



HAL
open science

ON JEVD OF SEMI-DEFINITE POSITIVE MATRICES AND CPD OF NONNEGATIVE TENSORS

Rémi André, Laurent Albera, Xavier Luciani, Eric Moreau

► **To cite this version:**

Rémi André, Laurent Albera, Xavier Luciani, Eric Moreau. ON JEVD OF SEMI-DEFINITE POSITIVE MATRICES AND CPD OF NONNEGATIVE TENSORS. IEEE SAM 2016, Jul 2016, Rio de Janeiro, Brazil. pp.7569738. hal-01862241

HAL Id: hal-01862241

<https://hal.science/hal-01862241v1>

Submitted on 27 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON JEVD OF SEMI-DEFINITE POSITIVE MATRICES AND CPD OF NONNEGATIVE TENSORS

Rémi André^(1,2), Laurent Albera^(3,4), Xavier Luciani^(1,2) and Eric Moreau^(1,2)

⁽¹⁾ Aix Marseille Université, CNRS, ENSAM, LSIS, UMR 7296, 13397 Marseille, France.

⁽²⁾ Université de Toulon, CNRS, LSIS, UMR 7296, 83957 La Garde, France.

⁽³⁾ Inserm, U1099, Rennes F-35000, France.

⁽⁴⁾ Université de Rennes 1, LTSI, Rennes F-35000, France.

ABSTRACT

In this paper, we mainly address the problem of Joint Eigen-Value Decomposition (JEVD) subject to nonnegative constraints on the eigenvalues of the matrices to be diagonalized. An efficient method based on the Alternating Direction Method of Multipliers (ADMM) is designed. ADMM provides an elegant approach for handling nonnegativity constraints, while taking advantage of the structure of the objective function. Numerical tests on simulated matrices show the interest of the proposed method for low Signal-to-Noise Ratio (SNR) values when the similarity transformation matrix is ill-conditioned. The ADMM was recently used for the Canonical Polyadic Decomposition (CPD) of nonnegative tensors leading to the ADMoM algorithm. We show through computer results that DIAG_+ , a semi-algebraic CPD method using our ADMM-based JEVD_+ algorithm, will give a better estimation of factors than ADMoM in the presence of swamps. DIAG_+ also appears to be less time-consuming than ADMoM when low-rank tensors of high dimensions are considered.

1. INTRODUCTION

The Canonical Polyadic Decomposition (CPD) of multi-way arrays plays an important role in various application areas such as medical diagnosis [1]. A large number of CPD methods were proposed [2–6]. Among them, semi-algebraic (or direct) methods showed a very good performance (e.g. fast convergence speed, robustness with respect to initialization and overfactoring) [7]. Semi-algebraic techniques aim at reformulating the CPD as a matrix factorization problem of lower dimension. For instance, the CFS (Closed Form Solution) [8] and DIAG (Direct ALgorithm) [7] algorithms formulate the CPD as a JEVD (Joint EVD) problem while the SSD-CP (CP decomposition based on Simultaneous Schur Decomposition) method formulates it as a problem of joint upper triangularization by an orthogonal similarity [9].

In some applications such as fluorescence spectroscopy [7], the multi-way array to be decomposed and its loading

matrices are nonnegative. Hence the need for efficient non-negative CPD methods. Note that the use of nonnegativity was shown to ensure the existence of optimal solutions to the CPD approximation problem [10]. Then we may wonder whether the semi-algebraic CPD methods can be modified in order to take into account the nonnegativity of the data. If tensors are nonnegative, then it appears that the corresponding matrices to be jointly diagonalized or triangularized have nonnegative eigenvalues. In other words, modifying the CFS and DIAG techniques in order to take into account nonnegativity requires to solve the following JEVD_+ problem:

Problem 1. Given K semi-definite positive matrices $M^{(k)}$ in $\mathbb{R}^{R \times R}$, factorize them as $M^{(k)} = \mathbf{A} \mathbf{D}^{(k)} \mathbf{A}^{-1}$ subject to:

$$\forall k \in \{1, \dots, K\}, \mathbf{D}^{(k)} \in \mathcal{D}^+$$

where \mathcal{D}^+ is the subset of diagonal matrices of $\mathbb{R}^{R \times R}$ such that $\text{Diag}(\mathcal{D}^+) = [0, +\infty]^R$.

