

# A Robust Approach to ARMA Factor Modeling

Lucia Falconi, Augusto Ferrante, Mattia Zorzi

**Abstract**—This paper deals with the dynamic factor analysis problem for an ARMA process. To robustly estimate the number of factors, we construct a confidence region centered in a finite sample estimate of the underlying model which contains the true model with a prescribed probability. In this confidence region, the problem, formulated as a rank minimization of a suitable spectral density, is efficiently approximated via a trace norm convex relaxation. The latter is addressed by resorting to the Lagrange duality theory, which allows to prove the existence of solutions. Finally, a numerical algorithm to solve the dual problem is presented. The effectiveness of the proposed estimator is assessed through simulation studies both with synthetic and real data.

**Index Terms**—Convex optimization, duality theory, dynamic factor analysis, nuclear norm.

## I. INTRODUCTION

We deal with the problem of constructing a dynamical model from a high dimensional stream of data that are assumed to be noisy observations of a process depending on a small number of hidden variables. In the static case, this problem is known as *factor analysis*. Its origins can be traced back to the beginning of the last century and the amount of literature produced on this topic is impressive: we refer the readers to the recent papers [1], [2], [3], [4] for an overview of the literature and a rich list of references. The solution of factor analysis problems may be obtained by decomposing the covariance matrix of the observed data as the sum of a diagonal positive definite matrix (accounting for the noise covariance) and a positive semidefinite matrix whose rank must be as small as possible since it equals the number of hidden variables in the model. The main problem of this solution is that it is inherently fragile; in fact, even a minuscule variation in the covariance matrix of the observed data usually leads to a substantial variation of the number of hidden variables which is the key feature of the modelling procedure. On the other hand such a matrix must be estimated and is therefore subject to errors. To address this fragility issue, a robust method has been recently proposed, [5], for the static factor analysis problem. The approach has been generalized with good results also to the dynamic framework of learning AR latent variable dynamic graphical models, [6], in which the problem of finding a “sparse plus low-rank” decomposition of the inverse power spectral density of the underlying process is considered.

Dynamic Factor Analysis (DFA) has been addressed much more recently, the first contribution in this field being apparently [7]. We refer to the surveys [8], [9] and to the recent

paper [10] for an overview of the literature on this subject. In [11] an interesting generalization is applied to modelization of dynamical systems.

In this paper, we address the dynamic autoregressive moving average (ARMA) case with the aim of extracting, from the observed data, a model featuring a small number of hidden variables. This is important both from the point of view of the model simplicity and to uncover the structure of the mechanism generating the data. The problem may be mathematically formulated as that of decomposing the spectral density of the process generating the data as the sum of a diagonal spectral density and a low-rank one. The fragility issue in this case is even more severe. We address the problem as follows:

- Given the observed data, we compute by standard methods (e.g. truncated periodogram) a raw estimate  $\hat{\Phi}$  of the spectral density  $\Phi$  generating the data.
- We compute a neighbourhood  $\mathcal{N}$  of  $\hat{\Phi}$  that contains  $\Phi$  with prescribed probability; clearly the size of  $\mathcal{N}$  depends on the sample size.
- We compute a refined estimate  $\Phi^\circ \in \mathcal{N}$  by imposing that it admits an additive decomposition as a diagonal spectral density and a spectral density with the lowest possible rank. To this end we set up an optimization problem that we address by resorting to duality. In particular, we prove existence of solutions and provide a numerical algorithm to compute a solution.

Our work may be cast in the rich stream of literature devoted to learning dynamic models having a topological structure describing the presence or the absence of interactions among the variables of the systems; see the former works [12], [13], [14] as well as their extensions to reciprocal processes [15], [16], sparse plus low rank graphical models [17], [18], [6], the Bayesian viewpoint proposed in [19], [20] and the case of oriented graphical models [21], [22]. The common aspects of these papers are a decomposition of the type “sparse + low rank” of a certain spectral density and/or the fact that an underlying graphical model (possibly with latent variables) is considered where the presence of an edge of the graph depends on the conditional dependence between the random process and is therefore related to the inverse of the spectral density of the observed process. In our work, instead, we consider a “diagonal + low rank” decomposition and we do not consider the presence of an underlying graphical model.

The contribution of this paper is twofold; first, we propose a procedure to estimate the number of latent factors in dynamic ARMA factor models: this is the most delicate aspect of factor analysis problems; second, we derive an identification method to estimate the parameters of a factor model describing the observed data.

Dynamic factor models has been deeply investigated in the last two decades in the fields of econometrics and statistics. We

L. Falconi, A. Ferrante, and M. Zorzi are with the Department of Information Engineering, University of Padova, Padova, Italy; e-mail: lucia.falconi@phd.unipd.it (L. Falconi); augusto@dei.unipd.it (A. Ferrante); zorzimatt@dei.unipd.it (M. Zorzi).

refer to [23] for an overview of the contributions provided by such a community over the recent years and a rich list of references. The fundamental results in this literature is to show how principal component techniques can be used to consistently estimate the latent components from the observable variables as the cross sectional dimension  $m$  and the sample size  $N$  both tend to infinity [24], [25], [26], [27]. An alternative likelihood-based method has been shown in [28], where the factor model is cast in state-space form and the likelihood is maximized using the Expectation Maximization (EM) algorithm. Note that in all these methods a preliminary first step consists in identifying the number of common factors. Bai and Ng [29] provide a class of information criteria to consistently estimate the number of factors in static models while Hallin and Liška [30] deal with the dynamic counterpart of the problem. Thus, in the econometrics and statistics literature, the solution to the dynamical factor analysis problem consists of 2-steps: first the estimation of the number of factors and then identification of the parameters of the model. Here, we present an alternative optimization-based approach which simultaneously estimates the number of dynamic latent factors and the parameters of an ARMA factor model. We also remark that, whereas the econometric literature focuses the attention on factor models with weakly correlated idiosyncratic component and with a divergent number of observed variables, in this paper we consider fixed cross-sectional dimension factor models where the idiosyncratic terms are assumed to be cross-sectionally uncorrelated.

The outline of the paper is as follows. In Section II we introduce the DFA problem for moving average (MA) models. Section III shows that such a problem admits solution by means of duality theory, while Section IV shows how to reconstruct the solution of the primal problem from the dual one. In Section V we propose an algorithm to compute the solution of the dual problem. In Section VI we extend the previous ideas to ARMA models. Section VII presents some numerical results. Finally, in Section VIII we draw the conclusions.

### A. Notation

Given a matrix  $M$ , we denote its transpose by  $M^\top$  and by  $M_{(i,j)}$  the element of  $M$  in the  $i$ -th row and  $j$ -th column. If  $M$  is a square matrix,  $\text{tr}(M)$ ,  $|M|$  and  $\sigma(M)$  denote its trace, its determinant and its spectrum, respectively. The symbol  $\|\cdot\|$  stands for the Frobenius norm. For  $A, B \in \mathbb{R}^{m \times m}$ , we define their inner product as  $\langle A, B \rangle := \text{tr}(A^\top B)$ . Let  $\mathbf{Q}_m$  be the space of real symmetric matrices of size  $m$ ; if  $M \in \mathbf{Q}_m$  is positive definite or positive semidefinite, then we write  $M \succ 0$  or  $M \succeq 0$ , respectively. We denote by  $(\cdot)^*$  the complex conjugate transpose.  $\Phi(e^{i\vartheta})$  for  $\vartheta \in [-\pi, \pi]$  denotes a function defined on the unit circle  $\{e^{i\vartheta} : \vartheta \in [-\pi, \pi]\}$ , and the dependence on  $\vartheta$  is dropped if necessary. If  $\Phi(e^{i\vartheta})$  is positive (semi-)definite  $\forall \vartheta \in [-\pi, \pi]$  we write  $\Phi(e^{i\vartheta}) \succ 0$  ( $\succeq 0$ ). Integrals are always defined from  $-\pi$  to  $\pi$  with respect to the normalized Lebesgue measure  $d\vartheta/2\pi$ .

## II. IDENTIFICATION OF MA FACTOR MODELS

Consider the MA factor model whose order is  $n$ :

$$y(t) = W_L u(t) + W_D w(t) \quad (1)$$

where

$$W_L(e^{i\vartheta}) = \sum_{k=0}^n W_{L,k} e^{-i\vartheta k}, \quad W_D(e^{i\vartheta}) = \sum_{k=0}^n W_{D,k} e^{-i\vartheta k},$$

$W_{L,k} \in \mathbb{R}^{m \times r}$ ,  $W_{D,k} \in \mathbb{R}^{m \times m}$  diagonal;  $u = \{u(t), t \in \mathbb{Z}\}$  and  $w = \{w(t), t \in \mathbb{Z}\}$  are normalized white Gaussian noises of dimension  $r$  and  $m$ , respectively, such that  $\mathbb{E}[u(t)w(s)^\top] = 0 \forall t, s$ . The aforementioned model has the following interpretation:  $u$  is the process which describes the  $r$  factors, with  $r \ll m$ , not accessible to observation;  $W_L$  is the factor loading transfer matrix;  $W_L u(t)$  is the latent variable;  $W_D w(t)$  is idiosyncratic noise. Accordingly,  $y = \{y(t), t \in \mathbb{Z}\}$  is a  $m$ -dimensional Gaussian stationary stochastic process with power spectral density

$$\Phi = \Phi_L + \Phi_D \quad (2)$$

where  $\Phi_L = W_L W_L^* \succeq 0$  and  $\Phi_D = W_D W_D^* \succeq 0$  belong to the finite dimensional space:

$$\mathcal{Q}_{m,n} = \left\{ \sum_{k=-n}^n R_k e^{-i\vartheta k}, R_k = R_{-k}^T \in \mathbb{R}^{m \times m} \right\}.$$

By construction,  $\text{rank}(\Phi_L) = r$ , where rank denotes the normal rank (i.e. the rank almost everywhere), and  $\Phi_D$  is diagonal. Hence,  $y$  represents a factor model if its spectral density can be decomposed as ‘‘low rank + diagonal’’ as in (2).

Assume to collect a finite length realization of  $y$  defined in (1), say  $y^N = \{y(1) \dots y(N)\}$  where the order  $n$  is known. We want to estimate the corresponding factor model, that is the decomposition in (2) as well as the number of factors  $r$ . To this aim, given our data  $y^N$ , we first compute the sample covariance lags  $\hat{R}_j$  as

$$\hat{R}_j = \frac{1}{N} \sum_{t=0}^{N-j} y(t+j)y(t)^\top, \quad j = 0 \dots n.$$

Then, an estimate  $\hat{\Phi}$  of  $\Phi$  is obtained by the truncated periodogram:

$$\hat{\Phi} = \sum_{k=-n}^n \hat{R}_k e^{i\vartheta k}. \quad (3)$$

Notice that  $\hat{\Phi}$  could be not positive definite for all  $\vartheta$ ; in that case, we can add  $\varepsilon I_m$  to the right side of Equation (3), with the constant  $\varepsilon > 0$  chosen in such a way as to ensure the positivity of  $\hat{\Phi}$ . On the other hand,  $\hat{\Phi}$  may not admit a low rank plus diagonal decomposition. Thus, we estimate directly

the two terms  $\Phi_L$  and  $\Phi_D$  of the decomposition (2) by solving the following optimization problem:

$$\begin{aligned} & \min_{\Phi, \Phi_L, \Phi_D \in \mathcal{Q}_{m,n}} \text{tr} \int \Phi_L \\ & \text{subject to } \Phi_L + \Phi_D = \Phi, \\ & \Phi \succ 0 \text{ a.e.}, \Phi_L, \Phi_D \succeq 0, \\ & \Phi_D \text{ diagonal}, \\ & \mathcal{S}_{IS}(\Phi || \hat{\Phi}) \leq \delta. \end{aligned} \quad (4)$$

