

# Blind Identification of Invertible Graph Filters with Multiple Sparse Inputs

Chang Ye, Rasoul Shafipour and Gonzalo Mateos

Dept. of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA

**Abstract**—This paper deals with the problem of blind identification of a graph filter and its sparse input signal, thus broadening the scope of classical blind deconvolution of temporal and spatial signals to irregular graph domains. While the observations are bilinear functions of the unknowns, a mild requirement on invertibility of the filter enables an efficient convex formulation, without relying on matrix lifting that can hinder applicability to large graphs. On top of scaling, it is argued that (non-cyclic) permutation ambiguities may arise with some particular graphs. Deterministic sufficient conditions under which the proposed convex relaxation can exactly recover the unknowns are stated, along with those guaranteeing identifiability under the Bernoulli-Gaussian model for the inputs. Numerical tests with synthetic and real-world networks illustrate the merits of the proposed algorithm, as well as the benefits of leveraging multiple signals to aid the (blind) localization of sources of diffusion.

**Index Terms**—Graph signal processing, network diffusion, bilinear equations, blind deconvolution, convex optimization.

## I. INTRODUCTION

Network processes such as neural activities at different regions of the brain [9], [10], vehicle trajectories over road networks [4], or spatial temperature profiles measured by a wireless sensor network [19], can be represented as signals supported on the nodes of a graph. Under the natural assumption that the signal properties are influenced by the graph topology (e.g., in a network diffusion or percolation process), the goal of graph signal processing (GSP) is to develop algorithms that exploit this relational structure. Accordingly, generalizations of fundamental signal processing tasks have been widely explored in recent work; see [14] for a comprehensive tutorial treatment. Notably graph filters – which generalize classical time-invariant systems – were conceived as information-processing operators acting on graph-valued signals [18]. Mathematically, graph filters are linear transformations that can be expressed as polynomials of the so-termed graph-shift operator (Section II). The graph shift offers an algebraic representation of network structure and can be viewed as a local diffusion operator. For the directed cycle graph representing e.g., periodic temporal signals, it boils down to the classical time-shift operator [18]. Given a shift, the polynomial coefficients fully determine the graph filter and are referred to as filter coefficients.

**Problem outline and envisioned applications.** In this paper, we revisit the blind identification of graph filters with sparse inputs, with emphasis on modeling diffusion processes and

localization of the sources of diffusion [22]. Specifically, given  $P$  observations of graph signals  $\{\mathbf{y}_i\}_{i=1}^P$  that we model as outputs of a diffusion filter (i.e., a polynomial in a known graph-shift operator), we seek to jointly identify the filter coefficients  $\mathbf{h}$  and the input signals  $\{\mathbf{x}_i\}_{i=1}^P$  that gave rise to the network observations. This inverse problem broadens the scope of classical blind deconvolution of temporal or spatial signals to graphs [1], [11]. Since the resulting bilinear inverse problem is ill-posed, we assume that the inputs are sparse – a well-motivated setting when few seeding nodes inject a signal that is diffused throughout a network [22]. Accordingly, envisioned application domains include environmental monitoring (where are the heat or seismic sources?), opinion formation in social networks (who started the rumor?), neural signal processing (which brain regions were activated?), and epidemiology (who is patient zero for the disease outbreak?). **Related work and contributions.** Different from most existing works dealing with source localization on graphs, e.g., [16], [20], [25], like [15] the advocated GSP approach is applicable even when a single snapshot of the diffused signal is available. Often the models of diffusion are probabilistic in nature, and resulting maximum-likelihood source estimators can only be optimal for particular (e.g., tree) graphs [16], or rendered scalable under restrictive dependency assumptions [5]. Relative to [9], [15], the proposed framework can accommodate signals defined on general undirected graphs and relies on a convex estimator of the sparse sources of diffusion. Furthermore, the setup where multiple output signals are observed (each one corresponding to a different sparse input), has not been thoroughly explored in convex-relaxation approaches to blind deconvolution of (non-graph) signals, e.g., [1], [13]; see [23] for a recent and inspiring alternative that we leverage here.

A noteworthy approach was put forth in [22], which casts the (bilinear) blind graph-filter identification task as a linear inverse problem in the “lifted” rank-one, row-sparse matrix  $\mathbf{x}\mathbf{h}^T$ . While the rank and sparsity minimization algorithms in [17], [22] can successfully recover sparse inputs along with low-order graph filters, reliance on matrix lifting can hinder applicability to large graphs. Beyond this computational consideration, the overarching assumption of [22] is that the inputs  $\{\mathbf{x}_i\}_{i=1}^P$  share a common support. Here instead we show how a mild requirement on invertibility of the graph filter facilitates an efficient convex formulation for the multi-signal case with arbitrary supports (Section III); see also [23] for a time-domain precursor. In Section IV we take a closer look at inherent scaling and (non-cyclic) permutation ambiguities

Work in this paper was supported by the NSF award CCF-1750428. Author emails: {cye7,rshafipo,gmateosb}@ur.rochester.edu.

arising with some particular graphs. We also briefly comment on identifiability under the Bernoulli-Gaussian model for the inputs [12], and state deterministic sufficient conditions under which the proposed convex relaxation can exactly recover the unknowns. Numerical tests with synthetic graphs and a structural brain network corroborate the effectiveness of the proposed approach in recovering the sparse input signals (Section V). Concluding remarks are given in Section VI.

