

Controllability of Bandlimited Graph Processes Over Random Time Varying Graphs

Fernando Gama, Elvin Isufi, Alejandro Ribeiro and Geert Leus

Abstract—Controllability of complex networks arises in many technological problems involving social, financial, road, communication, and smart grid networks. In many practical situations, the underlying topology might change randomly with time, due to link failures such as changing friendships, road blocks or sensor malfunctions. Thus, it leads to poorly controlled dynamics if randomness is not properly accounted for. We consider the problem of controlling the network state when the topology varies randomly with time. Our problem concerns target states that are *bandlimited* over the graph; these are states that have nonzero frequency content only on a specific graph frequency band. We thus leverage graph signal processing and exploit the bandlimited model to drive the network state from a fixed set of control nodes. When controlling the state from a few nodes, we observe that spurious, out-of-band frequency content is created. Therefore, we focus on controlling the network state over the desired frequency band, and then use a graph filter to get rid of the unwanted frequency content. To account for the topological randomness, we develop the concept of controllability in the mean, which consists of driving the expected network state towards the target state. A detailed mean squared error analysis is performed to quantify the statistical deviation between the final controlled state on a particular graph realization and the actual target state. Finally, we propose different control strategies and evaluate their effectiveness on synthetic network models and social networks.

Index Terms—Graph signal processing, random graphs, network controllability, graph signals, graph process, linear systems on graphs

I. INTRODUCTION

The controllability of complex networks plays a fundamental role in our understanding of natural and technological systems. Relevant examples involve the control of social, biological, financial, road, communication, and smart grid networks. Different works have highlighted the importance of the network structure when controlling a system evolving on top of that network [2]–[4]. Other works controlled said system from a few control or *driving* nodes [5]–[7]. As an illustrative example, consider the Zachary’s Karate club social network in Figure 1 [8]. The network state may be an opinion profile (e.g., members thoughts on a topic) and controlling, or *driving*, the network amounts to shaping those opinions towards a *desired* or *target* state. The social relationships between members affect the ability to control the opinion

profile and the objective is to sway all members opinions from a few influencing members (the *driving* nodes).

While providing seminal contributions on network controllability, the works in [2]–[7] ignore the coupling between the underlying topology and the target state (a.k.a. *the graph signal*). Recent evidence from graph signal processing (GSP) [9]–[11] has shown that this coupling can bring substantial benefits in graph signal sampling [12], interpolation [13], [14], adaptive reconstruction [15], and observability of diffusion processes [16]. A common point that unifies [10]–[16] is the so-called graph Fourier transform (GFT). The GFT expands the network state onto an orthonormal basis related to the underlying topology—formed by the eigenvectors of a matrix that represents the network, such as the adjacency or the Laplacian matrix—akin to how the discrete Fourier transform expands a time signal on the complex exponential orthonormal basis. The GFT basis vectors can be linked to different variability modes across the graph through the concept of total variation [10], [17]; hence, named the *graph oscillating modes*.

A particular class of graph signals is that of *bandlimited graph signals*, i.e., signals that can be expressed by only a few graph oscillating modes. The number of active modes forms the graph signal *bandwidth*. Likewise for time signals, purely bandlimited graph signals rarely exist. But the bandlimitedness prior poses a powerful and parsimonious model to develop practical tools. This prior is exploited in GSP for sampling [10], [11], [18], where signals arising in economic [19], handwritten digits [12], [14] or brain functional imaging [20], are approximately bandlimited. In the Zachary’s Karate club example, a bandlimited network state corresponds to opinions polarized into a few clusters of like-minded members [12]. Therefore, controlling the system to a bandlimited state implies imposing a similar opinion profile to members that influence each other; this influence is captured by the edge weight.

Network control towards a bandlimited graph signal is considered in [21]. The control signal is fed to a few driving nodes and is percolated through the graph until the target state is reached. The authors determined the trade-off between the control time and the number of driving nodes, provided conditions to reach any bandlimited state, and designed the control signals.

The work in [22] studied the challenge of controlling the network towards a bandlimited state with control signals of limited energy. The main result is the trade-off between the number of driving nodes and the control signal energy. But conditions to drive the network to any bandlimited state were not derived in this limited energy setting.

Along these lines, and in parallel with the shorter version

Part of this work is presented in [1]. Work in this paper is supported by NSF CCF 1717120, ARO W911NF1710438, ARL DCIST CRA W911NF-17-2-0181, ISTC-WAS and Intel DevCloud. F. Gama, and A. Ribeiro are with the Dept. of Electrical and Systems Eng., Univ. of Pennsylvania, USA. E. Isufi is with the Intelligent Systems Dept. and G. Leus is with the Dept. of Microelectronics, Delft Univ. of Technology, The Netherlands. E-mails: {fgama, aribeiro}@seas.upenn.edu, {e.isufi-1, g.j.t.leus}@tudelft.nl.

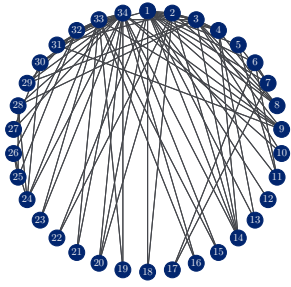


Figure 1. Zachary's Karate club network. The nodes represent club members and the edges capture social relationships between the members.

of this paper [1], the work in [23] used GSP to formulate the linear quadratic controller as an autoregressive moving average (ARMA) graph filter [24]. This ARMA formulation is used to control independently each graph mode to the desired state. But these control strategies require all nodes to act as driving nodes.

Altogether, the above works concern network control over time-invariant topologies. But in practice the network structure may change randomly over time due to link losses, or nodes that disappear with a given probability. This might be the case of: *i*) members that are not present a given day in the Zachary's Karate club; *ii*) communication links that are random in sensor networks; *iii*) power lines and buses that go down in smart grids due to local failures. In these situations, the graph structure is random and can be characterized by an expected graph with a variance around this expectation. Network controllability becomes, therefore, challenging since some connections cannot be exploited; hence, the derivations obtained in the deterministic setting may lead to a completely different state.

Motivated by the above observations, we study the possibilities to perform open loop control over random time varying networks. By exploring GSP tools as in [21]–[23], we propose a framework that accounts for the graph randomness in the analysis. Our main contributions are:

- 1) We study the problem of controlling dynamics over deterministic networks from a few driving nodes (Section III). We provide conditions on network controllability that relate the minimum number of driving nodes to the target control bandwidth and the control time. This result encompasses the three strategies proposed in [21].
- 2) We formulate the problem of controlling dynamics over random time varying networks from a few driving nodes. (Section IV). We develop the concept of controllability in the mean to drive the expected state towards a graph signal with a target bandwidth content on the expected graph. We also extend the conditions on network controllability from the deterministic to the stochastic setting.
- 3) We perform a mean squared error analysis to quantify the statistical deviation of the controlled state from the target state on a particular graph realization (Section IV-B). This analysis illustrates the role of the random graph model, the target state bandwidth, and the control signal.

- 4) We propose two control strategies to drive the expected state to the desired bandlimited state with the minimum mean squared error (Section V).
- 5) We corroborate the developed framework and study its performance on synthetic (Erdős-Rényi and geometric graphs) and real-world social networks (Facebook sub-graph and Zachary's Karate club) (Section VI).

To the best of our knowledge, this is the first contribution that approaches network controllability over random time varying topologies. We remark that while this work relies on the bandlimited prior to select the driving nodes, the reader might find of interest other parsimonious models [18], [25], [26] for such a task.

The remaining part of the paper proceeds as follows: Section II sets down the preliminary concepts and Section III contains our formulation of network controllability on deterministic graphs. Section IV formulates the controllability on random graphs. Section V develops the control strategies, while Section VI presents the numerical experiments. Finally, Section VII provides the concluding remarks. The proofs are collected in the appendix.

Notation. Normal letters a (or A) denote scalars, bold lower-case letters \mathbf{a} vectors, and bold uppercase letters \mathbf{A} matrices. The i th entry of a vector is $[\mathbf{a}]_i$ while the (i, j) th entry of a matrix is $[\mathbf{A}]_{i,j}$. Superscripts T and H denote the transpose and the Hermitian, respectively. The $N \times 1$ null vector is $\mathbf{0}_N$, the $N \times 1$ vector of all ones is $\mathbf{1}_N$ and the $N \times N$ identity matrix is \mathbf{I}_N . The diagonal operator denoted as $\text{diag}(\cdot)$ is defined such that $\mathbf{a} = \text{diag}(\mathbf{A})$ with $[\mathbf{a}]_i = [\mathbf{A}]_{i,i}$, and $\mathbf{A} = \text{diag}(\mathbf{a})$ is a diagonal matrix with vector \mathbf{a} on the main diagonal. The expectation operator is denoted as $\mathbb{E}[\cdot]$, the trace operator as $\text{tr}(\cdot)$, the rank of \mathbf{A} as $\text{rank}(\mathbf{A})$, the Kronecker product as \otimes , and the element-wise Hadamard product as \circ . The l_p vector and matrix norm is denoted as $\|\cdot\|_p$. The ceiling operator is denoted as $\lceil \cdot \rceil$ and the minimum and maximum operators as $\min\{\cdot\}$ and $\max\{\cdot\}$, respectively. If not otherwise stated, calligraphic letters \mathcal{A} indicate sets and the set cardinality is denoted as $|\mathcal{A}|$.

II. DIFFUSION PROCESSES ON GRAPHS

In this work, we consider controlling a diffusion process on random time varying graphs towards a desired state. To achieve this, we model diffusion processes through graph signal processing. We introduce the basics of GSP in Section II-A, define the random time varying graph model in Section II-B, and discuss diffusion processes in Section II-C.

A. Graph signal processing

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ denote a graph with $\mathcal{V} = \{v_1, \dots, v_N\}$ the set of N vertices, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ the set of edges, and $\mathcal{W} : \mathcal{E} \rightarrow \mathbb{R}_+$ a function that assigns positive weights to the edges. The graph serves as a mathematical representation of the network and its structure is captured by the graph shift operator (GSO) matrix $\mathbf{S} \in \mathbb{R}^{N \times N}$. The (i, j) th element of \mathbf{S} , $[\mathbf{S}]_{i,j}$, is nonzero only if $i = j$ or if $(v_j, v_i) \in \mathcal{E}$; so that \mathbf{S} respects the sparsity of \mathcal{G} . Standard choices for \mathbf{S} are the weighted graph adjacency matrix \mathbf{W} [17], [27], the graph Laplacian matrix \mathbf{L} [10], or

their respective generalizations [28, Chapter 8]. We consider \mathbf{S} admits an eigendecomposition $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$, where $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_N] \in \mathbb{C}^{N \times N}$ collects the orthonormal eigenvectors and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N) \in \mathbb{C}^{N \times N}$ contains the associated eigenvalues. This holds for all undirected graphs on which the graph Laplacian can be defined and also for the adjacency matrix of some directed graphs [17], [19].

A graph signal is a mapping from the vertex set to the field of real numbers, i.e., $x_i : v_i \rightarrow \mathbb{R}$, for $v_i \in \mathcal{V}$. An example of a graph signal is the opinion of members in Zachary's Karate club. We collect all node signals in the vector $\mathbf{x} \in \mathbb{R}^N$ with $[\mathbf{x}]_i = x_i$ being the value of node i [10].

The graph Fourier transform (GFT) is the projection of the graph signal \mathbf{x} on the eigenbasis \mathbf{V} and is denoted by $\tilde{\mathbf{x}} = \mathbf{V}^H \mathbf{x}$ [9], [10]. The elements $[\tilde{\mathbf{x}}]_k = \tilde{x}_k$ denote the graph Fourier coefficients of \mathbf{x} , whereas the eigenvectors \mathbf{v}_k form the basis of *graph oscillating modes*. Likewise, the inverse GFT is $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$, i.e., it writes \mathbf{x} as a linear combination of the graph oscillating modes weighed by the Fourier coefficients.

A graph signal is *bandlimited* if it has only a few nonzero Fourier coefficients. Without loss of generality, assume the first K elements of $\tilde{\mathbf{x}}$ are nonzero; so, we can write $\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_K^T, \mathbf{0}_{N-K}^T]^T$ where $\tilde{\mathbf{x}}_K \in \mathbb{C}^K$ and $K \leq N$. Then, \mathbf{x} is written in the compact form

$$\mathbf{x} = \mathbf{V}_K \tilde{\mathbf{x}}_K \quad (1)$$

where $\mathbf{V}_K \in \mathbb{C}^{N \times K}$ is the respective column-trimmed eigenvector matrix. The GFT $\tilde{\mathbf{x}}_K$ of $\tilde{\mathbf{x}}$ writes as $\tilde{\mathbf{x}}_K = \mathbf{V}_K^H \mathbf{x}$. This representation connects the signal bandwidth with the sampling and reconstruction strategies as shown in [12], [14], [15], [29], [30]. We will also exploit bandlimitedness in Section IV to control the network from a few driving nodes.

B. Random time varying graphs

We consider the following random graph model.

