LARGE-SCALE AUTOREGRESSIVE SYSTEM IDENTIFICATION USING KRONECKER PRODUCT EQUATIONS

Martijn Boussé[⋆] *and 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. {martijn.bousse, lieven.delathauwer}@kuleuven.be

ABSTRACT

By exploiting the intrinsic structure and/or sparsity of the system coefficients in large-scale system identification, one can enable efficient processing. In this paper, we employ this strategy for large-scale single-input multiple-output autoregressive system identification by assuming the coefficients can be well approximated by Kronecker products of smaller vectors. We show that the identification problem can be reformulated as the computation of a Kronecker product equation, allowing one to use optimization-based and algebraic solvers.

***Index Terms*—** system identification, autoregressive, Kronecker product, higher-order tensor, large-scale problems

1. INTRODUCTION

System identification is an important engineering problem in various applications [1]. Recently, there is a growing interest in *large-scale* system identification because of an increasing number and density of antennas or sensors in fields such as array processing, telecommunications, and (biomedical) signal processing [2–4]. In order to tackle such large-scale problems, the intrinsic structure and/or sparsity of the data can be exploited by means of parsimonious models.

Large-scale data is often *compressible*, or, in other words, it can often be described in terms of much fewer parameters than the total number of values [5]. Well-known examples are (exponential) polynomials, rational functions, and smooth and periodic functions [6–12]. Explicitly exploiting the intrinsic compactness of this type of data, enables efficient processing in large-scale applications. Popular compact models are low-rank matrix and tensor decompositions; see [13–15] and references therein. A well-known approach consists of reshaping a large-scale vector or matrix into a tensor which can then be modeled using a low-rank approximation [16]; this is illustrated for a sigmoid[1] in Figure 1. This approach has successfully allowed one to handle various large-scale applications in tensor-based scientific computing and (blind) system identification [6, 7, 11, 17–20].

Fig. 1. Low-rank matrix or tensor models can often provide a parsimonious and accurate representation for smooth data.

We adopt a similar strategy for autoregressive (AR) system identification [1,21], enabling large-scale applications. In this paper, we limit ourselves to single-input multiple-output (SIMO) AR models with Kronecker product constrained coefficients. Although this particular structure corresponds to a rank-1 model, as we will explain later, it can already provide an accurate and compact model while allowing us to explain the basic principles within the space restrictions of this paper. More specifically, we show that by explicitly exploiting the Kronecker structure, AR system identification can be reformulated as a type of *Kronecker product equation* (KPE). By the latter we mean a linear system of equations with a Kronecker product constrained solution, which has already been applied successfully in various applications [22–24]. A *generic* framework for this type of problems was developed in [24], allowing us to use optimization-based and algebraic solvers and formulate generic uniqueness conditions.

This research was supported by: Fonds de la Recherche Scientifique – FNRS and Fonds Wetenschappelijk Onderzoek – Vlaanderen under EOS Project no 30468160 (SeLMA), Research Council KU Leuven: C1 project C16/15/059-nD, FWO projects: G.0830.14N, G.0881.14N, 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.

[1]We evaluated a sigmoid of the form $f(\xi) = 1/(1 + e^{-20(\xi-1/2)})$ in 100 equidistant samples in $[0, 1]$ and then reshaped the vector of length 100 containing the values into a (10×10) matrix. We computed a low-rank model by truncating the singular value decomposition and the reconstructions are obtained by vectorizing the resulting rank-1 and rank-2 matrices.

In the remainder of this section we give an overview of the notation that is used in this paper, several basic definitions, and KPEs. We derive our method for large-scale SIMO AR system identification using KPEs in Section 2. In Section 3, we analyze our method via several numerical experiments. We conclude the paper and discuss future work in Section 4.

1.1. Notations and basic definitions

Vectors, matrices, and tensors are denoted by bold lowercase, bold uppercase, and calligraphic letters, respectively. The vectorization of an Nth-order tensor $\mathcal{A} \in \mathbb{K}^{I_1 \times I_2 \times \cdots \times I_N}$ ($\mathbb{K}$ meaning $\mathbb{R}$ or $\mathbb{C}$), denoted as $\text{vec}(\mathcal{A})$, 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 inverse operation of $\text{vec}(\cdot)$ is $\text{ten}(\cdot)$. We indicate the nth element in a sequence by a superscript between parentheses, e.g., $\{\mathbf{A}^{(n)}\}_{n=1}^{N}$. The outer and Kronecker product are denoted by $\otimes$ and $\otimes$, respectively. They are related through a vectorization: $\text{vec}(\mathbf{a} \otimes \mathbf{b}) = \mathbf{b} \otimes \mathbf{a}$.