## II. PRELIMINARIES AND PROBLEM STATEMENT

Consider a weighted and undirected network graph  $\mathcal{G} = (\mathcal{V}, \mathbf{A})$ , where  $\mathcal{V}$  is the set of vertices with cardinality  $|\mathcal{V}| = N$ , and  $\mathbf{A} \in \mathbb{R}^{N \times N}$  is the symmetric graph adjacency matrix whose entry  $A_{ij}$  denotes the edge weight between nodes  $i$  and  $j$ . As a more general algebraic descriptor of network structure, one can define a *graph-shift operator*  $\mathbf{S} \in \mathbb{R}^{N \times N}$  as any matrix having the same sparsity pattern as  $\mathcal{G}$  [18]. Accordingly,  $\mathbf{S}$  can be viewed as a local diffusion (or averaging) operator. Common choices are to set it to either  $\mathbf{A}$  (and its normalized counterparts) or variations of adjacency and Laplacian matrices [6], [14]. Since  $\mathbf{S}$  is real and symmetric, it is diagonalizable so that  $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ , with  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N)$ . Lastly, a graph signal  $\mathbf{x} : \mathcal{V} \mapsto \mathbb{R}^N$  is an  $N$ -dimensional vector, where entry  $x_i$  represents the signal value at node  $i \in \mathcal{V}$ .

### A. Graph-filter models of network diffusion processes

Let  $\mathbf{y}$  be a graph signal supported on  $\mathcal{G}$ , which is generated from an input graph signal  $\mathbf{x}$  via linear network dynamics of the form

$$\mathbf{y} = \alpha_0 \prod_{l=1}^{\infty} (\mathbf{I} - \alpha_l \mathbf{S}) \mathbf{x} = \sum_{l=0}^{\infty} \beta_l \mathbf{S}^l \mathbf{x}. \quad (1)$$

While  $\mathbf{S}$  encodes only one-hop interactions, each successive application of the shift in (1) diffuses  $\mathbf{x}$  over  $\mathcal{G}$ . Indeed, any process that can be understood as the linear propagation of a seed signal through a static graph can be written in the form in (1), and subsumes heat diffusion, consensus and the classic DeGroot model of opinion dynamics as special cases [3].

The diffusion expressions in (1) are polynomials on  $\mathbf{S}$  of possibly infinite degree, yet the Cayley-Hamilton theorem asserts they are equivalent to polynomials of degree smaller than  $N$ . Upon defining the vector of coefficients  $\mathbf{h} := [h_0, \dots, h_{L-1}]^T$  and the shift-invariant graph filter

$$\mathbf{H} := h_0 \mathbf{I}_N + h_1 \mathbf{S} + h_2 \mathbf{S}^2 + \dots + h_{L-1} \mathbf{S}^{L-1} = \sum_{l=0}^{L-1} h_l \mathbf{S}^l, \quad (2)$$

the signal model in (1) becomes  $\mathbf{y} = (\sum_{l=0}^{L-1} h_l \mathbf{S}^l) \mathbf{x} := \mathbf{H}\mathbf{x}$ , for some particular  $\mathbf{h}$  and  $L \leq N$ . Due to the local structure of  $\mathbf{S}$ , graph filters represent linear transformations that can be implemented in a distributed fashion [21], e.g., via  $L-1$  successive exchanges of information among neighbors.

Leveraging the spectral decomposition of  $\mathbf{S}$ , graph filters and signals can be represented in the frequency domain. Specifically, let us use the eigenvalues of  $\mathbf{S}$  to define the  $N \times L$  Vandermonde matrix  $\mathbf{\Psi}_L$ , where  $\Psi_{ij} := \lambda_i^{j-1}$ . The frequency representations of a signal  $\mathbf{x}$  and filter  $\mathbf{h}$  are defined as  $\tilde{\mathbf{x}} := \mathbf{V}^T \mathbf{x}$  and  $\tilde{\mathbf{h}} := \mathbf{\Psi}_L \mathbf{h}$ , respectively. The latter follows since the output  $\mathbf{y} = \mathbf{H}\mathbf{x}$  in the frequency domain is given by

$$\tilde{\mathbf{y}} = \text{diag}(\mathbf{\Psi}_L \mathbf{h}) \mathbf{V}^T \mathbf{x} = \text{diag}(\tilde{\mathbf{h}}) \tilde{\mathbf{x}} = \tilde{\mathbf{h}} \circ \tilde{\mathbf{x}}. \quad (3)$$

This identity can be seen as a counterpart of the convolution theorem for temporal signals, where  $\tilde{\mathbf{y}}$  is the elementwise product ( $\circ$ ) of  $\tilde{\mathbf{x}}$  and the filter's frequency response  $\tilde{\mathbf{h}} := \mathbf{\Psi}_L \mathbf{h}$ .

