

Distributed Optimal Control of Graph Symmetric Systems via Graph Filters

Fengjun Yang*, Fernando Gama[†], Somayeh Sojoudi[‡] and Nikolai Matni[§]

Abstract—Designing distributed optimal controllers subject to communication constraints is a difficult problem unless structural assumptions are imposed on the underlying dynamics and information exchange structure, e.g., sparsity, delay, or spatial invariance. In this paper, we borrow ideas from graph signal processing and define and analyze a class of *Graph Symmetric Systems* (GSSs), which are systems that are symmetric with respect to an underlying graph topology. We show that for linear quadratic problems subject to dynamics defined by a GSS, the optimal centralized controller is given by a novel class of graph filters with transfer function valued filter taps and can be implemented via distributed message passing. We then propose several methods for approximating the optimal centralized graph filter by a distributed controller only requiring communication with a small subset of neighboring subsystems. We further provide stability and suboptimality guarantees for the resulting distributed controllers. Finally, we empirically demonstrate that our approach allows for a principled tradeoff between communication cost and performance while guaranteeing stability. Our results can be viewed as a first step towards bridging the fields of distributed optimal control and graph signal processing.

I. INTRODUCTION

Computing a distributed optimal controller in which sub-controllers have access to subsets of global system information is in general a computationally intractable problem. Indeed, even when restricted to quadratic costs, Gaussian noise, and linear dynamics, the resulting optimal controller can be nonlinear and difficult to compute [1]. Nevertheless, significant progress has been made in distributed optimal controller synthesis over the past two decades by identifying structural assumptions on the underlying dynamics and information exchange structure such that the resulting distributed controller synthesis problem is convex.

One such structural assumption that has been shown to lead to tractable distributed optimal control problems is spatial invariance [2] (and other closely related notions of symmetry [3]). Such systems are invariant under subsystem permutations, and have been shown to have optimal centralized controllers that are approximately distributed. In particular, this allows for distributed controllers that enjoy stability and near-optimality guarantees to be computed by appropriately truncating the centralized controller.

*F. Yang is with the Dept. of Comput. and Info. Sci., University of Pennsylvania, Philadelphia, PA. [†]F. Gama is with the Dept. of Elect. and Comput. Eng., Rice University, Houston, TX. [‡]S. Sojoudi is with the Dept. of Elect. Eng. and Comput. Sci., University of California, Berkeley, CA. [§]N. Matni is with the Dept. of Elect. and Syst. Eng., Philadelphia, PA.

N.M. is supported by NSF awards CPS-2038873, CAREER award ECCS-2045834, and a Google Research Scholar award. F.Y. is in part supported by NSF CAREER award ECCS-2045834. F.G. and S.S. are supported by ONR and NSF grants.

Contributions: In this paper, inspired by results from graph signal processing, we introduce the notion of Graph Symmetric Systems (GSSs), which are systems that are symmetric with respect to an underlying graph topology (formalized in §II-A). We show that for such systems, the resulting Linear Quadratic (LQ) centralized optimal controller admits an efficient message passing implementation in the form of a novel class of graph filters defined by transfer function filter taps. We subsequently propose and analyze two complementary approaches to computing near-optimal distributed controllers by truncating the centralized optimal controller subject to stability constraints. By leveraging tools from robust System Level Synthesis (SLS) [4], [5], we show that these truncation algorithms can be solved via convex optimization, and that the resulting distributed controllers enjoy sub-optimality guarantees relative to the centralized optimal controller. These results constitute an important first step towards bridging the complementary, but traditionally disparate, fields of distributed optimal control and graph signal processing.

Related work: An alternative structural assumption for tractable distributed optimal control of linear systems can be specified in terms of the sparsity and delay patterns of the control system. In particular, it is possible to characterize conditions on the sparsity and delay patterns of the information exchanged between subcontrollers relative to the propagation of signals through sparse and delayed distributed plants such that distributed optimal control is tractable. The seminal paper [6] introduced the notion of quadratic invariance,¹ which built upon and generalized funnel causality [7], showed that so long as subcontrollers could communicate as quickly as control signals propagated through the plant, then the resulting distributed optimal control problem could be solved via convex optimization. This convex parameterization of sparse and delayed controllers has since been further generalized in the System Level Synthesis (SLS) [5] and Input-Output Parameterization (IOP) [8] frameworks, which allow for even richer classes of sparsity and delay patterns to be imposed on distributed controllers.

A related class of distributed controllers are those based on Graph Neural Networks (GNNs). GNNs can be viewed as graph filters followed by pointwise nonlinear activation functions [9], and among other favorable properties, enjoy stability to graph perturbations [10]. While recent use of GNNs for distributed control has shown promise [11]–[13],

¹We note that spatially invariant systems, as defined in [2], also satisfy quadratic invariance. We show in Appendix B that graph symmetric systems and controllers lead to optimal control problems satisfying quadratic invariance.

such results currently lack strong guarantees of stability. We believe the results in this paper are a first step towards addressing this gap in the literature, by explicitly connecting graph filters and distributed optimal controllers. The direct relationship between graph filters and GNNs suggests that understanding the former will give insight in the effects of the latter.

Notation: We use upper- and lower-case letters such as A and x to denote matrices and vectors respectively, although lower-case letters might also be used for scalars or functions (the distinction will be apparent from the context). For both upper- and lower-case letters, we use *boldface* such as \mathbf{A} and ϕ to denote transfer matrices or vector/scalar transfer functions.

II. THE LINEAR QUADRATIC REGULATOR PROBLEM FOR GRAPH SYMMETRIC SYSTEMS

Consider a discrete-time linear time-invariant (LTI) system composed of N interconnected scalar subsystems, each with state $x_i(t) \in \mathbb{R}$, control input $u_i(t) \in \mathbb{R}$ and which evolves under the dynamics

$$x_i(t+1) = \sum_{j=1}^N A_{ij}x_j(t) + \sum_{j=1}^N B_{ij}u_j(t) + w_i(t), \quad (1)$$

for suitable matrices A_{ij}, B_{ij} describing the interaction between subsystems. Here $w_i(t)$ is an i.i.d. zero-mean noise. We can compactly express the dynamics of the full system in terms of the joint states $x(t) = [x_1(t), \dots, x_N(t)]^\top$ and joint control actions $u(t) = [u_1(t), \dots, u_N(t)]^\top$ as

$$x(t+1) = Ax(t) + Bu(t) + w(t), \quad (2)$$

where (A, B) are defined such that the global dynamics (2) are consistent with the subsystem dynamics (1).