The rank of a tensor is equal to the minimal number of rank-1 tensors that generate the tensor as their sum. A rank-1 tensor is defined as the outer product of non-zero vectors.

1.2. Kronecker product equation

A KPE is a linear system of equations with a Kronecker product constrained solution that has been applied successfully in various domains [22–24]. In this paper, we limit ourselves to problems with the following Kronecker product structure:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad \text{with} \quad \mathbf{x} = \mathbf{v} \otimes \mathbf{u}, \tag{1}$$

in which $\mathbf{A} \in \mathbb{K}^{K \times Q}$, $\mathbf{x} \in \mathbb{K}^{Q}$, and $\mathbf{b} \in \mathbb{K}^{K}$. The solution $\mathbf{x}$ can be expressed as a Kronecker product $\mathbf{v} \otimes \mathbf{u}$ with $\mathbf{u} \in \mathbb{K}^{I}$ and $\mathbf{v} \in \mathbb{K}^{J}$ such that $Q = IJ$. More generally, $\mathbf{x}$ can be constrained by a Kronecker product of N non-zero vectors:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad \text{with} \quad \mathbf{x} = \bigotimes_{n=1}^{N} \mathbf{u}^{(n)},$$

in which $\mathbf{u}^{(n)} \in \mathbb{K}^{I_n}$ and $Q = I_1 I_2 \cdots I_N$. Importantly, a KPE is a special case of a linear system of equations with a tensor-decomposition constrained solution [24]. This type of problems could be solved by first solving the system without structure and subsequently computing a rank-1 model of the tensorized version of the solution. This approach works well if $\mathbf{A}$ has full column rank, but, in contrast to the methods in [24], fails when $\mathbf{A}$ is rank deficient or when $K < Q$, i.e., in the underdetermined case. The methods in [24] compute the least-squares (LS) solution of (1).

2. LARGE-SCALE SIMO AUTOREGRESSIVE SYSTEM IDENTIFICATION USING KPES

By exploiting the intrinsic structure or sparsity of a model, one can enable large-scale system identification. Here, we show that large-scale SIMO AR system identification can be reformulated as a particular type of KPE by exploiting the hypothesized Kronecker product structure of the coefficients. First, we define AR system identification and Kronecker constrained coefficients in Subsection 2.1 and Subsection 2.2, respectively. Next, we derive our KPE-based method for large-scale SIMO AR system identification in Subsection 2.3.

2.1. Autoregressive system identification

Consider a MIMO ARX model with Q outputs, P exogenous inputs, and system order L, that relates the outputs $y_q[k]$ using the following discrete difference equation:

$$\sum_{l=0}^{L} \sum_{q=1}^{Q} g_{pq}[l]\, y_q[k-l] = x_p[k] + n_p[k] \quad \text{for} \quad 1 \leq k \leq K \tag{2}$$

The AR coefficients are given by $g_{pq}[l]$ for $0 \leq l \leq L$, the pth exogenous input is denoted by $x_p[k]$ and the additive *white* noise is given by $n_p[k]$. Assuming we have $K + L$ samples, the model in (2) can be expressed in matrix form as follows:

$$\sum_{l=0}^{L} \mathbf{G}^{(l)} \mathbf{Y}^{(l)} = \mathbf{X} + \mathbf{N} \tag{3}$$

with $\mathbf{G}^{(l)}$ the lth $(P \times Q)$ coefficient matrix and $\mathbf{Y}^{(l)}$ the lth $(Q \times K)$ output matrix, which are defined element-wise as $g_{pq}^{(l)} = g_{pq}[l]$ and $y_{qk}^{(l)} = y_q[k-l]$, respectively, for $0 \leq l \leq L$. The input and noise matrix $\mathbf{X}$ and $\mathbf{N}$ both have dimensions $(P \times K)$. Note that one typically assumes $P = Q$ when considering the MIMO case; see, e.g., [1, 21] and references therein. The formulation in (3), however, is more general because we allow that $P \neq Q$. In this paper, we limit ourselves to *single-input multiple-output* systems, i.e., we have $P = 1$ and $Q > 1$. In that case, the ARX model in (3) reduces to:

$$\sum_{l=0}^{L} \mathbf{g}^{(l)^{\mathrm{T}}} \mathbf{Y}^{(l)} = \mathbf{x}^{\mathrm{T}} + \mathbf{n}^{\mathrm{T}}$$

with coefficients $\mathbf{g}^{(l)} \in \mathbb{K}^{Q}$ and input and noise $\mathbf{x}, \mathbf{n} \in \mathbb{K}^{K}$. The noise is omitted in the derivation of our method for notational convenience, but its influence is examined in Section 3.

2.2. Kronecker constrained system coefficients

Large-scale data can often be compactly modeled because of some intrinsic structure or sparsity of the data. In this paper, we take a similar approach as in [6, 7]: we assume the (large-scale) AR coefficients admit, or, can be well approximated by, a Kronecker product of N non-zero vectors, enabling a possibly very compact representation for large N. Consider the following Kronecker product structure for $\mathbf{g}^{(l)} \in \mathbb{K}^{Q}$:

$$\mathbf{g}^{(l)} = \mathbf{b}^{(l)} \otimes \mathbf{a}^{(l)}, \quad \text{for} \quad 0 \leq l \leq L, \tag{4}$$

with non-zero vectors $\mathbf{a}^{(l)} \in \mathbb{K}^{I^{(l)}}$ and $\mathbf{b}^{(l)} \in \mathbb{K}^{J^{(l)}}$ such that $Q = I^{(l)}J^{(l)}$, for $0 \leq l \leq L$. Clearly, this approach allows for a compact representation of the coefficients: we need only $(I^{(l)} + J^{(l)} - 1)$ values instead of $Q = I^{(l)}J^{(l)}$ to represent $\mathbf{g}^{(l)}$. Interestingly, constraint (4) corresponds to a rank-1 assumption on a matricized version of $\mathbf{g}^{(l)}$, i.e., we have: $\text{mat}\left(\mathbf{g}^{(l)}\right) = \mathbf{a}^{(l)^{\mathrm{T}}}\mathbf{b}^{(l)} = \mathbf{a}^{(l)} \otimes \mathbf{b}^{(l)}$. More generally, one can consider a Kronecker product of N non-zero vectors:

$$\mathbf{g}^{(l)} = \bigotimes_{n=1}^{N} \mathbf{u}^{(n,l)}, \quad \text{for} \quad 0 \leq l \leq L, \tag{5}$$

with $\mathbf{u}^{(n,l)} \in \mathbb{K}^{I_n^{(l)}}$ such that $Q = \prod_{n=1}^{N} I_n^{(l)}$, for $0 \leq l \leq L$. Increasing N, enables even more compact representations because we need only $\sum_{n=1}^{N} I_n^{(l)} - N + 1$ values instead of $Q = \prod_{n=1}^{N} I_n^{(l)}$ to represent $\mathbf{g}^{(l)}$. For example, if $I_n^{(l)} = I$ for $1 \leq n \leq N$ and $0 \leq l \leq L$, the number of unknown variables reduces from $\mathcal{O}(LI^N)$ to $\mathcal{O}(LNI)$. For $N > 2$, constraint (5) corresponds to a rank-1 assumption on a tensorized version of $\mathbf{g}^{(l)}$, i.e., we have: $\text{ten}\left(\mathbf{g}^{(l)}\right) = \mathbf{u}^{(1,l)} \otimes \mathbf{u}^{(2,l)} \otimes \cdots \otimes \mathbf{u}^{(N,l)}$. A detailed analysis on how to choose the dimensions of the vectors in the Kronecker product can be found in [6, 7].

2.3. Large-scale AR system identification as a KPE

By explicitly exploiting the Kronecker structure in the model, one can reformulate AR system identification as the computation of a (structured) KPE, allowing one to use optimization-based and algebraic solvers and formulate (generic) uniqueness conditions; see [24]. We illustrate this as follows.

