



**HAL**  
open science

# High-Dimensional MVAR Model Identification Based on Structured Sparsity Penalization

Z. Fang, L. Albera, A. Kachenoura, Huazhong Shu, Y Kang, Régine Le Bouquin Jeannès

► **To cite this version:**

Z. Fang, L. Albera, A. Kachenoura, Huazhong Shu, Y Kang, et al.. High-Dimensional MVAR Model Identification Based on Structured Sparsity Penalization. IEEE Signal Processing Letters, 2024, 31, pp.1975-1979. 10.1109/lsp.2024.3418712 . hal-04646587

**HAL Id: hal-04646587**

**<https://hal.science/hal-04646587v1>**

Submitted on 10 Oct 2024

**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.

# High-Dimensional MVAR Model Identification Based on Structured Sparsity Penalization

Z. Fang, L. Albera, *Senior, IEEE*, A. Kachenoura, H. Shu, Y. Kang, and R. Le Bouquin Jeannès, *Member, IEEE*

**Abstract**—Multivariate autoregressive modeling is widely considered in neuroscience, especially when effective connectivity is concerned. In high-dimensional space, the conventional least-squares estimation of the autoregressive coefficients is no more consistent, hence the interest in regularizing the solution. Therefore, regularized approaches have been developed such as the Least Absolute Shrinkage Selection Operator (LASSO) promoting sparsity, which performs well provided that it is combined with the extended Bayesian Information Criterion (eBIC) to jointly estimate the MVAR model order and penalty parameter. Unfortunately, this need for eBIC requires much more computation time, making such approaches unsuitable for identifying high-dimensional MVAR models. The method proposed in this paper, named SOCAR (Simultaneous identification of the Order and the Coefficients of multivariate AutoRegressive models), estimates both the order and the coefficients of the model by combining mixed-norm regularization with classical sparse priors (e.g. LASSO). Weighted penalties are minimized in order to drive more easily sparsity to the right place. Experiments carried out on simulated signals show that SOCAR offers a good compromise between performance and numerical complexity.

**Index Terms**—Multivariate autoregressive model, model order estimation, high dimension, sparsity, LASSO, mixed-norm.

## I. INTRODUCTION

AUTOREGRESSIVE modeling is widely considered in various domains [1], [2]. Some applications deal with a large amount of data, requiring MultiVariate AutoRegressive (MVAR) analysis [3], [4]. For instance, the MVAR model has been explored in ElectroEncephaloGraphy (EEG) to derive measures of effective connectivity [5], [6]. This connectivity makes it possible to identify directed brain networks involved in healthy functioning and in pathologies such as Parkinson's disease.

The Least Squares (LS) approach is the most well-known method to estimate the MVAR coefficients [7]. Burg and Yule-Walker methods are also widely used [8], [9]. However, these classical algorithms fail in managing moderate or high dimensions [10]. Consequently, regularized LS approaches have been introduced such as the Least Absolute Shrinkage Selection Operator (LASSO) promoting sparsity [11], [12], and many of its variants [13], [14], [15], [16], [17]. Note that this sparse structure of the MVAR model is relevant in applications such as EEG analysis [18], [19], [20], [21]. Now,

Z. Fang, L. Albera, A. Kachenoura and R. Le Bouquin Jeannès are with Univ-Rennes, INSERM, LTSI – UMR 1099, F-35000 Rennes, France and with Centre de Recherche en Information Biomédicale Sino-français (CRIBs), Université de Rennes, INSERM, Southeast University, Rennes, France, Nanjing, China. H. Shu is with CRIBs and Southeast University, Nanjing, China. Z. Fang and Y. Kang are with Shenzhen Technology University, Shenzhen, China. (corresponding authors: Laurent Albera and Yan Kang, e-mails: laurent.albera@univ-rennes.fr and kangyan@sztu.edu.cn).

for all these methods, the estimation of the MVAR model order is crucial. Choosing MVAR order using model selection criterion is quite well-established in literature [22], [23], [24]. Unfortunately, since these criteria need to identify the MVAR coefficients for a suitably refined set of candidate orders, a regularized strategy provides an unfeasible solution in terms of computational cost for high dimensions. Hence the need for a coupled approach estimating simultaneously the order and the coefficients of the MVAR model in order to reduce the numerical cost.

In this paper, we propose such a coupled approach, called SOCAR (Simultaneous identification of the Order and the Coefficients of multivariate AutoRegressive models), which promotes group sparsity in a different way from that generally used to infer sparse causal networks modeled by MVAR processes [25]. Indeed, the introduction of the mixed-norm [26] allows for an efficient estimation of the model order. Note that weights are inserted in order to drive more easily sparsity to the right place like in [17]. In addition, imposing sparsity of the MVAR model coefficients by means of the  $L_1$  norm [27] illustrates the possibility of combining our model order estimation strategy with classical regularized techniques assuming structural constraints [17]. An iterative optimization procedure based on the Alternating Direction Method of Multipliers (ADMM) [28] is implemented avoiding the calculus of sub-gradients [29], [30].

