NETWORK TOPOLOGY IDENTIFICATION FROM SPECTRAL TEMPLATES

Santiago Segarra[†], Antonio G. Marques[‡], Gonzalo Mateos[*] and Alejandro Ribeiro[†]

[†]Dept. of ESE, University of Pennsylvania, Philadelphia, PA, USA
[‡]Dept. of TSC, King Juan Carlos University, Madrid, Spain
[*]Dept. of ECE, University of Rochester, Rochester, NY, USA

ABSTRACT

Network topology inference is a cornerstone problem in statistical analyses of complex systems. In this context, the fresh look advocated here permeates benefits from convex optimization and graph signal processing, to identify the so-termed graph shift operator (encoding the network topology) given only the eigenvectors of the shift. These *spectral templates* can be obtained, for example, from principal component analysis of a set of graph signals defined on the particular network. The novel idea is to find a graph shift that while being consistent with the provided spectral information, it endows the network structure with certain desired properties such as sparsity. The focus is on developing efficient recovery algorithms along with identifiability conditions for two particular shifts, the adjacency matrix and the normalized graph Laplacian. Application domains include network topology identification from steady-state signals generated by a diffusion process, and design of a graph filter that facilitates the distributed implementation of a prescribed linear network operator. Numerical tests showcase the effectiveness of the proposed algorithms in recovering synthetic and structural brain networks.

Index Terms— Network topology inference, graph signal processing, spectral graph theory, principal component analysis

1. INTRODUCTION

Advancing a holistic theory of networks necessitates fundamental breakthorughs in modeling, identification, and controllability of distributed network processes – often conceptualized as *signals defined on the vertices of a graph* [1, 2]. 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 (GSP) is to develop algorithms that fruitfully leverage this relational structure [3, 4]. Instrumental to that end is the so-termed graph-shift operator (GSO) [4], a matrix capturing the graph's local topology and whose eigenbasis is central to defining graph Fourier transforms [5]. Most GSP works assume that the GSO (hence the graph) is known, and then analyze how the algebraic and spectral characteristics of the GSO impact the properties of the signals and filters defined on such a graph. Here instead we take the reverse path and investigate how to use information available from graph signals and filters to infer the underlying graph topology; see also [6,7]. In a nutshell, we follow a two-step approach whereby we first leverage results from GSP theory to identify the GSO's eigenbasis using the available information, and then rely on these *spectral templates* to recover the GSO itself.

Network topology inference from a set of (graph-signal) observations is a prominent problem in Network Science [2, 8, 9]. Since networks encode similarities between nodes, most approaches construct graphs whose edge weights correspond to the correlation, or coherence between signal profiles at incident nodes. This approach is not without merit and widely used in practice, but it exhibits several drawbacks, the main one being that links are formed taking into account only pairwise interactions, ignoring that the observed correlations can be due to latent network effects. Acknowledging these limitations, alternative methods rely on partial correlations [2, 10], Gaussian graphical models [11, 12], or, Granger causality [8, 13]. Differently, recent GSP-based network inference frameworks postulate that the network exists as a latent underlying structure, and that observations are generated as a result of a network process defined in such graph. For instance, network structure is estimated in [7] to unveil unknown relations among nodal time series adhering to an autoregressive model involving graph-filter dynamics. A factor analysis-based approach was put forth in [6] to infer graph Laplacians, seeking that input graph signals are smooth over the learned topologies. Different from [6,7] that operate on the graph domain, the goal here is to identify graphs that endow the given observations with desired spectral (frequency-domain) characteristics.

After surveying the required GSP background in Section 2, we formulate the GSO identification problem given spectral templates (Section 3). The novel idea is to search among all feasible networks for the one that endows the resulting graph-signal transforms with prescribed spectral properties, while the inferred graph also exhibits desirable structural characteristics. Optimization problems are formulated to recover two particular GSOs, namely the adjacency matrix (Section 3.1) and the normalized graph Laplacian (Section 3.2). Conditions under which the feasible set reduces to a singleton are derived, and convex relaxations leading to computationally-efficient algorithms are proposed for the general case. Section 4 outlines two pragmatic scenarios where acquisition of the GSO eigenbasis is well motivated, making the case that setups where the GSO is unknown but its spectral templates are available can be more the rule than the exception. Computer simulations highlight the effectiveness of the proposed algorithms in identifying both synthetic and real-world networks (Section 5). Concluding remarks are given in Section 6.

2. GRAPH SIGNALS AND GRAPH FILTERS

Here we formally introduce notation and terminology as well as different GSP tools that will be used throughout the paper.

