Blind Identification of Graph Filters with Sparse Inputs

Santiago Segarra, Gonzalo Mateos, Antonio G. Marques, and Alejandro Ribeiro

***Abstract*—Network processes are often conceptualized as signals defined on the vertices of a graph. To untangle the latent structure of such signals, one can be view them as outputs of *linear graph filters* modeling underlying network dynamics. This paper deals with the problem of blind graph filter identification, which finds applications in social and brain networks, to name a few. Given a graph signal $\mathbf{y}$ modeled as the output of a graph filter, the goal is to recover the vector of filter coefficients $\mathbf{h}$ and the input signal $\mathbf{x}$ which is assumed to be sparse. While the filtered graph signal is a bilinear function of $\mathbf{x}$ and $\mathbf{h}$, $\mathbf{y}$ is also a *linear combination* of the entries of the rank-one matrix $\mathbf{x}\mathbf{h}^T$. As with blind deconvolution of time (or spatial) domain signals, it is shown that the blind graph filter identification problem can be tackled via rank minimization subject to linear constraints. Graph-dependent conditions under which the solution set of the optimization problem includes only rank-one matrices are derived. Numerical tests with synthetic and real-world networks corroborate the effectiveness of the blind identification approach.**

***Keywords*—Graph signal processing, blind system identification, graph filter, graph process.**

I. Introduction

Coping with the challenges posed by fields such as network science and big data necessitates broadening the scope beyond classical time-varying signal analysis and processing, to also accommodate *signals defined on graphs* [1], [2], [12]. Under the assumption that the signal properties are related to the topology of the graph where they are supported, the goal of graph signal processing is to develop algorithms that fruitfully leverage this topology. A natural way to achieve this is to rely on the so-called graph-shift operator, which is a matrix that reflects the local connectivity of the graph [2].

We consider here that each node has a certain value, and these values are collected across nodes to form a graph signal. With this definition, graph filters – which are a generalization of classical time-invariant systems – are a specific class of operators whose input and output are graph signals. Mathematically, graph filters are linear transformations that can be expressed as polynomials of the graph-shift operator [3]. The polynomial coefficients determine completely the linear transformation and are referred to as *filter coefficients*. Such transformations can be implemented via local interactions among nodes, and may be used to model underlying processes, e.g., diffusion or percolation dynamics in the network.

This paper investigates the problem of blind identification of graph filters. More specifically, we are given a graph signal $\mathbf{y}$, postulate that the observed $\mathbf{y}$ can be modeled as the output of a graph filter, and seek to *jointly identify* the *filter* coefficients $\mathbf{h}$ and the *input signal* $\mathbf{x}$ that gave rise to $\mathbf{y}$. Since the problem is underdetermined, we assume that the length of $\mathbf{h}$ is small and that $\mathbf{x}$ is sparse. This is the case when e.g., a few seeding nodes inject a signal that is diffused throughout a network [4]. For most part of the analysis we consider that the support of the input signal $\mathbf{x}$ is known, however, the more challenging case of unknown support is also discussed. Applications of interest include opinion formation in social networks, inverse problems of biological signals supported on graphs, and modeling and estimation of diffusion processes.

Notation and the formal definition of a graph filter are introduced in Section II. Section III starts by formulating the blind graph filter identification problem as a bilinear optimization over $\mathbf{h}$ and $\mathbf{x}$. Based on the ideas in [5] for blind deconvolution of time (or spatial) domain signals and leveraging the particular structure of a graph filter, the problem is recast as a linear inverse problem to recover the rank-one matrix $\mathbf{x}\mathbf{h}^T$. The linear operator that maps the sought rank-one matrix to the output signal $\mathbf{y}$ is shown to depend on the spectral properties of the graph-shift operator. Conditions under which both problems are equivalent – which depend on the particular graph – are provided. In Section III-A, the rank minimization problem is relaxed using the convex nuclear-norm surrogate [8], and efficient solvers are briefly discussed. Extensions to identify the support of the input sparse signal $\mathbf{x}$ are discussed in Section III-B. Numerical tests illustrating the performance of the proposed algorithm on synthetic and real-world graphs are given in Section IV.[1]