Definition 1 (RES(p) graph model [31]). *Given an underlying graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a random edge sampling (RES) graph realization $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$ of \mathcal{G} consists of the same set of nodes \mathcal{V} and assumes the edge $(v_i, v_j) \in \mathcal{E}$ is sampled at time t (i.e., $(v_i, v_j) \in \mathcal{E}_t$) with a probability $0 < p \leq 1$. The edges are sampled independently over both the graph and the temporal dimension and are mutually independent from the graph signal if the latter has a stochastic nature.*

In other words, the RES(p) model states that the realization $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$ is drawn from the underlying graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the instantaneous edge set $\mathcal{E}_t \subseteq \mathcal{E}$ is generated via an independent Bernoulli process with probability p . Let us from now on denote with \mathbf{W} and $\mathbf{D} = \text{diag}(\mathbf{W}\mathbf{1}_N)$ the adjacency and degree matrix of \mathcal{G} , respectively. If the graph is undirected, we will also consider the unnormalized graph Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$. To ease the exposition, denote with \mathbf{W}_t , \mathbf{D}_t , and \mathbf{L}_t the respective matrices of the instantaneous graph \mathcal{G}_t and with $\bar{\mathbf{W}} = \mathbb{E}[\mathbf{W}_t]$, $\bar{\mathbf{D}} = \mathbb{E}[\mathbf{D}_t]$, and $\bar{\mathbf{L}} = \mathbb{E}[\mathbf{L}_t]$ those of the expected graph $\bar{\mathcal{G}}$. Under the RES(p) model, it holds that $\bar{\mathbf{W}} = p\mathbf{W}$, $\bar{\mathbf{D}} = p\mathbf{D}$, and $\bar{\mathbf{L}} = p\mathbf{L}$.

We assume the following.

Assumption 1. *The GSO of the underlying graph \mathcal{G} has an upper bounded spectral norm $\|\mathbf{S}\|_2 \leq \varrho$ for some $\varrho < \infty$.*

This assumption is generally met in practice and implies the graphs of interest have finite dimension and edge weights.

We also remark that more complex models than the RES(p) can be found in literature. In particular, many of these results can be readily extended to the model in which each edge (v_i, v_j) is sampled independently with a different probability p_{ij} . But for clarity of exposition, we will focus only on the RES(p) model.

C. Diffusion on graphs from a GSP perspective

The continuous-time diffusion of a signal \mathbf{x}_0 on a graph \mathcal{G} with Laplacian matrix \mathbf{L} is described by the differential equation [32], [33]

$$\frac{d\mathbf{x}(s)}{ds} = -\mathbf{L}\mathbf{x}(s), \quad \mathbf{x}(0) = \mathbf{x}_0. \quad (2)$$

This equation can be discretized as [34]

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1}, \quad \mathbf{A} = \mathbf{I} - \epsilon\mathbf{L}, \quad t \in \mathbb{N} \quad (3)$$

which is stable if $0 < \epsilon \leq 1/\|\mathbf{L}\|_2$. Alternatively, a diffusion process on a graph can be interpreted as the discrete-time shift of \mathbf{x}_0 through the graph edges [27], [35]

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1}, \quad \mathbf{A} = \mathbf{W}, \quad t \in \mathbb{N}. \quad (4)$$

Model (3) is used when the process is defined over a continuous space that has been discretized, usually in the form of a mesh, as for heat diffusion processes [33]. Model (4) is employed when the underlying support is naturally a graph, as for sensor network communications [35]. In essence, these are two examples of processes that describe the network state evolution by $\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1}$ and relate this state to the underlying time-invariant topology (the transition matrix \mathbf{A} depends on the shift operator). In this paper, we consider the more general case of random time varying topologies, i.e. $\mathbf{x}_t = \mathbf{A}_{t-1}\mathbf{x}_{t-1}$, and we abstract the relationship between the transition matrix and the underlying topology as follows.

Assumption 2. *Let \mathbf{A}_t be the time varying transition matrix of a diffusion process over a random time varying graph $\mathcal{G}_t \in \text{RES}(p)$. Then, $\mathbb{E}[\mathbf{A}_t]$ and \mathbf{S} share the same eigenvectors.*

That is, we consider diffusions on random graphs such that the eigenvectors of the expected transition matrix and the underlying GSO coincide. The following lemma shows this is the case for the diffusion models (3) and (4) on RES(p) graph realizations.

Lemma 1. *Let \mathcal{G} be a graph satisfying Assumption 1 and let \mathcal{G}_t be a RES(p) realization of it. For the diffusion models*

- (i) $\mathbf{S} = \mathbf{L}$ and $\mathbf{A}_t = \mathbf{I} - \epsilon\mathbf{L}_t$ [cf. (3)],
- (ii) $\mathbf{S} = \mathbf{W}$ and $\mathbf{A}_t = \mathbf{W}_t$ [cf. (4)],

Assumption 2 and $\|\mathbf{A}_t\|_2 \leq \varrho$ hold.

Other models that satisfy these conditions are the wave equation on graphs and graph-based ARMA models, see [16].

III. CONTROLLABILITY ON DETERMINISTIC GRAPHS

Consider the N -state linear system

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_{t-1} \quad (5)$$

where $\mathbf{x}_t \in \mathbb{R}^N$ denotes the state value on all nodes at time t , $\mathbf{u}_t \in \mathbb{R}^M$ is the control signal injected on $M \leq N$ nodes, and $\mathbf{A} \in \mathbb{R}^{N \times N}$ and $\mathbf{B} \in \mathbb{R}^{N \times M}$ are the transition and control input matrix, respectively. The relationship between the network state \mathbf{x}_t and the underlying topology is captured in (5) through the transition matrix \mathbf{A} ; it shares the eigenvectors with \mathbf{S} and this is the case for models (3) and (4).

System (5) is controllable if and only if the controllability matrix

$$\Omega = [\mathbf{B}, \mathbf{A}\mathbf{B}, \dots, \mathbf{A}^{T-1}\mathbf{B}] \quad (6)$$

has full rank N [36, Section 6.2.1]. While full rank of Ω guarantees the convergence of \mathbf{x}_t to any target signal \mathbf{x}^* , we focus on controlling the network state towards a bandlimited graph signal $\mathbf{x}^* = \mathbf{V}_K \tilde{\mathbf{x}}_K^*$. Here, $\tilde{\mathbf{x}}_K^* \in \mathbb{C}^K$ determines the desired frequency response over the K frequencies of interest; the *target bandwidth*. We thus define the bandwidth controllability as follows.

Definition 2 (Bandwidth controllability). *An N -state system on a graph is bandwidth controllable from $M \leq N$ nodes if, for any initial state \mathbf{x}_0 and some final time T , there exists a sequence of control signals $\{\mathbf{u}_t \in \mathbb{R}^M, t = 0, 1, \dots, T-1\}$ acting on a fixed set of M nodes that drive the network state to a value \mathbf{x}^* with any frequency content $\tilde{\mathbf{x}}_K^* = \mathbf{V}_K^H \mathbf{x}^*$ over the $K \leq N$ target bandwidth.*

Lead by the promising results of bandlimited graph signal reconstruction from samples on a few nodes [12], [14], [15], [29], [30], we aim to control \mathbf{x}_t through a fixed, time-invariant, set of nodes \mathcal{S} of cardinality $|\mathcal{S}| = M \leq N$. Let then $\mathbf{B} = \mathbf{C}^T$ denote a binary matrix that selects these nodes. More formally, \mathbf{C} belongs to the combinatorial set

$$\mathcal{C}_{M,N} = \{\mathbf{C} \in \{0,1\}^{M \times N} : \mathbf{C}\mathbf{1}_N = \mathbf{1}_M, \mathbf{C}^T\mathbf{1}_M \leq \mathbf{1}_N\} \quad (7)$$

that selects M out of N different nodes and the ordering relation \leq among vectors stands for the elementwise partial ordering [39, Example 2.23]. Observe that $\mathbf{C}\mathbf{C}^T = \mathbf{I}_M$ and $\mathbf{C}^T\mathbf{C} = \text{diag}(\mathbf{c})$ with $\mathbf{c} \in \{0,1\}^N$, such that $[\mathbf{c}]_i = 1$ if and only if node v_i belongs to \mathcal{S} .

With this in place, we write the linear system on graphs (5) in the GFT domain as

$$\begin{aligned} \tilde{\mathbf{x}}_t &= \mathbf{V}^H \mathbf{A} \mathbf{V} \tilde{\mathbf{x}}_{t-1} + \mathbf{V}^H \mathbf{C}^T \mathbf{u}_{t-1} \\ &\triangleq \tilde{\mathbf{A}} \tilde{\mathbf{x}}_{t-1} + \mathbf{V}^H \mathbf{C}^T \mathbf{u}_{t-1} \end{aligned} \quad (8)$$

where $\tilde{\mathbf{A}} = \mathbf{V}^H \mathbf{A} \mathbf{V}$ is a diagonal matrix containing the eigenvalues of \mathbf{A} [cf. Assumption 2]. For convenience, we write $\tilde{\mathbf{A}} = \text{diag}(\mathbf{a}) \in \mathbb{C}^{N \times N}$ with $\mathbf{a} \in \mathbb{C}^N$ the vector containing the eigenvalues of $\tilde{\mathbf{A}}$, known also as the spectral response of \mathbf{A} . Then, by splitting (8) into the K frequencies of interest we have

$$\begin{bmatrix} \tilde{\mathbf{x}}_{t,K} \\ \tilde{\mathbf{x}}_{t,N-K} \end{bmatrix} = \begin{bmatrix} \text{diag}(\mathbf{a}_K) \tilde{\mathbf{x}}_{t-1,K} \\ \text{diag}(\mathbf{a}_{N-K}) \tilde{\mathbf{x}}_{t-1,N-K} \end{bmatrix} + \begin{bmatrix} \mathbf{V}_K^H \mathbf{C}^T \mathbf{u}_{t-1} \\ \mathbf{V}_{N-K}^H \mathbf{C}^T \mathbf{u}_{t-1} \end{bmatrix} \quad (9)$$

where $\tilde{\mathbf{x}}_t = [\tilde{\mathbf{x}}_{t,K}^T, \tilde{\mathbf{x}}_{t,N-K}^T]^T$, $\mathbf{a} = [\mathbf{a}_K^T, \mathbf{a}_{N-K}^T]^T$ and $\mathbf{V} = [\mathbf{V}_K, \mathbf{V}_{N-K}]$.

For a system that is bandwidth controllable [cf. Def. 2], recursion (9) can drive the network state \mathbf{x}_t to any signal \mathbf{x}_T that has a desired frequency response $\tilde{\mathbf{x}}_K^*$ over the target bandwidth K . Thus, we can focus on the K frequencies of interest, determine the driving nodes through matrix \mathbf{C} , and design the control signals $\{\mathbf{u}_t\}$ such that the system

$$\tilde{\mathbf{x}}_{t,K} = \tilde{\mathbf{A}}_K \tilde{\mathbf{x}}_{t-1,K} + \mathbf{V}_K^H \mathbf{C}^T \mathbf{u}_{t-1}. \quad (10)$$

reaches $\tilde{\mathbf{x}}_K^*$ at time T , i.e. $\tilde{\mathbf{x}}_{T,K} = \tilde{\mathbf{x}}_K^*$. But we observe that focusing on these K frequencies leads to a non-zero value also on the $N-K$ remaining frequencies. One way to suppress this undesired content is to force $\tilde{\mathbf{x}}_{t,N-K} = \mathbf{0}_{N-K}$. This requires the design of a sampling matrix \mathbf{C} that satisfies $\mathbf{V}_{N-K}^H \mathbf{C}^T = \mathbf{0}_{(N-K) \times M}$ [cf. (9)]. The latter might be infeasible or might severely constraint the selection of driving nodes. Therefore, to avoid the out-of-band frequency content, we use a frequency-selective graph filter $\mathbf{H} = \mathbf{V}_K \mathbf{V}_K^H$ to force bandlimitedness on the final network state $\mathbf{x}^* = \mathbf{H} \mathbf{x}_T$. This filter can be implemented locally through a polynomial in the shift operator \mathbf{S} with degree at most N [37].

Hereinafter, the design variables are the sampling matrix \mathbf{C} and the control signals $\mathbf{u}_t \in \mathbb{R}^M$ for $t = 0, \dots, T-1$. And, since the initial state \mathbf{x}_0 is considered known we fix, without loss of generality, $\mathbf{x}_0 = \mathbf{0}_N$ as the common practice in control literature [22], [36, Section 2.3.2], [38, Section 2.1]. With this set down, we claim our first contribution, which we will also exploit for controllability on random graphs in Section IV.

Proposition 1. *Consider the linear system (10) describing a process over a deterministic graph \mathcal{G} . A necessary condition to control the system in a finite time T towards a target frequency content $\tilde{\mathbf{x}}_K^*$ over the K frequencies of interest is to select $M \geq \lceil K/T \rceil$ driving nodes.*

Proposition 1 provides a necessary condition on the minimum number of nodes to control system (10) in T instants. It shows the trade-off between the cardinality of the sampling set M , the signal bandwidth K , and the control time T . Thus, for $T \geq K$ there is the potential to control the network by acting on a single node. But this is not sufficient since the system controllability is affected by the network topology [2]–[4], i.e. the influence of the driving nodes on the frequency content, see also [12], [19], [21]. Hence, the driving nodes should be carefully picked to guarantee the controllability of (10). In what follows, we show the relation of Proposition 1 with [21] and [22].

a) *Relation with [21].* Proposition 1 encompasses under a single condition the three graph signal reconstruction strategies of [21]. In fact, $M \geq K$ and $T = 1$ covers the *multiple node-single time* seeding strategy; $M = 1$ and $T \geq K$ is a necessary condition to control the signal for the *single node-multiple time* seeding strategy; and $M \geq \lceil K/T \rceil$ covers the more involved *multiple node-multiple time* seeding approach. This is expected, since graph signal reconstruction through percolation is a particular case of system (5).

b) *Relation with [22]*. Differently from our approach, [22] focuses on designing the control signal $\mathbf{u} = [\mathbf{u}_{T-1}^\top, \dots, \mathbf{u}_1^\top, \mathbf{u}_0^\top]^\top \in \mathbb{R}^{MT \times 1}$ as a trade-off between sparsity in the vertex domain and signal energy. This problem writes as

$$\begin{aligned} & \underset{\mathbf{u}}{\text{minimize}} \quad \|\mathbf{u}\|_2^2 + \gamma \|\mathbf{u}\|_0 \\ & \text{subject to} \quad \mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t, \quad t = 0, \dots, T-1 \quad (11) \\ & \quad \mathbf{x}_0 = \mathbf{0}_N, \quad \mathbf{x}_T = \mathbf{x}^* \end{aligned}$$

where the constant γ trades the control signal energy $\|\mathbf{u}\|_2^2$ with sparsity $\|\mathbf{u}\|_0$. Problem (11) yields a sparse control signal \mathbf{u} across time, but the driving nodes are not necessarily fixed. In this regard, Proposition 1 imposes a minimum dimension on \mathbf{u} such that controllability is possible from a fixed set \mathcal{S} of driving nodes.

