Abstraction of Linear Consensus Networks with Guaranteed Systemic Performance Measures

Milad Siami[†] and Nader Motee[⋆]

Abstract—A proper abstraction of a large-scale linear consensus network with a dense coupling graph is one whose number of coupling links is proportional to its number of subsystems and its performance is comparable to the original network. Optimal design problems for an abstracted network are more amenable to efficient optimization algorithms. From the implementation point of view, maintaining such networks are usually more favorable and cost effective due to their reduced communication requirements across a network. Therefore, approximating a given dense linear consensus network by a suitable abstract network is an important analysis and synthesis problem. In this paper, we develop a framework to compute an abstraction of a given large-scale linear consensus network with guaranteed performance bounds using a nearly-linear time algorithm. First, the existence of abstractions of a given network is proven. Then, we present an efficient and fast algorithm for computing a proper abstraction of a given network. Finally, we illustrate the effectiveness of our theoretical findings via several numerical simulations.

I. INTRODUCTION

Reducing design complexity in interconnected networks of dynamical systems by means of abstraction are central in several real-world applications [1]–[6]. Various notions of abstractions for dynamical systems have been widely used by researchers in the context of control systems in past decades, see [7]–[10] and references in there, where the notion of *reduction* mainly implies projecting dynamics of a system to lower dimensional state spaces. In this paper, we employ a relevant notion of abstraction in the context of interconnected dynamical network: for a given dynamical network that is defined over a coupling graph, find another dynamical system whose coupling graph is significantly sparser and its performance quality remains close to that of the original network. In this definition, abstraction can be regarded as a notion of network reduction. There are several valid reasons why reduction in this sense is useful in design, maintenance, and implementation of dynamical networks. Real-time generation of state estimation in large-scale dynamical networks can be done much more efficiently and faster if proper abstractions are utilized. Optimal control problems that involve controller design, feedback gain adjustments, rewiring existing feedback loops, and etc. are more amenable to efficient computational tools that are specifically tailored for optimization problems with sparse structures. In security- or privacy-sensitive applications such as formation control of group of autonomous drones, it is usually required to minimize communication requirements across the network to reduce risk of external intrusions. In power network applications, network authorities periodically provide access to their network data and parameters for academic (or public) studies and evaluations. In order to reduce possibility of planned malicious attacks, network authorities can perform abstractions in order to hide actual values of parameters in their networks by preserving all other important characteristics of the network that interest researchers.

The goal of this paper is to address the abstraction problem for the class of linear consensus networks. In [1], we introduce a class of operators, so called systemic performance measure, for linear consensus networks that provides a unified framework for network-wide performance assessment. Several existing and popular performance measures in the literature, such as $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms of a consensus network from a disturbance input to its output, are examples of systemic performance measures. This class of operators is obtained through our close examination of functional properties of several existing gold standard measures of performance in the context of network engineering and science. An important contribution of this reference paper is that it enables us to optimize performance of a consensus network solely based on its intrinsic features. The authors formulate several optimal design problems, such as weight adjustment as well as rewiring of coupling links, with respect to this general class of systemic performance measures and propose efficient algorithms to solve them. In [11], [12], we quantify several fundamental tradeoffs between a $\mathcal{H}_2$-based performance measure and sparsity measures of a linear consensus network. The problem of sparse consensus network design has been considered before in [5], [13]–[15], where they formulate an ℓ_0-regularized $\mathcal{H}_2$ optimal control problem. The main common shortcoming of existing works in this area is that they are heavily relied on computational tools with no analytical performance guarantees for the resulting solution. More importantly, the proposed methods in these papers mainly suffer from high computational complexity as network size grows.

For a given linear consensus network with an undirected connected graph, the network abstraction problem seeks to construct a new network with a reasonably sparser graph compared to the original network such that the dynamical behavior of the two networks remains similar in an appropriately defined sense. We develop a methodology that computes abstractions of a given consensus network using a nearly-linear time $\tilde{\mathcal{O}}(m)^1$ algorithm with guaranteed systemic performance bounds, where m is the number of links. Unlike other existing work on this topic in the literature, our proposed framework: (i) works for a broad class of systemic performance measures including $\mathcal{H}_2$-based performance measures, (ii) does not involve any sort of relaxations such as ℓ_0 to ℓ_1,[2] (iii) provides guarantees for the existence of a sparse solution, (iv) can partially sparsify predetermined portions of a given network; and most importantly, (v) gives guaranteed levels of performance. While our approach is relied on several existing works in algebraic graph theory [16], [17], our control theoretic contributions are threefold. First, we show that there exist proper abstractions for every given linear consensus network. Second, we develop a framework to compute a proper abstraction of a network using a fast randomized algorithm. One of the main features of our method is that while the coupling graph of the abstracted network is a subset of the coupling graph of the original network, the link weights (the strength of each coupling) in the sparsified network

† M. Siami is with the Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA 02319. Email: `siami@mit.edu`.
⋆ N. Motee is with the Department of Mechanical Engineering and Mechanics, Packard Laboratory, Lehigh University, Bethlehem, PA 18015. Email: `motee@lehigh.edu`.