II. Graph signals and graph filters

Let $\mathcal{G}$ denote a directed graph with a set of nodes $\mathcal{N}$ (with cardinality N) and a set of links $\mathcal{E}$, such that if node i is connected to j, then $(i,j) \in \mathcal{E}$. Since $\mathcal{G}$ is directed, the set $\mathcal{N}_i : \{j \ \ |(j,i) \in \mathcal{E}\}$ stands for the (incoming) neighborhood of i. For any given $\mathcal{G}$ we define the adjacency matrix $\mathbf{A} \in \mathbb{R}^{N\times N}$ as a sparse matrix with nonzero elements A_{ji} if and only if $(i,j) \in \mathcal{E}$. The value of A_{ji} captures the strength of the connection from i to j. The focus of the paper is on analyzing (graph) signals defined on $\mathcal{N}$. These signals can be represented as vectors $\mathbf{x} = [x_1, ..., x_N]^T \in \mathbb{R}^N$, where the i-th component represents the value of the signal at node i.

The graph $\mathcal{G}$ can be endowed with the so-called *graph-shift operator* $\mathbf{S}$ [2], [3]. The shift $\mathbf{S} \in \mathbb{R}^{N\times N}$ is a matrix whose entry S_{ji} can be nonzero only if $i = j$ or if $(i,j) \in \mathcal{E}$. The sparsity pattern of the matrix $\mathbf{S}$ captures the local structure of

Work in this paper is supported by USA NSF CCF-1217963 and Spanish MINECO TEC2013-41604-R. S. Segarra and A. Ribeiro are with the Dept. of ESE., Univ. of Pennsylvania. G. Mateos is with the Dept. of ECE, Univ. of Rochester, A. G. Marques is with the Dept. of TSC, King Juan Carlos Univ. Emails: ssegarra@seas.upenn.edu, gmateosb@ece.rochester.edu, antonio.garcia.marques@urjc.es, and aribeiro@seas.upenn.edu.

[1] **Notation:** Generically, the entries of a matrix $\mathbf{X}$ and a (column) vector $\mathbf{x}$ will be denoted as X_{ij} and x_i; $\text{diag}(\mathbf{x})$ is a diagonal matrix satisfying $[\text{diag}(\mathbf{x})]_{ii} = [\mathbf{x}]_i$; $\mathbf{0}$ is the all-zero vector; and $\text{mod}_N(x)$ denotes the modulus (remainder) obtained after dividing x by N.

$\mathcal{G}$, but we make no specific assumptions on the values of its nonzero entries The intuition behind $\mathbf{S}$ is to capture a linear transformation that can be computed locally at the nodes of the graph. More rigorously, if $\mathbf{y}$ is defined as $\mathbf{y} = \mathbf{S}\mathbf{x}$, then node i can compute y_i provided that it has access to the value of x_j at $j \in \mathcal{N}_i$. Widely-used choices for $\mathbf{S}$ are the adjacency matrix $\mathbf{A}$ [2], [3] and the graph Laplacian [1]. We assume henceforth that $\mathbf{S}$ is diagonalizable, so that $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ with $\mathbf{\Lambda} \in \mathbb{R}^{N\times N}$ being diagonal. The shift $\mathbf{S}$ can be used to define graph-signal *operators* of the form

$$\mathbf{H} := \textstyle\sum_{l=0}^{L-1} h_l \mathbf{S}^l, \tag{1}$$

which are called *graph filters* [2]. For a given input $\mathbf{x}$, the output of the filter is simply $\mathbf{y} = \mathbf{H}\mathbf{x}$. The coefficients of the filter are collected into $\mathbf{h} := [h_0, \ldots, h_{L-1}]^T$, with $L-1$ denoting the filter degree. Graph filters are of particular interest because they represent linear transformations that can be implemented locally [4].

Leveraging the spectral decomposition of $\mathbf{S}$, graph filters and signals can be represented in the frequency domain. To be precise, let us use the eigenvectors of $\mathbf{S}$ to define the $N \times N$ matrix $\mathbf{U} := \mathbf{V}^{-1}$, and the eigenvalues of $\mathbf{S}$ to define the $N\times L$ Vandermonde matrix $\mathbf{\Psi}$, where $\Psi_{i,j} := (\Lambda_{ii})^{j-1}$. Using these conventions, the frequency representation of a *signal* $\mathbf{x}$ and of a *filter* $\mathbf{h}$ is defined as $\widehat{\mathbf{x}} := \mathbf{U}\mathbf{x}$ and $\widehat{\mathbf{h}} := \mathbf{\Psi}\mathbf{h}$, respectively.