Graphs. Let $\mathcal{G}$ denote an undirected 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 both (i, j) and (j, i) belong to $\mathcal{E}$. The set $\mathcal{N}_i := \{j \mid (j, i) \in \mathcal{E}\}$ stands for the 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 non-zero elements $A_{ij} = A_{ji}$ if and only if $(i, j) \in \mathcal{E}$. The values of A_{ji} can be binary, or real in the weighted case to

capture the strength of the connection from i to j. Matrix $\mathbf{A}$ can be used to define the degree vector $\mathbf{d} := \mathbf{A1}$ and the degree matrix $\mathbf{D} := \text{diag}(\mathbf{d})$. Moreover, the normalized Laplacian matrix is $\mathbf{L} := \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2} \in \mathbb{R}^{N \times N}$, which is positive semidefinite and has eigenvector $\sqrt{\mathbf{d}} := \mathbf{D}^{1/2}\mathbf{1}$ with associated eigenvalue zero. **Graph signals and shift operator.** Graph signals defined on the nodes of $\mathcal{G}$ are functions $f : \mathcal{N} \to \mathbb{R}$, equivalently represented as vectors $\mathbf{x} = [x_1, ..., x_N]^T \in \mathbb{R}^N$, where x_i denotes the signal value at node i. Since $\mathbf{x}$ does not account explicitly for the structure of the graph where the signal is defined, $\mathcal{G}$ is endowed with the so-called GSO $\mathbf{S}$ [4,5]. The shift $\mathbf{S} \in \mathbb{R}^{N \times N}$ is a matrix whose entry S_{ij} can be non-zero only if $i = j$ or if $(i, j) \in \mathcal{E}$. The sparsity pattern of $\mathbf{S}$ captures the local structure of $\mathcal{G}$, but we make no specific assumptions on the values of its non-zero entries. The shift $\mathbf{S}$ can also be understood as a linear transformation that can be computed locally at the nodes of the graph. More rigorously, if $\mathbf{y}$ is defined as $\mathbf{y} = \mathbf{Sx}$, then node i can compute y_i provided that it has access to the value of x_j at $j \in \mathcal{N}_i$. Typical choices for $\mathbf{S}$ are the adjacency matrix $\mathbf{A}$ [4,5], the (normalized) Laplacian $\mathbf{L}$ [3], and their respective generalizations [14]. We assume henceforth that $\mathbf{S}$ is symmetric, so that $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ with $\mathbf{\Lambda} \in \mathbb{R}^{N \times N}$ being diagonal, but our results hold for any normal GSO.

Graph filters. The shift $\mathbf{S}$ can be used to define linear, shift-invariant graph-signal *operators* of the form

$$\mathbf{H} := \sum_{l=0}^{L-1} h_l \mathbf{S}^l \qquad (1)$$

which are called *graph filters* [4]. For a given input $\mathbf{x}$, the output of the filter is simply $\mathbf{y} = \mathbf{Hx}$. 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 [15, 16].

Frequency domain representation. 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 $\mathbf{U} := \mathbf{V}^T$, and the eigenvalues of $\mathbf{S}$ to define the $N \times L$ Vandermonde matrix $\mathbf{\Psi}$, where $\Psi_{ij} := (\Lambda_{ii})^{j-1}$. Using these conventions, the frequency representations of a *signal* $\mathbf{x}$ and of a *filter* $\mathbf{h}$ are defined as $\widehat{\mathbf{x}} := \mathbf{Ux}$ and $\widehat{\mathbf{h}} := \mathbf{\Psi h}$, respectively [5]. Exploiting such representations, the filter's output $\mathbf{y} = \mathbf{Hx}$ in the frequency domain is ($\odot$ denotes Hadamard product)

$$\widehat{\mathbf{y}} = \text{diag}(\mathbf{\Psi h})\mathbf{Ux} = \text{diag}(\widehat{\mathbf{h}})\widehat{\mathbf{x}} = \widehat{\mathbf{h}} \odot \widehat{\mathbf{x}}. \qquad (2)$$

This identity is the counterpart of the convolution theorem for temporal signals; see e.g. [17] for a derivation. Interestingly, while in the time domain $\mathbf{U} = \mathbf{\Psi}$ and they correspond to the Discrete Fourier Transform (DFT) matrix, this is not true for general graphs.