Section II formulates the MVAR estimation problem and describes the SOCAR algorithm. Our method is applied to simulated, and compared with LS-like techniques: results are discussed in section III. A conclusion and some perspectives are provided in section IV.

## II. METHODOLOGY

### A. Problem formulation and classical approaches

Let  $\{\mathbf{x}(n)\}$  be a realization of an  $N$ -sample MVAR sequence of order  $p$ :

$$\mathbf{x}(n) = \sum_{\ell=1}^p \mathbf{A}_\ell^* \mathbf{x}(n-\ell) + \boldsymbol{\varepsilon}(n) \quad (1)$$

where  $\{\boldsymbol{\varepsilon}(n)\}$  is a realization of an  $M$ -dimensional white Gaussian noise sequence with zero mean and covariance matrix  $\sigma^2 \mathbf{I}$  and where  $\mathbf{A}_\ell^*$  is the  $\ell$ -th submatrix of the  $(M \times pM)$  coefficient matrix  $\mathbf{A}^* = [\mathbf{A}_1^*, \mathbf{A}_2^*, \dots, \mathbf{A}_p^*]$ . We assume in the sequel that the  $p$  matrices  $\mathbf{A}_\ell^*$  ( $1 \leq \ell \leq p$ ) are linearly independent and that the order  $p$  is lower than or equal to  $M^2$ . The matrix  $\mathbf{A}^*$  can be estimated using the LS method [7], which minimizes the Frobenius norm  $\|\mathbf{X} - \mathbf{A}\mathbf{B}\|_{\mathbb{F}}^2$  with respect to

the  $\mathbf{A}$  variable where  $\mathbf{X} = [\mathbf{x}(p+1), \mathbf{x}(p+2), \dots, \mathbf{x}(N)]$  and

$$\mathbf{B} = \begin{bmatrix} \mathbf{x}(p) & \mathbf{x}(p+1) & \cdots & \mathbf{x}(N-1) \\ \mathbf{x}(p-1) & \mathbf{x}(p) & \cdots & \mathbf{x}(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}(1) & \mathbf{x}(2) & \cdots & \mathbf{x}(N-p) \end{bmatrix}$$

provided that we have  $N \geq pM$ . Then, the LS solution is given by  $\mathbf{A}^* = \mathbf{X}\mathbf{B}^T(\mathbf{B}\mathbf{B}^T)^{-1}$  [31]. As mentioned above, the sparse regularized version of the LS method, LASSO, has been proposed [13]. It consists in minimizing the following objective function:

$$f(\mathbf{A}) = \|\mathbf{X} - \mathbf{A}\mathbf{B}\|_{\text{F}}^2 + \lambda\|\mathbf{A}\|_1 \quad (2)$$

where  $\lambda$  is a hyperparameter balancing between the data fitting term and the penalty term. Note that the  $L_1$  norm penalty term can be replaced with another term based on the  $L_{2,q}$  norm, which promotes a common zero among all MVAR-lag parameters when Granger Causality has to be measured [17]. In practice, the MVAR model order  $p$  is unknown and needs to be estimated.

### B. The SOCAR algorithm

We first introduce the  $(M^2 \times p)$  matrix  $\mathcal{A}^*$  given by  $\mathcal{A}^* = (\text{vec}(\mathbf{A}_1^*), \text{vec}(\mathbf{A}_2^*), \dots, \text{vec}(\mathbf{A}_p^*))$ . Note that the  $\text{vec}$  operator consists in vectorizing every  $(M \times M)$  matrix  $\mathbf{A}_\ell^*$  into an  $M^2$ -dimensional column vector obtained by stacking the columns of the matrix  $\mathbf{A}_\ell^*$  one below the other. Under the assumption that the  $p$  matrices  $\mathbf{A}_\ell^*$  ( $1 \leq \ell \leq p$ ) are linearly independent with  $p \leq M^2$ , the  $\mathcal{A}^*$  matrix is full column rank. So, estimating the order  $p$  of the MVAR model is equivalent to estimating the rank of the  $\mathcal{A}^*$  matrix. Then we can modify the LASSO objective function (2) by adding an additional term based on the nuclear norm, classically used to estimate the rank of a matrix. A link between the  $L_{1,2}$  mixed-norm and the nuclear norm can be derived from [32, proposition 1], showing that the  $L_{1,2}$  mixed-norm is a better candidate for the role of convex envelope of the rank when matrices are full column rank. Consequently, a solution aiming at estimating simultaneously the order and the coefficients of the MVAR model can be obtained by minimizing the following objective function:

$$g(\mathbf{A}) = \frac{1}{2}\|\mathbf{X} - \mathbf{A}\mathbf{B}\|_{\text{F}}^2 + \lambda_1\|\mathbf{A}\|_1 + \lambda_2\|\mathcal{A}\|_{1,2} \quad (3)$$