IV. CONTROLLABILITY ON RANDOM GRAPHS

When the network topology is time varying, system (5) should change to reflect the time dependency in the transition matrix. When the time variation is random, the controllability of the network state should follow a statistical approach. We propose a statistical framework in this section, where in Section IV-A we develop the concept of controllability in the mean and in Section IV-B we perform the mean squared error analysis.

A. Mean controllability

The dynamics of a time varying system on random graphs are given by

$$\mathbf{x}_t = \mathbf{A}_{t-1}\mathbf{x}_{t-1} + \mathbf{C}^\top \mathbf{u}_{t-1} \quad (12)$$

where under the $\text{RES}(p)$ model in Definition 1, $\{\mathbf{A}_t\}$ is a set of i.i.d. random matrices with $\mathbb{E}[\mathbf{A}_t] = \bar{\mathbf{A}}$. The deterministic design variables are contained in the second term of (12), $\mathbf{C}^\top \mathbf{u}_{t-1}$. The state \mathbf{x}_t depends on the random system matrices $\{\mathbf{A}_\tau\}_{\tau=0}^{t-1}$, which are independent from \mathbf{A}_t and from the deterministic design variables \mathbf{C} and $\{\mathbf{u}_\tau\}_{\tau=0}^{t-1}$. Note also that $\bar{\mathbf{A}}$ and \mathbf{S} share the same eigenvectors, i.e. $\bar{\mathbf{A}} = \mathbf{V}\text{diag}(\bar{\mathbf{a}})\mathbf{V}^\mathbf{H}$; thus, it captures in statistics the relation between the underlying topology and state \mathbf{x}_t . We can then write the mean evolution of (12) as

$$\boldsymbol{\mu}_t = \bar{\mathbf{A}}\boldsymbol{\mu}_{t-1} + \mathbf{C}^\top \mathbf{u}_{t-1} \quad (13)$$

where $\boldsymbol{\mu}_t = \mathbb{E}[\mathbf{x}_t]$. System (13) is a deterministic and time-invariant system analogous to (5). We develop the following controllability concept.

Definition 3 (Bandwidth controllability in the mean). *An N -state system on a random graph of the form in (12) with mean evolution in (13) is bandwidth controllable in the mean from $M \leq N$ nodes if, for any initial state \mathbf{x}_0 and some final time T , there exists a sequence of control signals $\{\mathbf{u}_t, t = 0, \dots, T-1\}$ acting on a fixed set of M nodes that drive the mean network state to a value \mathbf{x}^* with any frequency content $\mathbf{x}_K^* = \mathbf{V}_K^\mathbf{H} \mathbf{x}^*$ over the $K \leq N$ target bandwidth.*

Our goal is to control the mean system to a desired bandlimited graph signal $\mathbf{x}^* = \mathbf{V}_K \tilde{\mathbf{x}}_K^*$ in a finite time T

from a few nodes. We do so by designing the input signals $\{\mathbf{u}_t, t = 0, \dots, T-1\}$ and the node driving set \mathcal{S} (through matrix \mathbf{C}). In analogy to Section III, we focus on the K frequencies of interest of the mean system [cf. (10)]

$$\tilde{\boldsymbol{\mu}}_{t,K} = \tilde{\mathbf{A}}_K \tilde{\boldsymbol{\mu}}_{t-1,K} + \mathbf{V}_K^\mathbf{H} \mathbf{C}^\top \mathbf{u}_{t-1} \quad (14)$$

and drive it to the desired frequency content $\tilde{\mathbf{x}}_K^*$ [cf. Def. 3], with $\tilde{\mathbf{A}}_K = \tilde{\mathbf{A}}_K = \text{diag}(\bar{\mathbf{a}}_K) \in \mathbb{C}^{K \times K}$ containing the eigenvalues of $\bar{\mathbf{A}}$ which determine the spectral response of the system evolution on the expected graph. Then, we apply a (deterministic) linear filter $\mathbf{H} = \mathbf{V}_K \mathbf{V}_K^\mathbf{H}$ to keep only those K desired frequencies such that the mean network state $\boldsymbol{\mu}^* = \mathbb{E}[\mathbf{H}\mathbf{x}_T] = \mathbf{H}\boldsymbol{\mu}_T = \mathbf{x}^*$ results in a bandlimited graph signal.

Similarly to Proposition 1, we claim the following.

Proposition 2. *Consider the linear system (12) describing a process over a sequence of $\text{RES}(p)$ graphs \mathcal{G}_t with in-band mean evolution (14). A necessary condition to control the mean system in finite time T towards a target frequency content $\tilde{\mathbf{x}}_K^*$ over the K frequencies of interest is to select $M \geq \lceil K/T \rceil$ driving nodes.*

Like Proposition 1, Proposition 2 establishes a necessary condition to control a linear system, now, on random time varying graphs. As such, the same trade-off between the number of driving nodes M , the signal bandwidth K , and the control time T applies here. The next corollary extends this result to a sufficient condition under some restrictions on the eigenvector basis.

Corollary 1. *Under the hypothesis of Proposition 2 if there exists a set of driving nodes \mathcal{S} built by M nodes such that the corresponding M rows of \mathbf{V}_K are linearly independent vectors, then $M \geq K$ is a sufficient condition to control system (12) in the mean.*

An algorithm for finding such M nodes is readily available in [12, Algorithm 1].

B. Mean squared error analysis

In Section IV-A, we discussed that system (12) can be controlled in the mean. Therefore, it is paramount also to quantify the mean squared error (MSE) of the controlled state to gain statistical insight into how close the filtered final state on a specific graph realization $\mathbf{H}\mathbf{x}_T$ is to the actual desired signal \mathbf{x}^* . Towards this end, define $\Phi_{b,a} = \mathbf{A}_b \mathbf{A}_{b-1} \cdots \mathbf{A}_{a+1} \mathbf{A}_a$ as the state transition matrix between time instants $b \geq a$. The following theorem determines the MSE.

Theorem 1. *Let Assumptions 1 and 2 hold and let \mathbf{x}_t be a graph process defined over a sequence of $\text{RES}(p)$ graphs \mathcal{G}_t described by linear system (12). Given also a set of driving nodes \mathcal{S} characterized by selection matrix \mathbf{C} and a set of control signals $\{\mathbf{u}_t\}_{t=0}^{T-1}$ with initial state $\mathbf{x}_0 = \mathbf{0}_N$. The*

MSE between the filtered signal $\mathbf{H}\mathbf{x}_T$ on a particular graph realization and the actual desired signal is

$$\begin{aligned} \text{MSE}(T) &= \mathbb{E} \left[\|\mathbf{H}\mathbf{x}_T - \mathbf{x}^*\|_2^2 \right] \\ &= \alpha - 2 \sum_{\tau=0}^{T-1} \beta_\tau^\top \mathbf{C}^\top \mathbf{u}_\tau + \sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \mathbf{u}_\tau^\top \mathbf{C} \Gamma_{\tau,\tau'} \mathbf{C}^\top \mathbf{u}_{\tau'} \end{aligned} \quad (15)$$

which is a quadratic form on $\mathbf{C}^\top \mathbf{u}_\tau$ with coefficients $\alpha = \|\mathbf{x}^*\|_2^2 \in \mathbb{R}$, $\beta_\tau = (\tilde{\mathbf{A}}^{T-\tau-1})^\top \mathbf{H}^\top \mathbf{x}^* \in \mathbb{R}^{N \times 1}$ and $\Gamma_{\tau,\tau'} = \mathbb{E}[\Phi_{T-1,\tau+1}^\top \mathbf{H}^\top \mathbf{H} \Phi_{T-1,\tau'+1}] \in \mathbb{R}^{N \times N}$.

Given \mathbf{C} and $\{\mathbf{u}_t\}$, the MSE in (15) holds for any system described by (12), irrespective of their controllability. The MSE in (15) is a quadratic function in the design variables $\mathbf{C}^\top \mathbf{u}_\tau$ and the corresponding coefficients α , β_τ , and $\Gamma_{\tau,\tau'}$ depend on known quantities: the graph filter \mathbf{H} ; the desired state $\mathbf{x}^* = \mathbf{V}_K \tilde{\mathbf{x}}_K^*$; and the statistics (first and second order moments) of the underlying support through $\tilde{\mathbf{A}}$ and $\Gamma_{\tau,\tau'}$. This result highlights also the impact the driving nodes have on the overall performance and shows their connection with the underlying support and the target bandwidth. More precisely, the coefficient α provides the MSE floor if there is no control signal (i.e., given by the energy of the target state); β_τ takes into account the similarity between the target signal and the controlled signal evolution over the mean graph; and $\Gamma_{\tau,\tau'}$ accounts for the variability of the random graph. Finally, we note that the computation of $\Gamma_{\tau,\tau'}$ in (15) might be cumbersome for some transition matrices \mathbf{A}_t . We thus provide in the appendix two practical results that address this issue: first, we provide a general upper bound; second, we show how to exactly compute $\Gamma_{\tau,\tau'}$ for undirected graphs for the diffusion models in Lemma 1.

V. CONTROL STRATEGIES

In this section, we propose two control strategies (i.e., find \mathbf{C} and $\{\mathbf{u}_t\}$) for graph processes over random time varying graphs, where, depending on the scenario, one can be preferred over the other. In Section V-A we propose an unbiased control strategy, while in Section V-B we introduce a control strategy that leverages the bias-variance trade-off to minimize the MSE.

A. Unbiased controller

The mean state in (14) for $t = T$ can be expanded as

$$\tilde{\boldsymbol{\mu}}_{T,K} = \sum_{\tau=0}^{T-1} \tilde{\mathbf{A}}_K^{T-\tau-1} \mathbf{V}_K^\top \mathbf{C}^\top \mathbf{u}_\tau \quad (16)$$

where $\tilde{\mathbf{A}}_K = \tilde{\tilde{\mathbf{A}}}_K = \text{diag}(\tilde{\mathbf{a}}_K) \in \mathbb{C}^{K \times K}$ and $\tilde{\boldsymbol{\mu}}_{0,K} = \mathbf{0}_K$. For an unbiased controller, it must hold at final time T $\tilde{\boldsymbol{\mu}}_{T,K} = \tilde{\mathbf{x}}_K^*$. Combining (16) and $\tilde{\boldsymbol{\mu}}_{T,K} = \tilde{\mathbf{x}}_K^*$, we obtain

$$\tilde{\boldsymbol{\Omega}} \mathbf{u} = [\mathbf{I}_K, \tilde{\mathbf{A}}_K, \dots, \tilde{\mathbf{A}}_K^{T-1}] (\mathbf{I}_T \otimes \mathbf{V}_K^\top \mathbf{C}^\top) \mathbf{u} = \tilde{\mathbf{x}}_K^* \quad (17)$$

with in-band controllability matrix $\tilde{\boldsymbol{\Omega}} \in \mathbb{C}^{K \times TM}$ and input vector is $\mathbf{u} = [\mathbf{u}_{T-1}^\top, \mathbf{u}_{T-2}^\top, \dots, \mathbf{u}_1^\top, \mathbf{u}_0^\top]^\top \in \mathbb{R}^{MT \times 1}$. For $\tilde{\boldsymbol{\Omega}}$ being of full rank K (i.e. a controllable system), system (17) has infinite solutions on \mathbf{u} . Also often exists more than one set of nodes that guarantees controllability. We then select the

Algorithm 1 Constrained Greedy Approach.

Input: M : number of samples, T : time horizon
 $\tilde{\mathbf{x}}_K^*$: desired frequency response
 \mathbf{V}_K : frequency basis vectors of active frequencies
 $\tilde{\mathbf{A}}_K$: GFT of transition matrix
 $\text{MSE}(\cdot)$: function to compute MSE \triangleright See (15)
Output: \mathbf{C} : selected nodes, $\{\mathbf{u}_t\}$: control signals

```

1: procedure GREEDY( $M, T, \tilde{\mathbf{x}}_K^*, \mathbf{V}_K, \tilde{\mathbf{A}}_K, \text{MSE}(\cdot)$ )
2:   Set  $\mathcal{S} = \emptyset$   $\triangleright$  Selected nodes
3:   Set  $\mathcal{R} = \mathcal{V}$   $\triangleright$  Remaining nodes
4:   Set  $\text{bestMSE} \leftarrow \infty$ 
5:   for  $m = 1 : M$  do
6:     Set  $\text{bestNode} \leftarrow \emptyset$ 
7:     for  $n = 1 : N - m + 1$  do
8:       Select  $r_n \in \mathcal{R}$   $\triangleright$  Choose a remaining node
9:       Compute matrix  $\mathbf{C}$  for  $\mathcal{S} \cup \{r_n\}$ 
10:      Compute matrix  $\tilde{\boldsymbol{\Omega}}$   $\triangleright$  See (17)
11:      if  $\text{rank}(\tilde{\boldsymbol{\Omega}}) = \min\{K, T \min\{K, m\}\}$  then
12:        Solve  $\tilde{\boldsymbol{\Omega}} \mathbf{u} = \tilde{\mathbf{x}}_K^*$  for  $\mathbf{u}$ 
13:        Compute  $\text{MSE}(T)$   $\triangleright$  See (15)
14:        if  $\text{MSE}(T) < \text{bestMSE}$  then
15:          Set  $\text{bestNode} \leftarrow r_n$ 
16:          Set  $\text{bestMSE} \leftarrow \text{MSE}(T)$ 
17:        end if
18:      end if
19:    end for
20:    Set  $\mathcal{S} \leftarrow \mathcal{S} \cup \{\text{bestNode}\}$ 
21:    Set  $\mathcal{R} \leftarrow \mathcal{R} \setminus \{\text{bestNode}\}$ 
22:  end for
23: end procedure

```

set of nodes and design the control signals to minimize the $\text{MSE}(T)$, while guaranteeing the solution is unbiased.