Network diffusion processes. Graph filters can be used to model network diffusion processes. Specifically, the signal at node i during the step $(l + 1)$ of a linear diffusion process can be written as

$$x_i^{(l+1)} = \alpha_{i,i} x_i^{(l)} + \sum_{j \in \mathcal{N}_i} \alpha_{i,j} x_j^{(l)} \qquad (3)$$

where $\alpha_{i,j}$ are the diffusion coefficients; see e.g., [17]. Leveraging the GSP framework, (3) implies that the graph signal $\mathbf{x}^{(l+1)} = \mathbf{Sx}^{(l)}$ at iteration $l+1$ is the shifted version of $\mathbf{x}^{(l)}$, for a shift $\mathbf{S}$ with entries $S_{ij} = \alpha_{i,j}$ if either $i = j$ or $(i, j) \in \mathcal{E}$, and $S_{ij} = 0$ otherwise. For instance, if we set $\mathbf{S} = \mathbf{I} - \beta\mathbf{L}$ and let the signal of interest be $\mathbf{x} := \mathbf{x}^{(\infty)}$, then $\mathbf{x}$ solves the heat diffusion equation [18, 19]. However, more complex diffusion dynamics such as $\mathbf{x} = \Pi_{l=0}^{\infty}(\mathbf{I} - \beta_l\mathbf{S})\mathbf{x}^{(0)}$ and $\mathbf{x} = \sum_{l=0}^{\infty} \gamma_l\mathbf{x}^{(l)} = \sum_{l=0}^{\infty} \gamma_l\mathbf{S}^l\mathbf{x}^{(0)}$, could also be of interest.

The Cayley-Hamilton theorem guarantees that the aforementioned infinite-horizon processes can be equivalently described by a filter of degree N. Accordingly, several works have recognized that the steady-state signal $\mathbf{x}$ generated by a diffusion process can be modeled as the output of a graph filter $\mathbf{H} = \sum_{l=0}^{N-1} h_l\mathbf{S}^l$ with input (seed) $\mathbf{x}^{(0)}$ [18, 19]. This key insight can be used to relate the statistical and spectral properties of $\mathbf{x}$ and $\mathbf{H}$, which will be leveraged in the ensuing sections to identify $\mathbf{S}$ itself.

3. SHIFT INFERENCE FROM SPECTRAL TEMPLATES

Given a set of eigenvectors $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_N]$, also termed *spectral templates*, our goal is to find a graph shift $\mathbf{S}$ that is diagonalized by $\mathbf{V}$. As detailed in Section 4, the knowledge of $\mathbf{V}$ can for instance be obtained from the PCA decomposition of signals like the ones generated by the diffusion processes outlined in Section 2; and whose covariance depends on the structure of $\mathbf{S}$. Since the postulated problem has infinitely many solutions, we further impose conditions on $\mathbf{S}$ promoting desirable properties such as sparsity, or a priori information on the graph of interest such as non-negative edge weights. Note that by definition, $\mathbf{S}$ encodes the local structure of the graph it represents, thus, its recovery implies a successful identification of the graph topology of interest.

3.1. Identifying the adjacency matrix

Consider the unknown adjacency matrix $\mathbf{S} = \mathbf{A}$ of an unweighted and undirected graph, and suppose its spectral templates $\mathbf{V}$ are given. With $\boldsymbol{\lambda} = [\lambda_1, ..., \lambda_N]^T$ collecting the unknown eigenvalues of $\mathbf{S}$ (hence $\mathbf{\Lambda} = \text{diag}(\boldsymbol{\lambda})$), we aim at identifying $\mathbf{S}$ by solving

$$\min_{\{\mathbf{S}, \boldsymbol{\lambda}\}} \quad \|\mathbf{S}\|_0 \qquad (4)$$

$$\text{s. to} \quad \mathbf{S} = \sum_{k=1}^{N} \lambda_k \mathbf{v}_k \mathbf{v}_k^T, \qquad (5)$$

$$S_{ij} \in \{0,1\}, \;\; S_{ii} = 0, \;\; \mathbf{S} \in \mathcal{M}^N, \;\; \mathbf{S1} \geq \mathbf{1} \qquad (6)$$

where $\mathcal{M}^N$ denotes the set of $N \times N$ real and symmetric matrices. Among the potentially multiple feasible solutions, the cardinality function $\|\mathbf{S}\|_0$ in the objective selects as optimum the one that minimizes the number of edges. The role of constraint $\mathbf{S1} \geq \mathbf{1}$ (entrywise), which reasonably requires each node to have at least one neighbor, is to prevent the trivial solution $\mathbf{S} = \mathbf{0}$. As can be seen from (4)-(6), when all eigenvectors $\{\mathbf{v}_k\}_{k=1}^N$ are given the design of $\mathbf{S}$ amounts to finding the N eigenvalues in $\boldsymbol{\lambda}$. The constraint in (5) encodes the definition of a general graph shift $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$, while those in (6) incorporate the supplementary conditions implied by the fact that $\mathbf{S}$ is an adjacency matrix. Additional information can be incorporated in (6), such as a priori knowledge on particular entries of $\mathbf{S}$ or node degrees. Alternatively, the feasible set can be enlarged to accommodate for adjacency matrices with negative entries.

