A filter-based definition for stationary random graph signals defined over a directed graph

Antonio G. Marques, *Senior Member, IEEE*

***Abstract*—Most existing graph signal processing frameworks and stochastic graph signal models rely on undirected or normal graph operators, which admit orthogonal spectral decompositions and enable a well-grounded definition of graph stationarity. Directed graph operators are typically non-Hermitian and non-normal, leading to non-orthogonal spectral representations and complicating the modeling and estimation of random graph signals. Assuming that the matrix representation of the graph is diagonalizable, this paper introduces a model for random stationary signals over directed graphs based on a generative filtering interpretation, whereby stationary signals are obtained as the output of a graph filter driven by white noise. Under this model, the covariance matrix does not generally share the eigenbasis of the graph operator, resulting in a non-orthogonal power spectral representation. We show that stationarity nevertheless induces a rank-1 low-dimensional covariance structure, propose efficient covariance estimators based on rank-1 approximations, and analyze their performance.**

***Index Terms*—Random modeling of graph signals, digraphs, graph stationarity, covariance estimation, spectral estimation.**

I. Introduction

Modeling irregular domain data as signals on graphs has become a common approach in signal processing and machine learning, spurring the development of graph signal processing (GSP) and graph-based deep learning. The graph encodes relational structure and enables transforms, filters, and inference tools that explicitly exploit topology. Much of GSP has focused on undirected graphs, where standard shift operators (GSOs) such as the Laplacian are symmetric and admit orthonormal eigenbases, yielding a well-defined graph Fourier transform (GFT) and a direct analogy with classical harmonic analysis. Within this setting, stationary graph signals were introduced in [1]–[3], showing that wide-sense stationarity is characterized by the covariance and the GSO sharing an eigenbasis, and that stationary signals can be generated by filtering white noise on the graph.

Extending GSP to *directed graphs* is more challenging [4]. Directed GSOs such as the adjacency matrix or directed Laplacians are typically non-Hermitian and non-normal [5], leading to non-orthogonal and potentially ill-conditioned eigenbases, which complicates frequency interpretation, diagonalization of second-order statistics, and filter stability. Existing work has followed several routes. Some approaches replace the directed GSO with a Hermitian surrogate, restoring orthogonality at the cost of redefining node-to-node processing [6]–[9]. Others retain directed operators and adopt spectral representations based on GSO eigenvectors, yielding exact analysis under diagonalizability but non-orthogonal coordinates [5]. Additionally, when the GSO is not diagonalizable, boundary conditions and additional edges can restore a proper spectral representation [10], [11]. A further class of directed GSP works relies on self-adjoint diffusion operators, but is restricted to specific graph classes [12]. Despite these advances, no generally accepted stationary signal model exists for arbitrary directed graphs. Probabilistic modeling of graph data has also been studied, from diffusion-based generative models [13] to classical stationarity models imposing low-dimensional structure on second-order statistics [1]–[3]. In GSP, stationarity has proven effective for covariance estimation, signal denoising, and graph inference. Nonetheless, existing random models almost exclusively assume undirected or normal GSOs.

Inspired by the undirected case, we introduce a model for random stationary signals over directed graphs based on a generative filtering interpretation. We define a stationary directed-graph process as the output of a linear graph filter driven by a white input. Under this definition, the covariance matrix is generally no longer a polynomial of the GSO and does not share its eigenbasis, leading to a non-orthogonal spectral representation. Nonetheless, stationarity remains informative, as it constrains the covariance to a structured family that can be estimated from a limited number of samples. To exploit this structure, we assume the GSO is diagonalizable and propose a novel covariance estimation approach that first exploits the sample covariance to estimate the frequency response of the generating filter (up to unavoidable scaling and phase ambiguities) and then uses this estimate to denoise the covariance matrix. The resulting method amounts to solving a rank-1 approximation problem and leads to an efficient algorithm whose performance is characterized in terms of the number of samples and the orthogonality of the GSO's eigenbasis.

II. Stationary Random Signals over DiGraphs

Consider a directed graph with N nodes, described by a GSO $\mathbf{S} \in \mathbb{R}^{N\times N}$. We model the GSO, graph signals and graph filters as real-valued, but all results also hold for the complex-valued case. Moreover, we assume that $\mathbf{S}$ is diagonalizable

$$\mathbf{S} = \mathbf{V}\mathrm{diag}(\boldsymbol{\lambda})\mathbf{V}^{-1}, \tag{1}$$

where the eigenvalue vector $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_N]^\top$ may have complex values and $\mathbf{V} \in \mathbb{C}^{N\times N}$ is not necessarily unitary. This setting encompasses common choices of directed GSOs, including adjacency matrices and (random-walk) Laplacians.