In order to solve problem 1, we proposed an approach based on the Alternating Direction Method of Multipliers (ADMM) [11]. ADMM provides an elegant approach for handling nonnegativity constraints, while taking advantage of the structure of the objective function. Numerical tests on simulated matrices show the interest of the proposed method for low Signal-to-Noise Ratio (SNR) values when the similarity transformation matrix is ill-conditioned. The ADMM was recently used for the Canonical Polyadic Decomposition (CPD) of nonnegative tensors leading to the ADMoM algorithm [12]. We show through computer results that DIAG_+ , a semi-algebraic CPD method using our ADMM-based JEVD_+ algorithm, will give a better estimation of factors than ADMoM in the presence of swamps. DIAG_+ also appears to be less time-consuming than ADMoM when low-rank tensors of high dimensions are considered.

2. ALGORITHMS

The purpose of this section is first to show the interest in solving the JEVD_+ problem for the CPD of nonnegative tensors.

Secondly, an efficient ADMM-based JEVD₊ solution is proposed.

2.1. The CPD of nonnegative tensors

The Nonnegative CPD (NCPD) of a three-way array $\mathcal{T} \in [0, +\infty[^{N_1 \times N_2 \times N_3}$ is given by:

$$\mathcal{T} = \sum_{r=1}^R \mathbf{f}_r^{(1)} \circ \mathbf{f}_r^{(2)} \circ \mathbf{f}_r^{(3)} \quad (1)$$

where $R = \text{rank}(\mathcal{T})$ denotes the rank of \mathcal{T} and where:

$$\mathbf{F}^{(1)} = [\mathbf{f}_1^{(1)}, \dots, \mathbf{f}_R^{(1)}] \quad (2)$$

$$\mathbf{F}^{(2)} = [\mathbf{f}_1^{(2)}, \dots, \mathbf{f}_R^{(2)}] \quad (3)$$

$$\mathbf{F}^{(3)} = [\mathbf{f}_1^{(3)}, \dots, \mathbf{f}_R^{(3)}] \quad (4)$$

are the so-called nonnegative loading matrices of \mathcal{T} .

The first unfolding matrix $\mathbf{T}^{(1)}$ of size $(N_1 \times N_2 N_3)$ can be expressed as:

$$\mathbf{T}^{(1)} = \mathbf{F}^{(1)} \left(\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)} \right)^\top \quad (5)$$

with:

$$\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)} = [\mathbf{f}_1^{(3)} \otimes \mathbf{f}_1^{(2)}, \dots, \mathbf{f}_R^{(3)} \otimes \mathbf{f}_R^{(2)}] \quad (6)$$

$$= [\Phi^{(1)} \mathbf{F}^{(2)\top}, \dots, \Phi^{(N_3)} \mathbf{F}^{(2)\top}]^\top \quad (7)$$

according to [7] where $\Phi^{(n)}$ is a $(R \times R)$ diagonal matrix built from the n -th row of matrix $\mathbf{F}^{(3)}$. The Singular Value Decomposition (SVD) truncated at order R of $\mathbf{T}^{(1)}$ is given by:

$$\mathbf{T}^{(1)} = \mathbf{V} \Sigma \mathbf{W}^\top \quad (8)$$

Then there is a non-singular matrix \mathbf{A} of size $(R \times R)$ such that:

$$\mathbf{V} \Sigma = \mathbf{F}^{(1)} \mathbf{A}^{-1} \quad (9)$$

$$\mathbf{W}^\top = \mathbf{A} \left(\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)} \right)^\top \quad (10)$$