Let $\mathcal{C}_{M,N}^* = \{\mathbf{C} \in \mathcal{C}_{M,N} : \text{rank}(\tilde{\boldsymbol{\Omega}}) = K\}$ be the set of selection matrices that satisfy controllability. The optimal unbiased control strategy can be posed as

$$\min_{\mathbf{C} \in \mathcal{C}_{M,N}^*, \mathbf{u} \in \mathbb{R}^{TM}} \text{MSE}(T) \quad (18)$$

$$\text{s. t. } \tilde{\boldsymbol{\Omega}} \mathbf{u} = \tilde{\mathbf{x}}_K^*$$

$$\tilde{\boldsymbol{\Omega}} = [\mathbf{I}_K, \tilde{\mathbf{A}}_K, \dots, \tilde{\mathbf{A}}_K^{T-1}] (\mathbf{I}_T \otimes \mathbf{V}_K^\top \mathbf{C}^\top)$$

where $\text{MSE}(T)$ is given in (15). Oftentimes, we are interested in controlling the system with minimum energy [7], [22]. In such cases, the minimum energy control signal is [39, Section 6.2]

$$\mathbf{u}^* = \tilde{\boldsymbol{\Omega}}^\text{H} [\tilde{\boldsymbol{\Omega}} \tilde{\boldsymbol{\Omega}}^\text{H}]^{-1} \tilde{\mathbf{x}}_K^*. \quad (19)$$

Then, within the minimum energy framework, we select the nodes that minimize the $\text{MSE}(T)$ as follows

$$\min_{\mathbf{C} \in \mathcal{C}_{M,N}^*} \text{MSE}(T) \quad (20)$$

$$\text{s. t. } \mathbf{u}^* = \tilde{\boldsymbol{\Omega}}^\text{H} [\tilde{\boldsymbol{\Omega}} \tilde{\boldsymbol{\Omega}}^\text{H}]^{-1} \tilde{\mathbf{x}}_K^*,$$

$$\tilde{\boldsymbol{\Omega}} = [\mathbf{I}_K, \tilde{\mathbf{A}}_K, \dots, \tilde{\mathbf{A}}_K^{T-1}] (\mathbf{I}_T \otimes \mathbf{V}_K^\top \mathbf{C}^\top).$$

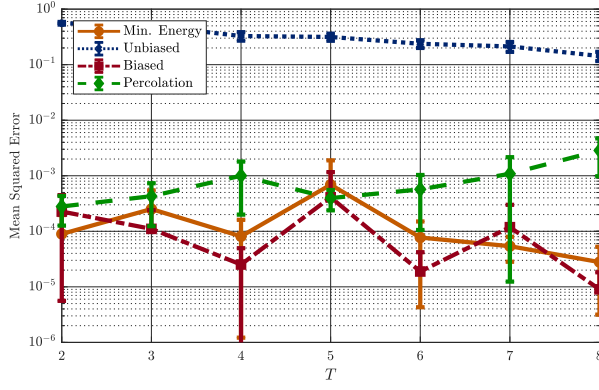
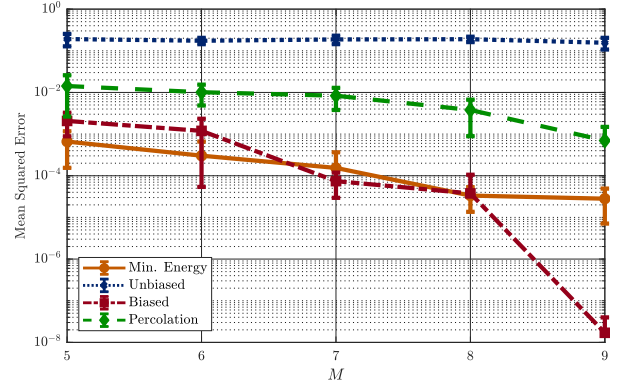

 (a) Sweeping time horizon T

 (b) Sweeping number of selected nodes M

Figure 2. Baseline for time-invariant networks on geometric graphs. (a) Parameter sweeping simulation as a function of time horizon T . (b) Parameter sweeping simulation as a function of number of selected samples M . The biased controller has a performance similar to the Min. Energy and performs slightly better than Percolation. The unbiased controller lags behind in terms of MSE. The error bars indicate the estimated standard deviation from the 500 graph realizations.

Problem (20) is non-convex due to the binary nature of the optimization variable \mathbf{C} . A heuristic solution is to follow a constrained greedy approach as described in Algorithm 1. The objective is to greedily select the nodes that improve the MSE while satisfying controllability. Specifically, for each candidate driving node, we check $\text{rank}(\tilde{\Omega})$ increases until we reach controllability as indicated in line 11 [cf. Proposition 2]. Since we are looking for the minimum energy controller, then line 12 entails computing $\mathbf{u}^* = (\tilde{\Omega}^H \tilde{\Omega})^{-1} \tilde{\Omega}^H \tilde{\mathbf{x}}_K^*$ if $\text{rank}(\tilde{\Omega}) = Tm$, and (19) if $\text{rank}(\tilde{\Omega}) = K$. While this constrained greedy approach has no theoretical guarantees [40], our numerical results in Section VI show that Algorithm 1 exhibits a performance close to the optimal solution.

B. Biased controller

When the requirement for an unbiased controller is not strict, we can leverage the bias-variance trade-off to further reduce the mean squared error of the controlled state. Given a fixed sampling set \mathbf{C} , the $\text{MSE}(T)$ (15) is a quadratic function on the control signals $\{\mathbf{u}_t, t = 0, 1, \dots, T-1\}$. Therefore, we express \mathbf{u}_t as a function of \mathbf{C} as follows.

First, the derivative of $\text{MSE}(T)$ in (15) w.r.t. \mathbf{u}_t is

$$\frac{\partial \text{MSE}(T)}{\partial \mathbf{u}_t} = -2\mathbf{C}\beta_t + 2 \sum_{\tau=0}^{T-1} \mathbf{C}\Gamma_{t,\tau} \mathbf{C}^T \mathbf{u}_\tau \quad (21)$$

for $t = 0, \dots, T-1$ with β_t and $\Gamma_{t,\tau}$ given in (15). By defining $\beta = [\beta_{T-1}^T, \dots, \beta_0^T]^T \in \mathbb{R}^{NT}$,

$$\Gamma = \begin{bmatrix} \Gamma_{T-1,T-1} & \Gamma_{T-1,T-2} & \cdots & \Gamma_{T-1,0} \\ \Gamma_{T-2,T-1} & \Gamma_{T-2,T-2} & \cdots & \Gamma_{T-2,0} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{0,T-1} & \Gamma_{0,T-2} & \cdots & \Gamma_{0,0} \end{bmatrix} \in \mathbb{R}^{NT \times NT} \quad (22)$$

and by setting (21) to zero, we can write

$$\Gamma_C \mathbf{u} = (\mathbf{I}_T \otimes \mathbf{C}) \Gamma (\mathbf{I}_T \otimes \mathbf{C}^T) \mathbf{u} = (\mathbf{I}_T \otimes \mathbf{C}) \beta = \beta_C \quad (23)$$

with $\Gamma_C \in \mathbb{R}^{MT \times MT}$ and $\beta_C \in \mathbb{R}^{MT \times 1}$. By construction Γ_C has rank MT since $\mathbf{C} \in \mathcal{C}_{M,N}$. Then, for a sampling matrix

\mathbf{C} such that $\text{rank}(\Gamma_C) \geq MT$, Γ_C is nonsingular and leads to the (parameterized) minimum $\text{MSE}(T)$ control signals

$$\mathbf{u}_C^* = \Gamma_C^{-1} \beta_C. \quad (24)$$

From this relation between the control signal and the sampling matrix, we consider a two-stage optimization approach [39, Section 4.1.3] to find \mathbf{C} that minimizes the $\text{MSE}(T)$. This optimization problem writes as

$$\begin{aligned} \min_{\mathbf{C} \in \mathcal{C}_{M,N}^*} \quad & \alpha - 2\beta_C^T \mathbf{u}_C^* + (\mathbf{u}_C^*)^T \Gamma_C \mathbf{u}_C^* \\ \text{s. t.} \quad & \mathbf{u}_C^* = \Gamma_C^{-1} \beta_C \\ & \Gamma_C = (\mathbf{I}_T \otimes \mathbf{C}) \Gamma (\mathbf{I}_T \otimes \mathbf{C}^T) \\ & \beta_C = (\mathbf{I}_T \otimes \mathbf{C}) \beta. \end{aligned} \quad (25)$$

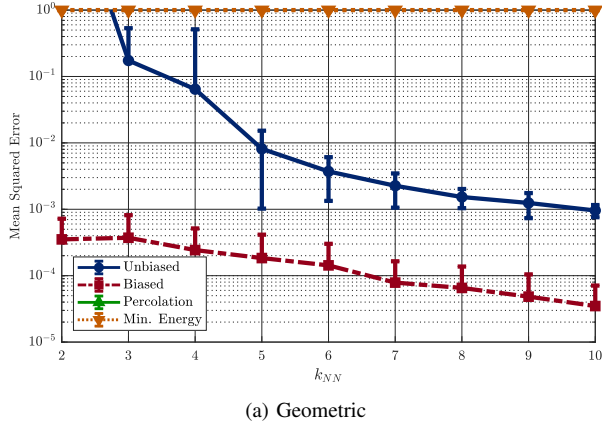
To deal with the non-convexity of (25), similarly to (20), we rely on a constrained greedy approach analogous to Algorithm 1. Specifically, we replace line 10 by the computation of Γ_C and β_C , line 11 by $\text{rank}(\Gamma_C) \geq mT$, and line 12 by (24).

VI. NUMERICAL EXPERIMENTS

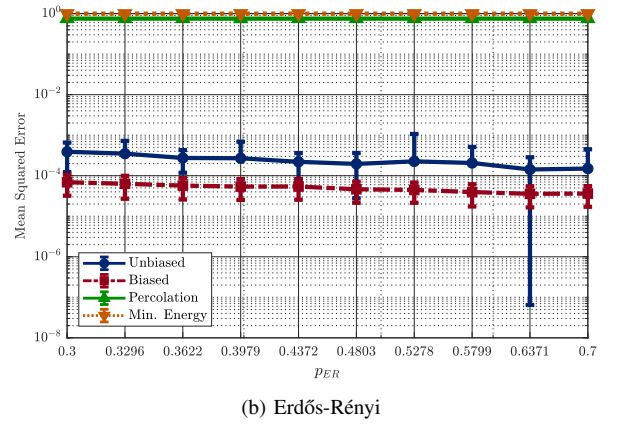
We evaluate the proposed control strategies on different scenarios to analyze the different trade-off when controlling the network. We compare the *unbiased* minimal energy controller (20) and the *biased* controller (25) with the *Percolation* control strategy of [21] and with the *Min. Energy* approach of [22]. Next, we consider synthetic network models, namely Erdős-Rényi (ER) graphs [41] and geometric graphs, while in Section VI-B we test the methods on real-world social networks, namely on the Zachary's Karate Club [8] and on a Facebook subnet [42].

A. Synthetic network models

The ER graph forms the edges between any two nodes randomly and independently with probability p_{ER} and has an average degree of $p_{\text{ER}}N$. The geometric graph draws nodes uniformly at random in the $[0, 1]^2$ plane and computes the

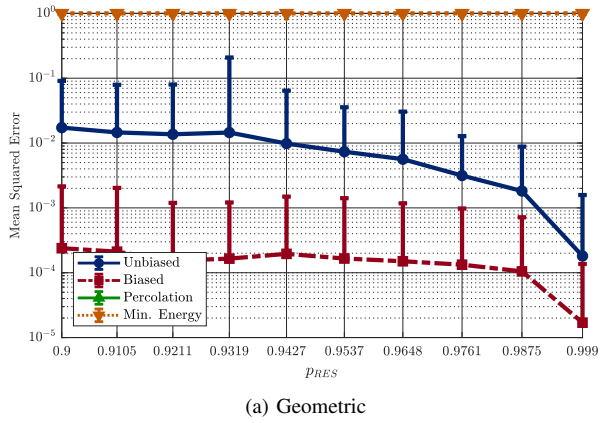


(a) Geometric

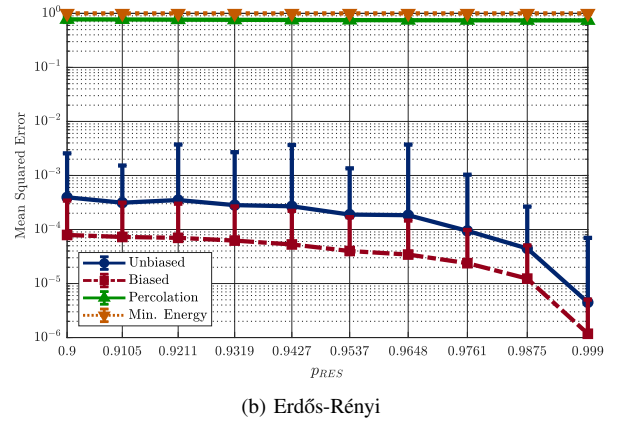


(b) Erdős-Rényi

Figure 3. Impact of graph connectivity measured by the average graph degree. (a) Geometric graph: note that when the average degree increases, the connectivity is higher, and as such there are more communication paths through which the signal can flow, and thus is less affected by link losses. (b) ER graph: the increase in connectivity does not lead to noticeable changes in the MSE since these graphs already have a large average degree (around 30). The error bars indicate $3\times$ the estimated standard deviation from the 500 graph realizations. We note that the y -axis limit has been set to $\text{MSE} = 1$ and that the Percolation method yields $\text{MSE} > 1$ and therefore is not shown in (a).