Our goal is to find a (potentially time-varying) state-feedback controller K_t that minimizes the cost

$$J(\{x(t)\}, \{u(t)\}) := \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_w [x(t)^\top Qx(t) + u(t)^\top Ru(t)] \quad (3)$$

where $u(t) = K_t(x(t))$ and $Q \succeq 0, R \succ 0$ are known symmetric $N \times N$ matrices. The Linear Quadratic Regulator (LQR) problem is then given by

$$\min_{\mathbf{K}} J(\{x(t)\}, \{u(t)\}) \quad \text{s. t. } u(t) = K_t(x(t)). \quad (4)$$

In the centralized setting where each subsystem has access to the global state, it is well-known that the controller that solves (4) is a linear static controller $u(t) = K^*x(t)$ where $K^* = -(R+B^\top PB)^{-1}B^\top PA$ and P is the unique solution to the discrete-time algebraic Ricatti equation:

$$P = A^\top PA - A^\top PB(R+B^\top PB)^{-1}B^\top PA + Q. \quad (5)$$

In this work, we consider a distributed variant of (4) where each subsystem can only exchange information with a small subset of subsystems. Specifically, this communication constraint is encoded as a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} =$

$\{v_1, \dots, v_N\}$ is the set of N components (nodes) and where $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of the corresponding interconnections (edges). It is assumed that the graph is undirected, i.e. $(v_i, v_j) \in \mathcal{E}$ if and only if $(v_j, v_i) \in \mathcal{E}$. As described in the introduction, general information exchange constraints can lead to non-convex optimal control problems [1]. However, as we show later, under suitable graph symmetry assumptions on the dynamics matrices A, B and the cost matrices Q, R , the optimal centralized controller admits a distributed message passing implementation allowing for a principled tradeoff between communication complexity and controller performance.

In the rest of this section, we borrow ideas from graph signal processing [14] and introduce the notion of *graph symmetric systems*. First, we introduce a convenient way to define operations that respect the underlying communication graph structure via the graph matrix description (GMD) $S \in \mathbb{R}^{N \times N}$. The matrix S is such that the (i, j) th entry is zero whenever there is no connection between components v_i and v_j , i.e. $[S]_{ij} = 0$ if $i \neq j$ and $(v_i, v_j) \notin \mathcal{E}$. Note that, since the graph is undirected, the matrix S is symmetric. Therefore, it has an eigendecomposition in terms of an orthonormal basis of eigenvectors $S = V\Lambda_S V^\top$ where $\Lambda_S \in \mathbb{R}^{N \times N}$ is a diagonal matrix with elements $\lambda_{S,i} \in \mathbb{R}$ such that $Sv_i = \lambda_{S,i}v_i$ for $v_i \in \mathbb{R}^N$ being the i th column of V . We now introduce the notion of a graph symmetric system.

Definition 1 (Graph Symmetric System). Given a GMD $S = V\Lambda_S V^\top$ for a graph \mathcal{G} , a linear system (2) is *graph symmetric* with respect to \mathcal{G} if the dynamics matrices A, B are simultaneously diagonalized by V , i.e.,

$$A = V\Lambda_A V^\top, \quad B = V\Lambda_B V^\top, \quad (6)$$

where Λ_A, Λ_B are diagonal.

Note that Definition 1 does not require the dynamics to be sparse. In fact, matrices A and B of the form in Definition 1 can be arbitrarily dense, i.e., the evolution of a subsystem state $x_i(t+1)$ can depend on subsystems that i cannot directly communicate with [15]. This is distinct from sparsity/delay constraints used in [5], [6], [8], and encodes a different notion of symmetry than that exploited in the distributed control of spatially invariant systems [2].

By well-known results in graph signal processing [14], simultaneous diagonalizability of the system matrices (A, B) and the GMD S implies² that they can be written as matrix polynomials of S of degrees at most $N-1$,

$$A = \sum_{k=0}^{N-1} h_{A,k} S^k, \quad B = \sum_{k=0}^{N-1} h_{B,k} S^k. \quad (7)$$

Matrices that can be expressed in this matrix polynomial forms are called *graph filters* [16] and the coefficients $h_{A,k}, h_{B,k}$ are referred to as the filter weights or *filter taps*.

²Under the assumption that S corresponds to a finite graph and has all distinct eigenvalues. On a high level, the result follows directly from the Cayley-Hamilton theorem.

We now give a message-passing interpretation of graph symmetric systems. First, it can be seen from the sparsity pattern of the GMD S that the output of $Sx(t)$ can be computed entirely as a linear combination of the states in nodes 1 hop away in \mathcal{G} . To see this, consider the operation $Sx(t)$ whose i^{th} entry yields

$$[Sx(t)]_i = \sum_{j:(v_j, v_i) \in \mathcal{E}} [S]_{ij} [x(t)]_j. \quad (8)$$

More generally, when considering polynomials, it is observed that $S^k x(t)$ is equivalent to exchanging k times information with one-hop neighbors. Therefore, if the system matrix A and the control matrix B are polynomials of S , then the evolution of the system can be computed entirely by means of exchanges with neighboring nodes. Hence, the system dynamics can be viewed as implementing distributed message passing [9]. Examples of such linear, distributed systems, include both discrete-time and continuous-time diffusions, solutions to the heat equation, among many others, see [15] and references therein.

For the rest of the paper, we also assume that the cost matrices for the LQR problem (4) can also be simultaneously diagonalized with the dynamics matrices. Formally, we make the following assumption.

Assumption 1. The system (2) defines a graph symmetric system with respect to a fixed GMD S , and the cost matrices (Q, R) defining the LQR problem (4) are graph symmetric with respect to S , i.e., they are simultaneously diagonalized by the ortho-bases V satisfying $S = V\Lambda_S V^\top$. In particular

$$Q = V\Lambda_Q V^\top, \quad R = V\Lambda_R V^\top,$$

where Λ_Q, Λ_R are symmetric.

III. OPTIMAL DISTRIBUTED LINEAR CONTROLLER VIA SYSTEM LEVEL SYNTHESIS

SLS provides a convex parameterization of achievable closed-loop system responses [5], [17], which can be leveraged to show that the optimal controller for graph symmetric systems under Assumption 1 can be written as a novel class of graph filters defined by transfer function valued filter taps.