### B. Problem formulation

For given shift operator  $\mathbf{S}$  and filter order  $L$ , suppose we observe  $P$  output signals collected in a matrix  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_P] \in \mathbb{R}^{N \times P}$  such that  $\mathbf{Y} = \mathbf{H}\mathbf{X}$ , where  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_P] \in \mathbb{R}^{N \times P}$  is sparse having at most  $S \ll N$  non-zero entries per column. The goal is to perform blind identification of the graph filter (and its input signals), which amounts to estimating sparse  $\mathbf{X}$  and the filter coefficients  $\mathbf{h}$  up to scaling and (possibly) permutation ambiguities; see Section IV. Sparsity is well motivated when the signals in  $\mathbf{Y}$  represent diffused versions of a few localized sources in  $\mathcal{G}$ , here indexed by  $\text{supp}(\mathbf{X}) := \{(i, j) \mid X_{ij} \neq 0\}$ . Moreover, the non-sparse formulation is ill-posed, since the number of unknowns  $NP + L$  in  $\{\mathbf{X}, \mathbf{h}\}$  exceeds the  $NP$  observations in  $\mathbf{Y}$ .

All in all, using (3) the diffused source localization task can be stated as a feasibility problem of the form

$$\text{find } \{\mathbf{X}, \mathbf{h}\} \text{ s. to } \mathbf{Y} = \mathbf{V} \text{diag}(\mathbf{\Psi}_L \mathbf{h}) \mathbf{V}^T \mathbf{X}, \|\mathbf{X}\|_0 \leq PS, \quad (4)$$

where the  $\ell_0$ - (pseudo) norm  $\|\mathbf{X}\|_0 := |\text{supp}(\mathbf{X})|$  counts the non-zero entries in  $\mathbf{X}$ . In words, the goal is to find the solution to a system of bilinear equations subject to a sparsity constraint in  $\mathbf{X}$ ; a hard problem due to the non-convex  $\ell_0$ -norm as well as the bilinear constraints. To deal with the latter, building on [23] we will henceforth assume that the filter  $\mathbf{H}$  is invertible.

## III. CONVEX RELAXATION FOR INVERTIBLE FILTERS

Here we show how to efficiently tackle the blind graph filter identification problem, through a convex relaxation of (4) when the diffusion filter is invertible.

To that end, note from (3) that graph filter  $\mathbf{H}$  is invertible if and only if  $\tilde{h}_i = \sum_{l=0}^{L-1} h_l \lambda_i^l \neq 0$ , for all  $i = 1, \dots, N$ . In words, the frequency response of the filter should not vanish at the graph frequencies  $\{\lambda_i\}$ . In such case one can show that the inverse operator  $\mathbf{G} := \mathbf{H}^{-1}$  is also a graph filter on  $\mathcal{G}$ , which can be uniquely represented as a polynomial in the shift  $\mathbf{S}$  of degree at most  $N-1$  [18, Theorem 4]. To be more specific, let  $\mathbf{g} \in \mathbb{R}^N$  be the vector of inverse-filter coefficients, i.e.,  $\mathbf{G} = \sum_{l=0}^{N-1} g_l \mathbf{S}^l$ . Then one can equivalently rewrite the generative model  $\mathbf{Y} = \mathbf{H}\mathbf{X}$  for the observations as

$$\mathbf{X} = \mathbf{G}\mathbf{Y} = \mathbf{V} \text{diag}(\tilde{\mathbf{g}}) \mathbf{V}^T \mathbf{Y}, \quad (5)$$

where  $\tilde{\mathbf{g}} := \mathbf{\Psi}_N \mathbf{g} \in \mathbb{R}^N$  is the inverse filter's frequency response and  $\mathbf{\Psi}_N \in \mathbb{R}^{N \times N}$  is Vandermonde. Naturally,  $\mathbf{G} = \mathbf{H}^{-1}$  implies the condition  $\tilde{\mathbf{g}} \circ \tilde{\mathbf{h}} = \mathbf{1}_N$  on the frequency responses, where  $\mathbf{1}_N$  denotes the  $N \times 1$  vector of all ones. Leveraging (5), one can recast (4) as a *linear* inverse problem

$$\min_{\{\mathbf{X}, \tilde{\mathbf{g}}\}} \|\mathbf{X}\|_0, \text{ s. to } \mathbf{X} = \mathbf{V} \text{diag}(\tilde{\mathbf{g}}) \mathbf{V}^T \mathbf{Y}, \mathbf{X} \neq \mathbf{0}. \quad (6)$$

This approach is markedly different from the matrix lifting technique used in [22] to handle the bilinear equations in (4).

The  $\ell_0$  norm in (6) makes the problem NP-hard to optimize. Over the last decade or so, convex-relaxation approaches to

**Algorithm 1** Iteratively-reweighted  $\ell_1$  minimization for (7)

- 1: **Input:** Matrix  $\mathbf{Y}^T \mathbf{V} \odot \mathbf{V}$ ,  $\delta > 0$  and  $\epsilon > 0$ .
- 2: **Initialize**  $t = 0$ ,  $\mathbf{w}^{(0)} = \mathbf{1}_{NP}$  and  $\mathbf{X}^{(0)} = \mathbf{0}$ .
- 3: **repeat**
- 4:   Solve
 
