization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 695–720, 2013.

[19] L. Sorber, M. Van Barel, and L. De Lathauwer, "Structured data fusion," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 4, pp. 586–600, June 2015.