Assuming the AR coefficients $\mathbf{g}^{(l)}$, for $0 \leq l \leq L$, can be modeled by a simple Kronecker product as in (4), we obtain:

$$\sum_{l=0}^{L} \left(\mathbf{b}^{(l)} \otimes \mathbf{a}^{(l)}\right)^{\mathrm{T}} \mathbf{Y}^{(l)} = \mathbf{x}^{\mathrm{T}}. \tag{6}$$

By taking the transpose, one can see that (6) reduces to:

$$\sum_{l=0}^{L} \mathbf{Y}^{(l)^{\mathrm{T}}} \left(\mathbf{b}^{(l)} \otimes \mathbf{a}^{(l)}\right) = \mathbf{x}. \tag{7}$$

For $L = 0$, this model reduces to a KPE of the form (1). For $L > 0$, the model in (7) is a straightforward generalization where the right-hand side equals a sum of $L+1$ matrix-times-Kronecker-product terms. More generally, one can consider a Kronecker product of N non-zero vectors as in (5), obtaining:

$$\sum_{l=0}^{L} \mathbf{Y}^{(l)^{\mathrm{T}}} \left(\bigotimes_{n=1}^{N} \mathbf{u}^{(n,l)}\right) = \mathbf{x},$$

which enables higher compression rates, as explained before. As such, large-scale AR system identification is reformulated as the computation of a particular KPE. Additionally, the matrix $\tilde{\mathbf{Y}} = \left[\mathbf{Y}^{(0)^{\mathrm{T}}}, \mathbf{Y}^{(1)^{\mathrm{T}}}, \cdots, \mathbf{Y}^{(L)^{\mathrm{T}}}\right] \in \mathbb{K}^{K \times (L+1)Q}$ has a block-Toeplitz structure due to the convolutive nature of the ARX model. This structure can be exploited to speed up KPE algorithms and relax uniqueness conditions; see [24].

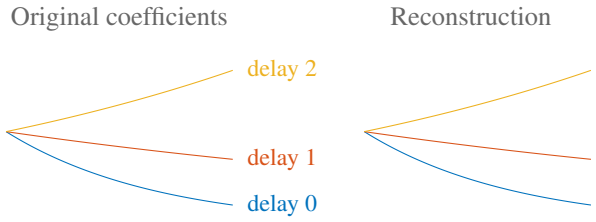

Fig. 2. By exploiting the rank-1 structure, the autoregressive coefficients are perfectly reconstructed (in the noiseless case).

3. EXPERIMENTS

We illustrate our method for various scenarios: 1) a proof-of-concept experiment in which we use exponentials as coefficient vectors, 2) an analysis of the influence of noise and sample size on the accuracy, and 3) an analysis of under- or overestimating the system order. For each experiment, we simulate a random SIMO ARX system by fixing both the coefficients and the outputs, and then constructing an input that satisfies the model in (3). We use i.i.d. zero-mean unit-variance Gaussian random outputs for each experiment and we further specify the particular coefficient definition in each experiment description. When considering the noisy case, we use Gaussian additive noise. We define the relative error $\epsilon_{\mathbf{A}}$ as the relative difference in Frobenius norm $\|\mathbf{A} - \hat{\mathbf{A}}\|_{\mathrm{F}}/\|\mathbf{A}\|_{\mathrm{F}}$. We use an adapted version of the non-linear LS algorithms with random initialization from [24] in order to solve KPEs.

3.1. Proof-of-concept experiment

Perfect reconstruction of the AR coefficients can be obtained in the noiseless case by exploiting the intrinsic rank-1 structure. We illustrate this for a large-scale SIMO ARX system of order $L = 2$ with $Q = 2500$ outputs and sample size $K = 600$. The coefficients $\mathbf{g}^{(l)}$ are defined as exponentials of length Q; more specifically, we have $g^{(0)}(\xi) = \frac{1}{10}\exp^{-3\xi/2}$, $g^{(1)}(\xi) = \frac{1}{10}(0.5)^{\xi/2}$, and $g^{(2)}(\xi) = \frac{1}{10}\exp^{(\xi/2)}$ uniformly sampled in $[0, 1]$. It is well-known that exponentials can be exactly represented by a rank-1 structure [6–8], validating the model in (7). We choose $N = 2$ and $I_1^{(l)} = I_2^{(l)} = I = 50$, for $0 \leq l \leq L$. Hence, we need only $(2I - 1) = 99$ values to model an AR coefficient vector instead of 2500, which amounts to a compression rate of $1 - \frac{I_{1+I_2-1}}{P} = 96.04\%$. The original coefficients and their reconstruction, up to machine precision, are shown in Figure 2.