A. Background: System-Level Synthesis

As noted in §4 of [5], we can compactly write the system dynamics (2) in the frequency domain as

$$(zI - A)\mathbf{x} = B\mathbf{u} + \mathbf{w},$$

where $\mathbf{x} = \sum_{t=0}^{\infty} z^{-t} x(t)$ is the signal $x(t)$ in the z -domain, and idem for \mathbf{u} and \mathbf{w} . For a (dynamic) linear state-feedback controller $\mathbf{u} = \mathbf{K}\mathbf{x}$, it follows immediately that

$$\begin{aligned} \mathbf{x} &= (zI - A - B\mathbf{K})^{-1} \mathbf{w} =: \Phi_x(z) \mathbf{w}, \\ \mathbf{u} &= \mathbf{K}(zI - A - B\mathbf{K})^{-1} \mathbf{w} =: \Phi_u(z) \mathbf{w}, \end{aligned} \quad (9)$$

where $\Phi_x(z) \in \mathbb{C}^{N \times N}$ and $\Phi_u(z) \in \mathbb{C}^{N \times N}$ are system responses that map the disturbance \mathbf{w} to state \mathbf{x} and control input \mathbf{u} , respectively. The following SLS theorem states that

all achievable responses lie in an affine subspace of strictly proper stable rational transfer functions $\frac{1}{z} \mathcal{RH}_\infty$.

Theorem 1. [5, Thm. 4.1] *For the LTI system evolving under the dynamics (2) and control policy $\mathbf{u} = \mathbf{K}\mathbf{x}$, the following statements are true:*

1) *The affine subspace defined by*

$$[zI - A \quad -B] \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} = I, \quad \Phi_x, \Phi_u \in \frac{1}{z} \mathcal{RH}_\infty \quad (10)$$

parameterizes all system responses from \mathbf{w} to (\mathbf{x}, \mathbf{u}) as defined in (9), achievable by an internally stabilizing state feedback controller \mathbf{K} .

2) *For any transfer matrices Φ_x, Φ_u satisfying (10), the controller $\mathbf{K} = \Phi_u \Phi_x^{-1}$ is internally stabilizing and achieves the desired system response in (9).*

For disturbance $w(t) \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, I)$, one can recast the optimization problem (4) in terms of system responses Φ_x and Φ_u as

$$\begin{aligned} \text{minimize}_{\Phi_x, \Phi_u} \quad & \left\| \mathbf{Q}^{1/2} \Phi_x \right\|_{\mathcal{H}_2}^2 + \left\| \mathbf{R}^{1/2} \Phi_u \right\|_{\mathcal{H}_2}^2 \\ \text{s.t.} \quad & \text{constraint (10)}. \end{aligned} \quad (11)$$

With a slight abuse of notation, we define $J(\Phi_x, \Phi_u)$ to be the LQR cost achieved by Φ_x and Φ_u in the objective of (11) and $J(\mathbf{K})$ as the cost (3) achieved by applying controller \mathbf{K} .

B. SLS for Graph Symmetric Systems

We now proceed to show that under Assumption 1, the optimal system response for a graph symmetric system that solves the LQR problem (11) can be written as a graph filter.

Theorem 2. *Given a GMD $S = V\Lambda_S V^\top$, consider an instance of the LQR problem (11) where the underlying system and cost satisfy Assumption 1. Then, there exists a global optimum (Φ_x^*, Φ_u^*) where both Φ_x^* and Φ_u^* are diagonalizable by V , i.e.,*

$$\Phi_x^* = V\Lambda_x^* V^\top, \quad \Phi_u^* = V\Lambda_u^* V^\top,$$

where Λ_x^* and Λ_u^* are diagonal transfer matrices. Hence, the optimal controller $\mathbf{K}^* = (\Phi_u^*)^{-1} (\Phi_x^*)$ can also be diagonalized by V .

Proof. See Appendix. \square

Remark 1. Note that the elements defined by the diagonal responses Λ_x^*, Λ_u^* are transfer functions $[\Lambda_x^*]_{ii}(z), [\Lambda_u^*]_{ii}(z)$. Thus the resulting graph filter taps are transfer functions as well, i.e., a transfer function $\Phi(z)$ that is simultaneously diagonalizable with the matrix S can be written as:

$$\Phi(z) = \sum_{k=0}^{N-1} \phi_k(z) S^k. \quad (12)$$

The main implication of Theorem 2 is that the optimal linear state-feedback controller for graph symmetric systems under Assumption 1 is a graph filter and can thus be implemented via distributed message passing. We note, however,

that the above result implies that the resulting optimal system response $\Phi^* := (\Phi_x^*, \Phi_u^*)$ could be dense, as S^{N-1} is dense if S defines a connected graph. This can be undesirable in practice, as it requires $N - 1$ communication exchanges with one-hop neighbors, potentially causing significant delays if the size N of the graph is large. In the next section, we leverage a robust variant of the SLS parameterization given in Theorem 1 to restrict the optimal system responses to only the first $F \ll N - 1$ filter taps while guaranteeing stability and near optimal performance.

We end this section by noting that in the graph signal processing literature [14], a controller of the form of (12) is known as a linear, shift-invariant (LSI) graph filter, and is analogous to an LTI filter. Note that $S\Phi^*x = \Phi^*Sx$, hence the name. In particular, Equation (12) is a spatially finite impulse response (FIR) graph filter [16] that is completely characterized by a finite set of N filter taps that can be conveniently described by a collection of transfer functions $\phi^* = [\phi_0^*, \dots, \phi_{N-1}^*]^\top \in \mathbb{R}^N$. We emphasize that the transfer functions themselves are not restricted to be temporally FIR. Spatially FIR graph filters are also known as convolutional graph filters [9] due to their sum-and-shift nature, understanding that the effect of the operation Sx is to shift the signal around the graph (thus, oftentimes, the GMD S is referred to as the graph shift operator). Furthermore, spatially FIR graph filters satisfy the convolution theorem that indicates that a convolution in the vertex domain can be computed by means of an elementwise multiplication in the spectrum domain [14]. Finally, in the context of finite graphs, it is observed that the space of FIR graph filters of the form (12), characterized by N filter taps, is equivalent to the space of spatially infinite impulse response (IIR) graph filters as well as autoregressive, moving average (ARMA) graph filters [18].