An interesting property of the proposed optimization is that the feasible set described by (5)-(6) is generally small. To be more precise, we define the matrix $\mathbf{W} := \mathbf{V} \odot \mathbf{V} \in \mathbb{R}_+^{N \times N}$ and denote by Q the number of singular values of $\mathbf{W}$ that are zero.

Proposition 1 *Assume that* (4)-(6) *is feasible, then it holds that:*
a) The nullspace of $\mathbf{W}$ has dimension at least one, so that $Q \geq 1$.
b) If $Q = 1$, the feasible set given by (5)-(6) *is a singleton.*

Proof (sketch): The key of the proof resides in noting that the columns of $\mathbf{W}$ correspond to the diagonal entries of $\mathbf{v}_k\mathbf{v}_k^T$ for $k = 1, \ldots, N$. Then, by combining (5) and (6) it follows that $\mathbf{W}\boldsymbol{\lambda} = \mathbf{0}$ for all feasible $\boldsymbol{\lambda}$, thus, feasibility implies $Q \geq 1$. When $Q = 1$, $\boldsymbol{\lambda}$

(and hence $\mathbf{S}$) is unique up to a scaling factor. Therefore using the fact that $\mathbf{S}$ is binary, statement b) follows. ∎

Relaxation and algorithmic discussion. The two sources of non-convexity that render the solution to (4)-(6) challenging are the presence of the ℓ_0 norm in the objective and the binary constraints $S_{ij} \in \{0,1\}$. We relax the former to an iteratively re-weighted ℓ_1 norm and the latter by replacing $\{0,1\}$ with its convex hull $[0,1]$. Specifically, with p denoting an iteration index, we aim to solve a sequence $p = 1, ..., P$ of ℓ_1-norm penalized problems

$$\min_{\{\mathbf{S}, \boldsymbol{\lambda}\}} \quad \sum_{i=1}^N \sum_{j=1}^N \omega_{ij}(p)|S_{ij}| \tag{7}$$

$$\text{s. to} \quad \mathbf{S} = \sum_{k=1}^N \lambda_k \mathbf{v}_k \mathbf{v}_k^T, \tag{8}$$

$$S_{ij} \in [0,1], \quad S_{ii} = 0, \quad \mathbf{S} \in \mathcal{M}^N, \quad \mathbf{S1} \geq \mathbf{1} \tag{9}$$

with weights $\omega_{ij}(p) := (|S_{ij}(p-1)| + \delta)^{-1}$ where δ is a small and positive constant. Intuitively, the goal of the re-weighted scheme in (7) is that if the value of $|S_{ij}(p-1)|$ is small, in the next iteration the penalization weight $\omega_{ij}(p)$ is large, promoting further shrinkage of S_{ij} towards zero; see, e.g., [20] for technical details. The recovery performance of the proposed scheme is analyzed in Section 5.

Building on Prop. 1, whenever $Q = 1$ the solutions of (7)-(9) and (4)-(6) coincide up to a scaling factor.

Proposition 2 *Let $\mathbf{S}_0^*$ and $\mathbf{S}_1^*$ denote solutions to (4)-(6) and (7)-(9), respectively. If $Q = 1$, then $\mathbf{S}_1^* = d_{\min}^{-1} \mathbf{S}_0^*$, with $d_{\min}$ being the minimum node degree in the graph to recover.*

Proof (sketch): Denoting by $\boldsymbol{\lambda}_0^*$ the unique solution to (4)-(6) (cf. Prop. 1), the solution to (7)-(9) is $\boldsymbol{\lambda}_1^* = \alpha \boldsymbol{\lambda}_0^*$ with α being the smallest positive number satisfying $\mathbf{S1} \geq \mathbf{1}$, which is $\alpha = d_{\min}^{-1}$. ∎