This function can be reformulated as follows:

$$g(\mathbf{A}) = \frac{1}{2}\|\mathbf{X} - \sum_{\ell=1}^q \mathbf{A}_\ell \mathbf{B}_\ell\|_{\text{F}}^2 + \lambda_1\|\mathbf{A}\|_1 + \lambda_2 \sum_{\ell=1}^q \|\mathbf{A}_\ell\|_{\text{F}} \quad (4)$$

where  $q$  is chosen sufficiently high to overestimate  $p$ . It is noteworthy that  $\mathbf{A}_\ell$  (resp.  $\mathbf{B}_\ell$ ) is the  $\ell$ -th matrix block of  $\mathbf{A}$  (resp.  $\mathbf{B}$ ), such as  $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_q]$  (resp.  $\mathbf{B} = [\mathbf{B}_1^T, \mathbf{B}_2^T, \dots, \mathbf{B}_q^T]^T$ ). Note also that using the Frobenius norm of each matrix block  $\mathbf{A}_\ell$  allows us to promote group sparsity within the whole matrix. According to equation (1), the  $q-p$  last matrix blocks  $\mathbf{A}_\ell$  of  $\mathbf{A}$  should be zero. This can be obtained using a suitable penalty parameter  $\lambda_2$ , which can

be difficult to get. Another solution proposed in [17] consists in using weights in order to drive more easily sparsity to the right place. In our case, an increasing sequence  $\{\alpha_\ell\}$  of weights should be chosen by the user in order to promote more sparsity in the last matrix blocks  $\mathbf{A}_\ell$  of  $\mathbf{A}$ , leading to the following cost function:

$$h(\mathbf{A}) = \frac{1}{2}\|\mathbf{X} - \sum_{\ell=1}^q \mathbf{A}_\ell \mathbf{B}_\ell\|_{\text{F}}^2 + \lambda_1\|\mathbf{A}\|_1 + \lambda_2 \sum_{\ell=1}^q \alpha_\ell \|\mathbf{A}_\ell\|_{\text{F}} \quad (5)$$

Now let's consider the strategy adopted to minimize equation (5), leading to the SOCAR method. The  $L_1$  and Frobenius norms are non-smooth and cannot be solved in a closed form. This would require the use of iterative proximal algorithms such as the ADMM technique, which is fast and accurate for convex optimization problems of large dimensions [28]. However, the ADMM convergence is not guaranteed when we split the objective function in more than two terms. Consequently, we replace the minimization of (5) with that of the following cost function:

$$\tilde{h}(\mathbf{A}) = \frac{1}{2}\|\mathbf{X} - \sum_{\ell=1}^q \mathbf{A}_\ell \mathbf{B}_\ell\|_{\text{F}}^2 + \lambda_2 \sum_{\ell=1}^q \beta_\ell \|\mathbf{A}_\ell\|_{\text{F}}^2 + \lambda_1\|\mathbf{A}\|_1 \quad (6)$$

with  $\beta_\ell = \alpha_\ell / \|\mathbf{A}_\ell\|_{\text{F}}$ . By assuming first that  $\beta_\ell$  does not depend on  $\mathbf{A}$ , the convex objective function (6) appears now as the sum of a differentiable term and a non-differentiable one, which can be minimized by ADMM. Next, by updating  $\beta_\ell$  from the last estimate of  $\mathbf{A}$ , it allows us to take into account their dependence. Such an iterative procedure has been proven to converge to minimize mixed norms in order to avoid the calculus of sub-gradients [29], [30], which is more complex and might not ensure the ADMM convergence as explained above. More particularly, in order to achieve an estimate of  $\mathbf{A}_\ell$  using the ADMM method, we define the following Lagrangian function:

$$\begin{aligned} \mathcal{L}(\mathbf{A}, \mathbf{C}, \mathbf{V}) &= \frac{1}{2}\|\mathbf{X} - \sum_{\ell=1}^q \mathbf{A}_\ell \mathbf{B}_\ell\|_{\text{F}}^2 + \sum_{\ell=1}^q (\lambda_1\|\mathbf{C}_\ell\|_1 \\ &+ \lambda_2\beta_\ell \|\mathbf{A}_\ell\|_{\text{F}}^2 + \frac{\rho}{2}\|\mathbf{C}_\ell - \mathbf{A}_\ell\|_{\text{F}}^2 + \langle \mathbf{V}_\ell, \mathbf{A}_\ell - \mathbf{C}_\ell \rangle) \end{aligned} \quad (7)$$

where  $\langle \mathbf{G}, \mathbf{H} \rangle = \text{trace}(\mathbf{G}^T \mathbf{H})$  denotes the inner product,  $\mathbf{V}_\ell$  is a Lagrangian multiplier, and  $\mathbf{C} = [\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_q]$ . Then, following the optimization strategy used in [29], [30], we derive  $\mathcal{L}$  with respect to each matrix  $\mathbf{A}_\ell$  by considering first that  $\beta_\ell$  is constant with respect to  $\mathbf{A}_\ell$ :