IV. LOCALIZED APPROXIMATIONS TO THE OPTIMAL DISTRIBUTED LINEAR CONTROLLER

In this section, we discuss several methods to approximate the optimal dense system response in the form of (12) with one that is localized and uses $F \ll N$ filter taps. We start with a projection method based on graph filter design. We then present a robust SLS formulation of the approximation problem that guarantees the stability of the resulting localized controller. Finally, we show how these two can be combined into a robust projection method that also ensures stability. In the following, we define $\Phi^* := (\Phi_x^*, \Phi_u^*)$, and recall that Φ^* can be written as a graph filter (12) defined by transfer function filter taps $\phi_k^*(z)$, $k = 0, \dots, N - 1$. We further recall that each transfer function filter tap $\phi_k^*(z)$ admits the following expansion in terms of its Markov parameters: $\phi_k^*(z) = \sum_{i=1}^{\infty} z^{-i} \phi_k^*[i]$.

A. Naive Projection

We propose an approach inspired by the graph signal processing literature, wherein we exploit the graph filter structure of the optimal system responses [16]. More specifically, we project the optimal system responses Φ^* onto graph

filters of order F in the \mathcal{H}_2 norm by solving the following optimization problem

$$\min_{\phi} \|\Phi - \Phi^*\|_{\mathcal{H}_2}^2. \quad (13)$$

Here $\phi := [\phi_0(z), \dots, \phi_{F-1}(z)]^\top \in \mathbb{C}^F$ collects the F transfer function filter taps defining $\Phi := \sum_{k=0}^{F-1} \phi_k(z)S^k$. If we further restrict each transfer function filter tap $\phi_k(z)$ to be FIR of order n , i.e., if we write $\phi_k(z) = \sum_{i=1}^n z^{-i} \phi_k[i]$, then this reduces to solving the following unconstrained quadratic program

$$\min_{\{\phi_k[i]\}} \sum_{i=1}^n \left\| \sum_{k=0}^{F-1} \phi_k[i]S^k - \sum_{l=0}^{N-1} \phi_l^*[i]S^l \right\|_{\mathbb{R}}^2. \quad (14)$$

Proposition 3 (Approximating filter taps). *If the eigenvalues $\{\lambda_i\}$ of the GMD S are all distinct, then the filter taps that solve (13) are given by*

$$\phi_k[i] = \phi_k^*[i] + \varepsilon_k[i] \quad (15)$$

where ε is the error vector computed as

$$\varepsilon[j] = \left(\sum_{i=1}^N \lambda_{iF} \lambda_{iF}^\top \right)^{-1} \left(\sum_{i=1}^N \lambda_{iF} \lambda_{i(N-F)}^\top \right) \phi_{N-F}^*[j] \quad (16)$$

with $\lambda_{iF} := [1, \lambda_i, \lambda_i^2, \dots, \lambda_i^{F-1}] \in \mathbb{R}^F$ is the collection of the first F powers of λ_i , $\lambda_{i(N-F)} := [\lambda_i^F, \dots, \lambda_i^{N-1}] \in \mathbb{R}^{N-F}$ is the collection of the remaining powers and $\phi_{N-F}^*[i] := [\phi_F^*[i], \dots, \phi_{N-1}^*[i]] \in \mathbb{R}^{N-F}$ collects the tail $N - F$ optimal filter taps.

Proof. It follows from using the convexity of (14), matrix calculus and properties of Vandermonde matrices. \square

Prop. 3 determines in closed-form how to compute the filter of order F that best approximates the optimal linear distributed controller in the \mathcal{H}_2 norm. It also shows that each filter tap transfer function $\phi_k(z)$ is obtained as the optimal tap $\phi_k^*(z)$ with an added corrective term that accounts for the $N - F$ taps that could not be included.

This approach is easy to implement computationally, as it only requires solving a least squares problem to minimize the projection cost. However, the resulting controller cannot be guaranteed to be stabilizing. As we show later via numerical simulation, approximations with a small number of filter taps F are often unstable. This motivates an approach that takes into account the stability of the resulting controller.

B. Localized Approximations via Robust SLS

Robust SLS [4], [5] offers a systematic way to reason about *approximate system responses*, i.e., system responses that do not exactly satisfy the achievability constraint (10). In particular, as shown in the following result, robust SLS allows for an explicit characterization of the effects of using approximate system responses for controller design.

Theorem 4 (Corollary 4.4 of [5]). *Let (Φ_x, Φ_u, Δ) be a solution to*

$$[zI - A \quad -B] \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} = I + \Delta, \quad \Phi_x, \Phi_u \in \frac{1}{z} \mathcal{RH}_\infty. \quad (17)$$

Then if $\|\Delta\|_{\mathcal{H}_\infty} < 1$ the controller $\mathbf{K} = \Phi_u \Phi_x^{-1}$ stabilizes the system (2), and the actual system response that is achieved is given by

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} (I + \Delta)^{-1} \mathbf{w}.$$

We leverage this result to provide an upper-bound on the amount of truncation that can be applied to Φ^* without destabilizing the system.

Corollary 5. *Let (Φ_x, Φ_u, Δ) be a solution to (17) and assume that $\|\Delta\|_{\mathcal{H}_\infty} < 1$. Then the controller $\mathbf{K} = \Phi_u \Phi_x^{-1}$ achieves an LQR cost (11) J that can be bounded as*

$$J(\mathbf{K}) \leq \frac{1}{1 - \|\Delta\|_{\mathcal{H}_\infty}} \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} \right\|_{\mathcal{H}_2}.$$

Proof. First, we note that by Theorem 4, the system responses $(\tilde{\Phi}_x, \tilde{\Phi}_u)$ achieved by \mathbf{K} are given by

$$\begin{bmatrix} \tilde{\Phi}_x \\ \tilde{\Phi}_u \end{bmatrix} = \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} (I + \Delta)^{-1}.$$