3.2. Identifying the Laplacian matrix and generalized shifts

Consider the unknown normalized Laplacian matrix $\mathbf{S} = \mathbf{L}$ of an unweighted and undirected graph. With $\mathcal{M}_+^N$ denoting the conical set of positive semidefinite $N \times N$ matrices, the problem to solve now is [cf. (4)-(6)]

$$\min_{\{\mathbf{S}, \boldsymbol{\lambda}\}} \quad \|\mathbf{S}\|_0 \tag{10}$$

$$\text{s. to} \quad \mathbf{S} = \sum_{k=1}^N \lambda_k \mathbf{v}_k \mathbf{v}_k^T, \quad \mathbf{S} \in \mathcal{M}_+^N, \quad \lambda_1 = 0, \tag{11}$$

$$S_{ij} \in [-1, 0] \text{ for } i \neq j, \quad S_{ii} = 1 \text{ for all } i. \tag{12}$$

In the above formulation we assumed that $\mathbf{v}_1$ corresponds to the eigenvector $\sqrt{\mathbf{d}}$ so that we can leverage the structure of normalized Laplacians and enforce $\lambda_1 = 0$. This, in turn, prevents the trivial solution $\mathbf{S} = \mathbf{I}$. Note also that the eigenvector $\sqrt{\mathbf{d}}$ can be easily identified from the N eigenvectors in $\mathbf{V}$ since it is the only one whose entries have all the same sign [21]. As was the case for (4)-(6), if more than one feasible solution exists, the identified Laplacian corresponds to the topology that minimizes the number of edges.

Interestingly, problem (10)-(12) also has a small feasible set. To state this formally, consider the matrix $\tilde{\mathbf{V}} = [\mathbf{1}, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_N]$, define $\tilde{\mathbf{W}} := \tilde{\mathbf{V}} \odot \tilde{\mathbf{V}}$ and denote by $\tilde{Q}$ the number of singular values of $\tilde{\mathbf{W}}$ that are zero. Then the following result, which is the counterpart of Prop. 1 for Laplacian identification, holds.

Proposition 3 *Assume that (10)-(12) is feasible, then it holds that:*
a) The nullspace of $\tilde{\mathbf{W}}$ has at least dimension one, so that $\tilde{Q} \geq 1$.
b) If $\tilde{Q} = 1$, the feasible set given by (11)-(12) is a singleton.

The result follows using arguments similar to those in the proof of Prop. 1. Regarding the design of efficient algorithms to solve (10)-(12), since (10) is the only source of non-convexity, the iteratively re-weighted scheme in (7) can also be applied here. Details are omitted due to space limitations.

Modifications to the proposed formulations can be made to accommodate for shifts that exhibit properties other than sparsity. For example, one could go after Laplacians with good mixing conditions, related to having a large second smallest eigenvalue [22]. This can be accomplished by introducing the optimization variable $\lambda_{\min}$, adding the constraints $\lambda_{\min} \leq \lambda_k$ for all $k \geq 2$ and augmenting the cost with the regularization $-\eta \lambda_{\min}$, with η being a tuning constant.

The identification of graph shifts different from the adjacency and the normalized Laplacian can be of interest, including the combinatorial Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$ and the random walk Laplacian $\mathbf{L} = \mathbf{D}^{-1}\mathbf{A}$. These would require minor modifications in the constraints (11)-(12). For the particular case of the combinatorial Laplacian, results in Prop. 3 can be extended directly provided that the degree vector $\mathbf{d}$, which contains the same information than the first eigenvector of the normalized Laplacian, is known. Specifically, the constraints $S_{ii} = 1$ in (12) need to be replaced with $S_{ii} = d_i$ and matrix $\tilde{\mathbf{V}}$, which is used in the definition of $\tilde{\mathbf{W}} = \tilde{\mathbf{V}} \odot \tilde{\mathbf{V}}$, has to be redefined as $\tilde{\mathbf{V}} = [\mathbf{d}, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_N]$.

4. OBTAINING THE GRAPH SPECTRAL TEMPLATES

The methods discussed in Section 3 are relevant in practice when we have access to the eigenvectors (i.e., the spectral templates) of $\mathbf{S}$, but not to the GSO itself. A couple of scenarios where that is indeed the case are presented next.

Signals generated through diffusion processes. Recall that graph filters can be used to model diffusion processes that depend on the topology of the graph (Section 2). Specifically, a diffused or steady-state signal $\mathbf{x}$ can be written as $\mathbf{x} = \mathbf{H}\mathbf{x}^{(0)}$ where $\mathbf{x}^{(0)}$ is a seed signal and $\mathbf{H} = \sum_{l=0}^{N-1} h_l \mathbf{S}^l$ is a graph filter. Under the assumption that $\mathbf{x}^{(0)}$ is white (with identity covariance matrix) and zero-mean, the covariance matrix of the output $\mathbf{C}_\mathbf{x} := \mathbb{E}[\mathbf{x}\mathbf{x}^T]$ is given by