As shown in [7], \mathbf{W} has the following matrix block structure:

$$\mathbf{W} = [\Gamma^{(1)}, \dots, \Gamma^{(N_3)}]^\top \quad (11)$$

where the N_3 matrices $\Gamma^{(n)}$ of size $(R \times N_2)$ are given by:

$$\Gamma^{(n)} = \mathbf{A} \Phi^{(n)} \mathbf{F}^{(2)\top} \quad (12)$$

Then we can define $N^{(3)}(N^{(3)} - 1)$ matrices $\mathbf{M}^{(n_1, n_2)}$ as follows:

$$\begin{aligned} \mathbf{M}^{(n_1, n_2)} &= \Gamma^{(n_1)} \Gamma^{(n_2)\#} \\ &= \mathbf{A} \Phi^{(n_1)} \mathbf{F}^{(2)\top} (\mathbf{F}^{(2)\top})^\# (\Phi^{(n_2)})^{-1} \mathbf{A}^{-1} \\ \mathbf{M}^{(n_1, n_2)} &= \mathbf{A} \mathbf{D}^{(n_1, n_2)} \mathbf{A}^{-1} \end{aligned} \quad (13)$$

where for each couple (n_1, n_2) , the $(R \times R)$ matrix $\mathbf{D}^{(n_1, n_2)}$ is given by:

$$\mathbf{D}^{(n_1, n_2)} = \Phi^{(n_1)} (\Phi^{(n_2)})^{-1} \quad (14)$$

is diagonal. Consequently, the DIAG algorithm [7] consists in reformulating the CPD problem (1) as the JEVD problem given by (13), which appears to be an optimization problem of smaller size. Once matrix \mathbf{A} is computed, the matrices $\mathbf{F}^{(1)}$ and $\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)}$ can be derived from the SVD of $\mathbf{T}^{(1)}$ using equations (9) and (10). As explained in [7], the column vectors of $\mathbf{F}^{(3)}$ and $\mathbf{F}^{(2)}$ can be derived from the columns of $\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)}$ using the HOSVD [13].

By considering nonnegative loading matrices, it clearly appears from (14) that the $N^{(3)}(N^{(3)} - 1)$ matrices $\mathbf{M}^{(n_1, n_2)}$ to be jointly diagonalized have nonnegative eigenvalues. Unfortunately, the JEVD methods [7, 14–16], such as the efficient Jacobi-like JDJM (Joint Diagonalization algorithm based on Targeting hyperbolic Matrices) technique based on the polar matrix factorization, did not take into account such a property. Consequently, let's see in the following subsection how to overcome this drawback by proposing a nonnegative JEVD method.

2.2. An ADMM-based JEVD₊ method

The JEVD₊ problem (see problem 1 defined in introduction) can be reformulated as the following minimization problem:

$$\underset{\mathbf{A}, \mathbf{B}, \mathbf{C}}{\text{minimize}} \Psi(\mathbf{A}, \mathbf{B}, \mathbf{C}) \quad \text{subject to} \quad \mathbf{C} = \tilde{\mathbf{C}} \quad (15)$$

with:

$$\Psi(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \|\mathbf{X}^{(1)} - \mathbf{A}(\mathbf{C} \odot \mathbf{B})^\top\|_F^2 + \frac{\alpha}{2} \|\mathbf{A}\mathbf{B}^\top - \mathbf{I}_R\|_F^2 + \iota_{\mathcal{C}^+}(\tilde{\mathbf{C}}) \quad (16)$$