Lemma 1 *The output* $\mathbf{y} = \mathbf{H}\mathbf{x}$ *of a graph filter in the frequency domain is given by*

$$\widehat{\mathbf{y}} = \mathrm{diag}\big(\mathbf{\Psi}\mathbf{h}\big)\mathbf{U}\mathbf{x} = \mathrm{diag}\big(\widehat{\mathbf{h}}\big)\widehat{\mathbf{x}}. \tag{2}$$

Proof : We first use $\mathbf{y} = \mathbf{H}\mathbf{x}$ to write $\mathbf{U}\mathbf{y} = \mathbf{U}\mathbf{H}\mathbf{x}$. Given the spectral decomposition of $\mathbf{S}$, it holds that $\mathbf{H} = \sum_{l=0}^{L-1} h_l \mathbf{S}^l = \mathbf{V}\big(\sum_{l=0}^{L-1} h_l \mathbf{\Lambda}^l\big)\mathbf{U}$. Combining both expressions yields $\mathbf{U}\mathbf{y} = \mathbf{U}\mathbf{H}\mathbf{x} = \big(\sum_{l=0}^{L-1} h_l \mathbf{\Lambda}^l\big)\mathbf{U}\mathbf{x}$. Since the diagonal matrix $\sum_{l=0}^{L-1} h_l \mathbf{\Lambda}^l$ can be written as $\mathrm{diag}(\mathbf{\Psi}\mathbf{h})$, it holds that $\mathbf{U}\mathbf{y} = \mathrm{diag}(\mathbf{\Psi}\mathbf{h})\mathbf{U}\mathbf{x}$. Using the definitions of $\widehat{\mathbf{y}}$, $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{h}}$, (2) follows. ∎

Lemma 1 shows that, similar to what happens in the classical time domain, $\widehat{\mathbf{y}}$ is the product of $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{h}}$. To establish further connections with time-varying signals, let us consider the directed cycle graph whose adjacency matrix $\mathbf{A}_{dc}$ is zero, except for entries $A_{i,j}$ such that $i = \mathrm{mod}_N(j) + 1$, which are one. If $\mathbf{S} = \mathbf{A}_{dc}$, it is easy to check that: i) $\mathbf{y} = \mathbf{H}\mathbf{x}$ can be found as the circular convolution of $\mathbf{h}$ and $\mathbf{x}$, and ii) both $\mathbf{U}$ and $\mathbf{\Psi}$ correspond to the Discrete Fourier Transform (DFT) matrix. Hence, while in the time domain $\mathbf{U} = \mathbf{\Psi}$, this is not true in the more general graph domain.

III. Blind identification of graph filters

Suppose we observe the output signal $\mathbf{y} = \mathbf{H}\mathbf{x}$, and the shift operator $\mathbf{S}$ as well as the degree of the filter $L-1$ are given. The present paper deals with *blind identification* of the graph filter (and its input signal), which in this context amounts to estimating $\mathbf{x}$ and the filter coefficients $\mathbf{h}$ from the given information. Such a challenging problem is a natural extension to graphs of classical blind system identification, or blind deconvolution of signals in the time or spatial domains.

As stated the problem is ill-posed, since the number of unknowns $N + L$ in $\{\mathbf{x}, \mathbf{h}\}$ exceed those N observations in $\mathbf{y}$. To make the problem feasible one could impose further structural constraints on $\mathbf{x}$, thus reducing the effective degrees of freedom. To that end we will henceforth assume that: **(As)** the input graph signal $\mathbf{x}$ is S-sparse, with known support $\mathrm{supp}(\mathbf{x}) := \{i \,|\, x_i \neq 0\}$. We will briefly discuss the unknown support case in Section III-B, but a thorough treatment is beyond the scope of this paper and will be reported elsewhere. Next, collect the S nonzero entries of $\mathbf{x}$ in the vector $\bar{\mathbf{x}} \in \mathbb{R}^S$, and denote by $\bar{\mathbf{U}} \in \mathbb{R}^{N\times S}$ the matrix formed by the columns of $\mathbf{U}$ indexed by $\mathrm{supp}(\mathbf{x})$. Under **(As)** the frequency representation of the sparse input signal is $\widehat{\mathbf{x}} = \mathbf{U}\mathbf{x} = \bar{\mathbf{U}}\bar{\mathbf{x}}$, and recovery of $\{\mathbf{x}, \mathbf{h}\}$ is equivalent to the recovery of $\{\bar{\mathbf{x}}, \mathbf{h}\}$.