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{A}, \mathbf{C}, \mathbf{V})}{\partial \mathbf{A}_\ell} &= \sum_{i \neq \ell}^q \mathbf{A}_i \mathbf{B}_i \mathbf{B}_\ell^T - \mathbf{X} \mathbf{B}_\ell^T + 2\lambda_2\beta_\ell \mathbf{A}_\ell \\ &+ \mathbf{V}_\ell + \mathbf{A}_\ell \mathbf{B}_\ell \mathbf{B}_\ell^T + \rho(\mathbf{A}_\ell - \mathbf{C}_\ell) \end{aligned} \quad (8)$$

From this derivative, we get an update rule for each  $\mathbf{A}_\ell$ :

$$\begin{aligned} \mathbf{A}_\ell^{(k+1)} &= (\mathbf{X} \mathbf{B}_\ell^T + \rho \mathbf{C}_\ell - \mathbf{V}_\ell - \sum_{i \neq \ell}^q \mathbf{A}_i^{(j)} \mathbf{B}_i \mathbf{B}_\ell^T) \\ &\times (\mathbf{B}_\ell \mathbf{B}_\ell^T + (2\lambda_2\beta_\ell + \rho)\mathbf{I})^{-1} \end{aligned} \quad (9)$$

with  $j = k + 1$  if  $i < \ell$  and  $j = k$  if  $i > \ell$ . Now, remembering that  $\beta_\ell$  depends on  $\mathbf{A}_\ell$ , we use the following update rule:

$$\beta_\ell^{(k+1)} = \alpha_\ell / \|\mathbf{A}_\ell^{(k+1)}\|_F \quad (10)$$

Regarding the update rule of  $\mathbf{C}$ , it is derived from the minimization of the Lagrangian function  $\mathcal{L}$  (7) with respect to  $\mathbf{C}$ , by solving the following problem:

$$\min_{\mathbf{C}} \frac{2\lambda_1}{\rho} \|\mathbf{C}\|_1 + \left\| \frac{1}{\rho} \mathbf{V} + (\mathbf{A} - \mathbf{C}) \right\|_F^2 \quad (11)$$

The solution of (11) is thus derived using the proximal operator of the  $L_1$  norm [33], which is the soft thresholding operator. The update of  $\mathbf{C}$  is then given by:

$$\mathbf{C}^{(k+1)} = \mathbf{S}_{\lambda_1/\rho}(\mathbf{A}^{(k+1)} + \mathbf{V}^{(k)}/\rho) \quad (12)$$

where  $\mathbf{S}$  is the soft-thresholding operator [33] defined by:

$$(\mathbf{S}_\eta(\mathbf{A}))_{i_1, i_2} = \begin{cases} A_{i_1, i_2} - \eta & \text{if } A_{i_1, i_2} > \eta \\ 0 & \text{if } |A_{i_1, i_2}| \leq \eta \\ A_{i_1, i_2} + \eta & \text{if } A_{i_1, i_2} < -\eta \end{cases} \quad (13)$$

The update of the Lagrangian multiplier is given by:

$$\mathbf{V}^{(k+1)} = \mathbf{V}^{(k)} + \rho(\mathbf{A}^{(k+1)} - \mathbf{C}^{(k+1)}) \quad (14)$$

All these variables are alternately updated until convergence. More particularly, the SOCAR algorithm stops when the primal and dual residuals, given by  $\mathbf{r}^{(k+1)} = \mathbf{A}^{(k+1)} - \mathbf{C}^{(k+1)}$  and  $\mathbf{s}^{(k+1)} = -\rho(\mathbf{C}^{(k+1)} - \mathbf{C}^{(k)})$ , respectively, are lower than  $\epsilon_{pri}^{(k+1)}$  and  $\epsilon_{dua}^{(k+1)}$ , respectively, with:

$$\begin{aligned} \epsilon_{pri}^{(k+1)} &= \sqrt{N} \epsilon_{abs} + \epsilon_{rel} \max\{\|\mathbf{A}^{(k+1)}\|_F, \|\mathbf{C}^{(k+1)}\|_F\} \\ \epsilon_{dua}^{(k+1)} &= \sqrt{N} \epsilon_{abs} + \epsilon_{rel} \|\mathbf{V}^{(k+1)}\|_F \end{aligned}$$