Here, the objective function promotes a solution for  $\Phi_L$  having low rank, see [17]. The first three constraints impose that  $\Phi_L$  and  $\Phi_D$  provide a genuine spectral density decomposition of type (2). The last constraint, in which  $\mathcal{S}_{IS}(\Phi || \hat{\Phi})$  is the Itakura-Saito divergence defined by

$$\mathcal{S}_{IS}(\Phi || \hat{\Phi}) = \int \log |\hat{\Phi} \Phi^{-1}| + \text{tr}[\hat{\Phi}^{-1} \Phi - I_m],$$

imposes that  $\Phi$  belongs to a set ‘‘centered’’ in the nominal spectral density  $\hat{\Phi}$  and with prescribed tolerance  $\delta$ . Notice that  $\Phi_D$  is uniquely determined by  $\Phi$  and  $\Phi_L$ . Thus, Problem (4) can be rewritten by removing  $\Phi_D$ :

$$\begin{aligned} (\Phi^\circ, \Phi_L^\circ) = & \arg \min_{\Phi, \Phi_L \in \mathcal{Q}_{m,n}} \text{tr} \int \Phi_L \\ & \text{subject to } \Phi \succ 0 \text{ a.e.}, \Phi_L, \Phi - \Phi_L \succeq 0, \\ & \Phi - \Phi_L \text{ diagonal}, \\ & \mathcal{S}_{IS}(\Phi || \hat{\Phi}) \leq \delta. \end{aligned} \quad (5)$$

#### A. The Choice of $\delta$

Before solving our problem, we deal with the choice of the tolerance parameter  $\delta$  appearing in the constraint of (5). This choice should reflect the accuracy of the estimate  $\hat{\Phi}$  of  $\Phi$ . This can be accomplished by choosing a desired probability  $\alpha \in (0, 1)$  and considering a ball of radius  $\delta_\alpha$  (in the Itakura-Saito topology) centered in  $\hat{\Phi}$  and containing the true spectrum  $\Phi$  with probability  $\alpha$ . The estimation of  $\delta_\alpha$  is not an easy task because we do not know the true power spectral density  $\Phi$ . Next, we propose a resampling-based method to estimate it. The idea is to approximate  $\Phi$  with  $\hat{\Phi}$ , and use this model to perform a resampling operation. Let

$$W(e^{i\vartheta}) = \sum_{k=0}^n W_k e^{-i\vartheta k}, \quad W_k \in \mathbb{R}^{m \times m}$$

be the minimum phase spectral factor of  $\hat{\Phi}$  and define the process  $\hat{y} = \{\hat{y}(t), t \in \mathbb{Z}\}$  as  $\hat{y}(t) := W(e^{i\vartheta})e(t)$ , where  $e(t)$  is an  $m$ -dimensional normalized white noise. The truncated periodogram (understood as *estimator*) based on a sample of the process  $\hat{y}$  of length  $N$  is

$$\hat{\Phi}_r(e^{i\vartheta}) = \sum_{k=-n}^n e^{-i\vartheta k} \frac{1}{N} \sum_{t=0}^{N-k} \hat{y}(t+k) \hat{y}(t)^T,$$

where the subscript ‘‘ $r$ ’’ stands for resampling, as it is the means by which we perform the resampling operation, and the boldface notation  $\hat{\Phi}_r$  is used to highlight that this is an *estimator*, namely a random matrix and must not be confused

with the corresponding *estimate* which is denoted by  $\hat{\Phi}_r$ . The latter is a deterministic matrix obtained by replacing the random process  $\hat{y}(t)$  with the corresponding realization  $\hat{y}(t)$ .

By generating a realization  $\hat{y}^N = \{\hat{y}(1) \dots \hat{y}(N)\}$  from  $\hat{\Phi}$  (i.e. by resampling the data), we can easily obtain a realization of the random variable  $\mathcal{S}_{IS}(\hat{\Phi} || \hat{\Phi}_r)$ . Accordingly, it is possible to compute numerically  $\delta_\alpha$  such that  $\Pr(\mathcal{S}_{IS}(\hat{\Phi} || \hat{\Phi}_r) \leq \delta_\alpha) = \alpha$  by a standard Monte Carlo procedure. Numerical simulations show that this technique indeed provides a good estimate of  $\delta$ .

It is worth noting that if the chosen  $\alpha$  is too large with respect to the data length  $N$ , the resulting  $\delta_\alpha$  may be too generous yielding to a diagonal  $\Phi$  obeying  $\mathcal{S}_{IS}(\Phi || \hat{\Phi}) \leq \delta_\alpha$ . In this case Problem (5) admits the trivial solution  $\Phi_L = 0$  and  $\Phi_D = \Phi$ . To rule out this trivial case,  $\delta$  in (5) must be strictly smaller than the upper bound

$$\delta_{\max} := \min_{\substack{\Phi \in \mathcal{S}_m^+ \\ \Phi \text{ diagonal}}} \mathcal{S}_{IS}(\Phi || \hat{\Phi})$$

where  $\mathcal{S}_m^+$  denotes the family of bounded and coercive functions defined on the unit circle and taking values in the cone of positive definite  $m \times m$  Hermitian matrices. Since  $\Phi$  must be diagonal, by denoting with  $\phi_i$  and by  $\hat{\gamma}_i$  the  $i$ -th element in the diagonal of  $\Phi$  and of  $\hat{\Phi}^{-1}$ , respectively, we have

$$\delta_{\max} = \left[ \sum_{i=1}^m \min_{\phi_i \in \mathcal{S}_1^+} \mathcal{S}_{IS}(\phi_i || \hat{\gamma}_i^{-1}) \right] + \int \log |\hat{\Phi} \text{diag}^2(\hat{\Phi}^{-1})|$$

where  $\text{diag}^2(\cdot)$  is the (orthogonal projection) operator mapping a square matrix  $M$  into a diagonal matrix of the same size having the same main diagonal of  $M$ . Therefore, since the Itakura-Saito divergence is nonnegative, the solution corresponds to  $\phi_i^{\text{opt}}(e^{i\vartheta}) = (\hat{\gamma}_i(e^{i\vartheta}))^{-1}$ ,  $i = 1, \dots, m$  for which  $\mathcal{S}_{IS}(\phi_i^{\text{opt}} || \hat{\gamma}_i^{-1}) = 0$ . Accordingly,

$$\delta_{\max} = \int \log |\hat{\Phi} \text{diag}^2(\hat{\Phi}^{-1})|. \quad (6)$$

The derivation of the aforementioned result is based on reasonings similar to [6, Section IV].

A more generous upper bound can be derived by assuming that  $\Phi$  is the spectrum of an MA process of order  $n$ . However, numerical experiments showed that  $\delta_{\max} \gg \delta_\alpha$  even in the case that  $N$  is relatively small.

### III. PROBLEM SOLUTION

In this section we first provide a finite dimensional matrix parametrization of Problem (5). The latter is then analyzed by resorting to the Lagrange duality theory, which allows us to prove the existence of a solution.

#### A. Matricial Reparametrization of the Problem

To study Problem (5) it is convenient to introduce the following matrix parametrization for  $\Phi$ ,  $\Phi_L$  and  $\Phi - \Phi_L$ :

$$\begin{aligned} \Phi &= \Delta X \Delta^* \in \mathcal{Q}_{m,n} \\ \Phi_L &= \Delta L \Delta^* \in \mathcal{Q}_{m,n} \\ \Phi - \Phi_L &= \Delta (X - L) \Delta^* \in \mathcal{Q}_{m,n} \end{aligned} \quad (7)$$

where  $\Delta(e^{i\theta})$  is the so-called shift operator:

$$\Delta(e^{i\theta}) := [I_m \quad e^{i\theta} I_m \quad \dots \quad e^{in\theta} I_m]; \quad (8)$$

$X$  and  $L$  are matrices in  $\mathbf{Q}_{m(n+1)}$  and  $X_{ij}$  denotes the block of  $X$  in position  $i, j$  with  $i, j = 0, \dots, n$ , so that

$$X = \begin{bmatrix} X_{00} & X_{01} & \dots & X_{0n} \\ X_{01}^\top & X_{11} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ X_{0n}^\top & X_{1n}^\top & \dots & X_{nn} \end{bmatrix}.$$

Moreover,  $\mathbf{M}_{m,n}$  denotes the vector space of matrices of the form

$$Y := [Y_0 \quad Y_1 \quad \dots \quad Y_n], \quad Y_0 \in \mathbf{Q}_m, \quad Y_1, \dots, Y_n \in \mathbb{R}^{m \times m}. \quad (9)$$

The linear mapping  $T : \mathbf{M}_{m,n} \rightarrow \mathbf{Q}_{m(n+1)}$  constructs a symmetric block-Toeplitz matrix from its first block row so that if  $Y$  is given by (9),

$$T(Y) = \begin{bmatrix} Y_0 & Y_1 & \dots & Y_n \\ Y_1^\top & Y_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & Y_1 \\ Y_n^\top & \dots & Y_1^\top & Y_0 \end{bmatrix}.$$

The adjoint of  $T$  is the mapping  $D : \mathbf{Q}_{m(n+1)} \rightarrow \mathbf{M}_{m,n}$  defined by  $D(X) = [[D(X)]_0 \quad \dots \quad [D(X)]_n]$  with

$$[D(X)]_0 = \sum_{h=0}^n X_{hh}, \quad [D(X)]_j = 2 \sum_{h=0}^{n-j} X_{hh+j}, \quad j = 1, \dots, n.$$

Next, the objective is to provide a more convenient formulation of Problem (5) in terms of  $X$  and  $L$ . To this end, we have to take into account the following points.

1) *Positivity Constraints*  $\Phi \succ 0$  a.e. and  $\Phi_L, \Phi - \Phi_D \succeq 0$ : It can be shown (see, for example, [17, Appendix A]) that, for any  $\Psi \in \mathcal{Q}_{m,n}$ ,  $\Psi \succeq 0$  if and only if there exists a matrix  $P \in \mathbf{Q}_{m(n+1)}$  such that  $\Delta P \Delta^*$  and  $P \succeq 0$ . Therefore, we replace the conditions  $\Phi_L \succeq 0$  with  $L \succeq 0$ , the condition  $\Phi - \Phi_L \succeq 0$  with  $X - L \succeq 0$ . Note that these conditions only guarantees  $X \succeq 0$  and thus  $\Phi$  to be positive semidefinite, however we will show that this is sufficient to guarantee that  $\Phi \succ 0$  a.e. at the optimum.

2) *Constraint  $\Phi - \Phi_L$  diagonal*: Let  $\text{ofd} : \mathbb{R}^{m \times m} \rightarrow \mathbb{R}^{m \times m}$  denote the linear operator defined as follows: given  $A \in \mathbb{R}^{m \times m}$ ,  $\text{ofd}(A)$  is the matrix in which each off-diagonal element is equal to the corresponding element of  $A$  and each diagonal element is zero. We define the ‘‘block ofd’’ linear operator  $\text{ofd}_B : \mathbf{M}_{m,n} \rightarrow \mathbf{M}_{m,n}$  as follows. Given  $Z = [Z_0 \quad Z_1 \quad \dots \quad Z_n] \in \mathbf{M}_{m,n}$ , then

$$\text{ofd}_B(Z) = [\text{ofd}(Z_0) \quad \text{ofd}(Z_1) \quad \dots \quad \text{ofd}(Z_n)].$$

It is not difficult that  $\text{ofd}_B$  is a self-adjoint operator, since  $\text{ofd}$  is self-adjoint as well. Then, it is easy to see that the condition  $\Phi - \Phi_L$  diagonal is equivalent to the condition  $[D(X - L)]_j$  diagonal for  $j = 0, \dots, n$ , that is  $\text{ofd}_B(D(X - L)) = 0$ .

3) *The Low Rank Regularizer*: We have

$$\text{tr} \int \Phi_L = \text{tr} \int \Delta L \Delta^* = \text{tr} \left( L \int \Delta^* \Delta \right) = \text{tr}(L)$$

where we exploited the fact that  $\int e^{ij\theta} = 1$  if  $j = 0$ , and  $\int e^{ij\theta} = 0$  otherwise.

4) *The Divergence Constraint*: A convenient matrix parameterization of the Itakura-Saito divergence  $\mathcal{S}_{IS}(\Phi || \hat{\Phi})$  can be obtained by making use of the following facts.