---

[1]We use $\tilde{\mathcal{O}}(.)$ to hide poly $\log \log$ terms from the asymptotic bounds. Thus, $f(n) \in \tilde{\mathcal{O}}(g(n))$ means that there exists $c > 0$ such that $f(n) \in \mathcal{O}\left(g(n) \log^c g(n)\right)$.

[2]We discuss some of the shortcomings of the ℓ_0/ℓ_1-regularization based sparsification methods in Section VIII.

are adjusted accordingly to reach predetermined levels of systemic performance. Third, we prove that our method can also be applied for partial abstraction of large-scale networks, which means that we can abstract a prespecified subgraph of the original network. This is practically plausible as our algorithm can obtain an abstraction using only spatially localized information. Moreover, this allows parallel implementation of the abstraction algorithm in order to achieve comparably lower time complexity.

II. NOTATION AND PRELIMINARIES

The set of real, positive real, and strictly positive real numbers are represented by $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{R}_{++}$, respectively. A matrix is generally represented by an upper case letter, say $X = [x_{ij}]$, where x_{ij} is the (i, j) element of matrix X and X^{T} indicates its transposition. We assume that $\mathbb{1}_n$ and I_n denote the $n \times 1$ vector of all ones and the $n \times n$ identity matrix, respectively. The centering matrix is defined by $M_n = I_n - \frac{1}{n}J_n$ in which J_n is the $n \times n$ matrix of all ones. Notation $X \succeq Y$ is equivalent to matrix $X - Y$ being positive semi-definite. A graph is represented by $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$, where $\mathcal{V}$ is the set of nodes, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of links, and $w : \mathcal{V} \times \mathcal{V} \to \mathbb{R}_+$ is the weight function. The value of the weight function is zero for $e \in \mathcal{V} \times \mathcal{V} \setminus \mathcal{E}$ and positive for $e \in \mathcal{E}$. The weighted degree of node $i \in \mathcal{V}$ is defined by

$$d_i := \sum_{e=\{i,j\} \in \mathcal{E}} w(e). \qquad (1)$$

The neighborhood of node i is denoted by set $\mathcal{N}(i)$ that consists of all adjacent nodes to i and its cardinality $|\mathcal{N}(i)|$ is equal to the number of neighbors of node i. In unweight graphs, $|\mathcal{N}(i)|$ is equal to the degree of node i. The adjacency matrix $A = [a_{ij}]$ of graph $\mathcal{G}$ is defined by setting $a_{ij} = w(e)$ if $e = \{i, j\} \in \mathcal{E}$, and $a_{ij} = 0$ otherwise. The Laplacian matrix of graph $\mathcal{G}$ with n nodes is defined by

$$L := \mathrm{diag}[d_1, \ldots, d_n] - A.$$

A n-by-m oriented incidence matrix $E = [\mathfrak{e}_{ij}]$ for $1 \leq i \leq n$ and $1 \leq j \leq m$ can be formed by assigning an arbitrary direction for every link of $\mathcal{G}$, labeling every link by a number $1 \leq j \leq m$, and letting $\mathfrak{e}_{ij} = 1$ whenever node i is the head of (directed) link j, $\mathfrak{e}_{ij} = -1$ if node i is the tail of (directed) link j, and $\mathfrak{e}_{ij} = 0$ when link j is not attached to node i for all possible orientations of links. The weight matrix $W = [w_{kk}]$ is the m-by-m diagonal matrix with diagonal elements $w_{kk} = w(e_k)$ for $1 \leq k \leq m$. It follows that

$$L = EWE^{\mathrm{T}}.$$

Assumption 1: All graphs in this paper are assumed to be finite, simple, undirected, and connected.

According to this assumption, every considered Laplacian matrix in this paper has exactly $n - 1$ positive eigenvalues and one zero eigenvalue, which allow us to index them in ascending order $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. The set of Laplacian matrices of all connected weighted graphs over n nodes is represented by $\mathfrak{L}_n$. The Moore-Penrose pseudo-inverse of L is denoted by $L^\dagger = [l^\dagger_{ji}]$ which is a square, symmetric, doubly-centered and positive semi-definite matrix. The corresponding resistance matrix $R = [r_{ij}]$ to Laplacian matrix L is defined by setting

$$r_{ij} = l^\dagger_{ii} + l^\dagger_{jj} - 2l^\dagger_{ij}$$

in which r_{ij} is called the effective resistance between nodes i and j. Moreover, we denote the effective resistance of link $e = \{i, j\}$ by $r(e) = r_{ij} = r_{ji}$. The ℓ_0 sparsity measure of matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is defined by

$$\|A\|_{\ell_0} := \mathrm{card}\{(i, j) \mid a_{ij} \neq 0\}. \qquad (2)$$

The $\mathcal{S}_{0,1}$ sparsity measure of matrix A is defined by

$$\|A\|_{\mathcal{S}_{0,1}} := \max\left\{ \max_{1 \leq i \leq n} \|A(i, .)\|_0, \max_{1 \leq j \leq n} \|A(., j)\|_0 \right\}, \qquad (3)$$

where $A(i, .)$ represents the i'th row and $A(., j)$ the j'th column of matrix A. The value of the $\mathcal{S}_{0,1}$-measure of a matrix is the maximum number of nonzero elements among all rows and columns of that matrix [18].