Thus, the cost achieved by the controller \mathbf{K} is bounded by

$$\begin{aligned} J(\mathbf{K}) &= J(\tilde{\Phi}_x, \tilde{\Phi}_u) \\ &= \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} (I + \Delta)^{-1} \right\|_{\mathcal{H}_2} \\ &\leq \|(I + \Delta)^{-1}\|_{\mathcal{H}_\infty} \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} \right\|_{\mathcal{H}_2} \\ &\leq \frac{1}{1 - \|\Delta\|_{\mathcal{H}_\infty}} \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} \right\|_{\mathcal{H}_2}, \end{aligned}$$

where we used Cauchy-Schwarz in the first inequality and the small gain theorem and the fact that $\|\Delta\|_{\mathcal{H}_\infty} < 1$ in the last step. \square

Corollary 5 offers a way to synthesize *stable* truncated system responses. Specifically, to synthesize a system response that uses only $F < N - 1$ filter taps while guaranteeing both stability and performance, we propose the following optimization problem:

$$\underset{\phi_x, \phi_u, \gamma \in (0,1)}{\text{minimize}} \quad \frac{1}{1 - \gamma} \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} \right\|_{\mathcal{H}_2} \quad (18a)$$

$$\text{s.t.} \quad [zI - A \quad -B] \begin{bmatrix} \Phi_x \\ \Phi_u \end{bmatrix} = I + \Delta,$$

$$\|\Delta\|_{\mathcal{H}_\infty} \leq \gamma,$$

$$\Phi_x(z) = \sum_{k=0}^{F-1} \phi_{x,k}(z) S^k, \quad \phi_{x,k}(z) \in \frac{1}{z} \mathcal{RH}_\infty,$$

$$\Phi_u(z) = \sum_{k=0}^{F-1} \phi_{u,k}(z) S^k, \quad \phi_{u,k}(z) \in \frac{1}{z} \mathcal{RH}_\infty. \quad (18b)$$

By Corollary 5, the solution to (18) defines a controller that is stabilizing. We further show in the next result that it enjoys guaranteed suboptimality bounds relative to the optimal controller defined by Φ^* . We first introduce the following notation: for a system response of the form $\Phi = \sum_{k=0}^{N-1} \phi_k(z) S^k$, we define the F -truncation $P_F(\Phi)$ and the F -tail $P_{F\perp}(\Phi)$ as

$$P_F(\Phi) := \sum_{k=0}^{F-1} \phi_k(z) S^k, \quad P_{F\perp}(\Phi) := \sum_{k=F}^{N-1} \phi_k(z) S^k.$$

Theorem 6. *Let $(\tilde{\Phi}_x, \tilde{\Phi}_u, \tilde{\gamma})$ be the optimal solution to the robust SLS problem (18). Let (Φ_x^*, Φ_u^*) be the optimal solution to the untruncated SLS problem (11). Suppose that*

$$\|\Delta^*\|_{\mathcal{H}_\infty} := \|(zI - A)P_{F\perp}(\Phi_x^*) - BP_{F\perp}(\Phi_u^*)\|_{\mathcal{H}_\infty} < 1.$$

Then, the controller $\mathbf{K} = \tilde{\Phi}_u \tilde{\Phi}_x^{-1}$ is stabilizing and

$$J(\mathbf{K}) \leq \frac{1}{1 - \|\Delta^*\|_{\mathcal{H}_\infty}} \left(J(\Phi_x^*, \Phi_u^*) + J(P_{F\perp}(\Phi_x^*), P_{F\perp}(\Phi_u^*)) \right) \quad (19)$$

Proof. By the constraints (18b) and Theorem 4, we immediately have that \mathbf{K} is stabilizing. To show the given suboptimality bound, we first note that there exist some γ such that $(P_F(\Phi_x^*), P_F(\Phi_u^*), \gamma)$ is a feasible solution to the robust optimization problem (18). To see this, observe that

$$\begin{aligned} & [zI - A \quad -B] \begin{bmatrix} P_F(\Phi_x^*) \\ P_F(\Phi_u^*) \end{bmatrix} \\ &= [zI - A \quad -B] \begin{bmatrix} \Phi_x^* \\ \Phi_u^* \end{bmatrix} - [zI - A \quad -B] \begin{bmatrix} P_{F\perp}(\Phi_x^*) \\ P_{F\perp}(\Phi_u^*) \end{bmatrix} \\ &= I - \Delta^*, \end{aligned}$$

where the last step follows from the achievability of the optimal response (Φ_x^*, Φ_u^*) and the definition of Δ^* . Thus, $(P_F(\Phi_x^*), P_F(\Phi_u^*), \|\Delta^*\|_{\mathcal{H}_\infty})$ is a feasible solution with our assumption that $\|\Delta^*\|_{\mathcal{H}_\infty} < 1$. Denote the robust SLS objective (18) as $\tilde{J}(\Phi_x, \Phi_u, \gamma)$. By the optimality of the solution $(\tilde{\Phi}_x, \tilde{\Phi}_u, \tilde{\gamma})$, we have that

$$\begin{aligned} \tilde{J}(\tilde{\Phi}_x, \tilde{\Phi}_u, \tilde{\gamma}) &\leq \tilde{J}(\Phi_x^*, \Phi_u^*, \|\Delta^*\|_{\mathcal{H}_\infty}) \\ &\leq \frac{1}{1 - \|\Delta^*\|_{\mathcal{H}_\infty}} \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} P_F(\Phi_x^*) \\ P_F(\Phi_u^*) \end{bmatrix} \right\|_{\mathcal{H}_2}, \end{aligned}$$

where we applied Corollary 5 in the second inequality. The desired result follows then from the fact that

$$\begin{bmatrix} P_F(\Phi_x^*) \\ P_F(\Phi_u^*) \end{bmatrix} = \begin{bmatrix} \Phi_x^* \\ \Phi_u^* \end{bmatrix} - \begin{bmatrix} P_{F\perp}(\Phi_x^*) \\ P_{F\perp}(\Phi_u^*) \end{bmatrix}$$

and an application of the triangle inequality. \square

This optimization problem is jointly quasi-convex and can be solved efficiently using bisection. Further, feasibility provides a stability certificate in the form of $\|\Delta\|_{\mathcal{H}_\infty} < 1$.

C. Robust Projection

Lastly, we can combine the robustness constraints used in robust SLS with the signal-processing-based projection method. Specifically, we solve the following optimization problem

$$\begin{aligned} & \underset{\Phi_x, \Phi_u, \gamma \in (0,1)}{\text{minimize}} && \|\Phi - \Phi^*\|_{\mathcal{H}_2}^2 \\ & \text{s.t.} && (\Phi_x, \Phi_u, \gamma) \text{ satisfy constraint (18b)}. \end{aligned} \quad (20)$$

We note that solving this problem does not give an upper bound on the cost of the resulting controller, but the robustness constraint ensures that the resulting controller is stabilizing.