First, since  $\Phi = \Delta X \Delta^*$  with  $X \succeq 0$ , there exists  $A \in \mathbb{R}^{m \times m(n+1)}$  such that  $X = A^\top A$ . Then, by using the Jensen-Kolmogorov formula we obtain

$$\int \log |\Phi| = \int \log |\Delta A^\top A \Delta^*| = \log |A_0^\top A_0| = \log |X_{00}| \quad (10)$$

which holds provided that  $X_{00} \succ 0$  and  $\Phi$  is coercive (i.e.  $|\Phi|$  is bounded away from zero on the unit circle). We need to generalize this result to spectral densities that may be singular on the unit circle. This is possible because the zeros of a rational spectral density, if any, have finite multiplicity so that the logarithm of the determinant of a rational spectral  $\Phi$  is integrable as long as the normal rank of  $\Phi$  is full.

*Lemma 3.1*: Consider a power spectral density  $\Phi \in \mathcal{Q}_{m,n}$  having full normal rank. Let  $X \in \mathbf{Q}_{m(n+1)}$  be such that  $X \succeq 0$ ,  $X_{00} \succ 0$ , and  $\Phi = \Delta X \Delta^*$ . Then

$$\int \log |\Phi| = \log |X_{00}|.$$

The proof is deferred to the appendix.

A second observation in order to conveniently parameterize the Itakura-Saito divergence constraint is that, by exploiting the cyclic property of the trace,

$$\begin{aligned} \int \text{tr}(\hat{\Phi}^{-1} \Phi) &= \int \text{tr}(\hat{\Phi}^{-1} \Delta X \Delta^*) \\ &= \text{tr} \left( X \int \Delta^* \hat{\Phi}^{-1} \Delta \right) = \langle X, T(\hat{P}) \rangle, \end{aligned}$$

where  $\hat{P}$  is defined from the expansion

$$\hat{\Phi}^{-1} = \sum_{k=-\infty}^{\infty} \hat{P}_k e^{-i\theta k}$$

as  $\hat{P} := [\hat{P}_0 \quad \dots \quad \hat{P}_n]$ .

Summing up, we get the following matrix reparametrization of Problem (5):

$$\begin{aligned} (X^\circ, L^\circ) &= \arg \min_{X, L \in \mathbf{Q}_{m(n+1)}} \text{tr}(L) \\ &\text{subject to } X_{00} \succ 0, L \succeq 0, X - L \succeq 0, \\ &\text{ofd}_B(D[X - L]) = 0, \\ &-\log |X_{00}| + \int \log |\hat{\Phi}| \\ &+ \langle X, T(\hat{P}) \rangle - m \leq \delta. \end{aligned} \quad (11)$$

We remark once again that to prove the equivalence between (5) and (11) we still need to show that  $\Phi \succ 0$  a.e. at the optimum: this fact will be established after the variational analysis.

## B. The Dual Problem

We reformulate the constrained minimization problem in (11) as an unconstrained problem by means of Duality Theory. If we use  $V, U \in \mathbf{Q}_{m(n+1)}$ ,  $V, U \succeq 0$  as the multipliers associated with the constraints on the positive semi-definiteness of  $X - L$  and  $L$ , respectively;  $Z \in \mathbf{M}_{m,n}$  as the multiplier associated with the constraint  $\text{ofd}_B(D(X - L)) = 0$  and  $\lambda \in \mathbb{R}, \lambda \geq 0$ , as the multiplier associated with the Itakura-Saito divergence, then the Lagrangian of Problem (11) is

$$\begin{aligned} \mathcal{L}(X, L, \lambda, U, V, Z) = & \text{tr}(L) - \langle V, X - L \rangle - \langle U, L \rangle + \\ & \langle Z, \text{ofd}_B(D(X - L)) \rangle + \lambda(-\log |X_{00}| + \\ & \int \log |\hat{\Phi}| + \langle X, T(\hat{P}) \rangle - m - \delta) \\ & = \langle L, I - U + V - T(\text{ofd}_B(Z)) \rangle + \\ & \langle X, T(\text{ofd}_B(Z)) - V + \lambda T(\hat{P}) \rangle - \\ & \lambda(\log |X_{00}| - \int \log |\hat{\Phi}| + m + \delta). \end{aligned} \quad (12)$$

Note that we have not included the constraint  $X_{00} \succ 0$  because, as we will show later on, this condition is automatically met by the solution of the dual problem.

The dual function is defined as the infimum of  $\mathcal{L}$  over  $X$  and  $L$ . Thanks to the convexity of the Lagrangian, we rely on standard variational methods to characterize the minimum.

- *Partial minimization with respect to  $L$* :  $\mathcal{L}$  depends on  $L$  only through  $\langle L, I - U + V - T(\text{ofd}_B(Z)) \rangle$  which is bounded below only if

$$I - U + V - T(\text{ofd}_B(Z)) = 0. \quad (13)$$

Thus, we get that

$$\inf_L \mathcal{L} = \begin{cases} \langle X, T(\text{ofd}_B(Z)) - V + \lambda T(\hat{P}) \rangle - \\ \lambda(\log |X_{00}| - \int \log |\hat{\Phi}| + m + \delta) & \text{if (13)} \\ -\infty & \text{otherwise.} \end{cases}$$

- *Partial minimization with respect to  $X$* : The terms in  $X_{00}$  are bounded below only if

$$\left[ T(\text{ofd}_B(Z)) - V + \lambda T(\hat{P}) \right]_{00} \succ 0 \quad (14)$$

and are minimized if  $\lambda > 0$  and

$$X_{00} = \left( [T(\hat{P}) + \lambda^{-1}(T(\text{ofd}_B(Z)) - V)]_{00} \right)^{-1}. \quad (15)$$

The Lagrangian is linear in the remaining variables  $X_{lh}$ , for  $(l, h) \neq (0, 0)$ , and therefore bounded below only if

$$\left[ T(\text{ofd}_B(Z)) - V + \lambda T(\hat{P}) \right]_{lh} = 0 \quad \forall (l, h) \neq (0, 0). \quad (16)$$

Therefore, the minimization of the Lagrangian with respect to  $X$  and  $L$  is finite if and only if (13), (14), and (16) hold in which case

$$\begin{aligned} \min_{X, L} \mathcal{L} = & -\lambda(-\log | [T(\hat{P}) + \lambda^{-1}(T(\text{ofd}_B(Z)) \\ & - V)]_{00} | - \int \log |\hat{\Phi}| + \delta). \end{aligned}$$

Otherwise the Lagrangian has no minimum and its infimum is  $-\infty$ .

To simplify the notation, let us define the vector space  $\mathcal{O}$  as:

$$\mathcal{O} := \{Z \in \mathbf{M}_{m,n} : \text{ofd}_B(Z) = Z, j = 0, \dots, n\};$$

since  $Z$  always appears in the form  $\text{ofd}_B(Z)$ , we can replace it with  $Z \in \mathcal{O}$ . Then, we can formulate the dual problem for the Lagrangian (12) as

$$\max_{(\lambda, U, V, Z) \in \tilde{\mathcal{C}}} \tilde{J} \quad (17)$$

where

$$\tilde{J} := \lambda \left( \log | [T(\hat{P}) + \lambda^{-1}(T(Z) - V)]_{00} | + \int \log |\hat{\Phi}| - \delta \right)$$

and the feasible set  $\tilde{\mathcal{C}}$  is given by:

$$\begin{aligned} \tilde{\mathcal{C}} := & \{(\lambda, U, V, Z) : U, V \in \mathbf{Q}_{m(n+1)}, U, V \succeq 0, Z \in \mathcal{O}, \\ & \lambda \in \mathbb{R}, \lambda > 0, I - U + V - T(Z) = 0, [\lambda T(\hat{P}) + T(Z) - \\ & V]_{00} \succ 0, [\lambda T(\hat{P}) + T(Z) - V]_{lh} = 0 \quad \forall (l, h) \neq (0, 0)\}. \end{aligned}$$

Note that the constraints  $I - U + V - T(Z) = 0$  and  $U \succeq 0$  are equivalent to the constraint  $I + V - T(Z) \succeq 0$ . Thus, we can eliminate the redundant variable  $U$ ; moreover, by changing the sign to the objective function  $\tilde{J}$  and observing that  $[T(\hat{P}) + \lambda^{-1}(T(Z) - V)]_{00} = \hat{P}_0 + \lambda^{-1}(Z_0 - V_{00})$ , we can rewrite (17) as a minimization problem:

$$\min_{(\lambda, V, Z) \in \mathcal{C}} J \quad (18)$$

where

$$J := \lambda \left( -\log | \hat{P}_0 + \lambda^{-1}(Z_0 - V_{00}) | - \int \log |\hat{\Phi}| + \delta \right)$$

and the corresponding feasible set  $\mathcal{C}$  is:

$$\begin{aligned} \mathcal{C} := & \{(\lambda, V, Z) : V \in \mathbf{Q}_{m(n+1)}, V \succeq 0, Z \in \mathcal{O}, \\ & I + V - T(Z) \succeq 0, \lambda \in \mathbb{R}, \lambda > 0, [\lambda \hat{P}_0 + Z_0 - V_{00}] \succ 0, \\ & [\lambda(T(\hat{P})) + T(Z) - V]_{lh} = 0 \quad \forall (l, h) \neq (0, 0)\}. \end{aligned}$$

## C. Existence of solutions

The aim of this section is to show that (18) admits solution. The set  $\mathcal{C}$  is not compact, as it is neither closed nor bounded. We show that we can restrict the search of the minimum of  $J$  over a compact set. Then, since the objective function is continuous over  $\mathcal{C}$  (and hence over the restricted compact set), we can use Weierstrass's Theorem to conclude that the problem does admit a minimum.

The first step consists in showing that we can restrict  $\mathcal{C}$  to a subset where  $\lambda \geq \varepsilon$  with  $\varepsilon > 0$  a positive constant.

*Proposition 3.1:* Let  $(\lambda^{(k)}, V^{(k)}, Z^{(k)})_{k \in \mathbb{N}}$  be a sequence of elements in  $\mathcal{C}$  such that

$$\lim_{k \rightarrow \infty} \lambda^{(k)} = 0.$$

Then, such a sequence cannot be an infimizing sequence.

The proof is essentially the same as the proof of Proposition 6.1 in [6] and it is therefore omitted.

As a consequence, minimizing the dual functional over the set  $\mathcal{C}$  is equivalent to minimize it over the set:

$$\mathcal{C}_1 := \{(\lambda, V, Z) : V \in \mathbf{Q}_{m(n+1)}, V \succeq 0, Z \in \mathcal{O}, \\ I + V - T(Z) \succeq 0, \lambda \in \mathbb{R}, \lambda \geq \varepsilon, [\lambda \hat{P}_0 + Z_0 - V_{00}] \succ 0, \\ [\lambda(T(\hat{P})) + T(Z) - V]_{lh} = 0 \forall (l, h) \neq (0, 0)\}.$$

Next we show that we can restrict  $\mathcal{C}_1$  to a subset in which both  $(T(Z) - V)$  and  $\lambda$  cannot diverge.

*Proposition 3.2:* Let  $(\lambda^{(k)}, V^{(k)}, Z^{(k)})_{k \in \mathbb{N}}$  be a sequence of elements in  $\mathcal{C}_1$  such that either

$$\lim_{k \rightarrow \infty} \|T(Z^{(k)}) - V^{(k)}\| = +\infty$$

or

$$\lim_{k \rightarrow \infty} \lambda^{(k)} = +\infty$$

or both. Then, such a sequence cannot be an infimizing sequence.

The above result is obtained by following arguments similar to the proof of Proposition 6.2 in [6] with a few small differences; we refer the interested reader to [31, Appendix C] for the detailed proof.

It follows from the previous proposition that there exists  $\beta \in \mathbb{R}$  with  $|\beta| < \infty$  such that  $T(Z) - V \succeq \beta I$ , and  $0 < \gamma < \infty$  such that  $\lambda \leq \gamma$ . Therefore, the set  $\mathcal{C}_1$  can be further restricted to the set:

$$\mathcal{C}_2 := \{(\lambda, V, Z) : V \in \mathbf{Q}_{m(n+1)}, V \succeq 0, Z \in \mathcal{O}, \lambda \in \mathbb{R}, \\ \beta I \preceq T(Z) - V \preceq I, \gamma \geq \lambda \geq \varepsilon, [\lambda \hat{P}_0 + Z_0 - V_{00}] \succ 0, \\ [\lambda(T(\hat{P})) + T(Z) - V]_{lh} = 0 \forall (l, h) \neq (0, 0)\}.$$

In addition, it is not possible for  $V$  and  $Z$  to diverge while keeping the difference  $T(Z) - V$  finite. Accordingly, we can further restrict the search for the optimal solution to a subset  $\mathcal{C}_3$  in which neither  $V$  nor  $Z$  can diverge:

*Proposition 3.3:* Let  $(\lambda^{(k)}, V^{(k)}, Z^{(k)})_{k \in \mathbb{N}}$  be a sequence of elements in  $\mathcal{C}_2$  such that

$$\lim_{k \rightarrow \infty} \|V^{(k)}\| = +\infty \quad (19)$$

or

$$\lim_{k \rightarrow \infty} \|Z^{(k)}\| = +\infty \quad (20)$$

or both. Then, such a sequence cannot be an infimizing sequence.