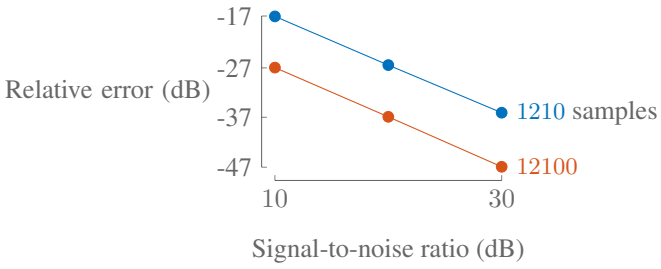

Fig. 3. While our method obtains accurate results with respect to the signal-to-noise ratio, increasing the number of samples further improves the accuracy of the coefficient estimates.

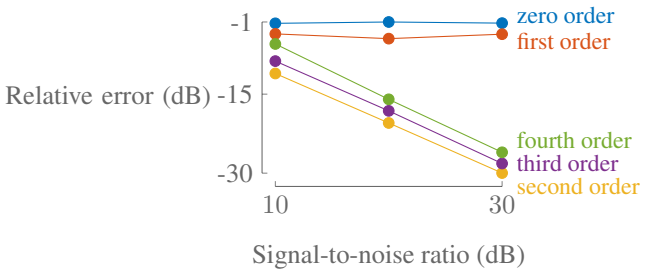

Fig. 4. While overestimation slightly reduces the accuracy, underestimating the system order fails to give accurate results.

3.2. Influence of noise and sample size on the accuracy

While increasing the sample size K improves the accuracy of the estimates, even a small number of samples can lead to accurate results. Also, the accuracy is quite high in comparison to the signal-to-noise ratio (SNR). We illustrate this for a large-scale SIMO ARX system of order $L = 5$ with $Q = 500$ outputs. We construct the rank-1 coefficient vectors as vectorized third-order rank-1 tensors using i.i.d. zero-mean unit-variance Gaussian random factor vectors. We use the following dimensions for the coefficient vectors: $(I_1^{(0)}, I_2^{(0)}, I_3^{(0)}) = (I_1^{(1)}, I_2^{(1)}, I_3^{(1)}) = (20, 5, 5)$, $(I_1^{(2)}, I_2^{(2)}, I_3^{(2)}) = (I_1^{(3)}, I_2^{(3)}, I_3^{(3)}) = (25, 5, 4)$, and $(I_1^{(4)}, I_2^{(4)}, I_3^{(4)}) = (I_1^{(5)}, I_2^{(5)}, I_3^{(5)}) = (50, 5, 2)$. We choose $K = 1210$ and 12100, which is equal to five and fifty times the number of unknown coefficients (242 values), and use an SNR equal to 10, 20, or 30 dB. The median results across fifty random experiments are illustrated in Figure 3.

3.3. Influence of the system order on the accuracy

Although the accuracy of the estimates slightly reduces, overestimating the system order L is not so critical. However, underestimating the order fails to give accurate results. We illustrate this for a large-scale SIMO ARX system of order $L = 2$ with $Q = 100$ outputs. We construct the rank-1 coefficient vectors as vectorized rank-1 matrices using i.i.d. zero-mean unit-variance Gaussian random factor vectors. We use the following dimensions for the coefficient vectors: $I_1^{(l)} = 20$ and $I_2^{(l)} = 5$, for $0 \leq l \leq L$. We choose $K = 50$, which is equal to twice the number of unknown coefficients. We use an SNR of 10, 20, or 30 dB. While estimating the coefficients, we vary the system order between zero and four. The median results across fifty random experiments are shown in Figure 4.