$$\tilde{\mathbf{g}}^{(t+1)} = \underset{\tilde{\mathbf{g}}}{\operatorname{argmin}} \left\| \mathbf{w}^{(t)} \circ [(\mathbf{Y}^T \mathbf{V} \odot \mathbf{V}) \tilde{\mathbf{g}}] \right\|_1$$
 s. to  $\mathbf{1}_N^T \tilde{\mathbf{g}} = 1$ .
- 5:   Form  $\mathbf{X}^{(t+1)} = (\mathbf{Y}^T \mathbf{V} \odot \mathbf{V}) \tilde{\mathbf{g}}^{(t+1)}$ .
- 6:   Update  $w_i^{(t+1)} = \frac{1}{|\operatorname{vec}(\mathbf{X}^{(t+1)})|_i + \delta}$ ,  $i = 1, 2, \dots, NP$ .
- 7:    $t \leftarrow t + 1$ .
- 8: **until**  $\|\mathbf{X}^{(t+1)} - \mathbf{X}^{(t)}\|_1 / \|\mathbf{X}^{(t)}\|_1 \leq \epsilon$
- 9: **return**  $\hat{\mathbf{g}} := \tilde{\mathbf{g}}^{(t+1)}$  and  $\hat{\mathbf{X}} := \mathbf{X}^{(t+1)}$

tackle sparsity minimization problems have enjoyed remarkable success, since they often entail no loss of optimality. Accordingly, we instead: (i) seek to minimize the  $\ell_1$ -norm convex surrogate of the cardinality function, that is  $\|\mathbf{X}\|_1 = \sum_{i,j} |X_{ij}|$ ; and (ii) express the filter in the graph spectral domain as in (5) to obtain the cost

$$\|\mathbf{X}\|_1 = \|\mathbf{G}\mathbf{Y}\|_1 = \|\mathbf{V} \operatorname{diag}(\tilde{\mathbf{g}}) \mathbf{V}^T \mathbf{Y}\|_1 = \|(\mathbf{Y}^T \mathbf{V} \odot \mathbf{V}) \tilde{\mathbf{g}}\|_1,$$

where  $\odot$  denotes the Khatri-Rao (i.e., columnwise Kronecker) product. This suggests solving the convex  $\ell_1$ -synthesis problem (in this case a linear program), e.g., [24], namely

$$\hat{\mathbf{g}} = \underset{\tilde{\mathbf{g}} \in \mathbb{R}^N}{\operatorname{argmin}} \|(\mathbf{Y}^T \mathbf{V} \odot \mathbf{V}) \tilde{\mathbf{g}}\|_1, \quad \text{s. to } \mathbf{1}_N^T \tilde{\mathbf{g}} = 1. \quad (7)$$

While the linear constraint in (7) avoids  $\hat{\mathbf{g}} = \mathbf{0}$ , it also serves to fix the scale of the solution.

As a result, under the pragmatic assumption that the diffusion filter is invertible, one can readily use e.g., an off-the-shelf interior-point method or a specialized sparsity-minimization algorithm to solve (7) efficiently. Different from the solvers in [17], [22], the aforementioned algorithmic alternatives are free of expensive singular-value decompositions per iteration. We have found that overall performance can be improved via the iteratively-reweighted  $\ell_1$ -norm minimization procedure tabulated under Algorithm 1; see also [2] for a justification of such refinement. In any case, notice that once the frequency response  $\hat{\mathbf{g}}$  of the inverse filter is recovered, one can readily reconstruct the sources via  $\hat{\mathbf{X}} = (\mathbf{Y}^T \mathbf{V} \odot \mathbf{V}) \hat{\mathbf{g}}$  as well as the filter  $\mathbf{H}$ , if desired.

In the next section we will take a closer look at the inherent ambiguities associated with the bilinear model  $\mathbf{Y} = \mathbf{H}\mathbf{X}$ , some of which are unique to the network setting dealt with here. These are of course important to delineate the scope of identifiability (i.e., uniqueness) results. We will complete our discussion with deterministic sufficient conditions under which the convex relaxation (7) is tight.

#### IV. IDENTIFIABILITY AND EXACT RECOVERY

To establish further connections with blind deconvolution of periodic discrete-time signals, recall these can be viewed as graph signals supported on the directed cycle graph (whose

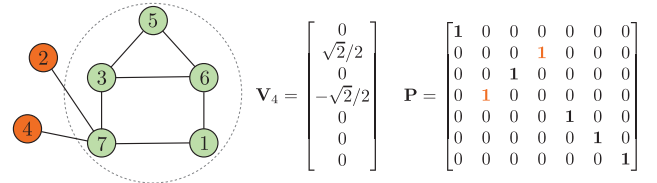


Fig. 1. Toy undirected graph (left) used to illustrate the symmetric permutation ambiguity between nodes 2 and 4. The fourth eigenvector  $\mathbf{v}_4$  of  $\mathbf{S} = \mathbf{A}$  (center) has the problematic form. Then if  $\{\mathbf{X}_0, \mathbf{h}_0\}$  satisfies the bilinear equations  $\mathbf{Y} = \mathbf{V} \operatorname{diag}(\tilde{\mathbf{h}}) \mathbf{V}^T \mathbf{X}$ , so does  $\{\mathbf{P}\mathbf{X}_0, \operatorname{diag}(\mathbf{p})\tilde{\mathbf{h}}_0\}$  for the shown permutation matrix  $\mathbf{P}$  (right) and  $\mathbf{p} = [1, 1, 1, -1, 1, 1, 1]^T$ .

circulant adjacency matrix is diagonalized by the DFT basis) [18]. In this special case, the blind identification task is known to suffer from unavoidable scaling and circulant-shift ambiguities; see e.g., [1], [23]. Here we examine more general symmetric permutation ambiguities arising with unweighted graphs, and briefly outline a relevant identifiability result as well as preliminary exact recovery conditions for (7).