The proof can be found in the appendix.

Thus, the minimization over  $\mathcal{C}_2$  is equivalent to the minimization over the subset:

$$\mathcal{C}_3 := \{(\lambda, V, Z) : V \in \mathbf{Q}_{m(n+1)}, \alpha I \succeq V \succeq 0, Z \in \mathcal{O}, \lambda \in \mathbb{R}, \\ \beta I \preceq T(Z) - V \preceq I, \gamma \geq \lambda \geq \varepsilon, [\lambda \hat{P}_0 + Z_0 - V_{00}] \succ 0, \\ [\lambda(T(\hat{P})) + T(Z) - V]_{lh} = 0 \forall (l, h) \neq (0, 0)\}$$

for a certain  $\alpha > 0$  positive constant.

Finally, we consider a sequence  $(\lambda^{(k)}, V^{(k)}, Z^{(k)})_{k \in \mathbb{Z}} \in \mathcal{C}_3$  such that  $[(\lambda^{(k)})^{-1}([Z^{(k)}]_0 - [V^{(k)}]_{00}) + \hat{P}_0]$  tends to be singular as  $k \rightarrow \infty$ . This implies that  $|(\lambda^{(k)})^{-1}([Z^{(k)}]_0 - [V^{(k)}]_{00}) + \hat{P}_0|$  tends to zero and hence  $J \rightarrow +\infty$ . Thus, such

a sequence cannot be an infimizing sequence. Therefore, the final set  $\mathcal{C}_C$  is:

$$\mathcal{C}_C := \{(\lambda, V, Z) : V \in \mathbf{Q}_{m(n+1)}, \alpha I \succeq V \succeq 0, Z \in \mathcal{O}, \beta I \preceq \\ T(Z) - V \preceq I, \lambda \in \mathbb{R}, \gamma \geq \lambda \geq \varepsilon, [\lambda \hat{P}_0 + Z_0 - V_{00}] \succeq \mu I, \\ [\lambda(T(\hat{P})) + T(Z) - V]_{lh} = 0 \forall (l, h) \neq (0, 0)\}$$

where  $\alpha, \beta, \gamma, \varepsilon$  and  $\mu$  such that  $|\alpha|, |\beta|, |\gamma|, |\varepsilon|$  and  $|\mu| < +\infty$ .

*Theorem 3.1:* Problem (18) is equivalent to

$$\min_{(\lambda, V, Z) \in \mathcal{C}_C} J(\lambda, V, Z)$$

and it admits solution.

*Proof.* Equivalence of the two problems has already been proven by the previous arguments. Since  $\mathcal{C}_C$  is closed and bounded, hence compact, and  $J$  is continuous over  $\mathcal{C}_C$ , by the Weierstrass's Theorem the minimum exists. ■

#### IV. SOLUTION OF THE PRIMAL PROBLEM

In this section, after proving that the primal problem (5) and its matrix reformulation (11) are equivalent, we show how to recover the solution of the primal problem.

Let  $(\lambda^\circ, V^\circ, Z^\circ)$  be a solution of (18) and  $(X^\circ, L^\circ)$  be the corresponding solution of (11). Since  $X_{00}^\circ$  is positive definite,  $\log |X_{00}^\circ|$  is finite. By Lemma 3.1, at the optimum  $\int \log |\Phi|$  must be finite as well; this implies that  $\Phi(e^{i\vartheta})$ ,  $\vartheta \in [-\pi, +\pi]$ , may be singular at most on a set of zero measure, or, in other terms,  $\Delta X^\circ \Delta^* \succ 0$  a.e.. This observation leads to the following proposition:

*Proposition 4.1:* Let  $(X^\circ, L^\circ)$  be a solution of (11). Then  $\Delta X^\circ \Delta^* \succ 0$  a.e.. Accordingly, (5) and (11) are equivalent.

Now we are ready to show how to recover the solution of the primal problem; to this aim we need the following result, see [32].

*Lemma 4.1:* Let  $Z \in \mathbf{M}_{m,n}$  and  $W \in \mathbf{Q}_m$ . If  $W \succ 0$  is such that

$$T(Z) \succeq \begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix} \quad (21)$$

then  $T(Z) \succ 0$ .

Exploiting the constraints  $[\lambda(T(\hat{P})) + T(Z) - V]_{lh} = 0, \forall (l, h) \neq (0, 0)$  and  $[\lambda \hat{P}_0 + Z_0 - V_{00}] \succ 0$ , it is not difficult to see that

$$V^\circ = \lambda^\circ T(\hat{P}) + T(Z^\circ) - \begin{bmatrix} W^\circ & 0 \\ 0 & 0 \end{bmatrix} \quad (22)$$

where

$$W^\circ := Z_0^\circ - V_{00}^\circ + \lambda^\circ \hat{P}_0 \succ 0. \quad (23)$$

Since  $V^\circ \succeq 0$  and in view of Lemma 4.1,  $\lambda^\circ T(\hat{P}) + T(Z^\circ) \succ 0$ . Hence,  $V^\circ$  has rank at least equal to  $mn$ .

Since the duality gap between (11) and (18) is equal to zero, we have that  $\langle V^\circ, X^\circ - L^\circ \rangle = 0$ , which in turn implies

$$V^\circ (X^\circ - L^\circ) = 0 \quad (24)$$

because  $V^\circ, X^\circ - L^\circ \succeq 0$ . Recalling that  $\text{rank}(V^\circ) \geq mn$ , in view of (24) the matrix  $X^\circ - L^\circ$  has rank at most equal to  $m$ .

Let  $\text{rank}(X^\circ - L^\circ) = \tilde{m} \leq m$ . Then, there exists a full-row rank matrix  $A \in \mathbb{R}^{\tilde{m} \times m(n+1)}$  such that

$$X^\circ - L^\circ = A^\top A. \quad (25)$$

By (24), it follows that  $V^\circ A^\top = 0$ . Let  $Y_D := [v_o \ v_1 \ \dots \ v_l] \in \mathbb{R}^{m(n+1) \times l}$  denote the matrix whose columns form a basis of  $\ker(V^\circ)$ . Note that the dimension  $l$  of the null space of  $V^\circ$  is at least  $\tilde{m}$  because  $\text{Im}(A^\top) \subseteq \ker(V^\circ)$  and  $\text{rank}(A^\top) = \tilde{m}$ ; also  $l \leq m$  because  $\text{rank}(V^\circ) \geq mn$ . Rewriting the matrix  $A^\top$  as  $A^\top = Y_D S$  with  $S \in \mathbb{R}^{l \times \tilde{m}}$ , from (25) we obtain

$$X^\circ - L^\circ = Y_D Q_D Y_D^\top, \quad (26)$$

with  $Q_D := S S^\top \in \mathbf{Q}_l$  unknown.

In a similar fashion, by the zero duality gap between (11) and (18), the *complementary slackness condition* for the multiplier associated to the positive semi-definiteness of  $L$  reads as  $\langle U^\circ, L^\circ \rangle = 0$ , which in turn implies  $U^\circ L^\circ = 0$ . Repeating the same reasoning as before, it can be seen that, if the dimension of the null space of  $U^\circ$  is  $\tilde{r}$  with  $\tilde{r} \geq r$  and  $Y_L := [u_o \ u_1 \ \dots \ u_{\tilde{r}}] \in \mathbb{R}^{m(n+1) \times \tilde{r}}$  is a matrix whose columns form a basis of  $\ker(U^\circ)$ , then  $L^\circ$  can be written as

$$L^\circ = Y_L Q_L Y_L^\top \quad (27)$$

with  $Q_L \in \mathbf{Q}_{\tilde{r}}$  unknown. Plugging (27) into (26), we then obtain

$$X^\circ - Y_L Q_L Y_L^\top = Y_D Q_D Y_D^\top. \quad (28)$$

Assume now that each block of  $X^\circ - L^\circ$  is diagonal, namely

$$\text{ofd}([Y_D Q_D Y_D^\top]_{hk}) = 0 \quad h, k = 0, \dots, n. \quad (29)$$

*Remark 1:* We can make the previous assumption without loss of generality. Indeed, let  $(\Phi^\circ, \Phi_L^\circ)$  be the solution of Problem (5) and  $\Phi_D^\circ = \Phi^\circ - \Phi_L^\circ$ ;  $X, L$  and  $D = X - L$  are any matrices in  $\mathbf{Q}_{m(n+1)}$  such that  $\Phi^\circ = \Delta X \Delta^*$ ,  $\Phi_L^\circ = \Delta L \Delta^*$  and  $\Phi_D^\circ = \Delta D \Delta^*$ . We can always consider a different matrix parametrization  $(\tilde{X}, \tilde{L}, \tilde{D})$  for  $\Phi^\circ, \Phi_L^\circ$  and  $\Phi_D^\circ$  as follows. First notice that there always exists a matrix  $\tilde{D}$  with all diagonal blocks such that  $\Phi_D^\circ = \Delta \tilde{D} \Delta^*$ ; in other words, we can always find  $\delta D \in \mathbf{Q}_{m(n+1)}$  such that  $\Delta \delta D \Delta^* = 0$  and  $\tilde{D} := D + \delta D$  satisfies  $\text{ofd}([\tilde{D}]_{hk}) = 0$  for  $h, k = 0, \dots, n$ . Now, let  $\delta X \in \mathbf{Q}_{m(n+1)}$  such that  $\Delta \delta X \Delta^* = 0$  and  $\tilde{X} := X + \delta X$  satisfies (15). Define  $\tilde{L} = \tilde{X} - \tilde{D} = X - D + \delta L$  with  $\delta L := \delta X - \delta D$ . It is easy to see that  $(\tilde{X}, \tilde{L})$  is still a solution of Problem (11) and it allows us to restrict to solutions of (11) for which (29) holds.

By applying the *ofd* operator to both sides of (28) and exploiting the assumption (29), it is not difficult to obtain:

$$\text{ofd}([Y_L Q_L Y_L^\top]_{00}) = \text{ofd}(X_{00}^\circ) \quad (30)$$

which is a system of  $m(m-1)/2$  linear equations in the  $\tilde{r}(\tilde{r}+1)/2$  unknowns  $Q_L$ . Notice that  $X_{00}$  is given by (15). Finally, once  $L^\circ$  is computed, in order to retrieve  $Q_D$  we exploit (29) and the following system of  $m(m+1)/2$  linear equations:

$$[Y_D Q_D Y_D^\top]_{00} = X_{00}^\circ - L_{00}^\circ. \quad (31)$$

Since both the dual and the primal problem admit solution, the resulting systems of equations (29), (30) and (31) do admit solutions.