The author is with the Ellis Madrid Unit and King Juan Carlos University.

This paper is supported by the Spanish AEI (AEI/10.13039/501100011033) grant PID2022-136887NB-I00, and the Community of Madrid via IDEA-CM (TEC-2024/COM-89) and the Ellis Madrid Unit.

ChatGPT & Claude were used to polish the writing and edit the code. All research ideas and contributions are from the authors.

The matrix $\mathbf{V}^{-1}$ defines the GFT [5], so that a graph signal $\mathbf{x} \in \mathbb{R}^N$ can be equivalently represented in the graph frequency domain as $\tilde{\mathbf{x}} \in \mathbb{C}^N$, with $\tilde{\mathbf{x}} = \mathbf{V}^{-1}\mathbf{x}$. The GSO also plays a key role in defining graph filters. We define a *graph filter* as any real-valued linear operator of the form

$$\mathbf{H} = \mathbf{V}\mathrm{diag}\big(\tilde{\mathbf{h}}\big)\mathbf{V}^{-1}, \tag{2}$$

where $\tilde{\mathbf{h}} \in \mathbb{C}^N$ is the *spectral response*, with $\tilde{h}_i$ quantifying the gain at the i-th graph frequency. For $\mathbf{H}$ to be real-valued, $\tilde{\mathbf{h}}$ must satisfy the conjugate-symmetry condition $\tilde{h}_k = \bar{\tilde{h}}_{k'}$ for every conjugate eigenvalue pair with $\lambda_k = \bar{\lambda}_{k'}$. An important subclass of graph filters consists of *polynomial filters*, where $\tilde{h}_i = \sum_{k=0}^{K} h_k \lambda_i^k$ for real coefficients $\{h_k\}_{k=0}^{K}$, yielding $\mathbf{H} = \sum_{k=0}^{K} h_k \mathbf{S}^k$.

Definition 1. *A random graph signal* $\mathbf{x} \in \mathbb{C}^N$ *is* stationary *with respect to the directed GSO* $\mathbf{S}$ *if it can be generated as*

$$\mathbf{x} = \mathbf{H}\mathbf{w}, \tag{3}$$

where $\mathbf{H}$ is a graph filter in $\mathbf{S}$ [cf. (2)] and $\mathbf{w}$ is a zero-mean white random vector, so that $\mathbb{E}[\mathbf{w}] = \mathbf{0}$ and $\mathbb{E}[\mathbf{w}\mathbf{w}^{\mathrm{H}}] = \sigma^2\mathbf{I}$.

This definition corresponds to one of the three definitions provided in [3], which were all equivalent for normal GSOs. Here, no normality assumption on $\mathbf{S}$ is imposed.

Covariance and spectral covariance: The covariance matrix of the stationary signal $\mathbf{x}$ is

$$\mathbf{C}_x := \mathbb{E}[\mathbf{x}\mathbf{x}^{\mathrm{H}}] = \sigma^2\mathbf{H}\mathbf{H}^{\mathrm{H}}. \tag{4}$$

By construction, $\mathbf{C}_x$ is Hermitian positive semidefinite. Using $\mathbf{H} = \mathbf{V}\mathrm{diag}(\tilde{\mathbf{h}})\mathbf{V}^{-1}$, an important difference with the undirected case is that, in general, $\mathbf{C}_x$ and $\mathbf{S}$ do not share an eigenbasis. As will be apparent next, this induces correlation among the spectral components of $\mathbf{x}$ and complicates the estimation of its second-order statistics, motivating the non-convex estimators proposed in this paper.

Consider now the spectral random process $\tilde{\mathbf{x}} = \mathbf{V}^{-1}\mathbf{x}$, whose covariance matrix is $\tilde{\mathbf{C}}_x := \mathbb{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{H}}]$. Matrix $\tilde{\mathbf{C}}_x$, which we will refer to as the *spectral covariance matrix*, is related to $\mathbf{C}_x$ and the generating filter $\mathbf{H} = \mathbf{V}\mathrm{diag}\big(\tilde{\mathbf{h}}\big)\mathbf{V}^{-1}$ as

$$\tilde{\mathbf{C}}_x = \mathbb{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{H}}] = \mathbf{V}^{-1}\mathbf{C}_x\mathbf{V}^{-\mathrm{H}} \tag{5}$$

$$\tilde{\mathbf{C}}_x = \sigma^2\mathrm{diag}\big(\tilde{\mathbf{h}}\big)\mathbf{W}\mathrm{diag}\big(\bar{\tilde{\mathbf{h}}}\big), \quad \text{with } \mathbf{W} := \mathbf{V}^{-1}\mathbf{V}^{-\mathrm{H}}. \tag{6}$$