Lemma 2 *Under* ***(As)*** *the graph filter output* $\mathbf{y} = \mathbf{H}\mathbf{x}$ *is*

$$\mathbf{y} = \mathbf{V}\big(\mathbf{\Psi}^T \odot \bar{\mathbf{U}}^T\big)^T \mathrm{vec}\big(\bar{\mathbf{x}}\mathbf{h}^T\big) \tag{3}$$

where $\odot$ *denotes the Khatri-Rao (i.e., columnwise Kronecker) product, and* $\mathrm{vec}(\cdot)$ *is the matrix vectorization operator.*

Proof : Let $\bar{\mathbf{u}}_i^T$ denote the i-th row of $\bar{\mathbf{U}}$, and likewise let $\boldsymbol{\psi}_i^T$ be the i-th row of $\mathbf{\Psi}$. Since $\widehat{\mathbf{x}} = \bar{\mathbf{U}}\bar{\mathbf{x}}$, it follows from (2) that $\widehat{y}_i = (\boldsymbol{\psi}_i^T\mathbf{h})(\bar{\mathbf{u}}_i^T\bar{\mathbf{x}}) = \big(\boldsymbol{\psi}_i^T \otimes \bar{\mathbf{u}}_i^T\big)\mathrm{vec}\big(\bar{\mathbf{x}}\mathbf{h}^T\big)$, where $\otimes$ denotes the Kronecker product. Upon stacking the entries $\widehat{y}_i$ for $i = 1, \ldots, N$ to form $\widehat{\mathbf{y}}$ and then using $\mathbf{y} = \mathbf{V}\widehat{\mathbf{y}}$, the result follows by identifying $\boldsymbol{\psi}_i^T \otimes \bar{\mathbf{u}}_i^T$ with the i-th row of $\big(\mathbf{\Psi}^T \odot \bar{\mathbf{U}}^T\big)^T$. ∎

While the filtered graph signal $\mathbf{y}$ is a bilinear function of $\bar{\mathbf{x}}$ and $\mathbf{h}$, Lemma 2 also shows that $\mathbf{y}$ is a *linear* combination of the entries of the rank-one matrix $\mathbf{Y} := \bar{\mathbf{x}}\mathbf{h}^T \in \mathbb{R}^{S\times L}$. In other words, there exists a linear mapping $\mathcal{M} : \mathbb{R}^{S\times L} \mapsto \mathbb{R}^N$ such that $\mathbf{y} = \mathcal{M}(\mathbf{Y})$. Note that $\mathcal{M}$ can be expressed in terms of a matrix multiplication with $\mathbf{M} := \mathbf{V}\big(\mathbf{\Psi}^T \odot \bar{\mathbf{U}}^T\big)^T \in \mathbb{R}^{N\times LS}$, since $\mathbf{y} = \mathbf{M}\mathrm{vec}(\mathbf{Y})$ as per (3). Building on the ideas in [5], one can thus pose the blind graph filter identification problem as a linear inverse problem, where the goal is to recover a rank-one $S \times L$ matrix $\mathbf{Y}$ from observations $\mathbf{y} = \mathcal{M}(\mathbf{Y})$. A natural formulation to tackle such inverse problem is

$$\min_{\mathbf{Y}} \mathrm{rank}(\mathbf{Y}), \quad \text{s. to } \mathbf{y} = \mathbf{V}\big(\mathbf{\Psi}^T \odot \bar{\mathbf{U}}^T\big)^T \mathrm{vec}\big(\mathbf{Y}\big). \tag{4}$$