## V. THE PROPOSED ALGORITHM

In this section we propose an algorithm to solve numerically the dual problem. To start with, as observed in Section IV, we rewrite (18) in a different fashion by getting rid of the slack variable  $V \in \mathbf{Q}_{m(n+1)}$ . This is done by introducing a new variable  $W \in \mathbf{Q}_m$  defined, similarly to (23), as

$$W := Z_0 - V_{00} + \lambda \hat{P}_0 \succ 0 \quad (32)$$

such that, as in (22), the variable  $V$  can be expressed as

$$V = \lambda T(\hat{P}) + T(Z) - \begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix}. \quad (33)$$

Accordingly, the dual problem (18) can be expressed in terms of the variables  $\lambda, W$  and  $Z$  as follows:

$$\min_{(\lambda, W, Z) \in \mathcal{C}} J \quad (34)$$

where

$$J := \lambda \left( -\log |\lambda^{-1} W| - \int \log |\hat{\Phi}| + \delta \right)$$

and the corresponding feasible set  $\mathcal{C}$  is:

$$\begin{aligned} \mathcal{C} := \{ & (\lambda, W, Z) : W \in \mathbf{Q}_m, W \succ 0, Z \in \mathcal{O}, \lambda \in \mathbb{R}, \\ & \lambda > 0, \lambda T(\hat{P}) + T(Z) - \begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix} \succeq 0, \\ & I + \lambda T(\hat{P}) - \begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix} \succeq 0 \}. \end{aligned}$$

We can further simplify our problem as follows. First, we observe that the constraint

$$V = \lambda T(\hat{P}) + T(Z) - \begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix} \succeq 0 \quad (35)$$

may be rewritten as

$$\lambda T(\hat{P}) + T(Z) \succeq \begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix}$$

and then, by Lemma 4.1,  $\lambda T(\hat{P}) + T(Z) \succ 0$ . Now, we can easily rewrite (35) recalling the characterization of a symmetric positive semidefinite matrix using the Schur complement. To this aim, it is convenient to introduce the linear operators  $T_{0,0} : \mathbf{M}_{m,n} \rightarrow \mathbf{Q}_m$ ,  $T_{0,1:n} : \mathbf{M}_{m,n} \rightarrow \mathbf{M}_{m,n-1}$  and  $T_{1:n,1:n} : \mathbf{M}_{m,n} \rightarrow \mathbf{Q}_{mn}$  that, for a given matrix  $H \in \mathbf{M}_{m,n}$  construct a symmetric block-Toeplitz matrix and extract the blocks in position  $(0,0)$ ,  $(0,1:n)$  and  $(1:n,1:n)$ , respectively. With this notation, we have

$$T(Z + \lambda \hat{P}) = \begin{bmatrix} T_{0,0}(Z + \lambda \hat{P}) & T_{0,1:n}(Z + \lambda \hat{P}) \\ T_{0,1:n}^\top(Z + \lambda \hat{P}) & T_{1:n,1:n}(Z + \lambda \hat{P}) \end{bmatrix}$$

and the constraint (35) is equivalent to require  $T_{1:n,1:n}(Z + \lambda \hat{P}) \succ 0$  and  $W \preceq Q(\lambda, Z)$  with

$$Q(\lambda, Z) := T_{0,0}(Z + \lambda \hat{P}) - T_{0,1:n}(Z + \lambda \hat{P}) \times T_{1:n,1:n}^{-1}(Z + \lambda \hat{P}) T_{0,1:n}^\top(Z + \lambda \hat{P}).$$

In a similar fashion, the last matrixial inequality constraint in  $\mathcal{C}$  can be equivalently expressed as  $W \preceq R(\lambda)$  where

$$R(\lambda) := I + T_{0,0}(\lambda\hat{P}) - T_{0,1:n}(\lambda\hat{P})(I + T_{1:n,1:n}(\lambda\hat{P}))^{-1} \times T_{0,1:n}^\top(\lambda\hat{P}).$$

Therefore, Problem (18) can be formulated as

$$\min_{(\lambda, W, Z) \in \mathcal{C}} J = \lambda \left( -\log |\lambda^{-1}W| - \int \log |\hat{\Phi}| + \delta \right) \quad (36)$$

where

$$\mathcal{C} := \{(\lambda, W, Z) : Z \in \mathcal{O}, \lambda \in \mathbb{R}, \lambda > 0, T_{1:n,1:n}(Z + \lambda\hat{P}) \succ 0, W \in \mathbf{Q}_m, W \succ 0, W \preceq Q(\lambda, Z), W \preceq R(\lambda)\}.$$

Before solving this problem, notice that  $J$  in (36) is jointly convex in  $(\lambda, W)$  and at each feasible point  $(\lambda_0, W_0)$  (where  $\lambda_0 \neq 0$ ) it is strictly convex in all directions except for the direction equal to the point itself, i.e.  $(\lambda_0, W_0)$ . Thus, if  $(\lambda_{\text{opt}}, W_{\text{opt}})$  is an optimal value for  $\lambda$  and  $W$ , all the other optimal values must lie on the non-strictly convex direction. As a consequence there exist  $\alpha_{\min} \leq 0$  and  $\alpha_{\max} \geq 0$  such that  $((1+\alpha)\lambda_{\text{opt}}, (1+\alpha)W_{\text{opt}})$  are optimal values for  $(\lambda, W)$  for any  $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ . Moreover, if  $\alpha_{\min} = \alpha_{\max} = 0$  then the pair  $(\lambda_{\text{opt}}, W_{\text{opt}})$  that, together with a certain  $Z$ , solves (36) is unique. This is indeed the case. In fact, at the optimum  $J$  is strictly negative so that at the optimum the derivative of  $J$  along the only non-strictly convex direction  $(\lambda_{\text{opt}}, W_{\text{opt}})$  is not equal to zero. In other words, at the optimum and along the only non-strictly convex direction  $J$  is not constant which implies  $\alpha_{\min} = \alpha_{\max} = 0$ .

Uniqueness of  $Z$  is a much more problematic issue. Indeed, we have observed in simulations that in some cases it may happen that there are different optimal values of  $Z$ . The corresponding number of identified latent factors, however, is not affected and the predictive powers of the identified models are essentially the same.

Solving Problem (36) simultaneously for  $\lambda$ ,  $W$ , and  $Z$  is not trivial because the inequality constraints  $W \preceq Q(\lambda, Z)$  and  $W \preceq R(\lambda)$  both depend on  $\lambda$ . On the other hand, once we fix the dual variable  $\lambda$  to a positive constant  $\bar{\lambda} > 0$ , the problem:

$$\min_{(W, Z) \in \mathcal{C}_{\bar{\lambda}}} J(\bar{\lambda}, W, Z) \quad (37)$$

with

$$\mathcal{C}_{\bar{\lambda}} := \{(W, Z) : Z \in \mathcal{O}, W \in \mathbf{Q}_m, T_{1:n,1:n}(Z + \bar{\lambda}\hat{P}) \succ 0, W \succ 0, W \preceq Q(\bar{\lambda}, Z), W \preceq R(\bar{\lambda})\}.$$

can be efficiently solved by resorting to the ADMM algorithm [33]. To this aim, we rewrite Problem (37) by introducing a new variable  $Y \in \mathbf{Q}_m$  defined as  $Y = Q(\bar{\lambda}, Z) - W$ :

$$\min_{\substack{(W, Z) \in \mathcal{C}_{W, Z}, \\ Y \in \mathbf{Q}_m^+}} J = \bar{\lambda} \left( -\log |\bar{\lambda}^{-1}W| - \int \log |\hat{\Phi}| + \delta \right) \quad (38)$$

$$\text{subject to } Y = Q(\bar{\lambda}, Z) - W$$

where

$$\mathcal{C}_{W, Z} := \{(W, Z) : Z \in \mathcal{O}, W \in \mathbf{Q}_m, W \succ 0, W \preceq R(\bar{\lambda}), T_{1:n,1:n}(Z + \bar{\lambda}\hat{P}) \succ 0\}$$

and  $\mathbf{Q}_m^+$  denotes the cone of symmetric positive semidefinite matrices of size  $m \times m$ . The *augmented Lagrangian* for (38) is:

$$\mathcal{L}_\rho(W, Z, Y, M) := \bar{\lambda} \left( -\log |\bar{\lambda}^{-1}W| - \int \log |\hat{\Phi}| + \delta \right) + \langle M, Y - Q(\bar{\lambda}, Z) + W \rangle + \frac{\rho}{2} \|Y - Q(\bar{\lambda}, Z) + W\|^2$$

where  $M \in \mathbf{Q}_m$  is the Lagrange multiplier, and  $\rho > 0$  is the *penalty parameter*. Accordingly, given the initial guesses  $W^{(0)}$ ,  $Z^{(0)}$ ,  $Y^{(0)}$  and  $M^{(0)}$ , the ADMM updates are:

$$(W^{(k+1)}, Z^{(k+1)}) = \arg \min_{(W, Z) \in \mathcal{C}_{W, Z}} \mathcal{L}_\rho(W, Z, Y^{(k)}, M^{(k)}) \quad (39)$$

$$Y^{(k+1)} = \arg \min_{Y \in \mathbf{Q}_m^+} \mathcal{L}_\rho(W^{(k+1)}, Z^{(k+1)}, Y, M^{(k)}) \quad (40)$$

$$M^{(k+1)} = M^{(k)} + \rho(Y^{(k+1)} - Q(\bar{\lambda}, Z^{(k+1)}) + W^{(k+1)}).$$

Problem (39) does not admit a closed form solution, therefore we approximate the optimal solution by a gradient projection step:

$$\begin{aligned} W^{(k+1)} &= \Pi(W^{(k)} - t_k \nabla_W \mathcal{L}_\rho(W^{(k)}, Z^{(k)}, Y^{(k)}, M^{(k)})) \\ Z^{(k+1)} &= \Pi_{\mathcal{O}}(Z^{(k)} - t_k \nabla_Z \mathcal{L}_\rho(W^{(k)}, Z^{(k)}, Y^{(k)}, M^{(k)})) \end{aligned}$$

where:

- $\nabla_W \mathcal{L}_\rho(W, Z, Y, M)$  denotes the gradient of the augmented Lagrangian with respect to  $W$ :

$$\nabla_W \mathcal{L}_\rho = -\bar{\lambda}W^{-1} + M + \rho(Y - Q + W).$$

- $\nabla_Z \mathcal{L}_\rho(W, Z, Y, M)$  denotes the gradient of the augmented Lagrangian with respect to  $Z$ :

$$\nabla_Z \mathcal{L}_\rho = D \left( \begin{bmatrix} I_m & \\ -T_{1:n,1:n}^{-1} T_{0,1:n}^\top & \end{bmatrix} (-M - \rho(Y - Q + W)) \begin{bmatrix} I_m & -T_{0,1:n} T_{1:n,1:n}^{-1} \end{bmatrix} \right)$$

where the omitted argument of the operators  $T_{0,1:n}$  and  $T_{1:n,1:n}$  is intended to be equal to  $(Z + \bar{\lambda}\hat{P})$ .

- $\Pi_{\mathcal{O}}$  denotes the projection operator onto  $\mathcal{O}$ :

$$\Pi_{\mathcal{O}}(A) = \text{ofd}_B(A).$$

- $\Pi$  denotes the projection operator onto the convex cone  $\{S \in \mathbf{Q}_m : S \preceq R(\bar{\lambda})\}$ . It is not difficult to see that

$$\Pi(A) = R(\bar{\lambda}) - \Pi_+(R(\bar{\lambda}) - A),$$

where  $\Pi_+$  is the projection operator onto the cone  $\mathbf{Q}_m^+$ .

- the step-size  $t_k$  is determined at each step  $k$  in an iterative fashion: we start by setting  $t_k = 1$  and we decrease it progressively of a factor  $\beta$ , with  $0 < \beta < 1$ , until the conditions  $W^{(k+1)} \succ 0$  and  $T_{1:n,1:n}(Z^{(k+1)} + \bar{\lambda}\hat{P}) \succ 0$  are met and the Armijo's condition [34] is satisfied.

Problem (40) admits a closed form solution, which can be easily computed as:

$$Y^{(k+1)} = \Pi_+ \left( Q(\bar{\lambda}, Z^{(k+1)}) - W^{(k+1)} - \frac{1}{\rho} M^{(k)} \right).$$

To define the stopping criterion, we need to introduce the following quantities

$$R^P = Y - Q(\bar{\lambda}, Z^{(k+1)}) + W^{(k+1)}$$

$$R^D = D \left( \begin{bmatrix} I_m & \\ -T_{1:n,1:n}^{-1} & T_{0,1:n}^\top \end{bmatrix} (\rho(Y^{(k+1)} - Y^{(k)})) \times \right. \\ \left. \begin{bmatrix} I_m & -T_{0,1:n} T_{1:n,1:n}^{-1} \end{bmatrix} \right)$$

which are referred to as the primal and dual residual, respectively. Notice that the omitted argument of the operators  $T_{0,1:n}$  and  $T_{1:n,1:n}$  is intended to be equal to  $(Z^{(k+1)} + \bar{\lambda}\hat{P})$ . Then, the algorithm stops when the following conditions are met:

$$\|R^P\| \leq m\varepsilon^{\text{ABS}} + \varepsilon^{\text{REL}} \max\{\|W^{(k)}\|, \|Q(\bar{\lambda}, Z^{(k)})\|, \|Y^{(k)}\|\}$$

$$\|R^D\| \leq m\sqrt{(n+1)}\varepsilon^{\text{ABS}} + \varepsilon^{\text{REL}} \|D \left( \begin{bmatrix} I_m & \\ -T_{1:n,1:n}^{-1} & T_{0,1:n}^\top \end{bmatrix} \times \right. \\ \left. M^{(k)} \begin{bmatrix} I_m & -T_{0,1:n} T_{1:n,1:n}^{-1} \end{bmatrix} \right)\|$$

where  $\varepsilon^{\text{ABS}}$  and  $\varepsilon^{\text{REL}}$  are the desired absolute and relative tolerances.