III. PROBLEM STATEMENT

A. Network model

We consider a class of consensus networks that consist of a group of subsystems whose state variables x_i, control inputs u_i, and output variables y_i are scalar and their dynamics evolve with time according to

$$\dot{x}_i(t) = u_i(t) + \xi_i(t) \qquad (4)$$
$$y_i(t) = x_i(t) - \bar{x}(t) \qquad (5)$$

for all $i = 1, \ldots, n$, where $x_i(0) = x_0^i$ is the initial condition and

$$\bar{x}(t) = \frac{1}{n}(x_1(t) + \ldots + x_n(t))$$

is the average of all states at time instant t. The impact of the uncertain environment on each agent's dynamics is modeled by the exogenous noise/disturbance input ξ_i. By applying the following linear feedback control law to the agents of this network

$$u_i(t) = \sum_{j=1}^{n} k_{ij}(x_j(t) - x_i(t)), \qquad (6)$$

where k_{ij} is the feedback gain between subsystems i and j, the closed-loop dynamics of network (4)-(6) can be written in the following compact form

$$\mathfrak{N}(L) : \begin{cases} \dot{x}(t) = -Lx(t) + \xi(t) \\ y(t) = M_n x(t) \end{cases} \qquad (7)$$

with initial condition $x(0) = x_0$, where x, ξ and y denote the state vector of the entire network, the exogenous disturbance input and the output vector of the network, respectively. The Laplacian matrix $L = [l_{ij}]$ is defined by

$$l_{ij} := \begin{cases} -k_{ij} & \text{if} \quad i \neq j \\ k_{i1} + \ldots + k_{in} & \text{if} \quad i = j \end{cases}. \qquad (8)$$

The coupling graph of the consensus network (7) is a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ with node set $\mathcal{V} = \{1, \ldots, n\}$, link set

$$\mathcal{E} = \left\{ \{i, j\} \mid \forall i, j \in \mathcal{V} : k_{ij} \neq 0 \right\}, \qquad (9)$$

and weight function

$$w(e) = \begin{cases} k_{ij} & \text{if} \quad e = \{i, j\} \in \mathcal{E} \\ 0 & \text{if} \quad e \notin \mathcal{E} \end{cases}. \qquad (10)$$

One may verify that the Laplacian matrix of graph $\mathcal{G}$ is equal to L.

Assumption 2: All feedback gains (weights) satisfy the following properties for all $i, j \in \mathcal{V}$:

(i) non-negativity: $k_{ij} \geq 0$,
(ii) symmetry: $k_{ij} = k_{ji}$,
(iii) simpleness: $k_{ii} = 0$.

Property (ii) implies that feedback gains are symmetric and (iii) means that there is no self-feedback loop in the network.

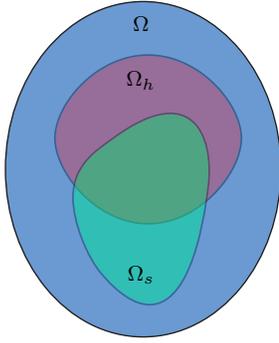*Assumption 3:* The coupling graph $\mathcal{G}$ of the consensus network (7) is time-invariant.

Fig. 1: A Venn diagram that shows the relationship among sets Ω, Ω_s, and Ω_h. The set of general systemic measures Ω is a superset of both the set of homogeneous systemic measures Ω_h and the set of spectral systemic measures Ω_s. While the intersection of sets Ω_s and Ω_h is nonempty, there are some systemic measures that belong only to one of these sets.

Based on Assumption 3, the corresponding eigenvector to the only marginally stable mode of the network is $\mathbb{1}_n$. This mode is unobservable from the performance output as the output matrix of the network satisfies $M_n \mathbb{1}_n = 0$.

B. Homogeneous Systemic Performance Measures

A systemic measure in this paper refers to a real-valued operator over the set of all consensus networks governed by (7) with the purpose of quantifying performance of this class of networks in presence exogenous uncertainties. Since every network with dynamics (7) is uniquely determined by its Laplacian matrix, it is reasonable to define a systemic performance measure as an operator on set $\mathfrak{L}_n$.