D. Implementation

We end this section by detailing practical implementation details for optimization problems (18) and (20). First, we note that computationally, one cannot directly optimize for the IIR system responses as is written in (18), (20). In practice, we use an FIR approximation of the strictly proper transfer functions $\phi_x(z)$ and $\phi_u(z)$, i.e., $\phi(z)$ is parameterized as

$$\phi(z) = \sum_{i=1}^n z^{-i} \phi[i]$$

for some given FIR order n . As shown in [5], the suboptimality incurred by such an FIR approximation decays exponentially in the horizon n .

The \mathcal{H}_∞ -norm constraints on the (also FIR) transfer matrix Δ can then be enforced via semidefinite programming (see Theorem 5.8 in [19]), potentially introducing a nontrivial computational burden. However, we note that one can replace the \mathcal{H}_∞ -norm constraint in optimization problems (18) and (20) with any induced norm constraint. A particularly appealing option is the $\ell_1 \rightarrow \ell_1$ induced norm (which defines the \mathcal{L}_1 -norm of the transpose system), as this norm decomposes columnwise. As shown in [5], [17], the resulting robustness constraints are linear and embarrassingly parallelizable. We defer this extension to future work.

V. NUMERICAL EXPERIMENTS

We show that our approach offers a principled way to trade off performance and communication complexity through numerical experiments. We also demonstrate the importance of the robustness constraints in synthesizing stable distributed controllers and compare the performance of robust SLS and projection-based methods on synthesizing localized controllers. All code needed to reproduce the examples found in this section is available at <https://github.com/unstable-zeros/graph-symmetric-systems>.

A. Setup

In the following experiments, we consider the distributed linear quadratic regulator (LQR) problem (4) over $N = 10$ scalar subsystems. We generate the GMD S and dynamic matrices A and B using a process similar to that in [11]. To generate a problem instance, we start by creating the communication network \mathcal{G} by randomly sampling N numbers

$\{u_i\}_{i=1}^N \sim U[0, 1]$, and creating a bi-directional link between v_i and each of its 3 nearest points as defined by the topology on the interval $[0, 1]$ under the metric $d(v_i, v_j) = |u_i - u_j|$. We then take S to be a symmetric matrix that shares the sparsity pattern of the Laplacian of \mathcal{G} , with its entry values sampled independently from $\mathcal{N}(0, 1)$. The GMD S is then normalized to have a spectral radius of 1. We generate the dynamics matrices A and B to share the same eigenvectors as S , and sample their eigenvalues i.i.d. from the standard normal distribution – hence both A and B are symmetric matrices. We take the cost matrices $Q = R = I_N$. For both of the following experiments, we randomly generate 50 problem instances using this process. We end by noting that \mathcal{G} generated this way have, on average, a diameter of 5.92 hops.

For the implementation of the optimization problems, we approximate the transfer functions $\phi(z)$ with an FIR horizon of $n = 10$. Further, for the robust SLS problem (18), instead of using bisection to determine the best value of γ , we fix $\gamma = 0.98$, as empirically the value of γ does not significantly affect the cost achieved by the controllers.

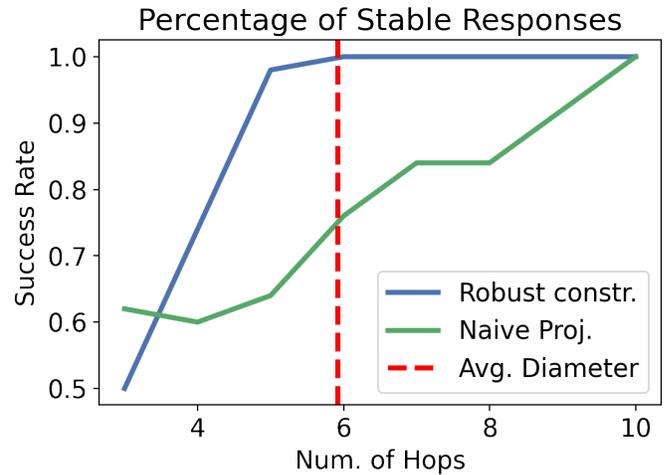


Fig. 1. The percentage of synthesized F -hop controllers that are stabilizing using naive projection and robust SLS-based synthesis across 50 random trials. The dashed red line denotes the average graph diameter of systems.

B. Importance of Stability Constraints

In this experiment, we demonstrate the importance of the robust SLS-based stability constraints in synthesizing stable distributed controllers. We vary the number of allowed filter taps and apply both the naive projection (13) and robust SLS (20) methods to 50 randomly generated problem instances. For the naive projection method, we report the percentage of resulting controllers that are stable. For robust SLS, we report the percentage of optimization problems (18) being feasible, as feasible solutions optimization problem (18) are guaranteed to be stabilizing. The results are shown in Figure 1.

First, we observe that as expected, a higher number of filter taps result in a higher probability of synthesizing stable

responses for both methods. However, the naive projection method has nonzero probability of resulting in an unstable controllers even with a large number of filter taps i.e., even when the projection error between the *stable* optimal system responses and the projected responses is small. On the other hand, the percentage of stable solutions resulting from the robust SLS problem increases monotonically with the number of filter taps, which is expected for a principled way of synthesizing stable controllers. Robust SLS also generally achieves a higher percentage of certifiably stable responses than that of naive projection, except for in the extremely sparse case of $F = 3$. This suggests that the robust constraint might be too restrictive for synthesizing extremely sparse responses. Combined with the low computation cost of naive projection, this suggests a potential benefit of applying both methods in the sparse regime.