The matrix $\mathbf{W}$ captures the non-orthogonality of the eigenvectors of $\mathbf{S}$. When $\mathbf{S}$ is normal, $\mathbf{V}$ is unitary and $\mathbf{W} = \mathbf{I}$, recovering the classical result that, for stationary graph signals defined over normal GSOs, $\tilde{\mathbf{C}}_x$ is diagonal [3].

In other words, for directed graphs, the spectral covariance $\tilde{\mathbf{C}}_x$ is typically *non-diagonal*, with its off-diagonal entries quantifying correlations between graph frequencies induced by the non-orthogonality of $\mathbf{V}$. In particular, strong correlations arise when the eigenvectors of $\mathbf{S}$ are nearly linearly dependent. Thus, the structure of the Hermitian positive semidefinite $\tilde{\mathbf{C}}_x$ jointly reflects the spectral content $\tilde{\mathbf{h}}$ of the filter and the geometry of the eigenspace of $\mathbf{S}$.

The lack of orthogonality among the rows of $\mathbf{V}$ implies that Parseval's identity does not hold, and $\tilde{\mathbf{C}}_x$ cannot be interpreted as an energy spectral density. Indeed, $\mathbb{E}[\|\mathbf{x}\|_2^2] = \mathrm{tr}(\mathbf{C}_x) = \mathrm{tr}(\tilde{\mathbf{C}}_x\mathbf{W}^{-1})$, highlighting the critical role of $\mathbf{W}$ in relating vertex- and spectral-domain energy representations. For notational convenience, in the numerical experiments we write $\mathbf{u} = \sigma\tilde{\mathbf{h}}$, so that $\mathbf{C}_x = \mathbf{V}(\mathbf{W} \circ \mathbf{u}\mathbf{u}^{\mathrm{H}})\mathbf{V}^{\mathrm{H}}$.

III. Estimation of the Spectral Covariance

Assume that R independent realizations $\{\mathbf{x}_r\}_{r=1}^{R}$ of the stationary signal $\mathbf{x}$ are observed. The sample covariance matrix is defined as $\hat{\mathbf{C}}_x = \frac{1}{R}\sum_{r=1}^{R}\mathbf{x}_r\mathbf{x}_r^{\mathrm{H}}$. This section addresses two related questions: a) how stationarity can be leveraged to improve covariance estimation, and b) how to assess the quality of the resulting estimate, identifying the key parameters that govern its performance.

We begin by defining the spectral counterpart of $\hat{\mathbf{C}}_x$ as

$$\hat{\tilde{\mathbf{C}}}_x = \mathbf{V}^{-1}\hat{\mathbf{C}}_x\mathbf{V}^{-\mathrm{H}}. \tag{7}$$

Note that $\hat{\tilde{\mathbf{C}}}_x$ can equivalently be computed by first transforming each signal to the frequency domain, $\tilde{\mathbf{x}}_r = \mathbf{V}^{-1}\mathbf{x}_r$ for all r, and then forming the sample covariance of $\{\tilde{\mathbf{x}}_r\}_{r=1}^{R}$.

The next step exploits (6) to establish the following structural property, where $\circ$ denotes the Hadamard product.

Property 1. *The spectral covariance matrix* $\tilde{\mathbf{C}}_x$ *admits the representation*

$$\tilde{\mathbf{C}}_x = \sigma^2\mathbf{W} \circ \tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}}. \tag{8}$$

Property 1, which follows from (6), shows that $\tilde{\mathbf{C}}_x$ is given by the Hadamard product of the positive definite matrix $\mathbf{W}$ and the rank-1 Hermitian positive semidefinite matrix $\sigma^2\tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}}$. This observation reduces the estimation of the spectral response $\tilde{\mathbf{h}}$ to a rank-1 matrix recovery problem.