where  $\epsilon_{abs}$  and  $\epsilon_{rel}$  are absolute and relative tolerances, respectively, and  $N$  is the number of time samples of the MVAR sequence  $\{\mathbf{x}(n)\}$ . If we call  $\mathbf{A}_\ell^{(c)}$  the estimate of  $\mathbf{A}_\ell^*$  after convergence, the estimate  $p^{(c)}$  of the order  $p$  is got by counting the first non-zero columns of  $\mathcal{A}^{(c)} = (\text{vec}(\mathbf{A}_1^{(c)}), \text{vec}(\mathbf{A}_2^{(c)}), \dots, \text{vec}(\mathbf{A}_q^{(c)}))$ . The different steps of SOCAR are summarized in Algorithm 1, which gives the estimate  $\mathbf{A}^{(c)}$  of  $\mathbf{A}^*$  as output. Parameters  $\lambda_1$  and  $\lambda_2$  are set experimentally. The penalty parameter  $\rho$  is fixed to 1. In practice, we propose to run SOCAR a second time based on the previously estimated values ( $\mathbf{A}^{(c)}$  and  $p^{(c)}$ ), using a small  $\lambda_2$  parameter in order to refine the solution. It is noteworthy that, as in [17], the convergence analysis of the proposed algorithm is left as an open question, while the ADMM-based optimization approach always converged in the experiments.

### III. RESULTS

In this section, we are interested in assessing the performance of the SOCAR method and comparing it with classical approaches such as LS [7], Yule-Walker [34], Burg [35] and LASSO [12]. Regarding LS, Yule-Walker and Burg, the MVAR model order is estimated using the eBIC criterion [24]. It is noteworthy that the latter was also used to estimate both the MVAR model order and the penalty parameter of regularized cost functions [17] and thus implemented in the LASSO method [12].

#### Algorithm 1 SOCAR

---

**Input:**  $\mathbf{X}, \mathbf{B}, \{\alpha_\ell\}_{1 \leq \ell \leq q}, \mathbf{A}^{(0)}, \mathbf{C}^{(0)}, \mathbf{V}^{(0)}, \rho, \lambda_1$  and  $\lambda_2$

- 1:  $\beta_\ell^{(0)} \leftarrow \alpha_\ell / \|\mathbf{A}_\ell^{(0)}\|_F$
- 2: **repeat**
- 3:    $k \leftarrow k + 1$
- 4:   **for**  $\ell = 1$  to  $q$  **do**
- 5:     Update  $\mathbf{A}_\ell^{(k+1)}$  using (9)
- 6:   **end for**
- 7:   Build  $\mathbf{A}^{(k+1)}$  from the  $q$  matrices  $\mathbf{A}_\ell^{(k+1)}$
- 8:   Update  $\beta_\ell^{(k+1)}$  using (10)
- 9:   Update  $\mathbf{C}^{(k+1)}$  using (12)
- 10:   Update  $\mathbf{V}^{(k+1)}$  using (14)
- 11: **until**  $\|\mathbf{r}^{(k+1)}\|_F \leq \epsilon_{pri}^{(k+1)}$  and  $\|\mathbf{s}^{(k+1)}\|_F \leq \epsilon_{dua}^{(k+1)}$
- 12:  $p^{(c)}$  given by counting the first non-zero columns of  $\mathcal{A}^{(c)}$

**Output:** MVAR coefficient matrix  $\mathbf{C}^{(k+1)}$

---

#### A. Performance criteria

In order to evaluate the robustness and accuracy of SOCAR, three performance criteria are considered. Firstly, we compare the estimated order with the true one. Secondly, the degree of similarity is computed as follows:

$$S(\mathbf{A}^*, \mathbf{A}^{(c)}) = \frac{\sum_i \sum_j (A_{i,j}^* - \bar{A}^*)(A_{i,j}^{(c)} - \bar{A}^{(c)})}{\sqrt{(\sum_i \sum_j (A_{i,j}^* - \bar{A}^*)^2)(\sum_i \sum_j (A_{i,j}^{(c)} - \bar{A}^{(c)})^2)}} \quad (15)$$

where  $\bar{K}$  denotes the mean of the components of matrix  $\mathbf{K}$ . Thirdly, the computation time is estimated for the five methods.