Definition 1: An operator $\rho : \mathfrak{L}_n \to \mathbb{R}_+$ is called a homogeneous systemic measure of order $-\alpha$, where $\alpha > 0$, if it satisfies the following properties for all matrices in $\mathfrak{L}_n$:

1. *Homogeneity:* For all $\kappa > 1$,
$$\rho(\kappa L) = \kappa^{-\alpha} \rho(L);$$

2. *Monotonicity:* If $L_2 \preceq L_1$, then
$$\rho(L_1) \leq \rho(L_2);$$

3. *Convexity:* For all $0 \leq c \leq 1$,
$$\rho(cL_1 + (1-c)L_2) \leq c\rho(L_1) + (1-c)\rho(L_2);$$

The set of all homogeneous systemic performance measures is denoted by Ω_h. We adopt an axiomatic approach to introduce and categorize a general class of performance measures that captures the quintessence of a meaningful measure of performance in large-scale dynamical networks [19]. Property 1 implies that intensifying the coupling weights by ratio $\kappa > 1$ results in κ^α times better performance. Property 2 guarantees that strengthening couplings in a consensus network never worsens the network performance with respect to a given systemic measure. The monotonicity property induces a partial ordering on all linear consensus networks with dynamics (7). Adding new coupling links or strengthening the existing couplings will result in better performance. Property 3 is imposed for the pure purpose of having favorable (convex) network design optimization problems.

The class of systemic performance measures can be classified based on their functional properties according to Definition 1. Let us denote the set of spectral systemic performance measures by Ω_s. This class consists of all measures that satisfy properties 2, 3 and orthogonal

invariance[3]. We refer to [20] for a comprehensive study of this class of performance measures. It is proven that all measures in Ω_s depend only on Laplacian eigenvalues. Let us represent the set of all general systemic performance measures that only satisfy properties 2 and 3 by Ω. Fig. 1 shows the relationship between the sets of spectral, homogeneous, and general systemic performance measures.

Definition 2: For a given linear consensus network $\mathfrak{N}(L)$ endowed with a homogeneous systemic measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$ of order $-\alpha$, its corresponding normalized performance index is defined by
$$\Pi_\rho(L) := \sqrt[\alpha]{\rho(L)}. \tag{11}$$

C. Network Abstraction Problem

Our goal is to develop a framework to compute an *abstraction* of a given linear consensus network with predetermined levels of performance and sparsity (i.e., link reduction).

Definition 3: Let us consider network $\mathfrak{N}(L)$ that is governed by (7). For a properly chosen pair of design parameters $d \in \mathbb{R}_{++}$ and $\epsilon \in (0,1)$, another $\mathfrak{N}(L_s)$ is said to be an (ϵ, d)-abstraction of $\mathfrak{N}(L)$ if and only if

(i) $\mathfrak{N}(L_s)$ has at most $dn/2$ feedback links;
(ii) $\mathfrak{N}(L_s)$ is an ϵ-approximation of $\mathfrak{N}(L)$ in the following sense
$$\left| \frac{\Pi_\rho(L) - \Pi_\rho(L_s)}{\Pi_\rho(L_s)} \right| \leq \epsilon \tag{12}$$

for every homogeneous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$.

Property (i) implies that the average number of neighbors for every node in $\mathfrak{N}(L_s)$ is less than d, i.e.,
$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} |\mathcal{N}(i)| = 2\frac{|\mathcal{E}_s|}{n} \leq d,$$

where $\mathcal{N}(i)$ and $\mathcal{E}_s$ denote the set of all adjacent nodes to i and the set of all links in the abstraction, respectively. Therefore, one can think of design parameter d as an upper bound on the desired average number of neighbors of nodes in the abstracted network which is independent of the network size. For Property (ii), inequality (12) indicates that the resulting abstracted network $\mathfrak{N}(L_s)$ has guaranteed performance bounds with respect to $\mathfrak{N}(L)$. The design constant ϵ is referred to as permissible performance loss parameter.

IV. EXAMPLES OF RELEVANT HOMOGENEOUS SYSTEMIC PERFORMANCE MEASURES

We now present some existing and widely-used systemic performance measures for linear consensus networks; a list of these measures are summarized in Table I.

A. Sum of Homogeneous Spectral Functions

This class of performance measures is generated by forming summation of a given function of Laplacian eigenvalues. For a given matrix $L \in \mathfrak{L}_n$, suppose that $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ is a decreasing homogeneous convex function. Then, the following spectral function
$$\rho(L) = \sum_{i=2}^{n} \varphi(\lambda_i) \tag{13}$$

is a homogeneous systemic measure [20]. Moreover, if φ is a homogeneous function of order $-\alpha$ where $\alpha > 1$, then its corresponding normalized index
$$\Pi_\rho(L) = \left(\sum_{i=2}^{n} \varphi(\lambda_i) \right)^{\frac{1}{\alpha}} \tag{14}$$

---

[3]A systemic measure is orthogonally invariant if $\rho(L) = \rho(ULU^T)$ for every orthogonal matrix U for which $UU^T = U^TU = I$.