Leveraging (8), several estimation strategies can be devised. When all entries of $\mathbf{W}$ are nonzero, a simple approach is to first construct the positive semidefinite matrix

$$\hat{\mathbf{G}} = \sigma^2\hat{\tilde{\mathbf{C}}}_x \oslash \mathbf{W}, \tag{9}$$

where $\oslash$ denotes Hadamard division. One can then compute

$$\hat{\tilde{\mathbf{h}}}_a = \arg\min_{\tilde{\mathbf{h}}} \|\hat{\mathbf{G}} - \tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}}\|_F^2, \tag{10}$$

which is obtained by performing the SVD of $\hat{\mathbf{G}}$ and retaining only the dominant singular value–vector pair [14], [15]. Once an estimate of the frequency response is available, the spectral covariance is reconstructed as

$$\hat{\tilde{\mathbf{C}}}_x(\hat{\tilde{\mathbf{h}}}) = \sigma^2\mathbf{W} \circ \hat{\tilde{\mathbf{h}}}\hat{\tilde{\mathbf{h}}}^{\mathrm{H}}. \tag{11}$$

If σ^2 is unknown, it can be absorbed into $\hat{\tilde{\mathbf{h}}}$, which is then recovered up to a scaling factor and a global phase ambiguity (due to the outer product structure).

When some entries of $\mathbf{W}$ are zero, the entrywise division $\hat{\tilde{\mathbf{C}}}_x \oslash \mathbf{W}$ is not well defined. For a threshold $\tau > 0$, define $\mathcal{W}_\tau := \{(i,j) : |\mathbf{W}_{ij}| \geq \tau\}$, and the matrix $\mathbf{G}_\tau$ with entries $[\mathbf{G}_\tau]_{ij} = \sigma^2[\tilde{\mathbf{C}}_x]_{ij}/[\mathbf{W}]_{ij}$ if $(i,j) \in \mathcal{W}_\tau$ and $[\mathbf{G}_\tau]_{ij} = 0$ otherwise. The estimate $\hat{\tilde{\mathbf{h}}}_b$ is then obtained by solving the masked rank-1 recovery problem

$$\hat{\tilde{\mathbf{h}}}_b = \arg\min_{\tilde{\mathbf{h}}\in\mathbb{C}^N} \sum\nolimits_{(i,j)\in\mathcal{W}_\tau} \big|[\hat{\mathbf{G}}_\tau]_{ij} - \tilde{h}_i\bar{\tilde{h}}_j\big|^2, \tag{12}$$

and substituting the result into (11). Since $\mathbf{G}_\tau$ typically contains substantial redundancy (up to N^2 measurements for N unknowns), this approach is expected to perform well provided that $[\hat{\mathbf{G}}_\tau]_{ij}$ is sufficiently close to its noiseless counterpart.

Finally, a more direct strategy is to solve the weighted rank-1 approximation

$$\hat{\tilde{\mathbf{h}}}_c = \arg\min_{\tilde{\mathbf{h}}\in\mathbb{C}^N} \sum\nolimits_{(i,j)\in\mathcal{W}_\tau} \left|[\hat{\tilde{\mathbf{C}}}_x]_{ij} - [\mathbf{W}]_{ij}\tilde{h}_i\bar{\tilde{h}}_j\right|^2, \quad (13)$$

and use the resulting estimate in (11).

All in all, three optimizations are proposed (10), (12), and (13). In all of them, we optimize over the spectral filter response and assume access to the GSO eigenvectors and the sample covariance matrix. The three optimizations are non-convex but relatively tractable. For the fully observed case in (10), the global optimum is found via SVD [15]. The masked and weighted formulations in (12) and (13) fall within the class of low-rank recovery problems with a benign optimization landscape and global optimality guarantees. More specifically, (12) and (13) can be efficiently addressed using low-rank factorization techniques. A common strategy consists of initializing via a spectral method (e.g., the dominant eigenvector of a thresholded version of $\hat{\tilde{\mathbf{C}}}_x$) followed by gradient-based refinement [16], [17]. For masked formulations such as (12), these approaches are closely related to low-rank matrix completion, for which it is known that, under mild incoherence conditions and sufficiently dense observation patterns, the associated nonconvex objectives exhibit no spurious local minima and gradient-based methods converge geometrically [17], [18]. Similar guarantees extend to weighted problems (13) (see, e.g., [19]).