We are also interested in comparing the true Power Spectral Densities (PSDs) with the estimated ones as explained afterwards. We calculate the true spectral matrix as follows:

$$\mathbf{P}(\nu) = \mathbf{H}(\nu) \boldsymbol{\Sigma}_x \mathbf{H}(\nu)^\dagger \quad (16)$$

where  $\boldsymbol{\Sigma}_x$  is the true residual covariance matrix of the MVAR model and  $\mathbf{H}(\nu)^\dagger$  is the Hermitian transpose of  $\mathbf{H}(\nu)$  with  $\mathbf{H}$  the transfer function of the system given by [31]:

$$\mathbf{H}(\nu) = (\mathbf{I} - \sum_{\ell=1}^p \mathbf{A}_\ell^* e^{-j2\pi\ell\nu})^{-1} \quad (17)$$

The true PSDs of the channels are computed from the diagonal components of the spectral matrices. The PSDs estimated from the five methods are obtained using (16) and (17) by replacing  $\mathbf{A}^*$  with its estimate and  $\boldsymbol{\Sigma}_x$  with the following estimated residual covariance matrix:

$$\hat{\boldsymbol{\Sigma}}_x = \frac{1}{N-p} \sum_{n=p+1}^N (\mathbf{x}(n) - \hat{\mathbf{x}}(n)) (\mathbf{x}(n) - \hat{\mathbf{x}}(n))^T \quad (18)$$

where  $\mathbf{x}(n)$  is given by (1) and  $\hat{\mathbf{x}}(n)$  is computed by replacing  $\mathbf{A}^*$  with  $\mathbf{A}^{(c)}$ .

#### B. Performance analysis

We generated 100-channel MVAR signals for two orders ( $p = 5$  and  $p = 15$ ) to test the robustness of our approach.

Method	$p = 5$		$p = 15$	
	$p^{(c)}$	$S$	$p^{(c)}$	$S$
SOCAR	5.00	0.994	15.00	0.997
LASSO	5.00	0.994	15.05	0.997
OLS	1.25	0.683	1.00	0.578
Yule-Walker	1.10	0.684	1.05	0.579
Burg	1.15	0.684	1.05	0.579

TABLE I: Quantitative estimation of the order and the coefficients for two MVAR models with 100 channels

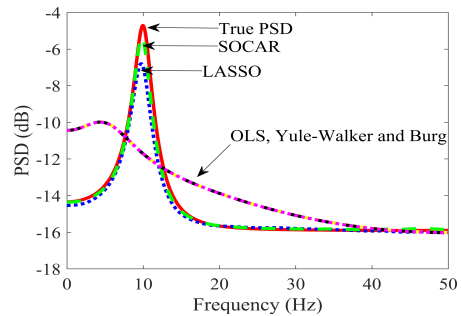
Method	$p = 5$	$p = 15$
SOCAR	17 mn	37 mn
LASSO	15 h	15 h
OLS	15 s	1 mn
Yule-Walker	16 s	1 mn
Burg	15 s	1 mn

TABLE II: Computation time to identify two MVAR models with 100 channels

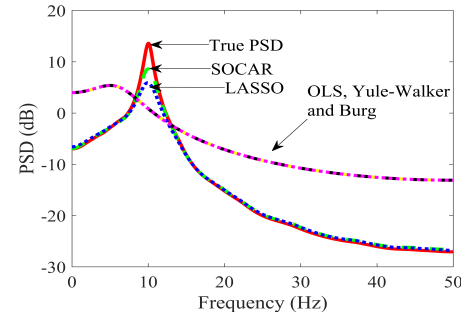
The sampling frequency was set to 1 kHz and signals were 10 seconds long. Sparsity was randomly applied on the coefficient matrices. The signals have been constructed in such a way that some of them exhibit oscillations in the [8 Hz - 12 Hz] band (well-known as the alpha band which is of interest in EEG). The stability of the model was verified before any experiment. The order and the degrees of similarity were estimated for the five aforementioned methods. Twenty independent Monte-Carlo simulations were carried out to get the results. Note that the degree of similarity was only computed from the signals exhibiting oscillations in the alpha band.

Quantitative results are given in Tables I and II. Unlike LS, Yule-Walker and Burg, LASSO and SOCAR are much more consistent, the estimated order being equal to the true one whatever its value ( $p = 5$  and  $p = 15$ ). Clearly, the LS, Yule-Walker and Burg methods underestimate the order. Regarding the degrees of similarity, LASSO and SOCAR are the only techniques which properly estimate the coefficient matrix of the MVAR model, allowing for an efficient reconstruction of the channel PSDs as shown in Table I. However, LASSO is much more time consuming than SOCAR, by a factor of 25 to 55 (see Table II). This is due to the introduction of eBIC to jointly estimate the MVAR model order and the LASSO penalty parameter. In the light of our experiments, this is the only guarantee of success for LASSO, which is very sensitive to the choice of its penalty parameter. Our approach, on the other hand, is much less sensitive to the choice of its penalty parameters, and therefore does not require the use of eBIC. Note that the poor behavior of the non regularized methods is due to the small  $k$ -ratio value of our scenario. Recall that the  $k$ -ratio is defined by  $k = N/Mp$  [36]. Such a  $k$ -ratio value is not large enough in our context [37] contrarily to some rules of thumb given for lower dimensional models [36].