where $\mathbf{X}^{(1)} = [\mathbf{M}^{(1)}, \dots, \mathbf{M}^{(K)}]$ is the first unfolding matrix of a third order tensor \mathcal{X} for which \mathbf{A} , \mathbf{B} and \mathbf{C} are the three loading matrices, where the k -th row of \mathbf{C} is the diagonal vector of $\mathbf{D}^{(k)}$ and where $\mathbf{B} = \mathbf{A}^{-\top}$ with $\mathbf{M}^{(k)} = \mathbf{A} \mathbf{D}^{(k)} \mathbf{A}^{-1}$ the k -th matrix to be jointly diagonalized (see problem 1). $\iota_{\mathcal{C}^+}$ is the indicator function of set of $(K \times R)$ matrices with nonnegative components. The indicator of a set \mathcal{S} of a Hilbert space \mathcal{H} is defined as $(\forall \mathbf{x} \in \mathcal{H}) \iota_{\mathcal{S}}(\mathbf{x}) = 0$ if $\mathbf{x} \in \mathcal{S}$, and $+\infty$ otherwise.

The augmented Lagrangian function associated with (15) is given by

$$\begin{aligned} \mathbf{L}_\rho(\mathbf{A}, \mathbf{B}, \mathbf{C}, \tilde{\mathbf{C}}, \Lambda) &= \Psi(\mathbf{A}, \mathbf{B}, \mathbf{C}) \\ &+ \|\Lambda \square (\mathbf{C} - \tilde{\mathbf{C}})\|_F^2 + \frac{\rho}{2} \|\mathbf{C} - \tilde{\mathbf{C}}\|_F^2 \end{aligned} \quad (17)$$

where the $(K \times R)$ matrix Λ is the matrix Lagrangian multiplier, where \square is Hadamard product operator and where the variable $\rho \in]0, +\infty[$ is a penalty parameter for the linear equality constraints.

An ADMM-like algorithm for solving (15) is obtained by successively i) minimizing the augmented Lagrangian function (17) with respect to \mathbf{A} , \mathbf{B} , \mathbf{C} and $\tilde{\mathbf{C}}$, one variable at a time while setting the others to their most recent values, and ii) updating the multiplier $\mathbf{\Lambda}$ using an ascent gradient rule. By using (16) and (17), it turns out that the alternating minimizations of (17) have the following closed form solutions:

$$\begin{aligned}\mathbf{A} &= (\mathbf{X}^{(1)}(\mathbf{C} \odot \mathbf{B}) + \alpha \mathbf{B}) \\ &\quad ((\mathbf{C} \odot \mathbf{B})^\top (\mathbf{C} \odot \mathbf{B}) + \alpha \mathbf{B}^\top \mathbf{B})^{-1} \\ \mathbf{B} &= (\mathbf{X}^{(2)}(\mathbf{C} \odot \mathbf{A}) + \alpha \mathbf{A}) \\ &\quad ((\mathbf{C} \odot \mathbf{A})^\top (\mathbf{C} \odot \mathbf{A}) + \alpha \mathbf{A}^\top \mathbf{A})^{-1} \\ \mathbf{C} &= (\mathbf{X}^{(3)}(\mathbf{B} \odot \mathbf{A}) + \rho \tilde{\mathbf{C}} - \mathbf{\Lambda}) \\ &\quad ((\mathbf{B} \odot \mathbf{A})^\top (\mathbf{B} \odot \mathbf{A}) + \rho \mathbf{I}_R)^{-1} \\ \tilde{\mathbf{C}} &= \text{P}_{\mathcal{C}^+}(\mathbf{C} + \frac{1}{\rho} \mathbf{\Lambda})\end{aligned}$$

where $\mathbf{X}^{(2)}$ and $\mathbf{X}^{(3)}$ are the second and the third unfolding matrix of \mathcal{X} , respectively, and where $\text{P}_{\mathcal{S}}$ denote the projection onto a convex set \mathcal{S} . The gradient of the augmented Lagrangian function L_ρ with respect to the multiplier $\mathbf{\Lambda}$ is given by $\rho(\mathbf{C} - \tilde{\mathbf{C}})$. Regarding the penalty parameter ρ , it is computed as explained in [12]:

$$\rho_{it+1} = \begin{cases} \tau^{(1)} \rho_{it} & \text{if } \|\mathbf{P}_{it}\|_F > \mu \|\mathbf{D}_{it}\|_F \\ \rho_{it} / \tau^{(2)} & \text{if } \|\mathbf{D}_{it}\|_F > \mu \|\mathbf{P}_{it}\|_F \\ \rho_{it} & \text{otherwise.} \end{cases} \quad (18)$$

where $\mathbf{P}_{it} = \mathbf{C}_{it} - \tilde{\mathbf{C}}_{it}$ and $\mathbf{D}_{it} = \rho(\tilde{\mathbf{C}}_{it} - \tilde{\mathbf{C}}_{it-1})$ with $\mu > 1$, $\tau^{(1)} > 1$ and $\tau^{(2)} > 1$. As far as α is concerned, we initialize it to 1 and we divide it by 2 every 10 iterations.

3. NUMERICAL RESULTS

Now, we investigate the behavior of the proposed JEVD₊ method and its use in the DIAG algorithm leading to the DIAG₊ technique.

To build simulated data sets, we firstly randomly generated $L = 100$ diagonal matrices $\mathbf{D}^{(k)}$ and a non-singular matrix \mathbf{A} of size $(R \times R)$ with $R = 3$ by using a uniform distribution over $[0, 1]$ and a standard normal distribution, respectively. Then we built L sets of K matrices $\mathbf{M}^{(k)}$ to be jointly diagonalized such that:

$$\forall k, \quad \mathbf{M}^{(k)} = \frac{\mathbf{A} \mathbf{D}^{(k)} \mathbf{A}^{-1}}{\|\mathbf{A} \mathbf{D}^{(k)} \mathbf{A}^{-1}\|_F} + \sigma \frac{\mathbf{N}^{(k)}}{\|\mathbf{N}^{(k)}\|_F} \quad (19)$$

where $\mathbf{N}^{(k)}$ models a Gaussian random noise and σ allows us to adjust the noise power to the desired SNR. We vary the SNR value from 0 to 70 dB. The condition number of \mathbf{A} was adjusted by means of the following β parameter:

$$\forall r > 1, \quad \mathbf{a}_r = \beta \mathbf{a}_1 + (1 - \beta) \mathbf{u}_{[0,1]} \quad (20)$$

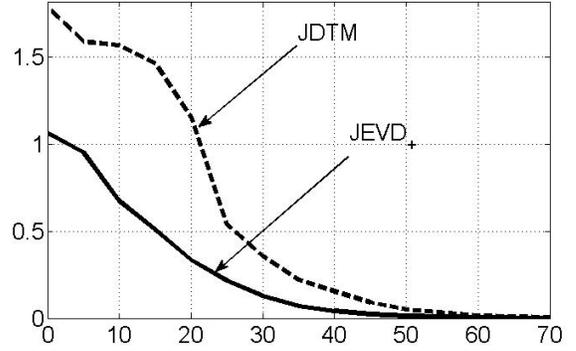


Fig. 1. Median value of the $m_{\mathcal{C}}$ criterion at the output of the JD TM and JEVD₊ methods as a function of SNR.

where \mathbf{a}_r denotes the r -th column vector of \mathbf{A} and where $\mathbf{u}_{[0,1]}$ is a R -dimensional vector generated using a uniform distribution over $[0, 1]$. The β parameter was fixed to 0.65.