Estimation guarantees: We study how the estimation error depends on R and on the spectral properties of the graph through $\mathbf{V}$ and $\mathbf{W}$. Assume that $\{\mathbf{x}_r\}_{r=1}^R$ are independent, zero-mean, sub-Gaussian random vectors with covariance $\mathbf{C}_x$. Standard covariance concentration results imply that, with probability at least $1-\delta$,

$$\|\hat{\mathbf{C}}_x - \mathbf{C}_x\|_F \le \kappa\,\|\mathbf{C}_x\|\left(\sqrt{\frac{N+\log(1/\delta)}{R}} + \frac{N+\log(1/\delta)}{R}\right),$$

for an absolute constant κ. The second term is negligible whenever $R \gtrsim N$. Passing to the spectral domain,

$$\begin{aligned}\|\hat{\tilde{\mathbf{C}}}_x - \tilde{\mathbf{C}}_x\|_F &\le \|\mathbf{V}^{-1}\|^2\,\|\hat{\mathbf{C}}_x - \mathbf{C}_x\|_F \\ &\le \kappa\,\|\mathbf{C}_x\|\,\|\mathbf{V}^{-1}\|^2\sqrt{\frac{N+\log(1/\delta)}{R}}. \quad (14)\end{aligned}$$

This is the first point where the directed nature of the GSO manifests itself in the statistics: the non-orthogonality of $\mathbf{V}$ amplifies covariance estimation error by a factor $\|\mathbf{V}^{-1}\|^2$. This factor will also appear when bounding the error in the spectral filter response. Clearly, $\|\mathbf{V}^{-1}\|^2$ can become large when the eigenvector matrix is ill-conditioned. In those cases, mitigation strategies such as those in [10], [11], which improve eigenvector conditioning by implementing small graph modifications, are well motivated.

Rank-1 structure and recovery of $\tilde{\mathbf{h}}$. When $\mathbf{W}$ has no zero entries, the rank-1 structure $\tilde{\mathbf{C}}_x = \mathbf{W}\circ(\tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}})$ can be exploited to estimate $\tilde{\mathbf{h}}$. Recall the definition of $\hat{\mathbf{G}}$ in (9), and $\mathbf{G} := \tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}}$ and $W_{\min} = \min_{ij}|[\mathbf{W}]_{ij}|$. We then have

$$\|\hat{\mathbf{G}} - \mathbf{G}\|_F = \|\mathbf{W}^{\circ(-1)}\circ(\hat{\tilde{\mathbf{C}}}_x - \tilde{\mathbf{C}}_x)\|_F \le \frac{\|\hat{\tilde{\mathbf{C}}}_x - \tilde{\mathbf{C}}_x\|_F}{W_{\min}}. \quad (15)$$

Since $\mathbf{G}$ is Hermitian, positive semidefinite, and rank-1, its largest eigenvalue is $\lambda_1(\mathbf{G}) = \|\tilde{\mathbf{h}}\|_2^2$ and all remaining eigenvalues are zero. Let $\hat{\tilde{\mathbf{h}}}_a$ denote the leading eigenvector of $\hat{\mathbf{G}}$. By the Davis–Kahan $\sin\Theta$ theorem,

$$\min_\theta \sin\angle(\hat{\tilde{\mathbf{h}}}_a, \tilde{\mathbf{h}}e^{\mathrm{j}\theta}) \;\le\; \frac{2\|\hat{\mathbf{G}} - \mathbf{G}\|}{\|\tilde{\mathbf{h}}\|_2^2} \;\le\; \frac{2\|\hat{\mathbf{G}} - \mathbf{G}\|_F}{\|\tilde{\mathbf{h}}\|_2^2}, \quad (16)$$

where we used $\|\cdot\| \le \|\cdot\|_F$. Combining (14)–(16) yields

$$\min_\theta \sin\angle(\hat{\tilde{\mathbf{h}}}_a, \tilde{\mathbf{h}}e^{\mathrm{j}\theta}) \le \frac{2\kappa\|\mathbf{C}_x\|\|\mathbf{V}^{-1}\|^2}{W_{\min}\|\tilde{\mathbf{h}}\|_2^2}\sqrt{\frac{N+\log(1/\delta)}{R}}. \quad (17)$$

For real-valued GSOs, filter responses at real eigenvalues satisfy $\tilde{h}_k \in \mathbb{R}$. This structural constraint, which follows directly from the conjugate-symmetry condition introduced in Section II, can be exploited to resolve the phase ambiguity at those indices and improve estimation reliability. By Perron–Frobenius, at least one eigenvalue of a non-negative GSO is always real and positive, so this constraint is always active in practice.