(a) Geometric



(b) Erdős-Rényi

Figure 4. Impact of the link loss. (a) Geometric graph. (b) ER graph. When p_{RES} increases, fewer links are lost and thus the network is easier to control, leading to a lower MSE. The error bars indicate $3\times$ the estimated standard deviation from the 500 graph realizations. We note that the y -axis limit has been set to $\text{MSE} = 1$ and that the Percolation method yields $\text{MSE} > 1$ and therefore is not shown in (a).

Euclidean distance d_{ij} between any pair of nodes. We assigned a Gaussian kernel edge weights $w_{ij} = \mathcal{W}(v_i, v_j) = e^{-d_{ij}^2}$ and kept only the k_{NN} nearest neighbors per node; the parameter k_{NN} controls the average degree. For both models, we considered realizations that result in connected graphs. To account for the randomness in the generative models and in the edge loss, we averaged the performance over 500 different underlying graphs where for each of them we accounted also for 5000 RES realizations.

Unless otherwise specified, we set $N = 100$, $p_{\text{ER}} = 0.5$, $k_{\text{NN}} = 5$, and the RES link loss probability to $p_{\text{RES}} = 0.95$. The control time is $T = 8$, the number of driving nodes is $M = 8$, and the initial state is $\mathbf{x}_0 = \mathbf{0}_N$. The target state \mathbf{x}^* has a bandwidth of $K = 10$ with GFT coefficients $\tilde{\mathbf{x}}_K^*$ decaying linearly as $[\tilde{\mathbf{x}}_K^*]_k = 1 - (k-1)/K$ for $k = 1, \dots, K$. We measured the controllability performance between the bandlimited controlled state and target one through the normalized MSE: $\text{MSE}(T) = \mathbb{E}[\|\mathbf{H}\mathbf{x}_T - \mathbf{x}^*\|^2 / \|\mathbf{x}^*\|^2]$.

Time-invariant network. To set a baseline, we first compared the biased (25) and the unbiased (20) controllers with

the Percolation [21] and Min. Energy [22] strategies on a fixed time-invariant network. This is the same control scenario that is considered in [21], [22] and is equivalent to setting $p_{\text{RES}} = 1$. We report the results for geometric graphs in Figure 2. Figure 2a is a parametric simulation as a function of time horizon T and Figure 2b is a parametric simulation as a function of the number of samples. In general, we observe that the biased estimator has a performance similar to the Min. Energy and slightly better than Percolation. The unbiased controller lags behind in terms of MSE.

Graph connectivity. In the first random time varying experiment, we studied the impact of the graph connectivity on the controllability performance. We accounted for the graph connectivity by changing the average degrees, i.e., $p_{\text{ER}}N$ for the ER graph and k_{NN} in the geometric graph. Figure 3 shows the MSE as the connectivity increases. From Figure 3a, we observe that the control on geometric networks improves with the average degree. This is intuitively satisfying since larger degrees lead to a higher connectivity between nodes; hence, it renders them more robust to the RES model. Contrarily, for the ER model in Figure 3b this behavior is not as much

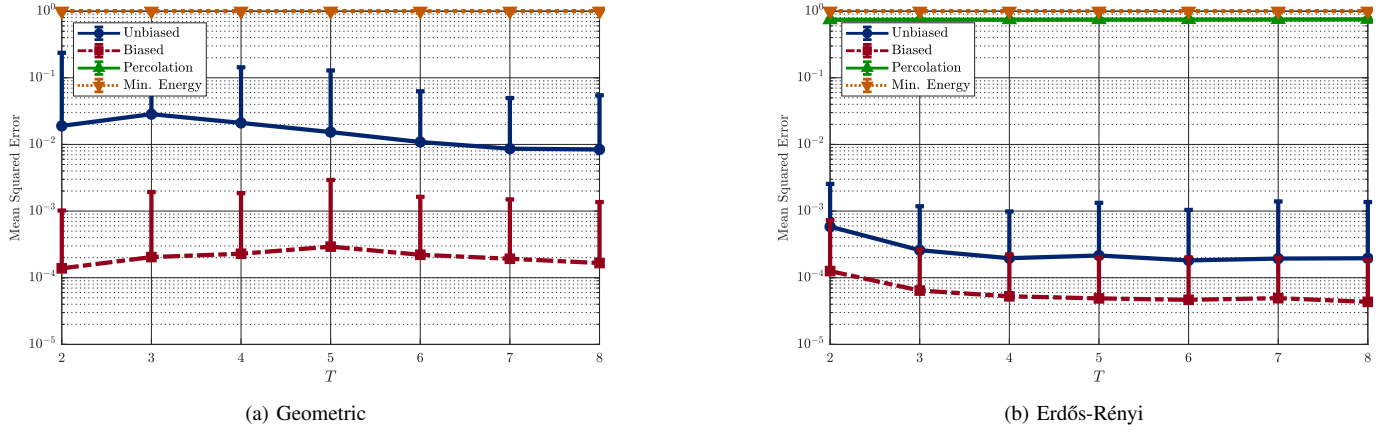


Figure 5. Impact of the control time. (a) Geometric graph. (b) ER graph. For the proposed strategies, an increase in control time yields a slightly lower MSE since there is more time to correct for the network evolution. The error bars indicate $3 \times$ the estimated standard deviation from the 500 graph realizations. We note that the y -axis limit has been set to $\text{MSE} = 1$ and that the Percolation method yields $\text{MSE} > 1$ and therefore is not shown in (a).

emphasized. We attribute this phenomenon to the large average degree of the ER graphs (above 30) and to the relatively high value of p_{RES} . That is, the loss of a few edges does not impact the overall ability to control the network.

Link loss. In the second experiment, we analyzed the impact of p_{RES} for a fixed average degree. From Figure 4, we note that as p_{RES} increases (fewer links are lost) the MSE reduces and leads to an easier to control network. This is because a higher p_{RES} yields realizations with fewer edge losses, thus more similar to the underlying (mean) graph.

Control time. In the third and last experiment, we analyzed the impact of the control time horizon T . From Figure 5, we observe that the proposed strategies are not significantly affected by changes in T as they only improve slightly.

From this set of experiments, we make three key observations. First, the proposed strategies offer the best performance. Second, the biased controller achieves the lowest MSE. This is expected since it leverages the bias-variance trade-off to minimize the overall MSE at expenses of a bias in the controlled state. Third, not accounting for the graph randomness affects seriously the performance, even for $p_{\text{RES}} = 0.999$. In fact, the deterministic alternatives of Percolation and Min. Energy have a worse performance by orders of magnitude compared with the proposed techniques. This contrast is particularly evident when comparing with the simulations for a time-invariant network in Figure 2. This could be explained by the fact that losing a link has a huge impact in the topology of the graph and severely affects the eigenbasis, thereby, changing the subspace of signals that are bandlimited on a given graph.

B. Real world graphs

We consider the formation of opinion profiles on two social networks, namely the Zachary's Karate Club [8] in Figure 1 of $N = 34$ nodes and a Facebook subnetwork [42] of $N = 234$ nodes.

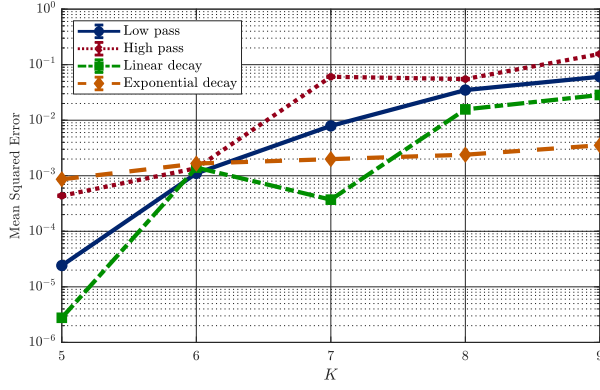
We set the control time to $T = 8$, the number of selected nodes to $M = \text{round}(0.08N)$, $p_{\text{RES}} = 0.95$, and $\mathbf{x}_0 = \mathbf{0}_N$. For simplicity of presentation, we focus only on the biased

controller strategy which has consistently yielded the best performance. Likewise, we do not compare it with the methods in [21], [22] given their poor performance on random time varying graphs. We again averaged the performance over 5000 different RES realizations.

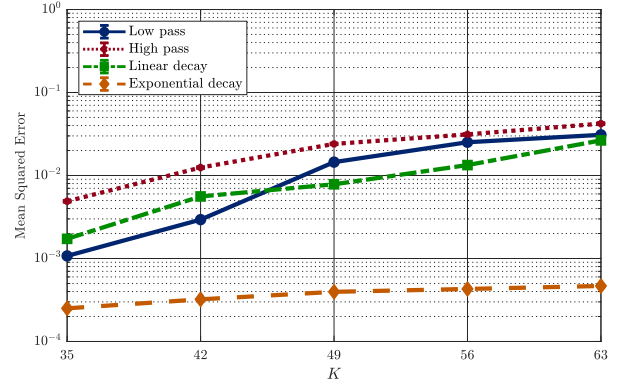
Bandwidth and spectrum of the desired state. In this experiment, we analyzed the impact that the desired state bandwidth and its GFT have on the controllability performance. We considered four different GFTs for the desired state, namely: (i) a step low-pass $[\tilde{\mathbf{x}}_K^*]_k = 1$ for $k = 1, \dots, K$; (ii) a step high-pass, where the active frequencies correspond to the K eigenvectors with highest total variation; (iii) a linear decay response given by $[\tilde{\mathbf{x}}_K^*]_k = 1 - (k - 1)/K$ for $k = 1, \dots, K$; and (iv) an exponential decay response with $[\tilde{\mathbf{x}}_K^*]_k = e^{1-k}$ for $k = 1, \dots, K$. For a fair comparison, we normalized all desired states to unit energy and analyzed different values of K ; a fraction of N between 0.15 and 0.27.

The results are depicted in Figure 6. First, we observe that controlling the system to a higher bandwidth state is harder since the set of graph frequencies to guarantee controllability increases. Second, we observe that the high-pass response is harder to achieve and responses that decay to zero (like the linear decay and the exponential decay) yield lower MSE. This is because high-pass responses are translated in the vertex domain as states having dissimilar values in adjacent nodes. They render the control of network dynamics to such states more challenging. Hence, we conclude that it is easier to drive the network state to a signal that varies smoothly on the nodes compared with a signal that has highly different values in connecting vertices; e.g., it is easier to convince someone to vote a conservative candidate, if she is surrounded by members that have the same political inclination.

Sampling heuristics. In the last experiment, we focused on the impact of the control nodes. We compared the proposed constrained greedy selecting heuristic in Algorithm 1 with the optimal combinatorial solution and a uniformly random sampling scheme. We fixed $K = 10$ and considered the linear decay desired state $\tilde{\mathbf{x}}_K^*$ such that $[\tilde{\mathbf{x}}_K^*]_k = 1 - (k - 1)/K$ for $k = 1, \dots, K$. The obtained results are shown in Figure 7.

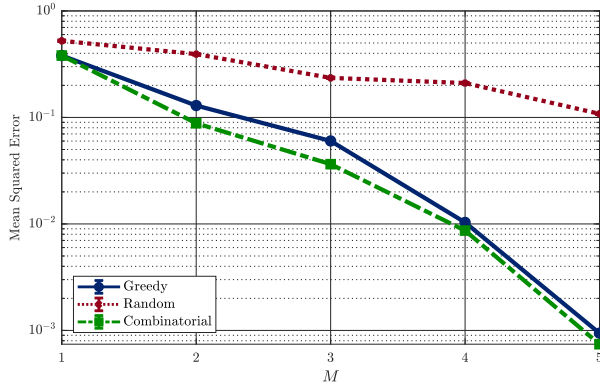


(a) Zachary's Karate Club

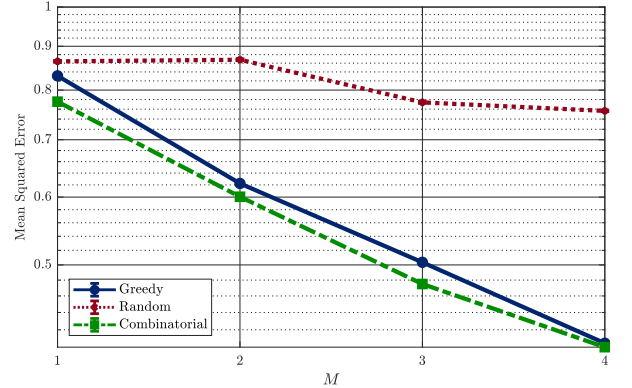


(b) Facebook

Figure 6. Impact of the bandwidth and the control signal shape. (a) Zachary's Karate Club social network. (b) Subnet of Facebook social network. An increased bandwidth (K) leads to a higher MSE, since these signals are harder to control for a fixed M and T . Higher graph frequency content signals are also harder to control.