It remains to determine the optimal value  $\lambda^\circ$  for  $\lambda$  which solves Problem (36). To this aim, we exploit the following result (see [34, pp.87-88]):

*Proposition 5.1:* If  $f$  is convex in  $(x, y)$  and  $\mathcal{C}$  is a convex non-empty set, then the function

$$g(x) = \inf_{y \in \mathcal{C}} f(x, y) \quad (41)$$

is convex in  $x$ , provided that  $g(x) > -\infty$  for some  $x$ . The domain of  $g$  is the projection of  $\text{dom}(f)$  on its  $x$ -coordinates.

This result guarantees that the function

$$g(\lambda) = \min_{(W, Z) \in \mathcal{C}_\lambda} J(\lambda, W, Z)$$

is convex in  $\lambda$ . Hence, in order to determine  $\lambda^\circ = \arg \min_{\lambda > 0} g(\lambda)$  we can choose an initial interval of uncertainty  $[a, b]$  containing  $\lambda^\circ$ , and we progressively reduce it by evaluating  $g(\lambda)$  at two points within the interval placed symmetrically, each at distance  $h > 0$  from the midpoint. This is repeated until the width of the uncertainty interval is smaller than a certain tolerance  $l > 0$ .

The overall procedure to solve the dual problem (36) is summarized in Algorithm 1.

## VI. IDENTIFICATION OF ARMA FACTOR MODELS

In this section we extend the proposed approach to ARMA processes. Consider the ARMA factor model:

$$y(t) = a^{-1}(W_L u(t) + W_D w(t)) \quad (42)$$

where

$$a(e^{i\vartheta}) = \sum_{k=0}^p a_k e^{-i\vartheta k}, \quad a_k \in \mathbb{R}$$

and  $W_L, W_D, u$  and  $w$  are defined analogously to (1). Notice that  $y_{MA}(t) := ay(t) = W_L u(t) + W_D w(t)$  is a MA process of order  $n$  whose spectral density  $\Phi = W_L W_L^* + W_D W_D^* \in \mathcal{Q}_{m,n}$  admits a low rank plus diagonal decomposition. Finally, it is worth noting that it is not restrictive to assume that the

### Algorithm 1

**Input:**  $b > a > 0, l > 0, h > 0$

**Output:**  $(\lambda^\circ, W^\circ, Z^\circ)$

- 1: **repeat**
- 2:    $\tilde{a} = (a + b)/2 - h; \tilde{b} = (a + b)/2 + h.$
- 3:   Compute  $g(\tilde{a})$  by applying the ADMM with  $\lambda = \tilde{a}$ .
- 4:   Compute  $g(\tilde{b})$  by applying the ADMM with  $\lambda = \tilde{b}$ .
- 5:   **if**  $g(\tilde{a}) < g(\tilde{b})$  **then**
- 6:      $b = \tilde{b}$
- 7:   **else**
- 8:      $a = \tilde{a}$
- 9:   **end if**
- 10: **until**  $b - a < l$
- 11:  $\lambda^\circ = (a + b)/2.$
- 12: Compute  $(W^\circ, Z^\circ)$  by applying the ADMM with  $\lambda = \lambda^\circ.$

autoregressive part in (42) is characterized by a scalar filter  $a$ ; Indeed, any ARMA factor model can be written in the form of (42).

Assume now to collect a realization  $y^N = \{y(1) \dots y(N)\}$  of numerosity  $N$  of the process  $y$ . Our aim is to estimate the factor model (42) and the number of factors  $r$ . Before proceeding, the following observation needs to be made: there is an identifiability issue in the problem. Indeed, if we multiply  $a(z)$ ,  $W_L$  and  $W_D$  by an arbitrary non-zero real number  $c$ , the model remains the same. We can easily eliminate this uninteresting degree of freedom by normalizing the polynomial  $a(z)$ , so that from now on we assume  $a_0 = 1$ . In addition, for identifiability of model (42), we assume that there are not zero/pole cancellations between the roots of the denominator  $a(z)$  and the numerator matrices  $W_L(z)$  and  $W_D(z)$ . In other words, we assume that (42) is a minimal representation of the ARMA model.

The idea is to estimate first  $a$ , and then  $\Phi_L$  and  $\Phi_D$  by preprocessing  $y^N$  through  $a$ . In more detail, the proposed solution consists of the following two steps:

- 1) The *AR dynamic estimation*. Given the realization  $y^N$ , we estimate the  $p$  parameters of the filter  $a$  by applying the maximum likelihood estimator proposed in [35, Section II.b]. In doing so, we are estimating an AR process whose spectral density is  $a^{-1}(a^{-1})^* I_m$ .
- 2) The *MA dynamic factor analysis*. Let  $y_{MA}^N$  be the finite length trajectory obtained by passing through the filter  $a^\circ(e^{i\vartheta})$  the trajectory  $y^N$  with zero initial conditions. After computing the truncated periodogram  $\hat{\Phi} \in \mathcal{Q}_{m,n}$  from  $y_{MA}^N$ , we solve Problem (5) with  $\hat{\Phi}$  in order to recover the number of latent factors.

Although the above procedure is suboptimal, the numerical simulations showed that the resulting estimator of the number of factors performs well, see Section VII-B.

## VII. NUMERICAL SIMULATIONS

In this section, we test the performance of the proposed approach both for MA and ARMA factor models. In all the simulations, the parameter  $\delta$  is computed according to the empirical procedure of Section II-A for  $\alpha = 0.5$ . Then,

	Proposed method	Hallin and Liška method
m = 20, r=2	0.32	0.72
m = 30, r=2	0.12	0.22
m = 40, r=2	0.02	0.02
m = 20, r=4	1.1	2.98
m = 30, r=4	0.62	2.04
m = 40, r=4	0.32	1.64

TABLE I: Mean absolute error between the estimated rank and the true rank  $r$ .

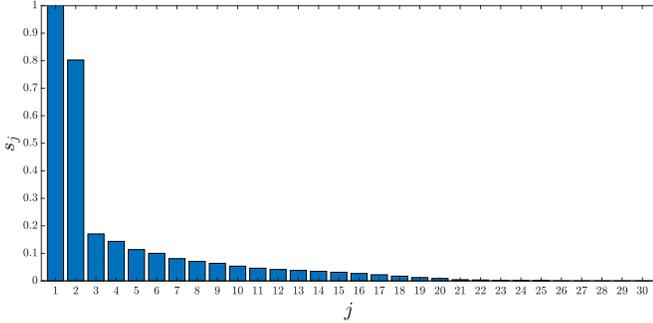


Fig. 1: Estimated MA factor model with  $n = 2$ ,  $m = 40$ , and  $r = 2$ . Integral over the unit circle of the first 30 normalized singular values of  $\Phi_L^\circ$  with  $N = 5000$ .

Problem (36) is solved by applying Algorithm 1 with  $l = 7$  and  $h = 3$ . In regard to the ADMM algorithm, we set  $\varepsilon^{\text{ABS}} = 10^{-4}$ ,  $\varepsilon^{\text{REL}} = 10^{-4}$  and the penalty term  $\rho = 0.05$ .

### A. Synthetic Example - MA factor models

We consider a Monte Carlo study composed by 50 experiments, where for each experiment:

- 1) We build an MA factor model (1) of order  $n = 2$ , with  $m$  manifest variables and  $r$  latent factors, computed by randomly generating the zeros of the transfer functions  $[W_L]_{(i,j)}$ 's and  $[W_D]_{(i,i)}$ 's for  $i = 1, \dots, m$ ,  $j = 1, \dots, r$  within the circle with center at the origin and radius 0.95 on the complex plane. The model is generated in such a way that  $\int \|\Phi_L(e^{i\theta})\| / \int \|\Phi_D(e^{i\theta})\| = 2$ , that is the idiosyncratic component is not negligible with respect to the latent variable.

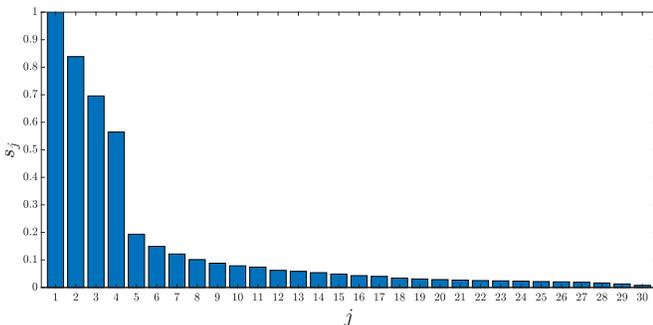


Fig. 2: Estimated MA factor model with  $n = 2$ ,  $m = 40$ , and  $r = 4$ . Integral over the unit circle of the first 30 normalized singular values of  $\Phi_L^\circ$  with  $N = 5000$

- 2) We generate from the model a sample  $y^N$  of length  $N = 5000$ .
- 3) We apply the proposed identification procedure to estimate the number of common factors. More precisely, we define

$$s_j := \int \frac{\sigma_j(\Phi_L^\circ(e^{i\theta}))}{\sigma_1(\Phi_L^\circ(e^{i\theta}))}$$

where  $\sigma_j(\Phi_L^\circ(e^{i\theta}))$  denotes the  $j$ -th largest eigenvalue of  $\Phi_L^\circ$  at frequency  $\theta$ . It is clear that  $s_j$  represents the integral of the  $j$ -th largest normalized singular value of  $\Phi_L^\circ$  over the unit circle. Let  $i_{max}$  be the first  $i$  such that  $s_{i+1} < 0.05$  and  $i_{min}$  the last  $i$  such that  $s_i > 0.3$ . Then, we define the “numerical rank” of  $\Phi_L^\circ$  as

$$r^\circ := \max_{i_{min} \leq i \leq i_{max}} s_i / s_{i+1}. \quad (43)$$

- 4) We compute the number of factors from the data sequence  $y^N$  by applying the method proposed by Hallin and Liška [30].
- 5) We assess the performance of the two estimators in terms of the mean absolute error

$$\bar{e} = \frac{1}{50} \sum_{i=1}^{50} |r - \hat{r}|$$

where  $\hat{r}$  is one of the two previous estimates and  $r$  is the true rank of the data generating process.

Table I shows the mean absolute error  $\bar{e}$  when  $r = 2, 4$  and  $m = 20, 30, 40$ .

We see that in the six Monte Carlo studies the proposed method outperforms the Hallin and Liška’s algorithm. In particular, the performances of the two estimators are comparable when the number  $m$  of observed variables is much higher than the number  $r$  of latent factors. The problem becomes more challenging when the ratio  $m/r$  decreases; improvements by our method are more sizable in these situations.

Figures 1 and 2 plot the quantities  $s_j$  obtained by applying our estimation method in one of the previous Monte Carlo experiment with  $m = 40$  manifest variables and  $r = 2$  and  $r = 4$  latent variables, respectively. We can notice that there is a knee point at  $j = 2$  and at  $j = 4$  in Figure 1 and Figure 2, respectively, so that we can recover the exact number of common factors in both cases.

### B. Synthetic Example - ARMA factor models

In order to test the robustness of the proposed algorithm also in the case of a more general model class, i.e. the ARMA model class considered in VI, we consider the following Monte Carlo simulation study composed of 50 experiments. We randomly build an ARMA factor model (42) with  $m = 40$ ,  $r = 2$ ,  $n = 2$  and  $p = 2$ ; without loss of generality we fix  $a_0 = 1$ . Then, for each Monte Carlo experiment a data sequence of length  $N = 5000$  is randomly generated from the model and the ARMA factor model identification procedure is performed. The boxplot of the quantities  $s_j$  for the estimated  $\Phi_L^\circ$ 's are shown in Figure 3 and it reveals that the proposed identification procedure is able to successfully recover the number of latent factors.