Error with respect to the enhanced covariance estimate. Let $\hat{\tilde{\mathbf{C}}}_x^c$ denote $\hat{\tilde{\mathbf{C}}}_x(\hat{\tilde{\mathbf{h}}}_c)$, i.e., the matrix obtained after substituting the minimizer in (13) into (11). Since the true $\tilde{\mathbf{C}}_x$ is feasible, optimality implies

$$\|\hat{\tilde{\mathbf{C}}}_x^c - \hat{\tilde{\mathbf{C}}}_x\|_F \;\le\; \|\tilde{\mathbf{C}}_x - \hat{\tilde{\mathbf{C}}}_x\|_F.$$

Therefore, by the triangle inequality,

$$\begin{aligned}\|\hat{\tilde{\mathbf{C}}}_x^c - \tilde{\mathbf{C}}_x\|_F &\le \|\hat{\tilde{\mathbf{C}}}_x^c - \hat{\tilde{\mathbf{C}}}_x\|_F + \|\hat{\tilde{\mathbf{C}}}_x - \tilde{\mathbf{C}}_x\|_F \\ &\le 2\|\hat{\tilde{\mathbf{C}}}_x - \tilde{\mathbf{C}}_x\|_F \le 2\kappa\|\mathbf{C}_x\|\|\mathbf{V}^{-1}\|^2\sqrt{\frac{N+\log(1/\delta)}{R}}. \quad (18)\end{aligned}$$

Hence, as $R\to\infty$, the error vanishes and the estimator is consistent. The bound in (18) is valid for any arbitrary error; a more challenging problem is to characterize performance when $\|\hat{\mathbf{C}}_x - \mathbf{C}_x\|_F$ is moderate but non-negligible. The behavior depends critically on $\mathbf{W}$ and on whether the goal is to estimate $\tilde{\mathbf{h}}$, $\tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}}$, or directly $\tilde{\mathbf{C}}_x$. Intuitively, estimating $\tilde{\mathbf{h}}$ becomes difficult (or impossible) when many entries of $\mathbf{W}$ are zero or very small. This phenomenon is well known for undirected graphs, where the power spectral density does not allow one to recover the sign of each spectral coefficient. In contrast, in the directed case, $\tilde{\mathbf{h}}$ can be recovered up to a global phase factor.

More formally, local performance is governed by the curvature of $f(\tilde{\mathbf{h}}) = \|\hat{\tilde{\mathbf{C}}}_x - \mathbf{W}\circ(\tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}})\|_F^2$ around the true solution. When the goal is to estimate $\tilde{\mathbf{C}}_x$ directly, zero entries in $\mathbf{W}$ are less problematic, since the corresponding entries of $\tilde{\mathbf{C}}_x$ are known to be zero as well. In this case, it is natural to analyze the curvature of $F(\mathbf{C}) = \frac{1}{2}\|\hat{\tilde{\mathbf{C}}}_x - \mathbf{C}\|_F^2$ restricted to the manifold $\mathcal{C} = \{\mathbf{W}\circ(\tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}}) : \tilde{\mathbf{h}}\in\mathbb{C}^N\}$. The local conditioning is controlled by the smallest nonzero singular value of the linear operator $J(\mathbf{d}) = \mathbf{W}\circ(\tilde{\mathbf{h}}\mathbf{d}^{\mathrm{H}} + \mathbf{d}\tilde{\mathbf{h}}^{\mathrm{H}})$. A detailed analysis of this curvature is left for future work.

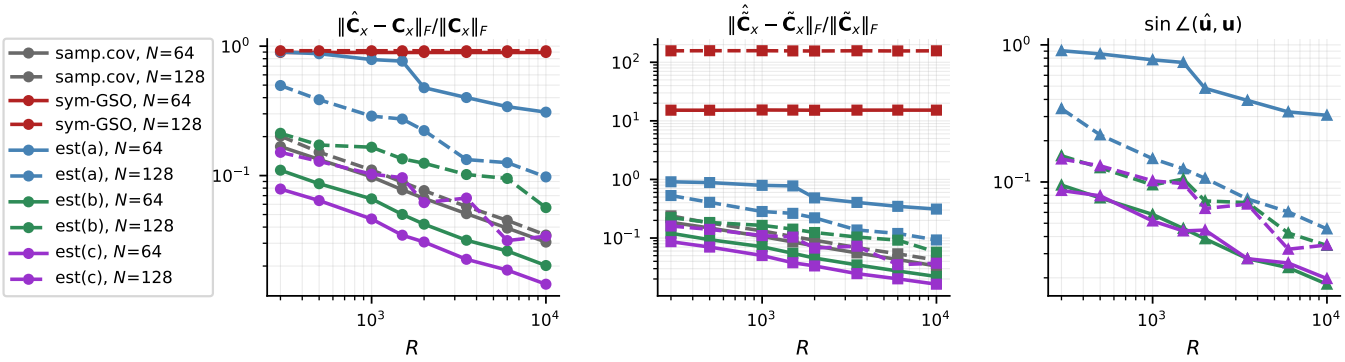

Fig. 1: Estimation performance as the number of realizations R increases. Five estimators are shown: sample covariance, sym-GSO baseline, est (a) (10), est (b) (12), and est (c) (13). Two graph sizes: $N = 64$ (solid) and $N = 128$ (dashed).