#### A. Permutation ambiguities for some unweighted graphs

In solving the bilinear inverse problem formulated in Section II, for some particular graphs in addition to scaling we may also encounter (non-cyclic shift) permutation ambiguities. We can resolve the scaling ambiguity by e.g., a fortiori setting  $\|\tilde{\mathbf{g}}_0\|_1 = 1$  as in the experiments of Section V, where  $\tilde{\mathbf{g}}_0$  is the ground-truth frequency response of the inverse filter. Inspired by the identifiability studies for sparsity-constrained bilinear problems [12], here we examine said permutation ambiguities for unweighted graphs with shift  $\mathbf{S} = \mathbf{V}\mathbf{A}\mathbf{V}^T$ .

Let  $\{\mathbf{X}_0, \tilde{\mathbf{h}}_0\}$  collect the ground-truth sparse input signals and the filter's frequency response, respectively. Let  $\mathbf{u}^{(i,j)} \in \mathbb{R}^N$  be a unit-norm vector with zero entries except for  $u_i^{(i,j)} = -u_j^{(i,j)} = 1/\sqrt{2}$ . As we show next, a permutation ambiguity arises if, say, the  $k$ th eigenvector of  $\mathbf{S}$  (i.e., the  $k$ th column of  $\mathbf{V}$ ) has the form  $\mathbf{u}^{(i,j)}$ . Indeed, in that case one could introduce a binary signed vector  $\mathbf{p} \in \{-1, 1\}^N$  with a single negative entry  $p_k = -1$ , to construct another solution of the form

$$\mathbf{X}_1 := \mathbf{P}\mathbf{X}_0, \quad \tilde{\mathbf{h}}_1 := \operatorname{diag}(\mathbf{p})\tilde{\mathbf{h}}_0, \quad (8)$$

where  $\mathbf{P} = \mathbf{I}_N - 2\mathbf{u}^{(i,j)}(\mathbf{u}^{(i,j)})^T = \mathbf{V} \operatorname{diag}(\mathbf{p}) \mathbf{V}^T$  is a symmetric permutation matrix that interchanges the signal values at nodes  $i$  ( $x_i$ ) and  $j$  ( $x_j$ ) when applied to the graph signal  $\mathbf{x}$ . It is immediate that the pair in (8) satisfies the generative model  $\mathbf{Y} = \mathbf{H}\mathbf{X} = \mathbf{V} \operatorname{diag}(\tilde{\mathbf{h}}) \mathbf{V}^T \mathbf{X}$ . So, if  $\mathbf{u}^{(i,j)}$  is an eigenvector of  $\mathbf{S}$  then we can not distinguish the values at nodes  $i$  and  $j$  and the problem remains non-identifiable.

To exemplify this situation, consider the toy graph illustrated in Fig 1-(left). One can consider the adjacency matrix as the shift ( $\mathbf{S} = \mathbf{A}$ ) and denote the corresponding eigenvectors as  $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_7]$ . Fig 1-(center) shows that  $\mathbf{v}_4 = \mathbf{u}^{(2,4)}$ . Then it follows that for the matrix  $\mathbf{P}$  in Fig 1-(right) and the vector  $\mathbf{p} = [1, 1, 1, -1, 1, 1, 1]^T$ , one can construct another solution  $\{\mathbf{X}_1, \tilde{\mathbf{h}}_1\} \neq \{\mathbf{X}_0, \tilde{\mathbf{h}}_0\}$  using (8). In other words, nodes 2 and 4 are indistinguishable.

While it is challenging to obtain a formal characterization of problematic graphs, in practice we have encountered issues with dense networks as well as with some very sparse graphs.

For (continuous-valued) weighted graphs such ambiguities effectively disappear. Before moving on to issues of exact recovery, a remark on identifiability of (4) for a simple but widely adopted (random) sparsity model is in order.

**Remark 1 (Identifiability for Bernoulli-Gaussian model)**

Because of its analytical tractability, the Bernoulli-Gaussian model is widely adopted to describe and generate random sparse matrices such as  $\mathbf{X} \in \mathbb{R}^{N \times P}$  (we also use it for the simulations in Section V). Sparse matrices adhering to the model are  $\mathbf{X} = \Omega \circ \mathbf{R}$ , where  $\Omega \in \mathbb{R}^{N \times P}$  is an i.i.d. Bernoulli matrix with parameter  $\theta$  (i.e.,  $\mathbb{P}[\Omega_{ij} = 1] = \theta$ ), and  $\mathbf{R} \in \mathbb{R}^{N \times P}$  is an independent random matrix with i.i.d. symmetric random variables drawn from a standard Gaussian distribution. Under the Bernoulli-Gaussian model, [12, Proposition 40] asserts that problem (6) is identifiable (up to scaling and symmetric permutation ambiguities) with probability at least  $1 - \exp(-c\theta P)$ , for  $\frac{1}{N} < \theta < \frac{1}{4}$  and  $P > cN \log(N)$ , where  $c > 0$  is a sufficiently large constant.