The good behavior of SOCAR and LASSO in terms of estimation accuracy is illustrated in Figure 1. More particularly, we compare the PSDs obtained from the five methods with the true one. As we can see, the LS, Yule-Walker and Burg methods do not bring out the information present in the low frequencies. Regarding the LASSO and SOCAR techniques, the difference with the reference is negligible, resulting in



(a) For a model order  $p = 5$



(b) For a model order  $p = 15$

Fig. 1: PSDs estimated from the five methods

more consistent and reliable estimates.

#### IV. CONCLUSION

In this paper, we proposed a coupled approach, called SOCAR, estimating simultaneously the order and the coefficients of an MVAR model. SOCAR amounts to the combination of two penalization strategies, namely the LASSO technique and the minimization of the mixed-norm which aims at making our method robust with respect to an overestimation of the model order. The simulations showed a better trade-off between performance and numerical complexity of SOCAR compared to classical approaches such as LS, Yule-Walker, Burg and LASSO. LASSO succeeded in achieving high performance like SOCAR provided that it was combined with eBIC to jointly estimate the MVAR model order and penalty parameter. In fact, unlike SOCAR, an arbitrary choice of penalty parameter cannot guarantee LASSO such good results. Unfortunately, this need for eBIC requires much more computation time, making LASSO unsuitable for identifying high-dimensional MVAR models. In addition, the spectral information is well preserved by our new approach.

In a future work, we plan to estimate adaptively both hyperparameters  $\lambda_1$  and  $\lambda_2$  guided by the estimation technique proposed in [38], which strikes a balance between the data fitting term and the regularization terms. The idea would be to achieve a closed form to update the hyperparameters in each iteration, forcing the solution to satisfy Morozov's discrepancy principle. Besides, we aim to replace the LASSO regularization of SOCAR by other ones proposed in the literature by promoting more specific structural priors about the MVAR model.