Homogeneous Systemic Performance Measure	Symbol	Representation
Spectral Riemann zeta function	$\zeta_q(L)$	$\left(\sum_{i=2}^{n} \lambda_i^{-q}\right)^{1/q}$
Gamma entropy	$I_\gamma(L)$	$\gamma^2 \sum_{i=2}^{n} \left(\lambda_i - \left(\lambda_i^2 - \gamma^{-2}\right)^{\frac{1}{2}}\right)$
System Hankel norm	$\eta(L)$	$\frac{1}{2}\lambda_2^{-1}$
Hardy-Schatten or $\mathcal{H}_p$ system norm	$\theta_p(L)$	$\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty}\sum_{k=1}^{n}\sigma_k(G(j\omega))^p\,d\omega\right\}^{1/p} = \alpha_0\left(\text{Tr}\left(L^\dagger\right)^{p-1}\right)^{\frac{1}{p}}$
Local Deviation Error for First Order Consensus Networks	$\Delta(L)$	$\frac{1}{2}\sum_{i=1}^{n}d_i^{-1}$
Local Deviation Error for Second Order Consensus Networks with $\beta > 0$	$\Upsilon(L)$	$\frac{1}{2\beta}\sum_{i=1}^{n}d_i^{-2}$
$\mathcal{H}_2$-norm of Second Order Consensus Networks with $\beta > 0$	$\Theta_2(L)$	$\left(\frac{1}{2\beta}\sum_{i=1}^{n}\lambda_i^{-2}\right)^{1/2}$

TABLE I: Some important examples of homogeneous systemic performance measures.

is also a homogeneous systemic performance measure [20]. Some notable examples of this class of measures are discussed in the following parts.

1) Spectral Riemann Zeta Measures: For a given network (7), its corresponding spectral Riemann zeta function of order $q \geq 1$ is defined by

$$\zeta_q(L) := \left(\sum_{i=2}^{n}\lambda_i^{-q}\right)^{1/q}, \tag{15}$$

where $\lambda_2,\ldots,\lambda_n$ are eigenvalues of L [21]. According to Assumption 3, all Laplacian eigenvalues are strictly positive and, as a result, function (15) is well-defined. According to the result presented in Subsection IV-A, since $\varphi(\lambda) = \lambda^{-q}$ for $q \geq 1$ is a decreasing homogeneous convex function, the spectral function (15) is a homogeneous systemic performance measure. The homogeneous systemic performance measure $\frac{1}{2}\zeta_1(L)$ is equal to the $\mathcal{H}_2$-norm squared of a first-order consensus network (7) and $\frac{1}{\sqrt{2}}\zeta_2(L)$ equal to the $\mathcal{H}_2$-norm of a second-order consensus model of a network of multiple agents (c.f. [11]).

2) Gamma Entropy: The notion of gamma entropy arises in various applications such as the design of minimum entropy controllers and interior point polynomial-time methods in convex programming with matrix norm constraints [22]. As it is shown in [23], the notion of gamma entropy can be interpreted as a performance measure for linear time-invariant systems with random feedback controllers by relating the gamma entropy to the mean-square value of the closed-loop gain of the system. The γ-entropy of network (7) is defined as

$$I_\gamma(L) := \begin{cases} \frac{-\gamma^2}{2\pi}\int_{-\infty}^{\infty}\log\det\left(I - \gamma^{-2}G(j\omega)G^*(j\omega)\right)d\omega \\ \qquad\qquad\qquad\qquad \text{for } \gamma \geq \|G\|_{\mathcal{H}_\infty} \\ \\ \infty \qquad\qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$

where $G(j\omega)$ is the transfer matrix of network (7) from ξ to y [23]. In [20], it is shown that the value of the γ-entropy for a given linear consensus network (7) can be explicitly computed in terms of Laplacian spectrum as follows

$$I_\gamma(L) = \begin{cases} \sum_{i=2}^{n} f_\gamma(\lambda_i) & \gamma \geq \lambda_2^{-1} \\ \\ \infty & \text{otherwise} \end{cases}, \tag{16}$$

where $f_\gamma(\lambda_i) = \gamma^2\left(\lambda_i - \left(\lambda_i^2 - \gamma^{-2}\right)^{\frac{1}{2}}\right)$. Furthermore, the γ-entropy $I_\gamma(L)$ is a homogeneous systemic performance measure.

B. Uncertainty volume

The uncertainty volume of the steady-state output covariance matrix of network (7) is defined by

$$|\Sigma| := \det\left(Y_\infty + \frac{1}{n}J_n\right) \tag{17}$$

in which

$$Y_\infty = \lim_{t\to\infty}\mathbb{E}\left[y(t)y^{\mathrm{T}}(t)\right].$$

This quantity is widely used as an indicator of the network performance [2] and [24]. Since $y(t)$ is the error vector that shows distance from consensus, the quantity (17) can be interpreted as volume of the steady-state error ellipsoid. It is straightforward to show this measure satisfies all properties of Definition 1.