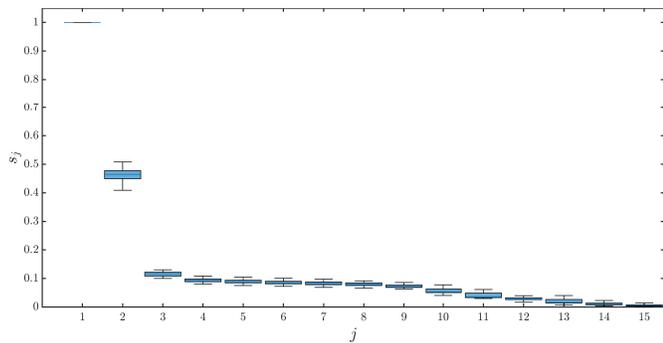


Fig. 3: Estimated ARMA factor model with  $m = 40$ ,  $r = 2$ ,  $n = 2$  and  $p = 2$ . Box-plot of the integral over the unit circle of the first 15 normalized singular values of  $\Phi_L^o$  with  $N = 5000$ .

### C. Smart Building Dataset

The SMLsystem is a house built in Valencia at the Universidad CEU Cardenal Herrera (CEU-UCH). It is a modular house that integrates a whole range of different technologies to improve energy efficiency, with the objective to construct a near zero-energy house. A complex monitoring system has been used in the SMLsystem: it has indoor sensors for temperature, humidity and carbon dioxide; outdoor sensors are also available for lighting measurements, wind speed, rain, sun irradiance and temperature. We refer the reader to [36] for a detailed description of the building and its monitoring system. Two datasets from the SMLsystem are available for download at the UCI Machine Learning repository <http://archive.ics.uci.edu/ml>. We take into account  $m = 17$  sensor signals extracted from these datasets: the indoor temperature (in  $^{\circ}\text{C}$ ) of the dinning-room and of the room, the weather forecast temperature (in  $^{\circ}\text{C}$ ), the carbon dioxide (in ppm) in the dinning room and in the room, the relative humidity (in %) in the dinning room and the room, the lighting in the dinning room and the room (in lx), the sun dusk, the wind (in cm/sec), the sun light (in klx) in the west, east and south facade, the sun irradiance (in dW), the outdoor temperature (in  $^{\circ}\text{C}$ ) and finally the outdoor relative humidity (in %). The data are sampled with a period of  $T = 15\text{min}$  and each sample is the mean of the last quarter, reducing in this way the signal noise. The first dataset  $y^{N_1} = \{y(1), \dots, y(N_1)\}$  was captured during March 2011 and has  $N_1 = 2764$  points ( $\approx 28$  days), while the second dataset  $y^{N_2} = \{y(N_1 + 1), \dots, y(N_1 + N_2)\}$  has  $N_2 = 1373$  points ( $\approx 14$  days) collected in June 2011.

It is reasonable to expect that the variability of the considered signals may be successfully explained by a smaller number of factors. Motivated by this reason, we apply the ARMA factor model identification procedure with parameters  $n = 2$  and  $p = 2$  using the realization  $y^{N_1}$ . As shown in Figure 4, we obtain an estimate of 4 latent factors.

For the sake of comparison, we also use the Matlab function `armax()` of the System Identification Toolbox to compute the prediction-error method (PEM) estimate for an ARMA model with polynomials  $A(z)$  and  $C(z)$ , diagonal, of order 2 from the realization  $y^{N_1}$ . It is well know that the PEM estimate

has guarantees of optimality, at least for large samples, for the identification of linear dynamic systems, so that it is interesting to use it as a benchmark to which to compare the prediction capability of our model.

We also compare our model with the factor model proposed in [28] where, however, the number  $r$  of latent factors is assumed to be a given input. We computed this input by preliminarily applying the Hallin and Liska's information criterion (IC). The paper [28] assumes an underlying AR factor model where the idiosyncratic components are temporarily and cross-sectionally uncorrelated and it employs the expectation maximization algorithm to compute the ML estimator. The selection of the model order both for our method and for the method proposed in [28] is obtained by applying the BIC criterion. Notice that since the IC has a random step, it gives different results when repeatedly applied to the same dataset  $y^{N_1}$ . In this particular case the IC method estimates either 1 or 4 latent factors: to be more than fair, we provided the method proposed in [28] with the input  $r = 4$  since the corresponding model explains the training data much better than the one corresponding to  $r = 1$ .

The second dataset  $y^{N_2}$  is used in the validation step to test the prediction capability of the three estimated models. The results are summarized in Figure 5 which displays for each output channel  $j = 1, \dots, m$  the fit (percentage) term:

$$J_{FIT,j} := 100 \left( 1 - \frac{\sqrt{\sum_{t=N_1+1}^{N_1+N_2} (y_j(t) - \hat{y}_j(t|t-1))^2}}{\sqrt{\sum_{t=N_1+1}^{N_1+N_2} (y_j(t) - \bar{y}_j)^2}} \right)$$

where  $\bar{y}_j := \frac{1}{N_2} \sum_{t=N_1+1}^{N_1+N_2} y_j(t)$  and  $\hat{y}_j(t|t-1)$  is the one-step ahead prediction at time  $t$  computed with zero initial conditions for the three estimated models. The figure shows that the ARMA factor model estimated with the proposed method matches quite well the measurement data  $y^{N_2}$ , reaching fit values that are essentially equal to the benchmark PEM estimate. This allows us to conclude that the available smart building dataset can be successfully modeled with the proposed method. It is a remarkable result since the factor model is parameterized by 257 coefficients, much less than the 612 coefficients of the PEM estimate. Not only the factor model is more parsimonious, but it is also able to organize the complex, high dimensional dataset in a suitable structured model, which is easier to understand and interpret. On the contrary, the PEM model is much more complex and does not give us any intuitive explanation of the underlying dynamics. Finally we notice the AR factor model obtained by applying the IC and the ML principle fails to provide good performances for several output channels.

We have repeated the numerical simulations with the SML-system dataset for different the values of the probability  $\alpha$ , specifically for  $\alpha = 0.2$  and  $\alpha = 0.8$ . We obtain that the estimated number of factors is still equal to 4 and the prediction capabilities of the model remain essentially the same: the algorithm appears to be robust with respect to the choice of  $\alpha$ .

We have also tested the proposed factor analysis method on the Smart Building dataset by changing the values of the tolerance

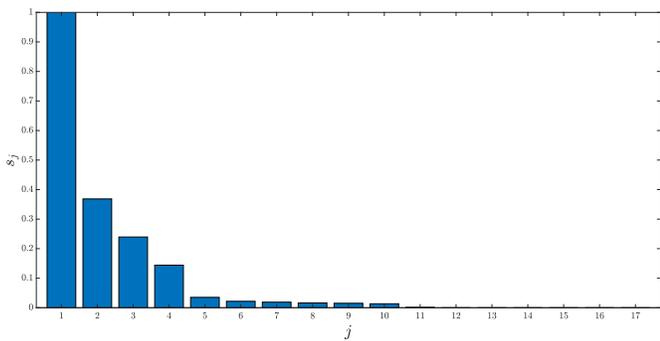


Fig. 4: Application of the ARMA factor models identification procedure by using the measurements  $y^{N_1}$  from the SMLsystem as training data. The figure shows the integral over the unit circle of the normalized singular values of  $\Phi_L^o$ .

parameters  $\varepsilon^{\text{ABS}}$  and  $\varepsilon^{\text{REL}}$ . By decreasing the value of the tolerances the computation time grows, whereas the resulting model remains essentially the same. On the other hand, for larger values of the tolerance parameters the ADMM algorithm may stop before reaching an accurate solution. As a matter of fact, by setting  $\varepsilon^{\text{REL}} = \varepsilon^{\text{ABS}} = 10^{-3}$ , the factor analysis procedure still recovers the exact number of latent factors, but the resulting model shows poorer performances.

*Remark.* Another important consequence of dealing with simple models such as those considered in our setting, is the possibility of identifying systems from a limited dataset. Indeed simple models have few parameters and hence highly reduce the risk of overfitting. To concretely show this advantage in our setting, we repeat the previous simulation by assuming that we have only access to the first 800 measurements ( $\approx 8$  days) of the Smart Building dataset  $y^{N_1}$  for the estimation step. We then compute the one-step ahead prediction capabilities of our ARMA factor model, the PEM model and the IC+ML model on the second dataset  $y^{N_2}$ . We find out that the proposed ARMA factor model provides the best performances, reaching an average fit term equal to 83.6% against the 45.9% and the 22.5% of the PEM and the IC+ML methods, respectively. In particular the PEM method completely fails to predict the channels number 11 and 12. We conclude that, differently from the PEM model, our ARMA factor model does not suffer from overfitting.

## VIII. CONCLUSION

A procedure to estimate the number of factors and to learn ARMA factor models has been proposed. This method is based on the solution of an optimization problem whose solution has been proven to exist via dual analysis. The simulations results applying the procedure both to synthetic and real data provide evidence of a good performance.

## APPENDIX

### Proof of Lemma 3.1

Since  $\Phi = \Delta X \Delta^*$  with  $X \succeq 0$ , there exists  $A \in \mathbb{R}^{m \times m(n+1)}$  such that  $X = A^T A$ . The matrix  $A$  is such that  $\Phi \succeq 0$  admits the spectral factorization  $\Phi = W W^*$

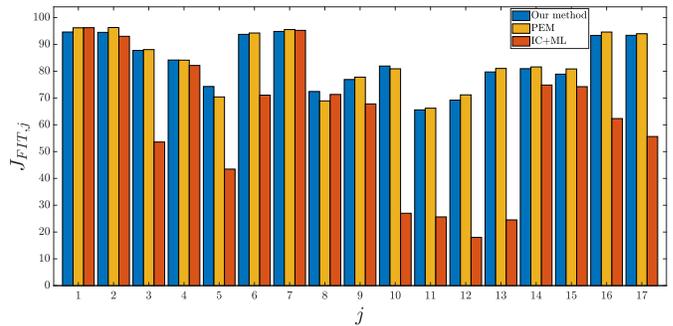


Fig. 5: Fit (in percentage) term  $J_{FIT,j}$  for each output channel for the model estimated via our factor analysis method, via PEM and via the IC+ ML approach. The fit values are computed by using the measurements  $y^{N_2}$  from the SMLsystem as validation data.

where  $W := \Delta A^T$ . Now, define  $\Phi_n := \Phi + \frac{1}{n}I$  with  $n \in \mathbb{N}$  and let  $W_n := \Delta A_n$  be a spectral factor of  $\Phi_n$  with  $A_n \in \mathbb{R}^{m \times m(n+1)}$ . Clearly,  $\lim_{n \rightarrow +\infty} \Phi_n = \Phi$ ; accordingly,  $\lim_{n \rightarrow +\infty} W_n = W$  and  $\lim_{n \rightarrow +\infty} A_n = A$ . Since  $\Phi_n \succ 0 \forall \vartheta$  we can exploit (10) to obtain

$$\int \log |\Phi_n| = \log |A_{n_0}^T A_{n_0}|.$$

Then, applying the limit operator to both sides, we have

$$\lim_{n \rightarrow +\infty} \int \log |\Phi_n| = \log |A_0^T A_0| = \log |X_{00}|.$$

To conclude the proof, it remains to show that in the left side of the previous equation it is possible to interchange the limit and the integral operators. To this aim, we introduce the sequence  $\{f_n\}_{n=1}^{+\infty}$  where  $f_n(t) := \log |\Phi_n(\vartheta)|$  and the function  $f(\vartheta) := \lim_{n \rightarrow +\infty} f_n(t) = \log |\Phi(\vartheta)|$ . Observe that, since the interval of integration  $[-\pi, \pi]$  is bounded and  $f_1(\vartheta) < +\infty$  for any  $\vartheta \in [-\pi, \pi]$ , then  $\int f_1(\vartheta) d\vartheta < +\infty$ . We also define the sequence  $\{g_n\}_{n=1}^{+\infty}$  as  $g_n(\vartheta) := f_n(\vartheta) - f_1(\vartheta)$  and  $g(\vartheta) := \lim_{n \rightarrow +\infty} g_n(\vartheta)$ .  $\{g_n\}$  is a pointwise non-increasing sequence of measurable non-positive functions,

$$\dots \leq g_2(\vartheta) \leq g_1(\vartheta) \leq 0, \quad \forall \vartheta \in [-\pi, +\pi]$$

converging to  $g(\vartheta)$  from above. Hence, it satisfies all the hypotheses of Beppo-Levi's monotone convergence theorem (applied with opposite signs), from which it immediately follows that

$$\lim_{n \rightarrow +\infty} \int g_n(\vartheta) = \int g(\vartheta),$$

and consequently

$$\lim_{n \rightarrow +\infty} \int f_n(\vartheta) = \int g(\vartheta) + \int f_1(\vartheta). \quad (44)$$

Now, since  $f_1(\vartheta) < +\infty$  for all  $\vartheta$ ,

$$g(\vartheta) = f(\vartheta) - f_1(\vartheta), \quad (45)$$

and, by plugging (45) into (44), we finally obtain

$$\lim_{n \rightarrow +\infty} \int f_n(\vartheta) = \int f(\vartheta). \quad \blacksquare$$

*Proof of Proposition 3.3*

Consider a sequence  $(\lambda^{(k)}, V^{(k)}, Z^{(k)})_{k \in \mathbb{N}}$  in  $\mathcal{C}_2$ .