A basic question is whether (4) is equivalent to the original blind identification problem. To give a rigorous answer, some definitions are introduced next. Given a set of row indices $\mathcal{I}$, define the complement set of indices $\bar{\mathcal{I}} := \{1, \ldots, N\}\backslash\mathcal{I}$ and the matrix $\bar{\mathbf{U}}_{\mathcal{I}}$ formed by the rows of $\bar{\mathbf{U}}$ indexed by $\mathcal{I}$. Moreover, for a given graph-shift operator $\mathbf{S}$ and $\mathrm{supp}(\mathbf{x})$ – fixed $\mathbf{V}$, $\mathbf{\Psi}$, and, $\bar{\mathbf{U}}$ – define the set $\mathcal{O}_{\mathbf{y}}$ of matrix minimizers of (4) as a function of $\mathbf{y}$. Then, the following result on the validity of the matrix problem formulation in (4) holds.

Proposition 1 *Let* $|\cdot|$ *denote the number of non-repeated elements of a set and* $\mathcal{I}_{S-1}$ *be the set of the row indices such that* $\mathrm{rank}(\bar{\mathbf{U}}_{\mathcal{I}_{S-1}}) \leq S-1$. *Then*

$$\mathcal{O}_y = \big\{\bar{\mathbf{x}}\mathbf{h}^T \,\big|\, \mathbf{y} = \textstyle\sum_{l=0}^{L-1} h_l \mathbf{S}^l \mathbf{x}\big\}, \tag{5}$$

for all $\mathbf{y}$ *if and only if*

$$\min_{\mathcal{I}_{S-1}} \big|\{\lambda_i\}_{i\in\bar{\mathcal{I}}_{S-1}}\big| > L-1. \tag{6}$$

Proof: If we show that (6) is violated if and only if there exists a rank-1 matrix $\mathbf{Y}=\bar{\mathbf{x}}\mathbf{h}^T$ such that $\mathbf{V}\big(\mathbf{\Psi}^T\odot\bar{\mathbf{U}}^T\big)^T\text{vec}\big(\mathbf{Y}\big)=\mathbf{0}$, then Corollary 1 in [6] completes the proof. Leveraging the fact that $\mathbf{V}$ is full-rank, the above equality can be written elementwise as $(\boldsymbol{\psi}_i^T\mathbf{h})(\bar{\mathbf{u}}_i^T\bar{\mathbf{x}})=0$ for $i=1,\ldots,N$, where $\boldsymbol{\psi}_i^T$ denotes the i-th row of $\mathbf{\Psi}$ and similarly for $\bar{\mathbf{U}}$. Since $\text{rank}(\bar{\mathbf{U}}_{\mathcal{I}_{S-1}})\leq S-1$, there exists $\bar{\mathbf{x}}\neq\mathbf{0}$ such that the elementwise equality holds for $i\in\mathcal{I}_{S-1}$. Exploiting the Vandermonde structure of $\mathbf{\Psi}$, it follows that $\mathbf{h}\neq\mathbf{0}$ satisfying the equality for $i\in\bar{\mathcal{I}}_{S-1}$ can be found if and only if (6) is violated, concluding the proof. ■

Ideally, when solving (4) for some output $\mathbf{y}$ one should recover the set of outer products of all possible combinations of inputs $\bar{\mathbf{x}}$ and filter coefficients $\mathbf{h}$ that can give rise to such output [cf. (5)]. This is not true in general [6, Theorem 1], however, Proposition 1 states conditions on the signal support and the graph-shift operator [cf. (6)] for the desired equivalence to hold. For the particular case of the directed cycle graph, we may select the support of $\mathbf{x}$ so that every choice of S rows of $\bar{\mathbf{U}}$ forms a full-rank matrix. Consequently, the cardinality in (6) is equal to $N-S+1$ entailing the following corollary.

Corollary 1 *If* $\mathbf{S}=\mathbf{A}_{dc}$ *and the support of* $\mathbf{x}$ *consists of either S adjacent or S equally spaced nodes then* (5) *holds if and only if* $N>L+S-2$.

Notice that condition (6) does not guarantee that the solution of (4) is unique, but rather that the outer product of the desired sparse signal and filter coefficients is contained in $\mathcal{O}_{\mathbf{y}}$. Conditions that guarantee identifiability of (4), i.e. uniqueness of solution, are left as future work.