(a) Zachary's Karate Club



(b) Facebook

Figure 7. Impact of the number of selected nodes and sampling strategy. More control nodes lead to a smaller MSE satisfying our intuition since the degrees of freedom increase. The proposed greedy heuristic is close to the optimal combinatorial solution and represents a substantial improvement compared the random node selection.

We observe that the MSE decreases as more control nodes are selected. We also observe that the greedy heuristic yields a performance similar to the optimal solution and represents a considerable improvement over random selection.

VII. CONCLUSIONS

In this paper, we studied the problem of controlling network states. We considered a random time varying network to be driven to a desired bandlimited state. To cope with the randomness in the underlying support, we introduced the concept of controllability in the mean, where we postulated to control the system as if it were running on the expected graph. We then carried out a detailed mean squared error analysis to quantify the deviation of the target signal, when the control is designed for the expected graph but ran on any given random network realization. We used this analysis to propose two different control strategies and evaluated their performance on both synthetic graph models and real-world social networks. We concluded that it is of paramount importance to take into account the random nature of the underlying topology. We leave as future work the analysis of more complex random network models, other parsimonious graph signal models, and other

control strategies that involve spectral or energetic constraints. Another direction worth investigating is the proposal of other heuristic solutions to the respective optimization problems.

APPENDIX A PROOF OF LEMMA 1

Proof. For model (i), $\mathbf{S} = \mathbf{L}$ and the system transition matrix is $\mathbf{A}_t = \mathbf{I} - \epsilon \mathbf{L}_t$ for $0 < \epsilon \leq 1/\|\mathbf{L}\|_2$. First, we prove that $\|\mathbf{L}_t\|_2 \leq \|\mathbf{L}\|_2 \leq \varrho$. Note that $\mathcal{G}_t \subseteq \mathcal{G}$ for every t and, therefore, from the Laplacian interlacing property [43] this condition always holds. The proof of Assumption 2 is straightforward, i.e., from $\mathbb{E}[\mathbf{A}_t] = \mathbf{I} - \epsilon \mathbb{E}[\mathbf{L}_t] = \mathbf{I} - \epsilon p \mathbf{L}$, which means that $\mathbb{E}[\mathbf{A}_t]$ and \mathbf{L} share the same eigenvectors. For the last condition, note that $\|\mathbf{A}_t\|_2 = \|\mathbf{I} - \epsilon \mathbf{L}_t\|_2 \leq 1$ since $\epsilon \leq 1/\|\mathbf{L}\|_2 \leq 1/\|\mathbf{L}_t\|_2$. Therefore, $\|\mathbf{A}_t\|_2$ is upper bounded by some finite ϱ .

For model (ii), $\mathbf{S} = \mathbf{W}$ and the system transition matrix is $\mathbf{A}_t = \mathbf{W}_t$. To prove that $\|\mathbf{W}_t\|_2 \leq \|\mathbf{W}\|_2$, recall that for connected graphs, the largest eigenvalue is positive and real [44, Theorem 0.2]. Then, since \mathbf{W} is considered to be normal and Assumption 1 holds, $\|\mathbf{W}\|_2 = \lambda_{\max}(\mathbf{W}) \leq \max \deg(\mathcal{G}) \leq \varrho < \infty$. Likewise, since $\mathcal{G}_t \subseteq \mathcal{G}$, then $\max \deg(\mathcal{G}_t) \leq$

$\max \deg(\mathcal{G}) < \infty$ and therefore $\|\mathbf{W}_t\|_2 \leq \varrho < \infty$ for all t . The proofs of the last two conditions are straightforward since $\mathbb{E}[\mathbf{A}_t] = p\mathbf{W}$ and $\|\mathbf{A}_t\|_2 = \|\mathbf{W}_t\|_2 \leq \varrho < \infty$. This completes the proof. \square

APPENDIX B

PROOF OF PROPOSITION 1 AND COROLLARY 1

Proof of Proposition 1 Recall that \mathcal{S} is the set of the selected M nodes and that $\mathbf{C}^T \mathbf{C} = \text{diag}(\mathbf{c})$, where $\mathbf{c} \in \{0, 1\}^N$ with $[\mathbf{c}]_i = 1$ if $v_i \in \mathcal{S}$ and $[\mathbf{c}]_i = 0$, otherwise. System (10) is equivalent to

$$\tilde{\mathbf{x}}_{t,K} = \tilde{\mathbf{A}}_K \tilde{\mathbf{x}}_{t-1,K} + \mathbf{V}_K^H \text{diag}(\mathbf{c}) \hat{\mathbf{u}}_{t-1} \quad (26)$$

where $\hat{\mathbf{u}}_t \in \mathbb{R}^{N \times 1}$ denotes the zero-extended control signal such that $[\hat{\mathbf{u}}_t]_i = [\mathbf{u}_t]_i$ if $v_i \in \mathcal{S}$ and $[\hat{\mathbf{u}}_t]_i = 0$, otherwise. Then, system (26) is controllable iff the $K \times TN$ matrix

$$\begin{aligned} \tilde{\Omega} &= [\mathbf{V}_K^H \text{diag}(\mathbf{c}), \tilde{\mathbf{A}}_K \mathbf{V}_K^H \text{diag}(\mathbf{c}), \dots, \tilde{\mathbf{A}}_K^{T-1} \mathbf{V}_K^H \text{diag}(\mathbf{c})] \\ &= [\mathbf{I}_K, \tilde{\mathbf{A}}_K, \dots, \tilde{\mathbf{A}}_K^{T-1}] (\mathbf{I}_T \otimes \mathbf{V}_K^H \text{diag}(\mathbf{c})) \end{aligned} \quad (27)$$

is full rank. Observe that

$$\begin{aligned} \text{rank}(\tilde{\Omega}) &\leq \min\{K, \text{rank}(\mathbf{I}_T \otimes \mathbf{V}_K^H \text{diag}(\mathbf{c}))\} \\ &\leq \min\{K, T \min\{K, M\}\} \end{aligned} \quad (28)$$

holds from $\text{rank}(\mathbf{AB}) \leq \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$. Therefore, to ensure the full rank K of $\tilde{\Omega}$, $M \geq \lceil K/T \rceil$ must hold, for some $T \geq 1$. This concludes the proof. \square

Proof of Corollary 1 Recall that, for two matrices $\mathbf{X} \in \mathbb{R}^{M \times N}$ and $\mathbf{Y} \in \mathbb{R}^{N \times K}$ [45, Section 0.4.6]

$$\text{if } \text{rank}(\mathbf{Y}) = N \Rightarrow \text{rank}(\mathbf{XY}) = \text{rank}(\mathbf{X}). \quad (29)$$

The mean system (14) is controllable, iff

$$\tilde{\Omega} = [\mathbf{I}_K, \tilde{\mathbf{A}}_K, \dots, \tilde{\mathbf{A}}_K^{T-1}] (\mathbf{I}_T \otimes \mathbf{V}_K^H \text{diag}(\mathbf{c})) \quad (30)$$

has rank K with $\tilde{\mathbf{A}}_K = \text{diag}(\bar{\mathbf{a}}_K)$.

The first term in (30) has rank

$$\begin{aligned} \mathbf{X} &= [\mathbf{I}_K, \tilde{\mathbf{A}}_K, \dots, \tilde{\mathbf{A}}_K^{T-1}] \in \mathbb{R}^{K \times TK} \\ \text{rank}(\mathbf{X}) &= K \end{aligned} \quad (31)$$

while the second term has rank

$$\begin{aligned} \mathbf{Y} &= (\mathbf{I}_T \otimes \mathbf{V}_K^H \text{diag}(\mathbf{c})) \in \mathbb{R}^{TK \times TN} \\ \text{rank}(\mathbf{Y}) &= T \text{rank}(\mathbf{V}_K^H \text{diag}(\mathbf{c})) \end{aligned} \quad (32)$$

since \mathbf{Y} consists of the Kronecker product of $\mathbf{V}_K^H \text{diag}(\mathbf{c})$ with an identity matrix. Note that $\mathbf{V}_K^H \text{diag}(\mathbf{c})$ selects indeed rows of \mathbf{V}_K .

Now, if $M \geq K$ and the node set \mathcal{S} are such that the selected M rows of \mathbf{V}_K form a set of K linearly independent vectors, then $\text{rank}(\mathbf{V}_K^H \text{diag}(\mathbf{c})) = K$. This implies that $\text{rank}(\mathbf{Y}) = TK$ and in virtue of (29), we obtain

$$\text{rank}(\tilde{\Omega}) = \text{rank}(\mathbf{XY}) = \text{rank}(\mathbf{X}) = K \quad (33)$$

yielding that the mean system (14) is controllable. \square

APPENDIX C

PROOF OF THEOREM 1

Proof. The MSE can be rewritten as

$$\begin{aligned} \text{MSE}(T) &= \mathbb{E}[\|\mathbf{H}\mathbf{x}_T - \mathbf{x}^*\|_2^2] \\ &= \mathbb{E}[\mathbf{x}_T^T \mathbf{H}^T \mathbf{H} \mathbf{x}_T] - 2(\mathbf{x}^*)^T \mathbf{H} \mathbb{E}[\mathbf{x}_T] + \|\mathbf{x}^*\|_2^2. \end{aligned} \quad (34)$$

where each term is computed next.

First, to compute $\mathbb{E}[\mathbf{x}_T]$ and $\mathbb{E}[\mathbf{x}_T^T \mathbf{H}^T \mathbf{H} \mathbf{x}_T]$, note that \mathbf{x}_T can be written as

$$\mathbf{x}_T = \sum_{\tau=0}^{T-1} \Phi_{T-1,\tau+1} \mathbf{B} \mathbf{u}_\tau \quad (35)$$

where $\Phi_{b,a} = \mathbf{A}_b \mathbf{A}_{b-1} \cdots \mathbf{A}_{a+1} \mathbf{A}_a$ is the state transition matrix in the interval $[a, b]$ for $b > a$. Since under the RES(p) model \mathbf{A}_t are i.i.d. matrices, $\mathbb{E}[\Phi_{b,a}] = \bar{\mathbf{A}}^{b-a+1}$. Thus, the expectation of (35) is

$$\mu_T = \mathbb{E}[\mathbf{x}_T] = \sum_{\tau=0}^{T-1} \bar{\mathbf{A}}^{T-\tau-1} \mathbf{C}^T \mathbf{u}_\tau. \quad (36)$$

For the second order moment $\mathbb{E}[\mathbf{x}_T^T \mathbf{H}^T \mathbf{H} \mathbf{x}_T]$, denote by $\mathbf{Q} = \mathbf{H}^T \mathbf{H}$ and by substituting (35) we have

$$\begin{aligned} \mathbb{E}[\mathbf{x}_T^T \mathbf{H}^T \mathbf{H} \mathbf{x}_T] &= \mathbb{E}[\mathbf{x}_T^T \mathbf{Q} \mathbf{x}_T] \\ &= \sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \mathbf{u}_\tau^T \mathbf{C} \mathbb{E}[\Phi_{T-1,\tau+1}^T \mathbf{Q} \Phi_{T-1,\tau'+1}] \mathbf{C}^T \mathbf{u}_{\tau'}. \end{aligned} \quad (37)$$

Define $\Gamma_{\tau,\tau'} = \mathbb{E}[\Phi_{t-1,\tau+1}^T \mathbf{Q} \Phi_{t-1,\tau'+1}] \in \mathbb{R}^{N \times N}$ [cf. (15)], so that (37) can be compactly written as

$$\mathbb{E}[\mathbf{x}_T^T \mathbf{H}^T \mathbf{H} \mathbf{x}_T] = \sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \mathbf{u}_\tau^T \mathbf{C} \Gamma_{\tau,\tau'} \mathbf{C}^T \mathbf{u}_{\tau'}. \quad (38)$$

Finally, by substituting (38) in the first term of the MSE (34) and (36) in the second term, we obtain the claimed expressions. This completes the proof. \square

APPENDIX D

SPECIAL CASES: USEFUL COMPUTATIONS OF THE QUADRATIC TERM $\Gamma_{\tau,\tau'}$ IN THEOREM 1

Computation of the quadratic term $\Gamma_{\tau,\tau'}$ in Theorem 1 can turn out to be quite cumbersome for arbitrary graph shift operators \mathbf{S}_t or transition matrices \mathbf{A}_t . In what follows, we offer two corollaries of Theorem 1 that address this issue. In particular, Corollary 2 gives an upper bound on the MSE that does not entail computation of second-order moments, while in Corollary 3 we show that, for the usually found case of undirected graphs, diffusion models in Lemma 1 admit an exact computation. Proofs follow after the statement of the corollaries.