$$\mathbf{C}_\mathbf{x} = \mathbf{H}\mathbb{E}\left[\mathbf{x}^{(0)}(\mathbf{x}^{(0)})^T\right]\mathbf{H}^T = \mathbf{H}\mathbf{H}^T \stackrel{\star}{=} \mathbf{V}\text{diag}(|\hat{\mathbf{h}}|^2)\mathbf{V}^T, \tag{13}$$

where for equality $\stackrel{\star}{=}$ we leveraged the frequency interpretation of a graph filter. Expression (13) reveals that the eigenvectors of the covariance matrix $\mathbf{C}_\mathbf{x}$ and those of $\mathbf{S}$ are the same. Thus, if $\mathbf{C}_\mathbf{x}$ is known, the spectral templates $\mathbf{V}$ can be readily obtained. More commonly, if $\mathbf{C}_\mathbf{x}$ is unknown but we have access to a set of M diffused signals $\{\mathbf{x}_m\}_{m=1}^M$, we may approximate $\mathbf{C}_\mathbf{x}$ with the sample covariance $\hat{\mathbf{C}}_\mathbf{x} = 1/M \sum_{m=1}^M \mathbf{x}_m \mathbf{x}_m^T$ and estimate the eigenvectors of $\mathbf{S}$. Applying the methods in Section 3, we can then use the estimated eigenvectors $\hat{\mathbf{V}}$ to recover $\mathbf{S}$ as illustrated in Section 5. Detailed analysis on the sensibility of the topology identification performance to the errors in estimating $\hat{\mathbf{V}}$ is left as future work.

Implementation of linear network operators. Matrices $\mathbf{B} \in \mathbb{R}^{N \times N}$ can be viewed as linear network operators since they are linear maps between graph signals. Conceivably, one might be interested in representing (or approximating) a pre-specified operator $\mathbf{B}$ as a graph filter $\mathbf{H}$, to facilitate its distributed implementation in the network. This problem is thoroughly investigated in [23], and one of the conditions derived for perfect implementation states that all the eigenvectors of $\mathbf{B}$ and $\mathbf{S}$ must coincide [23, Prop. 1]. Consequently, we can use the eigenvectors of a desired $\mathbf{B}$ as spectral templates to generate a graph shift $\mathbf{S}$. This way, $\mathbf{B}$ can be expressed as a graph filter and implemented via local agent interactions.
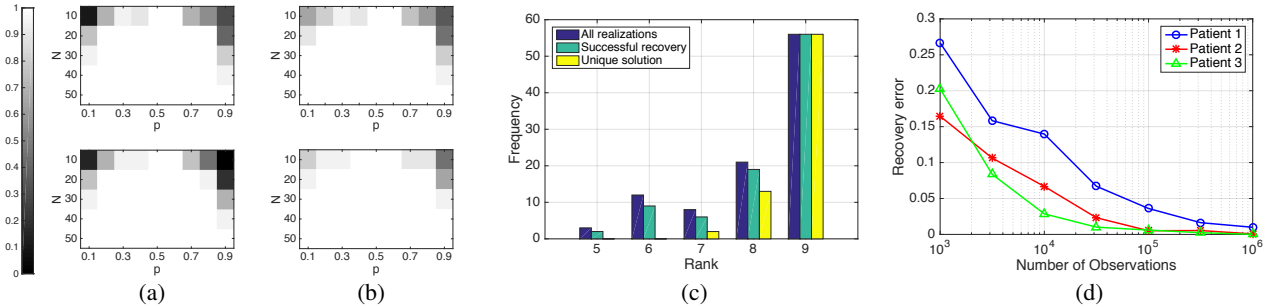
Fig. 1. (a) Proportion of topology identification problems with unique solution for Erdős-Rényi graphs as a function of their size N and the probability of edge appearance p for adjacency (top) and normalized Laplacian (bottom) matrices. (b) Recovery rate for the same set of graphs used in (a) when implementing the iteratively re-weighted approach. (c) Histogram of the rank of matrix $\tilde{\mathbf{W}}$ for $N = 10$ and $p = 0.2$. (d) Brain graph recovery error for three patients as a function of the number of signals observed in the estimation of the spectral templates.

5. NUMERICAL EXPERIMENTS

We illustrate the performance of the topology identification algorithm by solving (7)-(9), as well as its counterpart for Laplacian identification, for different synthetic and real-world graphs.