*B. Exact recovery conditions*

Suppose that (6) is identifiable and let  $\{\mathbf{X}_0, \tilde{\mathbf{g}}_0\}$  be the solution. The following proposition (that relies heavily on [24, Theorem 1]) offers sufficient conditions under which the convex relaxation (7) succeeds in exactly recovering  $\{\mathbf{X}_0, \tilde{\mathbf{g}}_0\}$ .

**Proposition 1** Let  $\mathcal{I} := \text{supp}(\text{vec}(\mathbf{X}_0))$  index the non-zero entries of vectorized  $\mathbf{X}_0$ , and let  $\mathcal{I}^c$  be the complement of  $\mathcal{I}$ . Moreover, define  $\mathbf{Z} := \mathbf{Y}^T \mathbf{V} \odot \mathbf{V} \in \mathbb{R}^{NP \times N}$  and let  $\mathbf{Z}_{\mathcal{S}}$  be the submatrix of  $\mathbf{Z}$  with rows indexed by  $\mathcal{S} \subset \{1, 2, \dots, NP\}$ . Then, the solution to (7) is unique and equal to  $\tilde{\mathbf{g}}_0$  if the two following conditions are satisfied:

- C1)**  $\text{rank}(\mathbf{Z}_{\mathcal{I}^c}) = N - 1$ ; and
- C2)** There exists a vector  $\mathbf{f} \in \mathbb{R}^{NP}$  such that  $\mathbf{Z}^T \mathbf{f} = \gamma \mathbf{1}_N$  for some  $\gamma \neq 0$ , such that  $\mathbf{f}_{\mathcal{I}} = \text{sign}(\mathbf{Z}_{\mathcal{I}} \tilde{\mathbf{g}}_0)$  and  $\|\mathbf{f}_{\mathcal{I}^c}\|_{\infty} < 1$ .

**Proof:** As per [24, Theorem 1],  $\tilde{\mathbf{g}}_0$  is the unique solution of (7) if  $\ker(\mathbf{Z}_{\mathcal{I}^c}) \cap \ker(\mathbf{1}_N) = \{\mathbf{0}\}$ . But since  $\tilde{\mathbf{g}}_0 \in \ker(\mathbf{Z}_{\mathcal{I}^c})$  and  $\tilde{\mathbf{g}}_0 \notin \ker(\mathbf{1}_N)$  because of the constraint in (7), then C1) ensures said intersection is  $\{\mathbf{0}\}$ . Optimality condition C2) essentially requires  $\mathbf{1}_N$  to belong to the set of subgradients of  $\|\mathbf{Z} \tilde{\mathbf{g}}\|_1$  at  $\tilde{\mathbf{g}}_0$ ; see [24, Theorem 1] for further details. ■

Naturally, a more insightful exact recovery and sample complexity result along the lines of the one in Remark 1 would be most valuable [i.e., when are C1)-C2) satisfied for the Bernoulli-Gaussian model?], but left as future work.

## V. NUMERICAL RESULTS

We assess the performance of our proposed approach by testing the iteratively-reweighted  $\ell_1$ -norm minimization procedure in Algorithm 1. The per-iteration sparse recovery problems are solved using CVX [7].

**Simulation setup.** In all cases we consider undirected graphs with graph-shift operator chosen as the normalized adjacency matrix  $\mathbf{S} = \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}$ , where  $\mathbf{D} := \text{diag}(\mathbf{A} \mathbf{1}_N)$  is a diagonal matrix of node degrees. The ground-truth sparse input matrix  $\mathbf{X}_0$  is drawn from a Bernoulli-Gaussian model as in Remark 1, for varying parameters  $N$ ,  $P$ , and sparsity level

(i.e., number of nonzero entries)  $S$ . Filter coefficients  $\mathbf{h}_0$  are generated according to  $\mathbf{h}_0 = (\mathbf{e}_1 + \alpha \mathbf{b}) / \|\mathbf{e}_1 + \alpha \mathbf{b}\|_1$ , where  $\mathbf{e}_1 = [1, 0, \dots, 0]^T \in \mathbb{R}^L$  is the first canonical basis vector and entries of  $\mathbf{b} \in \mathbb{R}^L$  are drawn independently from a standard Gaussian distribution. Such a model for  $\mathbf{h}_0$  is inspired by [23], and we later corroborate that the recovery performance improves as  $\alpha$  decreases. Also note that  $\mathbf{h}_0$  is normalized to unit  $\ell_1$ -norm to fix the scale of the problem. Finally, given  $\mathbf{X}_0$  and  $\mathbf{H}_0 = \mathbf{V} \text{diag}(\Psi_L \mathbf{h}_0) \mathbf{V}^T$ , the  $N \times P$  matrix of observations is generated as  $\mathbf{Y} = \mathbf{H}_0 \mathbf{X}_0$ .

The relative recovery error  $e_X = \|\hat{\mathbf{X}} - \mathbf{X}_0\| / \|\mathbf{X}_0\|$  is adopted as figure of merit to evaluate algorithmic performance. We estimate the rate of successful recovery for synthetic and real-world graphs under different parameters by defining a successful recovery as one with  $e_X < 0.01$ .