C. Hankel Norm

The Hankel norm of network (7) and transfer matrix $G(j\omega)$ from ξ to y is defined as the $\mathcal{L}_2$-gain from past inputs to the future outputs, i.e.,

$$\|G\|_H^2 := \sup_{\xi\in L_2(-\infty,0]}\frac{\int_0^\infty y^{\mathrm{T}}(t)y(t)dt}{\int_{-\infty}^{0}\xi^{\mathrm{T}}(t)\xi(t)dt}.$$

The value of the Hankel norm of network (7) can be equivalently computed using the Hankel norm of its disagreement form [3] that is given by

$$\dot{x}_d(t) = -L_d\,x_d(t) + M_n\,\xi(t), \tag{18}$$
$$y(t) = M_n\,x_d(t), \tag{19}$$

where the disagreement vector is defined by

$$x_d(t) := M_n\,x(t) = x(t) - \frac{1}{n}J_n\,x(t). \tag{20}$$

The disagreement network (18)-(19) is stable as the real part of every eigenvalue of the state matrix $-L_d = -(L+\frac{1}{n}J_n)$ is strictly negative. One can verify that the transfer matrices from $\xi(t)$ to $y(t)$ in both realizations are identical. Therefore, the Hankel norm of the system from $\xi(t)$ to $y(t)$ in both representations are well-defined and equal, and is given by [25]

$$\eta(L) := \|G\|_H = \sqrt{\lambda_{\max}(PQ)}, \tag{21}$$

where the controllability Gramian P is the unique solution of

$$\left(L + \frac{1}{n}J_n\right)P + P\left(L + \frac{1}{n}J_n\right) - M_n = 0$$

and the observability Gramian Q is the unique solution of

$$Q\left(L + \frac{1}{n}J_n\right) + \left(L + \frac{1}{n}J_n\right)Q - M_n = 0.$$

It is shown in [20] that the value of the Hankel norm of network (7) is equal to

$$\eta(L) = \frac{1}{2}\lambda_2^{-1}.$$

One can verify that this measure is a homogeneous systemic performance measure.

Remark 1: One may also consider the sum of the k largest eigenvalues of $L^\dagger$ as a performance measure. This is equivalent to evaluate the k slowest modes of the network, which are the most energetic modes. This measure satisfies properties of Definition 1 as it is convex and symmetric with respect to Laplacian eigenvalues (c.f. [26, Ch. 5.2] and [27]).

D. Hardy-Schatten or $\mathcal{H}_p$ System Norms

The $\mathcal{H}_p$-norm of networks (7) for $2 \le p \le \infty$ is defined by

$$\|G\|_{\mathcal{H}_p} := \left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\sum_{k=1}^{n}\sigma_k(G(j\omega))^p\,d\omega\right)^{\frac{1}{p}}, \qquad (22)$$

where G is the transfer matrix from $\xi(t)$ to $y(t)$ and $\sigma_k(j\omega)$ for $k = 1,\dots,n$ are singular values of $G(j\omega)$. To ensure well-definedness of performance measure (22), the marginally stable mode of the network must be unobservable through the output. Thus, this performance measure remains well-defined as long as the coupling graph of the network stays connected. This class of system norms captures several important performance and robustness features of linear control systems. For instance, a direct calculation reveals that the $\mathcal{H}_2$-norm of network (7) is

$$\|G\|_{\mathcal{H}_2} = \left(\frac{1}{2}\sum_{i=2}^{n}\lambda_i^{-1}\right)^{\frac{1}{2}}. \qquad (23)$$

This system norm quantifies the quality of noise propagation throughout the network [12]. The $\mathcal{H}_\infty$-norm of a network is an input-output system norm and its value for network (7) is

$$\|G\|_{\mathcal{H}_\infty} = \lambda_2^{-1}, \qquad (24)$$

where λ_2 is known as the algebraic connectivity of the network [3]. The value of $\mathcal{H}_\infty$-norm of network (7) can be interpreted as the worst attainable performance for all square integrable disturbance inputs.

In [20], the authors prove that the $\mathcal{H}_p$-norm of a given network $\mathfrak{N}(L)$ is given by

$$\theta_p(L) := \|G\|_{\mathcal{H}_p} = \alpha_0\left(\zeta_{P-1}(L)\right)^{1-\frac{1}{p}} \qquad (25)$$

in which $\alpha_0^{-1} = \sqrt[p]{-\beta(\frac{p}{2}, -\frac{1}{2})}$ and $\beta : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ is the well-known Beta function[4]. Moreover, this measure is a homogeneous systemic performance measure for all $2 \le p \le \infty$.