Random graphs. Consider Erdős-Rényi graphs [24] of varying size $N \in \{10, 20, \ldots, 50\}$ and different edge-formation probabilities $p \in \{0.1, 0.2, \ldots, 0.9\}$. For each combination of N and p we generate 100 graphs and try to recover their adjacency $\mathbf{A}$ and normalized Laplacian $\mathbf{L}$ matrices from the corresponding spectral templates $\mathbf{V}$. In Fig. 1(a) we plot the proportion of instances where the corresponding optimization problems – (4)-(6) for $\mathbf{A}$ and (10)-(12) for $\mathbf{L}$ – have unique solutions. Notice that multiple solutions are more frequent when the expected number of neighbors of a given node is close to either 1 or N. For intermediate values of p, the rank of both $\mathbf{W}$ and $\tilde{\mathbf{W}}$ is typically $N - 1$, guaranteeing that the solution to our optimization is unique (cf. Props. 1 and 3). Using the same set of graphs that those in Fig. 1(a), Fig. 1(b) shows the recovery rate when solving the iteratively re-weighted problem in (7)-(9) and its counterpart for the Laplacian. As expected, the rates in Fig. 1(b) dominate those in Fig. 1(a) since every instance with a unique solution is recovered successfully. Moreover, the improvement in the rates observed in Fig. 1(b) reflect the effect of the weighted ℓ_1 norm objective [cf. (7)] in recovering the true underlying graph.

As indicated by Props. 1 and 3, the rate of recovery is intimately related to the ranks of $\mathbf{W}$ and $\tilde{\mathbf{W}}$ for the adjacency and normalized Laplacian case, respectively. Fig. 1(c) further illustrates this relation via a histogram of the rank of $\tilde{\mathbf{W}}$ for the 100 graphs with $N = 10$ and $p = 0.2$. For more than half of the instances, the rank of $\tilde{\mathbf{W}}$ was equal to 9 (blue bar) and, as stated in Prop. 3, for all these graphs the solution was unique (yellow bar) and successfully recovered (cyan bar). We see that, as the rank of $\tilde{\mathbf{W}}$ degrades, uniqueness is no longer guaranteed but for most cases the true graph can still be recovered following the iteratively re-weighted scheme proposed. Only in 8 of the cases where $\mathrm{rank}(\tilde{\mathbf{W}}) < 9$ the recovery was not successful, entailing a recovery rate of 0.92, as reported in the corresponding entry ($N = 10$, $p = 0.2$) of the lower plot in Fig. 1(b).

Brain graphs. We consider the identification of unweighted and undirected graphs corresponding to human brains [25], consisting of $N = 66$ nodes or regions of interest (ROIs) and whose edges link ROIs with density of anatomical connections greater than a threshold. The threshold is chosen as the largest one that entails a connected graph. We test the recovery performance for noisy spectral templates $\hat{\mathbf{V}}$ obtained from sample covariances of signals generated through diffusion processes (cf. Section 4). Our framework can handle the recovery of $\mathbf{S}$ from noisy versions of $\mathbf{V}$ by rewriting the equality constraint (8) as an element-wise inequality bounding the absolute difference between S_{ij} and $\sum_{k=1}^{N} \lambda_k V_{ki} V_{kj}$ for all i, j. Denoting by $\hat{\mathbf{V}}_i$ the noisy spectral templates of patient $i \in \{1, 2, 3\}$ and by $\hat{\mathbf{A}}_i$ the adjacency matrices recovered from them, Fig. 1(d) plots the recovery error as a function of the number of signals observed in the computation of the sample covariance. The error is quantified as the proportion of edges misidentified, i.e., $\|\mathbf{A}_i - \hat{\mathbf{A}}_i\|_0 / \|\mathbf{A}_i\|_0$, and each point in Fig. 1(d) is the average across 50 random realizations. First notice that for an increasing number of observed signals we see a monotonous decrease in recovery error. For example, when going from 10^4 to 10^5 observations the error is (approximately) divided by seven, when averaged across patients. This is reasonable since a larger number of observations gives rise to a more reliable estimate of the covariance matrix entailing a less noisy version of the spectral templates. Notice that, even though the performance increases for all patients with the number of observations, the brain of patient 1 is consistently the hardest to identify. For instance, consider the errors for 10^5 observations, where for patient 1 we successfully recover 96.4% of the *edges*, whereas for patients 2 and 3 the average recovery rate is 99.5% and 99.4%, respectively. This points towards the fact that some graphs are inherently more robust for identification when given noisy spectral templates. A formal analysis of this phenomenon is left as future work.