**Random graphs.** Consider Erdős-Rényi random graphs with  $N = 50$  nodes, where edges are formed independently with probability  $p = 0.3$ . The rate of successful recovery is estimated for realizations of random graphs which are connected and do not give rise to permutation ambiguities (cf. Section IV-A). Figures 2(a) and 2(b) depict the recovery rates as a function of  $P$  and  $S$  for  $\alpha = 0.1$  and  $0.3$ , respectively, averaged over 100 realizations for (invertible) graph filters of order  $L = 5$ . As expected, in both cases recovery is more challenging for larger  $S$  and smaller  $P$ ; see the dark-gray area of low-success probability around the top-left corner. Moreover, decreasing  $\alpha$  makes successful recovery more likely. For instance, for  $\alpha = 0.1$  [Fig. 2(a)] we can successfully recover dense input signals with e.g.,  $S \approx N/2 = 25$  and only  $P = 10$  observations. For (effectively) lower-order filters resulting in more localized diffusion dynamics, one obtains favorable recovery performance.

We also compare the proposed approach against its less-scalable, matrix lifting-based precursor in [22]. Figures 2(c) and 2(d) respectively show the recovery rates for both methods as a function of sparsity  $S$  and filter order  $L$ , for  $N = 50$ ,  $p = 0.3$ ,  $P = 10$ , and  $\alpha = 0.5$  averaged over 20 realizations. Apparently, Algorithm 1 [Fig. 2(c)] can be successful over a larger range of values of  $L$ . Moreover, it uniformly outperforms the algorithm in [22, Problem (9)]; see Fig. 2(d).

**Brain graph.** We also consider a structural brain graph with  $N = 66$  nodes or neural regions of interest (ROIs), and edge weights given by the density of anatomical connections between regions [8]. The level of activity of each ROI can be represented by a graph signal  $\mathbf{x}$ , thus successive applications of  $\mathbf{S}$  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, then blind identification amounts to jointly estimating which regions were originally active, the activity in these regions and the coefficients of the linear combination.

We repeat the recovery-rate analysis performed for Erdős-Rényi graphs averaged over 20 realizations, and report the results in Figs. 2(e)–(h). Figures 2(e) and 2(f) showcase that our algorithm successfully identifies the initial excitation regions as well as the diffusion coefficients over a broad region in parameter space. By comparing Figs. 2(g) and 2(h), it is apparent that also in this setting the proposed approach

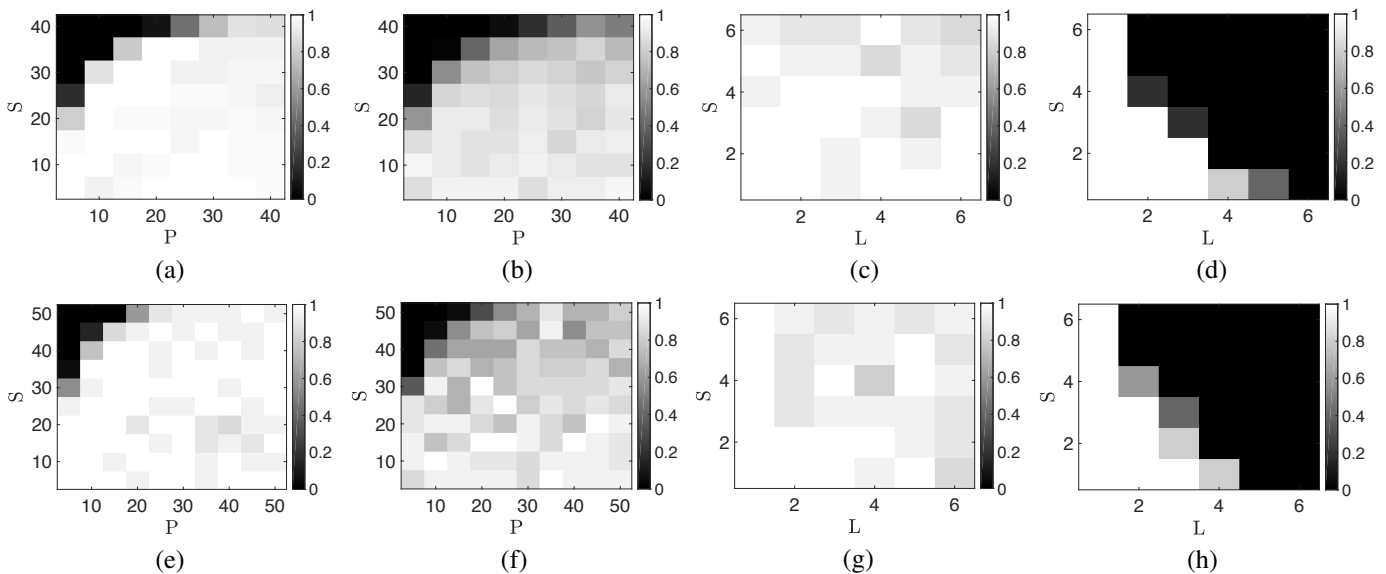


Fig. 2. Rate of recovery of  $\mathbf{X}$  as a function of  $S$  (number non-zero entries in  $\mathbf{X}$ ) and  $P$  (number of observations) in  $N = 50$ -node Erdős-Rényi graphs with  $p = 0.3$  (edge existence probability) for (a)  $\alpha = 0.1$  and (b)  $\alpha = 0.3$ , using Algorithm 1. Plots (e) and (f) are counterparts of (a) and (b), respectively, for the structural brain network in [8]. Recovery rate in Erdős-Rényi graphs ( $N = 50$ ,  $p = 0.3$ ) as a function of  $S$  and  $L$  (filter order) using (c) Algorithm 1 and (d) the matrix-lifting approach of [22]. Plots (g) and (h) are counterparts of (c) and (d), respectively, for the aforementioned structural brain network.

outperforms the state-of-the-art method in [22], corroborating the effectiveness of Algorithm 1.