A. Algorithmic approach via convex relaxation

Albeit natural, the rank minimization problem in (4) is NP-hard to optimize; see also [7]. The nuclear norm $\|\mathbf{Y}\|_*=\sum_k\sigma_k(\mathbf{Y})$, where $\sigma_k(\mathbf{Y})$ denotes the k-th singular value of $\mathbf{Y}$, is typically adopted as a convex surrogate to rank$(\mathbf{Y})$ [7], [8]. Accordingly, a convex heuristic is to solve

$$\min_{\mathbf{Y}}\|\mathbf{Y}\|_*,\quad\text{s. to }\mathbf{y}=\mathbf{V}\big(\mathbf{\Psi}^T\odot\bar{\mathbf{U}}^T\big)^T\text{vec}\big(\mathbf{Y}\big),\qquad(7)$$

hoping that the optimal solution has rank one, so that we can recover $\bar{\mathbf{x}}$ and $\mathbf{h}$ up to scaling. Being convex (7) is computationally appealing, in fact it can be shown that (7) is a semidefinite program (SDP) for which off-the-shelf solvers are available. Scalable algorithms for large-scale problems have also been developed; for instance the solver implemented to run the numerical tests in Section IV leverages the method of multipliers iterations described in [7, Sec. 5.3], and recently adopted in [5] for blind deconvolution of (non-graph) signals.

B. Unknown support

In various timely applications such as opinion formation in social networks, the challenge is not only the joint recovery of $\mathbf{h}$ and the non-zero coefficients $\bar{\mathbf{x}}$, but also the identification of supp$(\mathbf{x})$. Recovery of sparse $\mathbf{x}$ (which here also involves support identification) from the graph signal $\mathbf{y}$ is critical to unveil e.g., those influential actors or "seeds" leading to the network state represented by the observed graph signal $\mathbf{y}$. Here we briefly discuss such problem and outline a preliminary formulation that we are currently pursuing.

Going back to Lemma 2, the graph signal domain input-output relationship (3) can be rewritten in terms of the full signal $\mathbf{x}$ as $\mathbf{y}=\mathbf{V}\big(\mathbf{\Psi}^T\odot\mathbf{U}^T\big)^T\text{vec}\big(\mathbf{x}\mathbf{h}^T\big)$. While this again suggests recovery of $\mathbf{Y}=\mathbf{x}\mathbf{h}^T$ via rank minimization, sparsity of $\mathbf{x}$ translates into *row-wise sparsity* of $\mathbf{Y}$, i.e., rows $\mathbf{y}_i^T$ indexed by $\{1,\ldots,N\}\setminus\text{supp}(\mathbf{x})$ are identically zero. To leverage this dual structure in the sought matrix, we can solve

$$\min_{\mathbf{Y}}\|\mathbf{Y}\|_*+\lambda\|\mathbf{Y}\|_{2,1},\quad\text{s. to }\mathbf{y}=\mathbf{V}\big(\mathbf{\Psi}^T\odot\mathbf{U}^T\big)^T\text{vec}\big(\mathbf{Y}\big),\qquad(8)$$

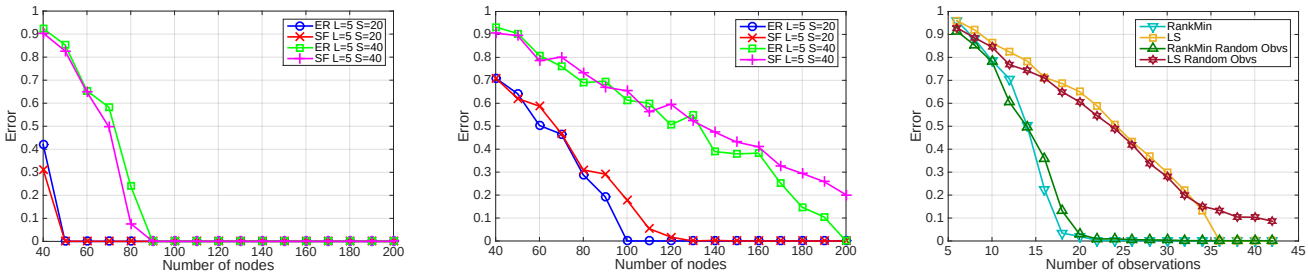
where $\|\mathbf{Y}\|_{2,1}:=\sum_{i=1}^N\|\mathbf{y}_i^T\|_2$ is the ℓ_2/ℓ_1 mixed norm encouraging row-wise (vector) sparsity in $\mathbf{Y}$, and λ is a tuning parameter. Proximal-splitting algorithms can be adopted to minimize the composite, non-differentiable cost in (8).