4. CONCLUSION AND FUTURE WORK

We have presented a method for AR system identification that enables large-scale applications by explicitly exploiting the hypothesized structure/sparsity of the system coefficients. In this paper, we have shown that the identification problem can be reformulated as the computation of a KPE, allowing one to use optimization-based solvers. Numerical experiments have shown that our method performs well in noisy conditions and that over-estimation of the system order is not so critical.

In a follow-up paper, we will address 1) the multiple-input multiple-output case, 2) the explicit exploitation of the block-Toeplitz structure in the computation and the uniqueness conditions, and 3) sum-of-Kronecker-products constrained coefficients. The latter means that we approximate a matricized or tensorized version of the coefficient vectors by a low-rank model instead of rank-1 model as explained in this paper.

5. REFERENCES

[1] L. Ljung, *System identification: Theory for the user*, Prentice hall, second edition, 1999.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[3] A. Holobar and D. Farina, "Blind source identification from the multichannel surface electromyogram," *Physiological Measurement*, vol. 35, no. 7, pp. R143, 2014.

[4] A. Bertrand, "Distributed signal processing for wireless EEG sensor networks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 6, pp. 923–935, Dec. 2015.

[5] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, SIAM, 1997.

[6] M. Boussé, O. Debals, and L. De Lathauwer, "A tensor-based method for large-scale blind source separation using segmentation," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 346–358, Jan. 2017.

[7] M. Boussé, O. Debals, and L. De Lathauwer, "Tensor-based large-scale blind system identification using seg-

mentation," *IEEE Transactions on Signal Processing*, vol. 65, no. 21, pp. 5770–5784, Nov. 2017.

[8] 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.

[9] O. Debals, M. Van Barel, and L. De Lathauwer, "Löwner-based blind signal separation of rational functions with applications," *IEEE Transactions on Signal Processing*, vol. 64, no. 8, pp. 1909–1918, 2016.

[10] L. Grasedyck, "Polynomial approximation in hierarchical Tucker format by vector-tensorization," Apr. 2010, Preprint 43, DFG/SPP1324, RWTH Aachen.

[11] 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.

[12] B. N. Khoromskij, "$\mathcal{O}(d\log N)$-quantics approximation of N-d tensors in high-dimensional numerical modeling," *Constructive Approximation*, vol. 34, no. 2, pp. 257–280, Oct. 2011.

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

[14] A. Cichocki, D. P. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. F. Caiafa, and A.-H. 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, Mar. 2015.

[15] N. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E.E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3551–3582, July 2017.

[16] W. Hackbusch, *Tensor spaces and numerical tensor calculus*, vol. 42, Springer, 2012.

[17] B. Sinquin and M. Verhaegen, "Quarks: Identification of large-scale Kronecker vector-autoregressive models," *IEEE Transactions on Automatic Control*, June 2018.

[18] B. Sinquin and M. Verhaegen, "K4SID: Large-scale subspace identification with Kronecker modeling," *IEEE Transactions on Automatic Control*, May 2018.

[19] L. N. Ribeiro, A. L. F. de Almeida, and J. C. M. Mota, "Identification of separable systems using trilinear filtering," in *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2015, Cancun, Mexico)*, Dec. 2015, pp. 189–192.

[20] C. Paleologu, J. Benesty, and S. Ciochina, "Linear system identification based on a kronecker product decomposition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, May 2018.

[21] H. Lütkepohl, *New introduction to multiple time series analysis*, Springer, 2005.

[22] M. Boussé, N. Vervliet, O. Debals, and L. De Lathauwer, "Face recognition as a Kronecker product equation," in *IEEE 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2017, Curaçao, Dutch Antilles)*, Dec. 2017, pp. 276–280.

[23] M. Boussé, G. Goovaerts, N. Vervliet, O. Debals, S. Van Huffel, and L. De Lathauwer, "Irregular heartbeat classification using Kronecker product equations," in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2017, Jeju Island, South-Korea)*, July 2017, pp. 438–441.

[24] M. Boussé, N. Vervliet, I. Domanov, O. Debals, and L. De Lathauwer, "Linear systems with a canonical polyadic decomposition constrained solution: Algorithms and applications," *Numerical Linear Algebra with Applications*, pp. 1–18, Aug. 2018.