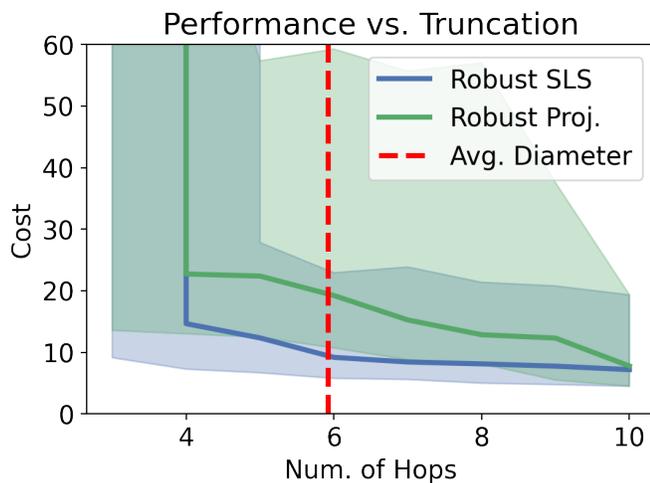


Fig. 2. Cost achieved by F -hop controllers across 50 randomly generated systems. Solid lines denote the median cost achieved by each method, and the shaded regions show the 25-th and 75-th percentile of the costs.

C. Truncation Performance

In this experiment, we compare the performance of robust SLS and robust projection for different filter tap numbers F on the same 50 randomly generated problem instances. We show the median (solid lines), 25-th and 75-th percentile (shaded regions) of the costs achieved by both methods in Figure 2.

First, we note that the median costs decreases monotonically for both methods as the number of hops increase. This shows that that the optimization problems can leverage the increase in expressivity of the graph filters to achieve better performance, which matches our intuition. Second, we note that the robust SLS-based method achieves a lower cost than robust Projection over for all numbers of filter taps considered. We also note that to the left of 4 hops, the upper boundary of the shaded region, which represents the 75-th percentile of the cost, is infinite, indicating that at least 25% of the robust synthesis problems are infeasible. This again suggests a need to develop more flexible methods in the sparse regime.

VI. CONCLUSION

In this works, we introduced the notion of graph symmetric systems and showed that for linear quadratic problems, the optimal system response for graph symmetric systems can be written as (potentially dense) graph filters. We then proposed three methods to approximate the optimal responses with localized responses and validated their performance in numerical simulation. Directions of future work include relaxing the GSS constraints, applying the results on \mathcal{L}_1 norm to enable distributed computation, and understanding how this can better inform GNN-based controllers with nonlinear activation functions.

REFERENCES

- [1] Hans S Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.
- [2] Bassam Bamieh, Fernando Paganini, and Munther A Dahleh, "Distributed control of spatially invariant systems," *IEEE Transactions on automatic control*, vol. 47, no. 7, pp. 1091–1107, 2002.
- [3] Paolo Massioni and Michel Verhaegen, "Distributed control for identical dynamically coupled systems: A decomposition approach," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 124–135, 2009.
- [4] Nikolai Matni, Yuh-Shyang Wang, and James Anderson, "Scalable system level synthesis for virtually localizable systems," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 3473–3480.
- [5] James Anderson, John C. Doyle, Steven H. Low, and Nikolai Matni, "System level synthesis," *Annu. Rev. Control*, vol. 47, pp. 364 – 393, 2019.
- [6] Michael Rotkowitz and Sanjay Lall, "A characterization of convex problems in decentralized control," *IEEE transactions on Automatic Control*, vol. 50, no. 12, pp. 1984–1996, 2005.
- [7] Bassam Bamieh and Petros G Voulgaris, "A convex characterization of distributed control problems in spatially invariant systems with communication constraints," *Systems & control letters*, vol. 54, no. 6, pp. 575–583, 2005.
- [8] Luca Frieri, Yang Zheng, Antonis Papachristodoulou, and Maryam Kamgarpour, "An input–output parametrization of stabilizing controllers: Amidst youla and system level synthesis," *IEEE Control Systems Letters*, vol. 3, no. 4, pp. 1014–1019, 2019.
- [9] Luana Ruiz, Fernando Gama, and Alejandro Ribeiro, "Graph neural networks: Architectures, stability and transferability," *Proc. IEEE*, vol. 109, no. 5, pp. 660–682, May 2021.
- [10] Fernando Gama, Joan Bruna, and Alejandro Ribeiro, "Stability properties of graph neural networks," *IEEE Trans. Signal Process.*, vol. 68, pp. 5680–5695, 25 Sep. 2020.
- [11] Fernando Gama and Somayeh Sojoudi, "Distributed linear-quadratic control with graph neural networks," *Signal Process.*, vol. 196, no. 108506, pp. 1–14, July 2022.
- [12] Fernando Gama, Qingbiao Li, Ekaterina Tolstaya, Amanda Prorok, and Alejandro Ribeiro, "Synthesizing decentralized controllers with graph neural networks and imitation learning," *arXiv:2012.14906v4 [cs.LG]*, 23 March 2022, to appear in IEEE Trans. Signal Process.
- [13] Fengjun Yang and Nikolai Matni, "Communication topology co-design in graph recurrent neural network based distributed control," in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 3619–3626.
- [14] Aliaksei Sandryhaila and José M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, 11 Jan. 2013.
- [15] Fernando Gama, Elvin Isufi, Alejandro Ribeiro, and Geert Leus, "Controllability of bandlimited graph processes over random time varying graphs," *IEEE Trans. Signal Process.*, vol. 67, no. 24, pp. 6440–6454, 6 Nov. 2019.
- [16] Santiago Segarra, Antonio G. Marques, and Alejandro Ribeiro, "Optimal graph-filter design and applications to distributed linear networks operators," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4117–4131, 11 May 2017.

- [17] Y. Wang, N. Matni, and J. C. Doyle, "Separable and localized system-level synthesis for large-scale systems," *IEEE Trans. Autom. Control*, vol. 63, no. 12, pp. 4234–4249, Dec. 2018.
- [18] Elvin Isufi, Andreas Loukas, Andrea Simonetto, and Geert Leus, "Autoregressive moving average graph filtering," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 274–288, 3 Oct. 2016.
- [19] Bogdan Dumitrescu, *Positive trigonometric polynomials and signal processing applications*, vol. 103, Springer, 2007.