Secondly, we randomly generated $L = 100$ loading matrices $\mathbf{F}^{(1)}$, $\mathbf{F}^{(2)}$ and $\mathbf{F}^{(3)}$ of size $(N^{(1)} \times R)$, $(N^{(2)} \times R)$ and $(N^{(3)} \times R)$, respectively, by using a uniform distribution over $[0, 1]$ with $R = 2$ and $N^{(1)} = N^{(2)} = N^{(3)} = 100$. The condition number was also adjusted using the strategy described by (20). Two β values were considered: $\beta = 0$ and $\beta = 0.45$. Then we built L sets of third order tensors \mathcal{T} to be canonically decomposed such that their first loading matrix $\mathbf{T}^{(1)}$ is defined by :

$$\mathbf{T}^{(1)} = \frac{\mathbf{F}^{(1)} (\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)})^\top}{\|\mathbf{F}^{(1)} (\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)})^\top\|_F} + \sigma \frac{\mathbf{Q}^{(k)}}{\|\mathbf{Q}^{(k)}\|_F} \quad (21)$$

where $\mathbf{Q}^{(k)}$ models a Gaussian random noise and σ allows us to adjust the noise power to the desired SNR. The SNR value was fixed to 100 dB.

Regarding the first performance criterion, we used an error between the calculated and actual matrices. More particularly, we computed the normalized Frobenius norm $m_{\mathcal{G}}$ of the difference between the actual matrix \mathbf{G} and its estimate $\hat{\mathbf{G}}$. Note that the indetermination up to scale and permutation was corrected before the computation of $m_{\mathcal{G}}$. The second performance criterion was the CPU time. It is noteworthy that all experiments were run with Matlab R2014a on Intel® Xeon® 4-Core 2.8GHz, equipped with 32GB RAM.

Figures 1 and 2 display the median of the $m_{\mathcal{C}}$ and $m_{\mathcal{A}}$ criteria, respectively, at the output of the JD TM and JEVD₊ methods as a function of the SNR when the similarity transformation matrix is ill-conditioned. Thus figures 1 and 2 clearly show that both the eigenvalues and the eigenvectors are better estimated for low SNR values by taking into account the nonnegativity of the eigenvalues during the optimization procedure. More particularly, our ADMM-based JEVD₊ method gives better results than JD TM for SNR values below 40 dB.

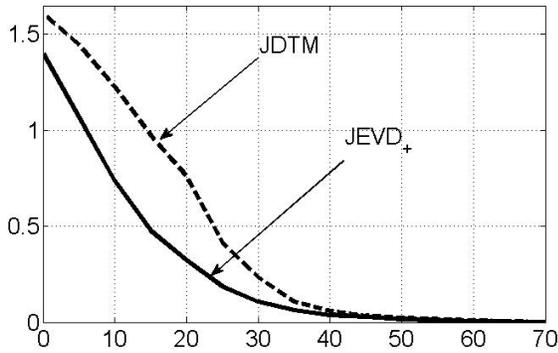


Fig. 2. Median value of the m_A criterion at the output of the JD TM and JEVD₊ methods as a function of SNR.

Table 1 compares the DIAG₊ method based on our ADMM-based JEVD₊ technique with ADMoM [12] in terms of performance for low-rank tensors of high dimensions. Note that the same stopping criteria were used for both ADMM procedures, say a threshold of 10^{-6} and a maximum number of iterations of 1000. When the three loading matrices are well-conditioned, both methods give a similar estimation accuracy while DIAG₊ is ten times faster than ADMoM. When the three loading matrices are ill-conditioned, not only DIAG₊ is still faster than ADMoM, but it estimates the loading matrices with a better accuracy. In fact, the DIAG procedure allows us to algebraically reformulate the original CPD problem as an optimization problem of smaller size, hence the gain of time. But it also allows us to reformulate the original CPD problem as a well-conditioned problem, hence the better robustness with respect to ill-conditioned loading matrices.

methods	factor collinearity index (β)	Error (10^{-3})	CPU time (s)
ADMoM	0	0.0003	3.72
DIAG ₊	0	0.0004	0.29
ADMoM	0.45	0.181	13
DIAG ₊	0.45	0.0017	0.85

Table 1. Performance of ADMoM versus DIAG₊.