IV. Numerical Results

We assess spectral estimation using synthetic data generated from real-valued directed graphs. The GSO $\mathbf{S}$ is constructed as an unweighted directed adjacency matrix obtained from an Erdős–Rényi graph, followed by a random orientation of a subset of edges and the addition of new directed edges. Signals are real-valued and generated as independent realizations of $\mathcal{N}(\mathbf{0}, \mathbf{C}_x)$ with $\mathbf{C}_x = \mathbf{V}(\mathbf{W} \circ \mathbf{u}\mathbf{u}^{\mathrm{H}})\mathbf{V}^{\mathrm{H}}$, where $\mathbf{u} = \sigma\tilde{\mathbf{h}} \in \mathbb{C}^N$ satisfies the conjugate-symmetry condition $u_k = \bar{u}_{k'}$ for every conjugate eigenvalue pair with $\lambda_k = \bar{\lambda}_{k'}$, ensuring $\mathbf{C}_x$ is real-valued. In the simulations, we set the edge density to 0.25, the orientation fraction to 0.50, and the addition fraction to 0.40.

We compare the following estimators: the baseline sample covariance estimator (7); a symmetrized-GSO (sym-GSO) baseline that applies the undirected estimator of [3] to $(\mathbf{S} + \mathbf{S}^\top)/2$; the rank-1 estimator in (10); the masked estimator in (12) (initialized with (10)); and the weighted gradient-based estimator in (13) (initialized with (12)). Performance in Fig. 1 is evaluated in terms of $\|\hat{\tilde{\mathbf{C}}}_x - \tilde{\mathbf{C}}_x\|_F/\|\tilde{\mathbf{C}}_x\|_F$ (central panel) as R varies. To gain insights, $\|\hat{\mathbf{C}}_x - \mathbf{C}_x\|_F/\|\mathbf{C}_x\|_F$ (left panel) and $\sin\angle(\hat{\mathbf{u}}, \mathbf{u})$ (right panel) are also plotted. Results are averages over 30 Monte Carlo realizations for $N = 64$ and $N = 128$. Additional results, and details on initialization, step-size and hyperparameter selection can be found in the online repository [20].

Discussion: The results highlight several consistent trends. First, the sym-GSO baseline, which ignores the directed structure by symmetrizing $\mathbf{S}$, yields large errors that do not decrease with R, confirming that the undirected stationarity assumption is fundamentally mismatched when the graph is directed, and motivating the framework developed in this letter. Second, the rank-1 estimator exhibits noticeably poorer performance than the masked one, indicating that small entries in $\mathbf{W}$ critically affect naïve rank-1 estimation. Third, the best spectral estimates (central panel) are generated by the weighted gradient-based estimator and, to a slightly lesser extent, its masked-gradient predecessor, demonstrating the benefit of explicitly incorporating the weights into the cost. The right panel reveals that the ranking of the estimators in the spectral domain (central panel) closely mirrors that of $\mathbf{u}$ estimation, confirming that spectral error is driven by the quality of $\hat{\mathbf{u}}$. Fourth, while the qualitative behavior of the estimators is stable, the relative performance gaps depend moderately on the specific graph realization and sampling regime, reflecting differences in conditioning and sparsity patterns induced by $\mathbf{W}$. Finally, since the proposed formulations explicitly target $\tilde{\mathbf{C}}_x$ rather than $\mathbf{C}_x$, the gap between the structured estimators and the sample covariance baseline is more pronounced in the spectral domain. If $\mathbf{C}_x$ were the primary objective, one could reformulate the estimation problems to directly penalize the vertex-domain error (see next section).

V. Extensions and Future Directions

This letter focused on non-parametric estimation of the spectral covariance $\tilde{\mathbf{C}}_x$ for a general filter response $\tilde{\mathbf{h}}$. We outline several natural directions for future work.

Parametric models. When the filter belongs to a low-dimensional parametric family $\tilde{h}_i = g_{\boldsymbol{\theta}}(\lambda_i)$, one substitutes $\tilde{h}_i\bar{\tilde{h}}_j$ with $g_{\boldsymbol{\theta}}(\lambda_i)\bar{g}_{\boldsymbol{\theta}}(\lambda_j)$ and optimizes over $\boldsymbol{\theta}$, significantly reducing sample complexity. A particularly relevant example is *polynomial kernels*, which naturally generalize MA/AR/ARMA models to digraphs [21].