We first show that  $[Z^{(k)}]_0$  cannot diverge. Indeed, assume by contradiction that  $\lim_{k \rightarrow \infty} \|[Z^{(k)}]_0\| = +\infty$ . Since it is a symmetric and traceless matrix, this implies

$$\lim_{k \rightarrow \infty} \min_{\alpha^{(k)} \in \sigma([Z^{(k)}]_0)} \alpha^{(k)} = -\infty. \quad (46)$$

In view of (46), since  $\lambda^{(k)} \hat{P}_0$  is bounded and  $V^{(k)}$  positive semidefinite  $\forall k$ , then  $(\lambda^{(k)} \hat{P}_0 + [Z^{(k)}]_0 - [V^{(k)}]_{00})$  has at least a negative eigenvalue for  $k$  sufficiently large, so that the sequence  $(\lambda^{(k)}, V^{(k)}, Z^{(k)})$  is not in  $\mathcal{C}_2$ . We conclude that

$$\lim_{k \rightarrow \infty} \|[Z^{(k)}]_0\| < \infty.$$

As a consequence, since  $\beta I \preceq T(Z^{(k)}) - V^{(k)} \preceq I$  (which is one of the condition for the sequence to be in  $\mathcal{C}_2$ ), and  $[T(Z^{(k)})]_{hh} = [Z^{(k)}]_0$  by construction, it holds that  $\forall k$

$$\|[V^{(k)}]_{hh}\| < \infty, \quad h = 0, \dots, n.$$

Then, from  $V^{(k)} \succeq 0$  it follows that also the off-diagonal blocks of  $V^{(k)}$  must be bounded  $\forall k$ , i.e.

$$\|[V^{(k)}]_{hl}\| < \infty, \quad l \neq h, \quad l, h = 0, \dots, n. \quad (47)$$

Finally, by the boundedness of  $(T(Z^{(k)}) - V^{(k)})$  and by (47) we obtain that  $\forall k$

$$\|[Z^{(k)}]_h\| < \infty \quad h = 1, \dots, n, \quad (48)$$

which concludes the proof. ■

REFERENCES

- [1] L. Ning, T. T. Georgiou, A. Tannenbaum, and S. P. Boyd, "Linear models based on noisy data and the Frisch scheme," *SIAM Review*, vol. 57, no. 2, pp. 167–197, 2015.
- [2] D. Bertsimas, M. S. Copenhaver, and R. Mazumder, "Certifiably optimal low rank factor analysis," *Journal of Machine Learning Research*, vol. 18, no. 29, pp. 1–53, 2017.
- [3] V. Ciccone, A. Ferrante, and M. Zorzi, "Learning latent variable dynamic graphical models by confidence sets selection," *Kybernetika*, vol. 55, no. 4, pp. 74–754, 2019.
- [4] M. Zorzi and R. Sepulchre, "Factor analysis of moving average processes," in *European Control Conference (ECC)*, Linz, 2015, pp. 3579–3584.
- [5] V. Ciccone, A. Ferrante, and M. Zorzi, "Factor models with real data: A robust estimation of the number of factors," *IEEE Transactions on Automatic Control*, vol. 64, no. 6, pp. 2412–2425, June 2019.
- [6] —, "Learning latent variable dynamic graphical models by confidence sets selection," *IEEE Transactions on Automatic Control*, vol. 65, no. 12, pp. 5130–5143, 2020.
- [7] J. Geweke, "The dynamic factor analysis of economic time series," in *Latent variables in socio-economic models*, D. Aigner and A. Goldberger, Eds. Amsterdam: North-Holland, 1977.
- [8] M. Deistler and C. Zinner, "Modelling high-dimensional time series by generalized linear dynamic factor models: An introductory survey," *Communications in Information & Systems*, vol. 7, no. 2, pp. 153–166, 2007.
- [9] J. Stock and M. Watson, "Dynamic factor models," 2010, internal report. [Online]. Available: [https://www.princeton.edu/~mwatson/papers/dfm\\_oup\\_4.pdf](https://www.princeton.edu/~mwatson/papers/dfm_oup_4.pdf)
- [10] G. Figà-Talamanca, S. Focardi, and M. Patacca, "Common dynamic factors for cryptocurrencies and multiple pair-trading statistical arbitrages," *Decisions in Economics and Finance*, 2021.
- [11] G. Bottegal and G. Picci, "Modeling complex systems by generalized factor analysis," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 759–774, 2015.
- [12] J. Songsiri and L. Vandenberghe, "Topology selection in graphical models of autoregressive processes," *Journal of Machine Learning Research*, vol. 11, no. Oct, pp. 2671–2705, 2010.
- [13] E. Avventi, A. Lindquist, and B. Wahlberg, "ARMA identification of graphical models," *IEEE Trans. Autom. Control*, vol. 58, no. 5, pp. 1167–1178, May 2013.
- [14] S. Maanan, B. Dumitrescu, and C. Giurcăneanu, "Conditional independence graphs for multivariate autoregressive models by convex optimization: Efficient algorithms," *Signal Processing*, vol. 133, pp. 122–134, 2017.
- [15] D. Alpagò, M. Zorzi, and A. Ferrante, "Identification of sparse reciprocal graphical models," *IEEE Control Systems Letters*, vol. 2, no. 4, pp. 659–664, Oct 2018.
- [16] —, "A scalable strategy for the identification of latent-variable graphical models," *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3349–3362, 2022.
- [17] M. Zorzi and R. Sepulchre, "AR identification of latent-variable graphical models," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2327–2340, Sept 2016.
- [18] S. Maanan, B. Dumitrescu, and C. Giurcăneanu, "Maximum entropy expectation-maximization algorithm for fitting latent-variable graphical models to multivariate time series," *Entropy*, vol. 20, p. 76, 01 2018.
- [19] M. Zorzi, "Empirical Bayesian learning in AR graphical models," *Automatica*, vol. 109, 2019.
- [20] M. Zorzi, "Autoregressive identification of Kronecker graphical models," *Automatica*, vol. 119, p. 109053, 2020.
- [21] M. S. Veedu, H. Doddi, and M. V. Salapaka, "Topology learning of linear dynamical systems with latent nodes using matrix decomposition," 2020.
- [22] M. S. Veedu and M. V. Salapaka, "Topology identification under spatially correlated noise," 2020.
- [23] M. Lippi, M. Deistler, and B. Anderson, "High-dimensional dynamic factor models: A selective survey and lines of future research," *Econometrics and Statistics*, 2022.
- [24] M. Forni, M. Hallin, M. Lippi, and L. Reichlin, "The generalized dynamic-factor model: Identification and estimation," *The Review of Economics and Statistics*, vol. 82, no. 4, pp. 540–554, 2000. [Online]. Available: <http://www.jstor.org/stable/2646650>
- [25] —, "The generalized dynamic factor model: One-sided estimation and forecasting," *Journal of the American Statistical Association*, vol. 100, no. 471, pp. 830–840, 2005. [Online]. Available: <http://www.jstor.org/stable/27590616>
- [26] M. Forni, M. Hallin, M. Lippi, and P. Zaffaroni, "Dynamic factor models with infinite-dimensional factor spaces: One-sided representations," *Journal of Econometrics*, vol. 185, no. 2, pp. 359–371, 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0304407614002693>
- [27] —, "Dynamic factor models with infinite-dimensional factor space: Asymptotic analysis," *Journal of Econometrics*, vol. 199, no. 1, pp. 74–92, 2017.
- [28] C. Doz, D. Giannone, and L. Reichlin, "A quasi—maximum likelihood approach for large, approximate dynamic factor models," *The Review of Economics and Statistics*, vol. 94, no. 4, pp. 1014–1024, 2012. [Online]. Available: <http://www.jstor.org/stable/23355337>
- [29] J. Bai and S. Ng, "Determining the number of factors in approximate factor models," *Econometrica*, vol. 70, no. 1, pp. 191–221, 2002.
- [30] M. Hallin and R. Liška, "Determining the number of factors in the general dynamic factor model," *Journal of the American Statistical Association*, vol. 102, no. 478, pp. 603–617, 2007. [Online]. Available: <https://doi.org/10.1198/016214506000001275>
- [31] L. Falconi, "Robust factor analysis of moving average processes," 2021, master thesis. [Online]. Available: <http://tesi.cab.unipd.it/65215/>
- [32] J. Songsiri, J. Dahl, and L. Vandenberghe, "Graphical models of autoregressive processes," *Convex optimization in signal processing and communications*, pp. 89–116, 2010.
- [33] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011. [Online]. Available: <http://dx.doi.org/10.1561/22000000016>
- [34] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge, United Kingdom: Cambridge University Press, 2004.
- [35] F. Crescente, L. Falconi, F. Rozzi, A. Ferrante, and M. Zorzi, "Learning ar factor models," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 274–279.
- [36] F. Zamora-Martínez, P. Romeu, P. Botella-Rocamora, and J. Pardo, "On-line learning of indoor temperature forecasting models towards energy efficiency," *Energy and Buildings*, vol. 83, pp. 162 – 172,

2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778814003569>



**Lucia Falconi** received the M.S. degree in Automation Engineering (*cum laude*) from the University of Padova, Padova, Italy, in 2020. She is currently pursuing the Ph.D. degree in Information Engineering from the University of Padova. Her research interests are in the areas of systems identification, robust estimation and stochastic systems.



**Augusto Ferrante** was born in Piove di Sacco, Italy, on August 5, 1967. He received the “Laurea” degree, *cum laude*, in Electrical Engineering in 1991 and the Ph.D. degree in Control Systems Engineering in 1995, both from the University of Padova.

He has been a faculty member of the Colleges of Engineering of the University of Udine and of the “Politecnico di Milano”. He is presently Professor in the “Department of Information Engineering” of the University of Padova.

His research interests are in the areas of linear systems, spectral estimation, optimal control and optimal filtering, quantum control, and stochastic realization.



**Mattia Zorzi** received the M.S. degree in Automation Engineering and the Ph.D. degree in Information Engineering from the University of Padova, Padova, Italy, in 2009 and 2013, respectively. He held Postdoctoral appointments with the Department of Electrical Engineering and Computer Science, University of Liege, Liege, Belgium, and with the Human Inspired Technology Research Centre, University of Padova, Padova, Italy. He held visiting positions with the Department of Electrical and Computer Engineering, University of California, Davis, USA, and with the Department of Engineering, University of Cambridge, Cambridge, U.K., in 2011 and 2013-2014, respectively. He is currently an Associate Professor with the Department of Information Engineering, University of Padova. His current research interests include machine learning, robust estimation, identification theory.

Dr. Zorzi has been an Associate Editor of IEEE Control Systems Letters since 2019. He serves as an Associate Editor on the IEEE Control System Society Conference Editorial Board since 2017 and on the International Program Committee of IFAC World Congress 2020. He is a member of the IFAC Technical Committee on Modelling, Identification and Signal Processing.