Recovery of simultaneously low-rank and row-sparse matrices from noisy compressive measurements was also proposed in [9] for hyperspectral image reconstruction. Recent theoretical results on recovery of simultaneously structured matrix models suggest that other (possibly nonconvex) criteria could be appealing as well [10]. In any case, existing results should be reexamined in the graph signal processing setting advocated here.

IV. Numerical results

Here we illustrate the performance of the blind graph filter identification algorithm, by solving (7) for different graphs $\mathcal{G}$ and varying the parameters N, L and S. Obtained estimates will be henceforth denoted as $\{\tilde{\mathbf{x}},\tilde{\mathbf{h}}\}$. In all cases we define the graph-shift operator $\mathbf{S}=\mathbf{A}$, where $\mathbf{A}$ is the adjacency matrix of $\mathcal{G}$. For given signal dimensionality parameters, synthetic observations $\mathbf{y}$ are generated according to the model (3). The "true" vectors $\bar{\mathbf{x}}_0$ and $\mathbf{h}_0$ are drawn from standard multivariate Gaussian distributions, and are normalized to unit length. The root-mean-square error RMSE $:=\|\tilde{\mathbf{x}}\tilde{\mathbf{h}}^T-\bar{\mathbf{x}}_0\mathbf{h}_0^T\|_F$ is adopted as figure of merit to assess recovery performance, and the median error over 100 graph signal and topology realizations is reported in the plots. In all test cases, the rank-minimization algorithm is compared against a naive least-squares (LS) solution of the linear system of equations (3).

Random graphs. For $40\leq N\leq 200$, shift operators are generated from the adjacency matrices of Erdös-Rényi (with edge presence probability $p=0.1$) and scale-free preferential attachment (Barabási-Albert) random graphs [12]. We explicitly check that the resulting network graphs are connected. For each N, we generate graphs signals under two settings of increasing problem difficulty, namely: (i) $L=5$, $S=20$; and (ii) $L=5$, $S=40$. Figures 1a and 1b depict the RMSE as a function of N for the rank minimization algorithm and LS, respectively. While LS is expected to succeed only for $N\geq N_0:=S\times L$, Fig. 1b shows this is (tightly) the case for the Erdös-Rényi random graph, whereas for scale-free graphs there is a nontrivial gap between N_0 and the value of N for which the RMSE first vanishes. This suggests that for increasingly structured graphs, matrix $\mathbf{M}:=\mathbf{V}\big(\mathbf{\Psi}^T\odot\bar{\mathbf{U}}^T\big)^T$ is prone to losing rank. Fig. 1a shows the performance of the nuclear-norm minimization algorithm, which is less sensitive to



(a) (b) (c)

Fig. 1: Recovery performance for different graphs, filters and input signals. The adopted figure of merit is the RMSE $:= \|\tilde{\mathbf{x}}\tilde{\mathbf{h}}^T - \bar{\mathbf{x}}_0\mathbf{h}_0^T\|_F$. (a) Results for the nuclear-norm minimization in (7) using an Erdös-Rényi (ER), a Scale-Free (SF) Barabási-Albert graph and two different combinations of (L, S). Perfect recovery is obtained when $N \geq 2(L+S)$. (b) Counterpart of (a) when the recovery is performed using LS. Note that for SF graphs perfect recovery is not guaranteed even if $N > LS$. (c) Results for the brain network in [11]. The approach in (7) outperforms LS, is robust to the sampling scheme and yields good results even when the number of observations is close to $L+S$.

the graph topology. Most importantly, the algorithm accurately recovers the graph filter coefficients and the input signal even when the signal length N is markedly below N_0. By inspection of Fig. 1a one can also examine how far is the recovery threshold from the fundamental lower bound of $L+S$ required observations, which is dictated by the degrees of freedom in the blind graph filter identification problem.