Approximate GSOs. When eigenvector conditioning is poor, a practical alternative is to estimate the spectral covariance using a surrogate GSO with better spectral behavior. Candidates include the graph modifications of [10], [11], which either improve conditioning or enable diagonalization while preserving directed structure, as well as symmetrized operators such as $(\mathbf{S}+\mathbf{S}^\top)/2$ or $(\mathbf{S}\mathbf{S}^\top+\mathbf{S}^\top\mathbf{S})/2$, which reduce the problem to the undirected setting [3]. Note that symmetrized surrogates are only expected to be reliable when the graph is weakly directed; as illustrated in the numerical experiments, they can fail significantly when the directed structure is pronounced. An important open question is which spectral profiles (low-pass, bandlimited, smooth) are best suited to each surrogate.

Further directions. Additional open problems include *vertex-domain reformulations* that directly penalize the Frobenius error of $\mathbf{C}_x$ rather than $\tilde{\mathbf{C}}_x$; *Wiener filter design* for directed graphs; generalizations to directed hypergraphs; and *real-data validation* on directed network datasets.

REFERENCES

[1] B. Girault, "Stationary graph signals using an isometric graph translation," in *Eur. Signal Process. Conf. (EUSIPCO)*, 2015, pp. 1516–1520.

[2] N. Perraudin and P. Vandergheynst, "Stationary signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3462–3477, 2017.

[3] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5911–5926, 2017.

[4] A. G. Marques, S. Segarra, and G. Mateos, "Signal processing on directed graphs: The role of edge directionality when processing and learning from network data," *IEEE Signal Process. Mag.*, vol. 37, no. 6, pp. 99–116, 2020.

[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, 2014.

[6] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, "On the graph Fourier transform for directed graphs," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 6, pp. 796–811, 2017.

[7] R. Shafipour, A. Khodabakhsh, G. Mateos, and E. Nikolova, "A directed graph Fourier transform with spread frequency components," *IEEE Trans. Signal Process.*, vol. 67, no. 4, pp. 946–960, 2018.

[8] C. Chan, A. Cionca, and D. V. D. Ville, "Hilbert transform on graphs: Let there be phase," *IEEE Signal Process. Lett.*, pp. 1725–1729, 2025.

[9] D. B. Tay, "Dual frames of localized spectral graph wavelets," *Digital Signal Processing*, vol. 161, p. 105094, 2025.

[10] B. Seifert and M. Püschel, "Digraph signal processing with generalized boundary conditions," *IEEE Trans. Signal Process.*, vol. 69, pp. 1422–1437, 2021.

[11] A. B. Bardi, T. Yazdanpanah, M. Daković, and L. Stanković, "Graph fourier transform enhancement through envelope extensions," *IEEE Trans. Signal Process.*, vol. 73, pp. 3598–3613, 2025.

[12] H. Sevi, G. Rilling, and P. Borgnat, "Modeling signals over directed graphs through filtering," in *IEEE Global Conf. Signal and Info. Process. (GlobalSIP)*, 2018, pp. 718–722.

[13] S. Rozada, A. Cavallo, A. G. Marques, H. Jamali-Rad, E. Isufi *et al.*, "Graph-aware diffusion for signal generation," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2026.

[14] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[15] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Johns Hopkins Univ. Press, 2013.

[16] R. H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from a few entries," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2980–2998, 2010.

[17] R. Ge, C. Jin, and Y. Zheng, "No spurious local minima in nonconvex low rank problems: A unified geometric analysis," in *Intl. Conf. Mach. Learn. (ICML)*, 2017, pp. 1233–1242.

[18] S. Bhojanapalli, B. Neyshabur, and N. Srebro, "Global optimality of local search for low rank matrix recovery," in *Adv. Neural Inf. Process. Systems (NIPS)*, vol. 29, 2016.

[19] N. Srebro and T. Jaakkola, "Weighted low-rank approximations," in *Intl. Conf. Mach. Learn. (ICML)*, 2003, pp. 720–727.

[20] A. G. Marques. (2026) Stationary signals on directed graphs: Code repository. [Online]. Available: https://github.com/agmarques/stationarysignals_digraphs

[21] E. Isufi, A. Loukas, A. Simonetto, and G. Leus, "Autoregressive moving average graph filtering," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 274–288, 2016.