Traditional methods like graphical lasso [10] fail to recover $\mathbf{S}$ from the sample covariance of filtered white signals. This occurs because, in general, the filter $\mathbf{H}$ introduces conditional dependence between signal values more than one hop apart. Even when focusing on filters of the form $\mathbf{H} = h_0\mathbf{I} + h_1\mathbf{S}$, graphical lasso underperforms compared to the method presented. More precisely, based on 10^5 observations, the recovery error of graphical lasso averaged over 50 realizations and with optimal tuning parameters is 0.303, 0.350, and 0.270 for patients 1, 2, and 3, respectively [cf. Fig. 1(d)].

6. CONCLUSIONS

We studied the problem of identifying a graph shift operator $\mathbf{S}$, i.e., the topology of a graph $\mathcal{G}$ of interest, given its eigenbasis $\mathbf{V}$. The focus was on setups where $\mathbf{S}$ represents the adjacency or the normalized Laplacian of $\mathcal{G}$. We formulated optimization problems to recover $\mathbf{S}$, presented their convex relaxations, and derived conditions under which the relaxed problem is guaranteed to recover the desired shift. To highlight the practical relevance of the proposed schemes we outlined scenarios where $\mathbf{V}$ is known but $\mathbf{S}$ is not, and carried out numerical tests showing the effectiveness in recovering synthetic and brain graphs, even from imperfect spectral templates.

7. REFERENCES

[1] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamic Processes on Complex Networks*, Cambridge University Press, Cambridge, UK, 2012.

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

[3] D.I. Shuman, S.K. 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.

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

[5] A. Sandryhaila and J.M.F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, June 2014.

[6] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, "Learning laplacian matrix in smooth graph signal representations," *arXiv preprint arXiv:1406.7842v2 [cs.LG]*, 2015.

[7] J. Mei and J. M. F. Moura, "Signal processing on graphs: Modeling (causal) relations in big data," *arXiv preprint arXiv:1503.00173v1 [cs.IT]*, 2015.

[8] O. Sporns, *Discovering the Human Connectome*, MIT Press, Boston, MA, 2012.

[9] X. Cai, J. A. Bazerque, and G. B. Giannakis, "Sparse structural equation modeling for inference of gene regulatory networks exploiting genetic perturbations," *PLoS, Computational Biology*, June 2013.

[10] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, no. 3, pp. 432–441, 2008.

[11] M. Slawski and M. Hein, "Estimation of positive definite m-matrices and structure learning for attractive gaussian markov random fields," *Linear Algebra and its Applications*, vol. 473, pp. 145–179, 2015.

[12] N. Meinshausen and P. Buhlmann, "High-dimensional graphs and variable selection with the lasso," *Ann. Stat.*, vol. 34, pp. 1436–1462, 2006.

[13] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura, and S. L. Bressler, "Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by granger causality," *PNAS*, vol. 101, pp. 98499854, 2004.

[14] C. Godsil and G. Royle, *Algebraic Graph Theory*, vol. 207, Springer-Verlag, Graduate Texts in Mathematics, 2001.

[15] 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.

[16] S. Segarra, A. G. Marques, and A. Ribeiro, "Distributed implementation of linear network operators using graph filters," in *53rd Allerton Conf. on Commun. Control and Computing*, Univ. of Illinois at U-C, Monticello, IL, Sept. 30- Oct. 2, 2015.

[17] S. Segarra, G. Mateos, A. G. Marques, and A. Ribeiro, "Blind identification of graph filters with sparse inputs," in *IEEE Intl. Wrksp. Computat. Advances Multi-Sensor Adaptive Process. (CAMSAP)*, Cancun, Mexico, Dec. 13-16, 2015.

[18] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, "Reconstruction of graph signals through percolation from seeding nodes," *arXiv preprint arXiv:1507.08364*, 2015.

[19] S. Segarra, G. Mateos, A. G. Marques, and A. Ribeiro, "Blind identification of graph filters with sparse inputs," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2015 IEEE 6th International Workshop on*. IEEE, 2015.

[20] E. J. Candes, M. B. Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimzation," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, Dec. 2008.

[21] T. Biyikoğu, J. Leydold, and P. Stadler, *Laplacian Eigenvectors of Graphs: Perron-Frobenius and Faber-Krahn Type Theorems*, Springer-Verlag Berlin Heidelberg, 2007.

[22] F. Chung, *Spectral Graph Theory*, vol. 92, American Mathematical Soc., 1997.

[23] S. Segarra, A. G. Marques, and A. Ribeiro, "Distributed linear network operators using graph filters," *arXiv preprint arXiv:1510.03947*, 2015.

[24] B. Bollobás, *Random Graphs*, Springer, 1998.

[25] 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, pp. e159, 2008.