APPENDIX

A. Proof for Theorem 2

We show that for any optimal system response Φ_x^* , Φ_u^* of the optimization problem (11) that is not diagonalizable with V , i.e.,

$$\Lambda_x^* = V^\top \Phi_x^* V, \quad \Lambda_u^* = V^\top \Phi_u^* V$$

are not diagonal, we can construct a simultaneously diagonalizable system response Φ'_x , Φ'_u that is equally optimal. In particular, we construct such a system response as follows:

$$\begin{aligned} \Phi'_x &= V \Lambda'_x V^\top, & \Phi'_u &= V \Lambda'_u V^\top \\ [\Lambda'_x]_{ij} &= \begin{cases} [\Lambda_x^*]_{ij} & i = j \\ 0, & \text{o.w.} \end{cases}, & [\Lambda'_u]_{ij} &= \begin{cases} [\Lambda_u^*]_{ij} & i = j \\ 0 & \text{o.w.} \end{cases}. \end{aligned} \quad (21)$$

$$(22)$$

We first show that Φ'_x , Φ'_u are feasible solutions of (11). From the achievability condition on Φ_x^* , Φ_u^* , we have that

$$(zI - A)\Phi_x^* - B\Phi_u^* = I. \quad (23)$$

Using the simultaneous diagonalizability of A and B , we have

$$\begin{aligned} V(zI - \Lambda_A)V^\top V \Lambda_x^* V^\top - V \Lambda_B V^\top V \Lambda_u^* V^\top &= I \\ \implies (zI - \Lambda_A)\Lambda_x^* - \Lambda_B \Lambda_u^* &= I. \end{aligned} \quad (24)$$

Since the matrices $(zI - \Lambda_A)$, Λ_B and I are diagonal, we have

$$(zI - \Lambda_A)\Lambda'_x - \Lambda_B \Lambda'_u = \text{diag}[(zI - \Lambda_A)\Lambda_x^* - \Lambda_B \Lambda_u^*] = I, \quad (25)$$

where $\text{diag}[\mathbf{M}]$ projects a matrix \mathbf{M} onto its diagonal elements. Therefore, Φ'_x , Φ'_u is also feasible.

Now, we show that Φ'_x , Φ'_u gives a cost at least as good as that of Φ_x^* , Φ_u^* . By the simultaneous diagonalizability of the matrices Q and R , and the fact that the \mathcal{H}_2 -norm is invariant under unitary transformations, we have that

$$\begin{aligned} &\|Q^{1/2}\Phi'_x\|_{\mathcal{H}_2}^2 + \|R^{1/2}\Phi'_u\|_{\mathcal{H}_2}^2 \\ &= \left\| V \Lambda_Q^{1/2} V^\top V \Lambda'_x V^\top \right\|_{\mathcal{H}_2} + \left\| V \Lambda_R^{1/2} V^\top V \Lambda'_u V^\top \right\|_{\mathcal{H}_2} \\ &= \left\| \Lambda_Q^{1/2} \Lambda'_x \right\|_{\mathcal{H}_2}^2 + \left\| \Lambda_R^{1/2} \Lambda'_u \right\|_{\mathcal{H}_2}^2. \end{aligned} \quad (26)$$

Denoting the i -th eigenvalue of \mathbf{Q} and \mathbf{R} with $\lambda_{Q,i}$, $\lambda_{R,i}$, respectively, we have the inequality

$$\begin{aligned} &\left\| \Lambda_Q^{1/2} \Lambda'_x \right\|_{\mathcal{H}_2}^2 + \left\| \Lambda_R^{1/2} \Lambda'_u \right\|_{\mathcal{H}_2}^2 \\ &= \sum_{i,j} \lambda_{Q,i} \|[\Lambda'_x]_{ji}\|_{\mathcal{H}_2}^2 + \sum_{i,j} \lambda_{R,i} \|[\Lambda'_u]_{ji}\|_{\mathcal{H}_2}^2 \\ &\leq \sum_{i,j} \lambda_{Q,i} \|[\Lambda_x^*]_{ji}\|_{\mathcal{H}_2}^2 + \sum_{i,j} \lambda_{R,i} \|[\Lambda_u^*]_{ji}\|_{\mathcal{H}_2}^2 \\ &= \left\| \Lambda_Q^{1/2} \Lambda_x^* \right\|_{\mathcal{H}_2}^2 + \left\| \Lambda_R^{1/2} \Lambda_u^* \right\|_{\mathcal{H}_2}^2, \end{aligned} \quad (27)$$

which follows from the definition of Λ'_x and Λ'_u in equation (22). Reversing the steps in (26), we see that Φ'_x , Φ'_u achieves a cost at least as good as that of Φ_x^* , Φ_u^* . We can thus conclude that there always exists an optimal simultaneously diagonalizable system response to the LQR problem (11). The controller $\mathbf{K}' = (\Phi'_u)(\Phi'_x)^{-1}$ is thus optimal and simultaneously diagonalizable by V .

B. GSS and Controllers Satisfy Quadratic Invariance

Here we show that optimal control problems over graph symmetric systems and controllers satisfy quadratic invariance [6]. Before proceeding, we remark that the analysis of LQR optimal control problem over GSSs does not require quadratic invariance. In particular, in Theorem 2 we analyze the *unconstrained optimal control problem* and show that the resulting *unconstrained optimal controller* satisfies a corresponding notion of graph symmetry. However, in the interest of completeness, we show that if such a constraint were imposed on the controller during synthesis, the resulting problem satisfies quadratic invariance.

To that end, the corresponding constrained controller synthesis problem can be stated as

$$\begin{aligned} &\underset{\mathbf{K}}{\text{minimize}} && J(\mathbf{K}) \\ &\text{subject to} && \mathbf{K} \text{ stabilizes system (2)} \\ &&& \mathbf{K} \in \mathcal{S} \end{aligned}$$

where $\mathcal{S} := \{S \in \mathcal{RH}_\infty \mid S \text{ is diagonalizable by } V\}$. Denoting the plant input-output transfer function as

$$\mathbf{G}(z) = (zI - A)^{-1}B,$$

we have the following proposition.

Proposition 7. *The set of graph symmetric controllers \mathcal{S} is quadratically invariant under \mathbf{G} if system (2) is graph symmetric.*

Proof. The proof follows directly from the definition of quadratic invariance (see [6, Def. 2]) by straightforward calculation. First, we note that

$$\mathbf{G}(z) = V(zI - \Lambda_A)^{-1} \Lambda_B V^\top$$

is diagonalizable by V . For any controller $\mathbf{K} \in \mathcal{S}$, it then follows immediately that

$$\mathbf{K} \mathbf{G} \mathbf{K} \in \mathcal{S}$$

as the product of simultaneously diagonalizable matrices is also simultaneously diagonalizable, proving the claim. \square