4. CONCLUSION

In this paper, we showed how the ADMM algorithm could be used to compute the solution of the JEVD₊ problem and to handle the nonnegativity constraint of the eigenvalues of semi-definite positive matrices. Numerical tests have shown the good behavior of the proposed ADMM-based JEVD₊ approach for low SNR values when the similarity transformation matrix is ill-conditioned. The ADMM was recently used for the CPD of nonnegative tensors leading to the ADMoM algorithm [12]. We showed through computer results that

DIAG₊, a semi-algebraic CPD method using our ADMM-based JEVD₊ algorithm, gave a better estimation of factors than ADMoM in the presence of swamps. DIAG₊ also appeared to be less time-consuming than ADMoM when low-rank tensors of high dimensions were considered.

5. REFERENCES

- [1] H. BECKER, L. ALBERA, P. COMON, R. GRIBONVAL, F. WENDLING, and I. MERLET, “Brain source imaging: from sparse to tensor models,” *IEEE Signal Processing Magazine, special issue on Brain-Computer Interfaces*, vol. 32, no. 6, pp. 100–112, November 2015.
- [2] L. DE LATHAUWER, B. DE MOOR, and J. VANDEWALLE, “Computation of the canonical decomposition by means of a simultaneous generalized schur decomposition,” *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 2, pp. 295–327, 2004.
- [3] T. G. KOLDA and B. W. BADER, “Tensor decompositions and applications,” *SIAM Review*, vol. 51, no. 3, pp. 455–500, September 2009.
- [4] P. COMON, X. LUCIANI, and A. L. F. DE ALMEIDA, “Tensor decompositions, alternating least squares and other tales,” *Journal of Chemometrics*, vol. 23, April 2009.
- [5] P. COMON, “Tensors: a brief introduction,” *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 44–53, May 2014.
- [6] A.-H. PHAN, P. TICHAVSKY, and A. CICHOCKI, “CANDECOMP/PARAFAC decomposition of high-order tensors through tensor reshaping,” *IEEE Transactions On Signal Processing*, vol. 61, no. 19, pp. 4847–4860, 2013.
- [7] X. LUCIANI and L. ALBERA, “Canonical polyadic decomposition based on joint eigenvalue decomposition,” *Chemometrics and Intelligent Laboratory Systems*, vol. 132, pp. 152–167, March 2014.
- [8] F. ROEMER and M. HAARDT, “A semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalization (SECSI),” *Elsevier Signal Processing*, vol. 93, pp. 2462–2473, 2013.
- [9] S. HAJIPOUR SARDOUIE, L. ALBERA, M. B. SHAMSOLLAHI, and I. MERLET, “Canonical polyadic decomposition of complex-valued multi-ways arrays based on simultaneous schur decomposition,” in *ICASSP 13, 2013 IEEE International Conference on Acoustics Speech and Signal Processing*, Vancouver, Canada, May 26–31 2013, pp. 4178–4182.

- [10] L.-H. LIM and P. COMON, “Nonnegative approximations of nonnegative tensors,” *Journal of chemometrics*, pp. 432–441, 2009.
- [11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [12] A. P. LIAVAS and N. D. SIDIROPOULOS, “Parallel algorithms for constrained tensor factorization via alternating direction method of multipliers,” *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5450–5463, 2015.
- [13] L. DE LATHAUWER, B. DE MOOR, and J. VANDERWALLE, “A multilinear singular value decomposition,” *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.
- [14] T. FU and X. GAO, “Simultaneous diagonalization with similarity transformation for non-defective matrices,” vol. 4, pp. 1137–1140, May 2006.
- [15] R. IFERROUDJENE, K. ABED MERAÏM, and A. BELOUCHRANI, “A new jacobi-like method for joint diagonalization of arbitrary non-defective matrices,” *Applied Mathematics and Computation*, vol. 2, pp. 363–373, 2009.
- [16] X. LUCIANI and L. ALBERA, “Joint eigenvalue decomposition of non-defective matrices based on the lu factorization with application to ICA,” *IEEE Transactions On Signal Processing*, vol. 63, no. 17, pp. 4594–4608, September 2015.