REFERENCES

- [1] C.-S. Ouyang, R.-C. Yang, C.-T. Chiang, R.-C. Wu, and L.-C. Lin, "EEG autoregressive modeling analysis: A diagnostic tool for patients with epilepsy without epileptiform discharges," *Clinical Neurophysiology*, vol. 131, no. 8, pp. 1902–1908, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1388245720303497>
- [2] R. Annamma George and R. Periyasamy, "Modelling of electrocardiogram using autoregressive moving average model and linear predictive coefficient a comparative study," in *2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*, 2021, pp. 1–4.
- [3] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: journal of the Econometric Society*, pp. 424–438, 1969.
- [4] O. Paiss and G. F. Inbar, "Autoregressive modeling of surface emg and its spectrum with application to fatigue," *IEEE Transactions on Biomedical Engineering*, vol. BME-34, no. 10, pp. 761–770, 1987.
- [5] C. Yang, R. Le Bouquin Jeannès, G. Faucon, and H. Shu, "Extracting information on flow direction in multivariate time series," *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 251–254, 2011.
- [6] A. K. Abbas, G. Azemi, S. Amiri, S. Ravanshadi, and A. Omidvarnia, "Effective connectivity in brain networks estimated using eeg signals is altered in children with adhd," *Computers in Biology and Medicine*, vol. 134, p. 104515, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0010482521003097>
- [7] P. Jantana and P. Sudasna-na-Ayudhya, "Least squares and discounted least squares in autoregressive process," *Silpakorn University Open Journal Systems*, vol. 2, pp. 122–135, 2002.
- [8] R. A. Roberts and C. T. Mullis, *Digital signal processing*. Addison-Wesley Longman Publishing Co., Inc., 1987.
- [9] S. M. Kay, *Modern spectral estimation: theory and application*. Pearson Education India, 1988.
- [10] Y. Wang, C.-M. Ting, and H. Ombao, "Modeling effective connectivity in high-dimensional cortical source signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 7, pp. 1315–1325, 2016.
- [11] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.
- [12] P. A. Valdés-Sosa, J. M. Sánchez-Bornot, A. Lage-Castellanos, M. Vega-Hernández, J. Bosch-Bayard, L. Melie-García, and E. Canales-Rodríguez, "Estimating brain functional connectivity with sparse multivariate autoregression," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1457, pp. 969–981, 2005.
- [13] S. Haufe, K.-R. Müller, G. Nolte, and N. Krämer, "Sparse causal discovery in multivariate time series," in *causality: objectives and assessment*. PMLR, 2010, pp. 97–106.
- [14] A. C. Lozano, N. Abe, Y. Liu, and S. Rosset, "Grouped graphical granger modeling for gene expression regulatory networks discovery," *Bioinformatics*, vol. 25, no. 12, pp. i110–i118, 2009.
- [15] J. Songsiri, "Sparse autoregressive model estimation for learning granger causality in time series," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 3198–3202.
- [16] A. Shojaie and G. Michailidis, "Discovering graphical granger causality using the truncating lasso penalty," *Bioinformatics*, vol. 26, no. 18, pp. i517–i523, 2010.
- [17] P. Manomaisaowapak and J. Songsiri, "Joint learning of multiple granger causal networks via non-convex regularizations: Inference of group-level brain connectivity," *Neural Networks*, vol. 149, pp. 157–171, 2022.
- [18] T. R. Mullen, C. A. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, T.-P. Jung, and G. Cauwenberghs, "Real-time neuroimaging and cognitive monitoring using wearable dry eeg," *IEEE transactions on biomedical engineering*, vol. 62, no. 11, pp. 2553–2567, 2015.
- [19] Y. Antonacci, L. Astolfi, G. Nollo, and L. Faes, "Information transfer in linear multivariate processes assessed through penalized regression techniques: validation and application to physiological networks," *Entropy*, vol. 22, no. 7, p. 732, 2020.
- [20] M. Billinger, C. Brunner, and G. R. Müller-Putz, "Single-trial connectivity estimation for classification of motor imagery data," *Journal of neural engineering*, vol. 10, no. 4, p. 046006, 2013.
- [21] Y. Antonacci, L. Minati, L. Faes, R. Pernice, G. Nollo, J. Toppi, A. Pietrabissa, and L. Astolfi, "Estimation of granger causality through artificial neural networks: applications to physiological systems and chaotic electronic oscillators," *PeerJ Computer Science*, vol. 7, p. e429, 2021.
- [22] H. Akaike, "A new look at the statistical model identification," *IEEE transactions on automatic control*, vol. 19, no. 6, pp. 716–723, 1974.
- [23] G. Schwartz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, pp. 461–464, 1978.
- [24] J. Chen and Z. Chen, "Extended bayesian information criteria for model selection with large model spaces," *Biometrika*, vol. 95, no. 3, pp. 759–771, 2008.
- [25] A. Bolstad, B. D. Van Veen, and R. Nowak, "Causal network inference via group sparse regularization," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2628–2641, 2011.
- [26] X. Han, "Robust low-rank tensor approximations using group sparsity," Ph.D. dissertation, Rennes 1, 2019.
- [27] X. Han, L. Albera, A. Kachenoura, L. Senhadji, and H. Shu, "Low rank canonical polyadic decomposition of tensors based on group sparsity," in *2017 25th European Signal Processing Conference (EUSIPCO)*, 2017, pp. 668–672.
- [28] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "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.
- [29] D. Kong, J. Liu, B. Liu, and X. Bao, "Uncorrelated group LASSO," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1, 2016.
- [30] Z. Ma, F. Nie, Y. Yang, J. R. Uijlings, and N. Sebe, "Web image annotation via subspace-sparsity collaborated feature selection," *IEEE Transactions on Multimedia*, vol. 14, no. 4, pp. 1021–1030, 2012.
- [31] E. Parzen, "Autoregressive spectral estimation," in *Time Series in the Frequency Domain*, ser. Handbook of Statistics. Elsevier, 1983, vol. 3, pp. 221–247. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0169716183030138>
- [32] X. Shu, F. Porikli, and N. Ahuja, "Robust orthonormal subspace learning: Efficient recovery of corrupted low-rank matrices," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3874–3881.
- [33] N. Parikh and S. Boyd, "Foundations and trends in optimization," *Proximal algorithms*, vol. 1, no. 3, pp. 127–239, 2014.
- [34] G. Eshel, "The yule walker equations for the ar coefficients," *Internet resource*, vol. 2, pp. 68–73, 2003.
- [35] J. P. Burg, "A new analysis technique for time series data," *paper presented at Advanced Study Institute on signal Processing, NATO Enschede, Netherlands*, 1968.
- [36] Y. Antonacci, J. Toppi, A. Pietrabissa, A. Anzolin, and L. Astolfi, "Measuring connectivity in linear multivariate processes with penalized regression techniques," *IEEE Access*, 2024.
- [37] C. M. Endemann, B. M. Krause, K. V. Nourski, M. I. Banks, and B. Van Veen, "Multivariate autoregressive model estimation for high-dimensional intracranial electrophysiological data," *NeuroImage*, vol. 254, p. 119057, 2022.
- [38] C. He, C. Hu, W. Zhang, and B. Shi, "A fast adaptive parameter estimation for total variation image restoration," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 4954–4967, 2014.