Corollary 2. *Under the same conditions of Theorem 1 and from Lemma 1 the MSE (15) can be upper bounded by*

$$\begin{aligned} \text{MSE}(T) &\leq \|\mathbf{x}^*\|_2^2 - 2 \sum_{\tau=0}^{T-1} (\mathbf{x}^*)^T \mathbf{H} \bar{\mathbf{A}}^{T-\tau-1} \mathbf{C}^T \mathbf{u}_\tau \\ &\quad + \sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \varrho^{2(T-\tau'+1)} \langle \mathbf{u}_{\tau'}, \mathbf{u}_\tau \rangle. \end{aligned} \quad (39)$$

The result of Corollary 2 can be interpreted as the worst case scenario to account for the variability in the topology. In fact, (39) shows only first order dependence from the RES(p) model, but it does not show dependence from the second order moment. For the models in Lemma 1, we can consider for model (i) $\varrho = 1$ ($\bar{\mathbf{A}} = \mathbf{I} - \epsilon p \mathbf{L}$), while for model (ii) $\varrho = \max \deg(\mathcal{G})$ ($\bar{\mathbf{A}} = p \mathbf{W}$). Further insight then on the role of the graph variability is given by Corollary 3, which shows the explicit dependence on the link activation probability p .

Corollary 3. *Under the same conditions of Theorem 1 and additionally given that \mathcal{G} is an undirected graph, for the diffusion models in Lemma 1 the following holds:*

$$\begin{aligned} \Gamma_{\tau, \tau'} &= (\bar{\mathbf{A}}^{\tau' - \tau})^\top \mathbf{Q}_{T - \tau' - 1} \\ \mathbf{Q}_a &= \mathbb{E} [\mathbf{A}_{T-a}^\top \mathbf{Q}_{a-1} \mathbf{A}_{T-a}], \quad a = 1, \dots, T - \tau' - 1 \end{aligned} \quad (40)$$

for $\tau \leq \tau'$, $\mathbf{Q}_0 = \mathbf{H}^\top \mathbf{H}$, and $\bar{\mathbf{A}}^{\tau' - \tau} = (\mathbb{E}[\mathbf{A}_t])^{\tau' - \tau}$.

For model (i) in Lemma 1 ($\bar{\mathbf{A}} = \mathbf{I} - \epsilon p \mathbf{L}$), \mathbf{Q}_a is

$$\begin{aligned} \mathbf{Q}_a &= \epsilon^2 p^2 \mathbf{W}^\top \mathbf{Q}_{a-1} \mathbf{W} \\ &\quad + \epsilon^2 p(1-p) \\ &\quad \cdot (\mathbf{W}^\top \circ \mathbf{Q}_{a-1} \circ \mathbf{W} - \text{diag}(\mathbf{W}^\top \circ \mathbf{Q}_{a-1} \circ \mathbf{W})) \\ &\quad + \epsilon^2 p(1-p) \text{diag}(\mathbf{W}^\top \text{diag}(\mathbf{Q}_{a-1}) \mathbf{W}) \\ &\quad + 2\epsilon p(\mathbf{I} - \epsilon p \mathbf{D}) \text{diag}(\mathbf{Q}_{a-1} \mathbf{W}) \\ &\quad + ((\mathbf{I} - \epsilon p \mathbf{D})^2 + \epsilon^2 p(1-p) \mathbf{W}^\top \mathbf{W}) \circ \text{diag}(\mathbf{Q}_{a-1}). \end{aligned} \quad (41)$$

For model (ii) in Lemma 1 ($\bar{\mathbf{A}} = p \mathbf{W}$), \mathbf{Q}_a is

$$\begin{aligned} \mathbf{Q}_a &= p^2 \mathbf{W}^\top \mathbf{Q}_{a-1} \mathbf{W} \\ &\quad + p(1-p) (\mathbf{W}^\top \circ \mathbf{Q}_{a-1} \circ \mathbf{W} - \text{diag}(\mathbf{W}^\top \circ \mathbf{Q}_{a-1} \circ \mathbf{W})) \\ &\quad + p(1-p) \text{diag}(\mathbf{W}^\top \text{diag}(\mathbf{Q}_{a-1}) \mathbf{W}). \end{aligned} \quad (42)$$

Proof of Corollary 2 From the first term in (34), we have

$$\mathbb{E} [\mathbf{x}_T^\top \mathbf{H}^\top \mathbf{H} \mathbf{x}_T] = \mathbb{E} [\text{tr} [\mathbf{H} \mathbf{x}_T \mathbf{x}_T^\top \mathbf{H}^\top]]. \quad (43)$$

The trace argument in (43) can be expanded as

$$\mathbf{H} \mathbf{x}_T \mathbf{x}_T^\top \mathbf{H}^\top = \sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \mathbf{H} \Phi_{T-1, \tau+1} \mathbf{C}^\top \mathbf{u}_\tau \mathbf{u}_{\tau'}^\top \mathbf{C} \Phi_{T-1, \tau'+1}^\top \mathbf{H}^\top. \quad (44)$$

From Lemma 1 we have $\|\mathbf{A}_t\|_2 \leq \varrho$; so that, by the submultiplicativity of the spectral norm, we can write

$$\|\Phi_{b,a}\|_2 = \|\mathbf{A}_b \mathbf{A}_{b-1} \cdots \mathbf{A}_a\|_2 \leq \|\mathbf{A}_b\|_2 \cdots \|\mathbf{A}_a\|_2 \leq \varrho^{b-a+1}. \quad (45)$$

Also, observe that for any square matrix \mathbf{X} and positive semidefinite matrix \mathbf{Y} , it holds that $\text{tr}[\mathbf{X}\mathbf{Y}] \leq \|\mathbf{X}\|_2 \text{tr}[\mathbf{Y}]$ (46).

Given that the filter \mathbf{H} does not amplify any frequency (i.e. $\|\mathbf{H}\|_2 = 1$), then plugging back (44) and (45) into (43) yields

$$\begin{aligned} &\mathbb{E} \left[\sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \text{tr} [\mathbf{H} \Phi_{T-1, \tau+1} \mathbf{C}^\top \mathbf{u}_\tau \mathbf{u}_{\tau'}^\top \mathbf{C} \Phi_{T-1, \tau'+1}^\top \mathbf{H}^\top] \right] \\ &= \sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \mathbb{E} \left[\text{tr} \left[\left(\Phi_{T-1, \tau'+1}^\top \mathbf{H}^\top \mathbf{H} \Phi_{T-1, \tau+1} \right) (\mathbf{C}^\top \mathbf{u}_\tau \mathbf{u}_{\tau'}^\top \mathbf{C}) \right] \right] \\ &\leq \sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \mathbb{E} \left[\|\Phi_{T-1, \tau'+1}^\top\|_2 \|\Phi_{T-1, \tau+1}\|_2 \text{tr} [\mathbf{C}^\top \mathbf{u}_\tau \mathbf{u}_{\tau'}^\top \mathbf{C}] \right] \\ &\leq \sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \varrho^{2(t-\tau'+1)} \text{tr} [\mathbf{C} \mathbf{C}^\top \mathbf{u}_\tau \mathbf{u}_{\tau'}^\top] \\ &= \sum_{\tau=0}^{T-1} \sum_{\tau'=0}^{T-1} \varrho^{2(t-\tau'+1)} \langle \mathbf{u}_{\tau'}, \mathbf{u}_\tau \rangle. \end{aligned} \quad (46)$$

Using (46) to bound (43) and replacing it in (15) yields (39). \square

Proof of Corollary 3 Consider $\tau \leq \tau'$. From (15), we have

$$\begin{aligned} \Gamma_{\tau, \tau'} &= \mathbb{E} \left[(\mathbf{A}_{T-1} \cdots \mathbf{A}_{\tau+1})^\top \mathbf{Q} (\mathbf{A}_{T-1} \cdots \mathbf{A}_{\tau'+1}) \right] \\ &= \mathbb{E} [\mathbf{A}_{\tau+1}^\top \mathbf{A}_{\tau+2}^\top \cdots \mathbf{A}_{T-1}^\top \mathbf{Q} \mathbf{A}_{T-1} \cdots \mathbf{A}_{\tau'+2} \mathbf{A}_{\tau'+1}]. \end{aligned} \quad (47)$$

Then, since for two random matrices \mathbf{X}, \mathbf{Y} it holds that $\mathbb{E}[\mathbf{X}] = \mathbb{E}[\mathbb{E}[\mathbf{X}|\mathbf{Y}]]$ (47, Theorem 34.4), (47) becomes

$$\begin{aligned} \Gamma_{\tau, \tau'} &= \mathbb{E} [\mathbb{E} [\mathbf{A}_{\tau+1}^\top \cdots \mathbf{A}_{T-1}^\top \mathbf{Q} \mathbf{A}_{T-1} \cdots \mathbf{A}_{\tau'+1} | \mathbf{A}_{T-2}, \dots, \mathbf{A}_{\tau+1}]] \\ &= \mathbb{E} [\mathbf{A}_{\tau+1}^\top \cdots \mathbb{E} [\mathbf{A}_{T-1}^\top \mathbf{Q} \mathbf{A}_{T-1} | \mathbf{A}_{T-2}, \dots, \mathbf{A}_{\tau+1}] \cdots \mathbf{A}_{\tau'+1}]. \end{aligned} \quad (48)$$

which under the RES(p) model (i.e., matrices \mathbf{A}_a are i.i.d.) can be written as

$$\begin{aligned} \Gamma_{\tau, \tau'} &= \mathbb{E} [\mathbf{A}_{\tau+1}^\top \cdots \mathbb{E} [\mathbf{A}_{T-1}^\top \mathbf{Q} \mathbf{A}_{T-1}] \cdots \mathbf{A}_{\tau'+1}] \\ &= (\bar{\mathbf{A}}^{\tau' - \tau})^\top \mathbb{E} [\mathbf{A}_{\tau+1}^\top \mathbb{E} [\cdots \mathbb{E} [\mathbf{A}_{T-1}^\top \mathbf{Q} \mathbf{A}_{T-1}] \cdots] \mathbf{A}_{\tau'+1}]. \end{aligned} \quad (49)$$

Further, for $a \geq 1$ and assuming for now (to be proven later on) that \mathbf{Q}_{a-1} is symmetric and positive semidefinite, we proceed to compute the (i, j) entry of matrix $\mathbf{Q}_a = \mathbb{E}[\mathbf{A}_{T-a}^\top \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]$. Towards this end, denote simply by $[\mathbf{Q}_{a-1}]_{ij} = q_{ij}$ and $[\mathbf{A}_{T-a}]_{ij} = a_{ij}$ for $i, j = 1, \dots, N$.

For $i \neq j$ and since \mathbf{A}_{T-a} is symmetric, the (i, j) element of \mathbf{Q}_a becomes

$$\begin{aligned} \mathbb{E} [[\mathbf{A}_{T-a}^\top \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ij}] &= \sum_{k=1}^N \sum_{\ell=1}^N \mathbb{E} [a_{ki} a_{\ell j}] q_{k\ell} \\ &= \sum_{k=1; k \neq j}^N \sum_{\ell=1; \ell \neq i}^N \mathbb{E} [a_{ki}] \mathbb{E} [a_{\ell j}] q_{k\ell} + \mathbb{E} [a_{ji} a_{ij}] q_{ji} \end{aligned} \quad (50)$$

where we have used the independence of the distinct elements in \mathbf{A}_{T-a} . The second term of (50) groups the element (i, j)

together with (j, i) due to symmetry of \mathbf{A}_{T-a} . Analogously, for the diagonal elements $i = j$, we get

$$\begin{aligned} \mathbb{E} [\mathbf{A}_{T-a}^T \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ii} &= \sum_{k=1}^N \sum_{\ell=1}^N \mathbb{E}[a_{ki} a_{\ell i}] q_{k\ell} \\ &= \sum_{k=1}^N \sum_{\ell=1; \ell \neq k}^N \mathbb{E}[a_{ki}] \mathbb{E}[a_{\ell i}] q_{k\ell} + \sum_{k=1}^n \mathbb{E}[a_{ki}^2] q_{kk}. \end{aligned} \quad (51)$$

With this in place, let us first consider the simpler model (ii) in Lemma 1 where $\mathbf{S}_t = \mathbf{W}_t$ and $\mathbf{A}_t = \mathbf{W}_t$. For this case, we have $[\mathbf{A}_t]_{ij} = B w_{ij}$, where B is a Bernoulli random variable of parameter p and $w_{ij} = [\mathbf{W}]_{ij}$. Then, by substituting $\mathbb{E}[a_{ij}] = p w_{ij}$ and $\mathbb{E}[a_{ij}^2] = (p^2 + p(1-p)) w_{ij}^2$ in (50), we get

$$\begin{aligned} p^2 \sum_{k=1; k \neq j}^N \sum_{\ell=1; \ell \neq j}^N w_{ki} w_{\ell j} q_{k\ell} + (p^2 + p(1-p)) w_{ji}^2 q_{ji} \\ = p^2 \sum_{k=1}^N \sum_{\ell=1}^N w_{ki} w_{\ell j} q_{k\ell} + p(1-p) w_{ji}^2 q_{ji} \end{aligned} \quad (52)$$

which can be written in the compact form

$$\begin{aligned} \mathbb{E} [\mathbf{A}_{T-a}^T \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ij} \\ = p^2 [\mathbf{W}^T \mathbf{Q}_{a-1} \mathbf{W}]_{ij} + p(1-p) [\mathbf{W}^T \circ \mathbf{Q}_{a-1} \circ \mathbf{W}]_{ij}. \end{aligned} \quad (53)$$

Likewise, for $i = j$ (51) becomes

$$\begin{aligned} p^2 \sum_{k=1}^N \sum_{\ell=1; \ell \neq k}^N w_{ki} w_{\ell i} q_{k\ell} + (p^2 + p(1-p)) \sum_{k=1}^N w_{ki}^2 q_{kk} \\ = p^2 \sum_{k=1}^N \sum_{\ell=1}^N w_{ki} w_{\ell i} q_{k\ell} + p(1-p) \sum_{k=1}^N w_{ki}^2 q_{kk} \end{aligned} \quad (54)$$

which can also be written compactly as

$$\begin{aligned} \mathbb{E} [\mathbf{A}_{T-a}^T \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ii} \\ = p^2 [\mathbf{W}^T \mathbf{Q}_{a-1} \mathbf{W}]_{ii} + p(1-p) [\mathbf{W}^T \text{diag}(\mathbf{Q}_{a-1}) \mathbf{W}]_{ii}. \end{aligned} \quad (55)$$

By combining (53) and (55) yields (42). Finally, note that if \mathbf{Q}_{a-1} is symmetric and positive semidefinite, then so is \mathbf{Q}_a . To complete the proof, observe that $\mathbf{Q}_0 = \mathbf{H}^T \mathbf{H}$ is symmetric and positive semidefinite, thus (42) holds for all $a \geq 1$.