E. Local deviation error:

In network (7), the local deviation of subsystem i is equal to the deviation of the state of subsystem i from the weighted average of

---

[4] $\beta(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$ for $\mathrm{Re}\{x\}, \mathrm{Re}\{y\} > 0$.

states of its immediate neighbors, which can be formally defined by

$$\varepsilon_i(t) := x_i(t) - \frac{1}{d_i}\sum_{e=\{i,j\}\in\mathcal{E}}w(e)\,x_j(t). \qquad (26)$$

The expected cumulative local deviation is then defined by

$$\Delta(L) = \lim_{t\to\infty}\mathbb{E}\left[\sum_{i=1}^{n}\varepsilon_i(t)^2\right] \qquad (27)$$

with respect to input ξ being a white noise process with identity covariance. The notion of local deviation can be extended and defined for velocity variables in the second-order consensus network (96)-(97) (c.f., [11]) as follows

$$\varsigma_i(t) := v_i(t) - \frac{1}{d_i}\sum_{e=\{i,j\}\in\mathcal{E}}w(e)\,v_j(t) \qquad (28)$$

that is equal to the deviation of the velocity of subsystem i from the weighted average of velocities of its neighbors. The expected cumulative local deviation is then given by

$$\Upsilon(L) = \lim_{t\to\infty}\mathbb{E}\left[\sum_{i=1}^{n}\varsigma_i(t)^2\right], \qquad (29)$$

where it is assumed that input ξ in network model (96)-(97) is a white noise process with identity covariance.

Theorem 1: The operators $\Delta, \Upsilon : \mathfrak{L}_n \to \mathbb{R}_+$ defined by (27) and (29) are homogeneous systemic performance measures. Moreover, they can also be characterized as

$$\Delta(L) = \frac{1}{2}\sum_{i=1}^{n}d_i^{-1} \qquad (30)$$

and

$$\Upsilon(L) = \frac{1}{2\beta}\sum_{i=1}^{n}d_i^{-2} \qquad (31)$$

in which d_i is the degree of node $i \in \mathcal{V}$.

Proof: Let us define the total local deviation at time t by

$$\varepsilon_{\text{total}}(t) := \sum_{i\in\mathcal{V}}\varepsilon_i(t)^2. \qquad (32)$$

We reformulate (26) as

$$\begin{aligned}\varepsilon_i(t) &= d_i^{-1}\left(d_i x_i(t) - \sum_{e=\{i,j\}\in\mathcal{E}}w(e)\,x_j(t)\right) \\ &= d_i^{-1}\sum_{e=\{i,j\}\in\mathcal{E}}w(e)\,(x_i(t) - x_j(t)). \end{aligned} \qquad (33)$$

Therefore, we get

$$\varepsilon(t) = \text{diag}\left[d_1^{-1}, \cdots, d_n^{-1}\right]L\,x(t).$$

where $\varepsilon(t)$ is concatenation of elements $\varepsilon_i(t)$ for $i = 1,\dots,n$. Also, we can rewrite (32) as follows

$$\varepsilon_{\text{total}}(t) = \varepsilon^{\text{T}}(t)\varepsilon(t) = x^{\text{T}}(t)Qx(t)$$

with

$$Q = L\,\text{diag}\left[d_1^{-2}, \cdots, d_n^{-2}\right]L.$$

Thus, according to [12, Thm. 5] the steady-state of $\varepsilon_{\text{total}}$ is given by

$$\begin{aligned}\Delta(L) &= \lim_{t\to\infty}\mathbb{E}\left[\varepsilon_{\text{total}}(t)\right] = \frac{1}{2}\,\text{Tr}\left(L^\dagger Q\right) \\ &= \frac{1}{2}\sum_{i\in\mathcal{V}}d_i^{-1}. \end{aligned} \qquad (34)$$

Now we show this measure is a homogeneous systemic performance measure. We first show that (34) has property 1, which means

$$\Delta(\kappa L) \;=\; \frac{1}{2}\sum_{i\in\mathcal{V}}(\kappa d_i)^{-1} \;=\; \kappa^{-1}\Delta(L).$$

Furthermore, it is monotone, because if $L_1 \preceq L_2$ then we have

$$e_i^{\mathrm{T}} L_1 e_i \leq e_i^{\mathrm{T}} L_2 e_i,$$

where e_i for $i = 1, \ldots, n$ are the standard basis for the n-dimensional Euclidean space. Therefore, we have $L_1(i,i) \leq L_2(i,i)$ that guarantees the monotonicity of Δ. Moreover, its convexity follows from convexity of function $1/x$ for all $x \in \mathbb{R}_+$. Because consider two Laplacian matrices L_1 and L_2 with node degrees $d_i^{(1)}$ and $d_i^{(2)}$, respectively, for $i = 1, \ldots, n$. Then, we get