Brain graph. In this test case we consider a weighted undirected graph $\mathcal{G}$ of the human brain, consisting of $N = 66$ nodes or regions of interest (ROIs) and whose edge weights are given by the density of anatomical connections between regions [11]. The level of activity of each ROI can be represented by a graph signal $\mathbf{x}$, where larger values represent higher levels of activity. Successive applications of $\mathbf{S}$ thus model a linear evolution of the brain activity pattern. Supposing we observe a linear combination (filter) of the evolving states of an originally sparse brain signal (i.e., where only a few regions are active), then blind identification amounts to jointly estimating the original brain signal and the coefficients of the linear combination. Since the graph topology – hence N – is fixed here, for $L = 6$ and $S = 6$ we sample the filtered signal $\mathbf{y} \in \mathbb{R}^{66}$ and vary the sampling rate so that the number of observations fed to the algorithms is $6 \leq N_{obs} \leq 42$. In other words, the sampled output signal processed by the algorithm is $\mathbf{y}_c = \mathbf{C}\mathbf{y} \in \mathbb{R}^{N_{obs}}$, where $\mathbf{C} \in \mathbb{R}^{N_{obs} \times N}$ is a sampling matrix that selects N_{obs} entries from the vector $\mathbf{y}$. We consider two sampling schemes: (i) the elements of $\mathbf{y}_c$ are selected randomly from $\mathbf{y}$; and (ii) we (deterministically) select those entries of $\mathbf{y}$ indexed by $\text{supp}(\mathbf{x})$, and the other $N_{obs} - S$ elements in $\mathbf{y}_c$ are selected at random from the remaining $N - S$ entries in $\mathbf{y}$. Figure 1c depicts the recovery error attained by the nuclear-norm minimization algorithm and LS as a function of N_{obs}, under both aforementioned sampling schemes. Once more, the proposed method markedly outperforms the naive LS solver, in the sense that it exactly recovers the unknown vectors even when $N_{obs} < N_0 = 36$ and not too far above $L + S = 12$. Also notice that, in contrast to LS, the developed algorithm is quite robust to the sampling scheme adopted.

V. Conclusions

We studied the problem of blind graph filter identification, which extends blind deconvolution of time (or spatial) domain signals to graphs. The developed rank minimization approach is robust to sampling and outperforms naive LS. It also offers satisfactory performance even when the number of observations is close to $L + S$, i.e., the fundamental lower bound dictated by the degrees of freedom in the problem. Ongoing research addresses the questions left open with regards to identifiability and theoretical guarantees of the convex nuclear-norm relaxation. Also of interest will be to investigate stable recovery in the presence of noise.

References

[1] D. Shuman, S. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, Mar. 2013.

[2] A. Sandryhaila and J. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.

[3] ——, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, June 2014.

[4] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, "Interpolation of graph signals using shift-invariant graph filters," in *European Signal Process. Conf. (EUSIPCO)*, Nice, France, Aug. 31 - Sept. 4 2015.

[5] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1711–1732, March 2014.

[6] S. Choudhary and U. Mitra, "Identifiability scaling laws in bilinear inverse problems," *arXiv preprint arXiv:1402.2637*, 2014.

[7] B. Recht, M. Fazel, and P. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, August 2010.

[8] M. Fazel, H. Hindi, and S. P. Boyd, "A rank minimization heuristic with application to minimum order system approximation," in *Proc. of American Control Conference*, vol. 6, 2001, pp. 4734–4739.

[9] M. Golbabaee and P. Vandergheynst, "Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, 2012, pp. 2741–2744.

[10] S. Oymak, A. Jalali, M. Fazel, Y. C. Eldar, and B. Hassibi, "Simultaneously structured models with application to sparse and low-rank matrices," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2886–2908, May 2015.

[11] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the structural core of human cerebral cortex," *PLoS Biol*, vol. 6, no. 7, p. e159, 2008.

[12] E. D. Kolaczyk, *Statistical Analysis of Network Data: Methods and Models*. New York, NY: Springer, 2009.