For model (i) in Lemma 1 we proceed in an analogous way. In this case, $\mathbf{S}_t = \mathbf{L}_t$ and $\mathbf{A}_t = \mathbf{I} - \epsilon \mathbf{L}_t = (\mathbf{I} - \epsilon \mathbf{D}_t) + \epsilon \mathbf{W}_t$, where $\mathbf{D}_t = \text{diag}(\mathbf{W}_t \mathbf{1})$ is the degree matrix. This means that $[\mathbf{A}_t]_{ij} = a_{ij} = \epsilon B w_{ij}$ if $i \neq j$ and $[\mathbf{A}_t]_{ii} = a_{ii} = 1 - \epsilon \sum_{k=1}^N B_k w_{ik}$ with B_k being i.i.d. Bernoulli random variables with probability p .

Then, for $i \neq j$, from (50) we have

$$\begin{aligned} \mathbb{E} [\mathbf{A}_{T-a}^T \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ij} &= \\ &= \sum_{k=1; k \neq j, i}^N \sum_{\ell=1; \ell \neq i, j}^N \mathbb{E}[a_{ki}] \mathbb{E}[a_{\ell j}] q_{k\ell} \\ &\quad + \mathbb{E}[a_{ji} a_{ij}] q_{ji} + \mathbb{E}[a_{ii}] \mathbb{E}[a_{jj}] q_{ij} \\ &= \epsilon^2 p^2 \sum_{k=1; k \neq j, i}^N \sum_{\ell=1; \ell \neq i, j}^N w_{ki} w_{\ell j} q_{k\ell} \\ &\quad + \epsilon^2 (p^2 + p(1-p)) w_{ji}^2 q_{ji} + (1 - \epsilon p d_i) q_{ij}. \end{aligned} \quad (56)$$

Now, recalling that $w_{ii} = 0$ (i.e., no self-loops), (56) becomes

$$\begin{aligned} \mathbb{E} [\mathbf{A}_{T-a}^T \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ij} &= \epsilon^2 p^2 \sum_{k=1}^N \sum_{\ell=1}^N w_{ki} w_{\ell j} q_{k\ell} \\ &\quad + \epsilon^2 p(1-p) w_{ji}^2 q_{ji} + (1 - \epsilon p d_i) q_{ij}, \end{aligned} \quad (57)$$

which can be further written in the compact form

$$\begin{aligned} \mathbb{E} [\mathbf{A}_{T-a}^T \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ij} &= \epsilon^2 p^2 [\mathbf{W}^T \mathbf{Q}_{a-1} \mathbf{W}]_{ij} \\ &\quad + \epsilon^2 p(1-p) [\mathbf{W}^T \circ \mathbf{Q}_{a-1} \circ \mathbf{W}]_{ij} + [(1 - \epsilon p \mathbf{D}) \mathbf{Q}_{a-1}]_{ij} \end{aligned} \quad (58)$$

For $i = j$, we start with (51)

$$\begin{aligned} \mathbb{E} [\mathbf{A}_{T-a}^T \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ii} \\ &= \sum_{k=1; k \neq i}^N \sum_{\ell=1; \ell \neq k, i}^N \mathbb{E}[a_{ki}] \mathbb{E}[a_{\ell i}] q_{k\ell} \\ &\quad + \sum_{k=1; k \neq i}^N \mathbb{E}[a_{ki}^2] q_{kk} + 2 \mathbb{E}[a_{ii}] \sum_{\ell=1; \ell \neq i}^N \mathbb{E}[a_{\ell i}] q_{i\ell} + \mathbb{E}[a_{ii}^2] q_{ii}. \end{aligned} \quad (59)$$

Then, recalling that $w_{ii} = 0$, we replace the first and second order moments for each a_{ij} and obtain

$$\begin{aligned} \mathbb{E} [\mathbf{A}_{T-a}^T \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ii} &= \epsilon^2 p^2 \sum_{k=1}^N \sum_{\ell=1; \ell \neq k}^N w_{ki} w_{\ell i} q_{k\ell} \\ &\quad + \epsilon^2 (p^2 + p(1-p)) \sum_{k=1}^N w_{ki}^2 q_{kk} + 2(1 - \epsilon p d_i) \epsilon p \sum_{\ell=1}^N w_{\ell i} q_{i\ell} \\ &\quad + q_{ii} \left(1 - 2 \epsilon p d_i + \epsilon^2 p^2 \sum_{k=1}^N \sum_{\ell=1}^N w_{ik} w_{i\ell} \right. \\ &\quad \left. + \epsilon^2 p(1-p) \sum_{k=1}^N w_{ik}^2 \right). \end{aligned} \quad (60)$$

Finally, this can be rewritten as

$$\begin{aligned} \mathbb{E} [\mathbf{A}_{T-a}^T \mathbf{Q}_{a-1} \mathbf{A}_{T-a}]_{ii} \\ &= \epsilon^2 p^2 [\mathbf{W}^T \mathbf{Q}_{a-1} \mathbf{W}]_{ii} + \epsilon^2 p(1-p) [\mathbf{W}^T \text{diag}(\mathbf{Q}_{a-1}) \mathbf{W}]_{ii} \\ &\quad + 2 \epsilon p [(1 - \epsilon p \mathbf{D}) \text{diag}(\mathbf{Q}_{a-1} \mathbf{W})]_{ii} \\ &\quad + [((1 - \epsilon p \mathbf{D})^2 + \epsilon^2 p(1-p) \mathbf{W}^T \mathbf{W}) \circ \text{diag}(\mathbf{Q}_{a-1})]_{ii} \end{aligned} \quad (61)$$

completing the proof. \square

REFERENCES

- [1] F. Gama, E. Isufi, G. Leus, and A. Ribeiro, "Control of graph signals over random time-varying graphs," in *43rd IEEE Int. Conf. Acoust., Speech and Signal Process.* Calgary, AB: IEEE, 15-20 Apr. 2018, pp. 4169-4173.
- [2] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, pp. 268-276, March 2001.
- [3] M. Newman, A.-L. Barabási, and D. J. Watts, *The Structure and Dynamics of Networks*, ser. Princeton Studies in Complexity. Princeton, NJ: Princeton University Press, 2006.
- [4] A. Lombardi and M. Hörnquist, "Controllability analysis of networks," *Physical Rev. E*, vol. 75, no. 056110, pp. 1-5, May 2007.
- [5] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, pp. 167-173, May 2011.
- [6] Z. Yuan, C. Zhao, Z. Di, and W.-X. Wang, "Exact controllability of complex networks," *Nature Commun.*, vol. 4, no. 2447, pp. 1-9, Sep. 2013.

- [7] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Trans. Control Network Syst.*, vol. 1, no. 1, pp. 40–52, March 2014.
- [8] W. W. Zachary, "An information flow model for conflict and fission in small groups," *J. Anthropol. Res.*, vol. 33, no. 1, pp. 452–473, 1977.
- [9] G. Taubin, "Geometric signal processing on polygonal meshes," in *Eurographics 2000*. Interlaken, Switzerland: Eur. Assoc. Comput. Graph., 21–25 Aug. 2000, pp. 1–11.
- [10] 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, May 2013.
- [11] A. Sandryhaila and J. M. F. Moura, "Big data analysis with signal processing on graphs: Representation processing of massive data sets with irregular structure," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80–90, Sep. 2014.
- [12] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, Dec. 2015.
- [13] S. K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *38th IEEE Int. Conf. Acoust., Speech and Signal Process.* Vancouver, BC: IEEE, 26–31 May 2013, pp. 5445–5449.
- [14] F. Gama, A. G. Marques, G. Mateos, and A. Ribeiro, "Rethinking Sketching as Sampling: A Graph Signal Processing Approach," *arXiv:1611.00119v2 [cs.IT]*, 11 Sep. 2018. [Online]. Available: <http://arxiv.org/abs/1611.00119>
- [15] P. Di Lorenzo, P. Banelli, E. Isufi, S. Barbarossa, and G. Leus, "Adaptive graph signal processing: Algorithms and optimal sampling strategies," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3584–3598, July 2018.
- [16] E. Isufi, P. Banelli, P. Di Lorenzo, and G. Leus, "Observing and tracking bandlimited graph processes," *arXiv:1712.00404v2 [eess.SP]*, 19 Oct. 2018. [Online]. Available: <http://arxiv.org/abs/1712.00404>
- [17] 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.
- [18] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, May 2018.
- [19] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Sampling of graph signals with successive local aggregations," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1832–1843, Apr. 2016.
- [20] W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Graph frequency analysis of brain signals," *IEEE J. Select. Topics Signal Process.*, vol. 10, no. 7, pp. 1189–1203, Oct. 2016.
- [21] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, "Reconstruction of graph signals through percolation from seeding nodes," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4363–4378, Aug. 2016.
- [22] S. Barbarossa, S. Sardellitti, and A. Farina, "On sparse controllability of graph signals," in *41st IEEE Int. Conf. Acoust., Speech and Signal Process.* Shanghai, China: IEEE, 20–25 March 2016, pp. 4104–4108.
- [23] J. A. Bazerque and P. Monzón, "Control of networked systems in the graph-frequency domain," in *51st Asilomar Conf. Signals, Systems and Comput.* Pacific Grove, CA: IEEE, 29 Oct.–1 Nov. 2017, pp. 1444–1448.
- [24] 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, Jan. 2017.
- [25] S. Chen, R. Varma, A. Singh, and J. Kovačević, "Representations of piecewise smooth signals on graphs," in *41st IEEE Int. Conf. Acoust., Speech and Signal Process.* Shanghai, China: IEEE, 20–25 March 2016, pp. 6370–6374.
- [26] A. Jung and N. Tran, "Localized linear regression in networked data," *IEEE Signal Process. Lett.*, vol. 26, no. 7, pp. 1090–1094, July 2019.
- [27] 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.
- [28] C. Godsil and G. Royle, *Algebraic Graph Theory*, ser. Graduate Texts in Mathematics. New York, NY: Springer, 2001, vol. 207.
- [29] I. Pesenson, "Sampling in Paley-Wiener spaces on combinatorial graphs," *Trans. Amer. Math. Soc.*, vol. 360, no. 10, pp. 5603–5627, Oct. 2008.
- [30] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4845–4860, Sep. 2016.
- [31] E. Isufi, A. Loukas, A. Simonetto, and G. Leus, "Filtering random graph processes over random time-varying graphs," *IEEE Trans. Signal Process.*, vol. 65, no. 16, pp. 4406–4421, Aug. 2017.
- [32] R. I. Kondor and J. D. Lafferty, "Diffusion kernels on graphs and other discrete input spaces," in *19th Int. Conf. Mach. Learning*, Sydney, Australia, 8–12 July 2002, pp. 315–322.
- [33] D. Thanou, X. Dong, D. Kressner, and P. Frossard, "Learning heat diffusion graphs," *IEEE Trans. Signal, Inform. Process. Networks*, vol. 3, no. 3, pp. 484–499, Sep. 2017.
- [34] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.
- [35] F. Gama and A. Ribeiro, "Ergodicity in stationary graph processes: A weak law of large numbers," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2761–2774, May 2019.
- [36] T. Kailath, *Linear Systems*, ser. Prentice-Hall Inform. Syst. Series. Englewood Cliffs, NJ: Prentice-Hall, 1980.
- [37] 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.
- [38] F. Lewis and V. L. Syrmos, *Optimal Control*, 2nd ed. New York, NY: John Wiley & Sons, 1995.
- [39] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [40] L. F. O. Chamon and A. Ribeiro, "Greedy sampling of graph signals," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 34–47, Jan. 2018.
- [41] P. Erdős and A. Rényi, "On random graphs I," *Publicationes Mathematicae Debrecen*, vol. 6, pp. 290–297, 1959.
- [42] J. McAuley and J. Leskovec, "Learning to discover social circles in Ego networks," in *26th Conf. Neural Inform. Process. Syst.* Stateline, TX: Neural Inform. Process. Syst. Foundation, 3–8 Dec. 2012.
- [43] G. Chen, G. Davis, F. Hall, Z. Li, K. Patel, and M. Stewart, "An interlacing result on normalized laplacians," *SIAM J. Discrete Math.*, vol. 18, no. 2, pp. 353–361, Aug. 2016.
- [44] D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs: Theory and Applications*. New York, NY: Academic Press, 1979.
- [45] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.
- [46] S.-D. Wang, T.-S. Kuo, and C.-F. Hsu, "Trace bounds on the solution of algebraic matrix Riccati and Lyapunov equation," *IEEE Trans. Autom. Control*, vol. 31, no. 7, pp. 654–656, July 1986.
- [47] P. Billingsley, *Probability and Measure*, 3rd ed., ser. Wiley Series in Probability and Mathematical Statistics. New York, NY: John Wiley & Sons, 1995.