## VI. CONCLUSION

We studied the problem of blind graph filter identification, which extends blind deconvolution of time (or spatial) domain signals to graphs. By introducing a mild assumption on invertibility of the graph filter, we obtained a computationally simpler convex relaxation for (diffused) source localization in the multi-signal case. Ongoing work includes deriving suitable graph-dependent conditions under which exact (and stable) recovery can be guaranteed, even when only a fraction of nodes is observed. This is a challenging problem, since the favorable (circulant) structure of time-domain filters is no longer present in the network-centric setting dealt with here.

## REFERENCES

- [1] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1711–1732, 2014.
- [2] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted  $\ell_1$  minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 877–905, 2008.
- [3] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, pp. 118–121, 1974.
- [4] J. A. Deri and J. M. F. Moura, "New York City taxi analysis with graph signal processing," in *Proc. IEEE Global Conf. on Signal and Information Process.*, Dec. 2016, pp. 1275–1279.
- [5] S. Feizi, M. Médard, G. Quon, M. Kellis, and K. Duffy, "Network infusion to infer information sources in networks," *arXiv preprint arXiv:1606.07383 [cs.SI]*, 2016.
- [6] A. Gavili and X.-P. Zhang, "On the shift operator, graph frequency, and optimal filtering in graph signal processing," *IEEE Trans. Signal Process.*, vol. 65, no. 23, pp. 6303–6318, 2017.
- [7] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx>, Mar. 2014.
- [8] 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 Biology*, vol. 6, no. 7, p. e159, 2008.
- [9] C. Hu, X. Hua, J. Ying, P. M. Thompson, G. E. Fakhri, and Q. Li, "Localizing sources of brain disease progression with network diffusion model," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 7, pp. 1214–1225, 2016.
- [10] W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Graph frequency analysis of brain signals," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 7, pp. 1189–1203, Oct. 2016.
- [11] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding blind deconvolution algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2354–2367, 2011.
- [12] Y. Li, K. Lee, and Y. Bresler, "Identifiability in bilinear inverse problems with applications to subspace or sparsity-constrained blind gain and phase calibration," *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 822–842, Feb 2017.
- [13] S. Ling and T. Strohmer, "Self-calibration and biconvex compressive sensing," *Inverse Problems*, vol. 31, no. 11, p. 115002, 2015.
- [14] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing," *arXiv preprint arXiv:1712.00468 [eess.SP]*, 2017.
- [15] R. Pena, X. Bresson, and P. Vandergheynst, "Source localization on graphs via  $\ell_1$  recovery and spectral graph theory," in *Proc. IEEE Image, Video, and Multidimensional Signal Process. Workshop*, 2016, pp. 1–5.
- [16] P. C. Pinto, P. Thiran, and M. Vetterli, "Locating the source of diffusion in large-scale networks," *Physical Review Letters*, vol. 109, no. 6, p. 068702, 2012.
- [17] D. Ramírez, A. G. Marques, and S. Segarra, "Graph-signal reconstruction and blind deconvolution for diffused sparse inputs," in *Proc. Int. Conf. Acoustics, Speech, Signal Process.*, Mar. 2017, pp. 4104–4108.
- [18] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, 2013.
- [19] I. D. Schizas, G. Mateos, and G. B. Giannakis, "Distributed LMS for consensus-based in-network adaptive processing," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2365–2381, Mar. 2009.
- [20] E. Sefer and C. Kingsford, "Diffusion archeology for diffusion progression history reconstruction," *Knowledge and information systems*, vol. 49, no. 2, pp. 403–427, 2016.
- [21] S. Segarra, A. G. Marques, and A. Ribeiro, "Optimal graph-filter design and applications to distributed linear network operators," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4117–4131, Aug 2017.
- [22] S. Segarra, G. Mateos, A. G. Marques, and A. Ribeiro, "Blind identification of graph filters," *IEEE Trans. Signal Process.*, vol. 65, no. 5, pp. 1146–1159, Mar. 2017.
- [23] L. Wang and Y. Chi, "Blind deconvolution from multiple sparse inputs," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1384–1388, 2016.
- [24] H. Zhang, M. Yan, and W. Yin, "One condition for solution uniqueness and robustness of both  $\ell_1$ -synthesis and  $\ell_1$ -analysis minimizations," *Adv. Comput. Math.*, vol. 42, no. 6, pp. 1381–1399, 2016.
- [25] P. Zhang, J. He, G. Long, G. Huang, and C. Zhang, "Towards anomalous diffusion sources detection in a large network," *ACM T. Internet Techn.*, vol. 16, no. 1, p. 2, 2016.