$$
\begin{aligned}
\Delta(cL_1 + (1-c)L_2) &= \sum_{i\in\mathcal{V}} \frac{1}{c\,d_i^{(1)} + (1-c)\,d_i^{(2)}} \\
&\leq \sum_{i\in\mathcal{V}}\left(\frac{c}{d_i^{(1)}} + \frac{1-c}{d_i^{(2)}} \right) \\
&= \sum_{i\in\mathcal{V}} \frac{c}{d_i^{(1)}} + \sum_{i\in\mathcal{V}} \frac{1-c}{d_i^{(2)}} \\
&= c\Delta(L_1) + (1-c)\Delta(L_2)
\end{aligned}
$$

for all $0 \leq c \leq 1$. This completes the proof of the first part. For the second part, let us define the total local deviation error at time t by

$$\varsigma_{\text{total}}(t) \;:=\; \sum_{i\in\mathcal{V}} \varsigma_i(t)^2. \tag{35}$$

We similarly reformulate (28) as

$$
\begin{aligned}
\varsigma_i(t) &= d_i^{-1}\left(d_i v_i(t) - \sum_{e=\{i,j\}\in\mathcal{E}} w(e)\, v_j(t) \right) \\
&= d_i^{-1} \sum_{e=\{i,j\}\in\mathcal{E}} w(e)\,(v_i(t) - v_j(t)).
\end{aligned}
$$

Therefore, we have

$$\varsigma(t) = \mathrm{diag}\big[d_1^{-1}, \cdots, d_n^{-1}\big]\, L\, v(t),$$

where $\varsigma(t)$ is concatenation of elements $\varsigma_i(t)$ for all $1 \leq i \leq n$. Moreover, we can rewrite (35) as follows

$$
\begin{aligned}
\varsigma_{\text{total}}(t) &= \varsigma^{\mathrm{T}}(t)\varsigma(t) \\
&= v^{\mathrm{T}}(t)Q v(t), \tag{36}
\end{aligned}
$$

where Q is given by $Q = L\,\mathrm{diag}\big[d_1^{-2}, \cdots, d_n^{-2}\big]\,L$. Therefore, the steady-state of ς_{total} can be characterized as

$$
\begin{aligned}
\Upsilon(L) &= \lim_{t\to\infty} \mathbb{E}\left[\varsigma_{\text{total}}(t)\right] = \frac{1}{2\beta}\,\mathrm{Tr}\left((L^{\dagger})^2 Q \right) \\
&= \frac{1}{2\beta}\sum_{i\in\mathcal{V}} d_i^{-2}. \tag{37}
\end{aligned}
$$

This measure is a homogeneous systemic performance measure. It is straightforward to show that (37) satisfies property 1 by verifying that
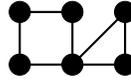
$$\Upsilon(\kappa L) = \frac{1}{2\beta}\sum_{i\in\mathcal{V}}(\kappa d_i)^{-2} = \kappa^{-2}\Upsilon(L).$$

It is monotone, as if $L_1 \preceq L_2$, then we have

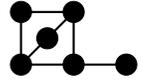$$e_i^{\mathrm{T}} L_1 e_i \leq e_i^{\mathrm{T}} L_2 e_i.$$

As a result, it follows that $L_1(i,i) \leq L_2(i,i)$ that guarantees the monotonicity of Υ. Finally, its convexity can be concluded from convexity of function $1/x^2$ for all $x \in \mathbb{R}_+$. ∎

Remark 2: For first-order consensus network (7) that are defined

(a)

(b)

Fig. 2: Two isospectral graphs with six nodes [29].

over d-regular coupling graphs, the corresponding microscopic measure (30) scales linearly with network size. For regular lattices that are d-regular graphs, our result assumes the reported result of [28] as its special case.

Remark 3: Fig. 2 shows example of two isospectral[5] graphs that are not isometric[6]. While the value of a spectral systemic performance measure is equal for both graphs, the value of an expected cumulative local deviation measure is different for each of these graphs and depend on their specific interconnection topology. This simple observation implies that systemic performance measures (30) and (31) are suitable tools to differentiate among networks with isospectral coupling graphs.

V. ABSTRACTION WITH GUARANTEED BOUNDS

In this section, we develop a fast abstraction algorithm for the class of linear consensus networks (7) with guaranteed bounds with respect to the class of homogeneous systemic performance measures.

A. Intrinsic Tradeoffs on the Best Achievable Abstractions

The abstraction goals are to reduce the number of feedback links while preserving a desired level of performance. From notation (3), one can easily verify that the value of $\mathcal{S}_{0,1}$-measure is equal to the maximum of $|\mathcal{N}(i)|$ for all nodes $i = 1, \ldots, n$, which makes it a suitable surrogate for design parameter d. The next result reveals an inherent interplay between sparsity and performance.

Theorem 2: For a given network (7) that is endowed with a homogenous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$ of order $-\alpha$, suppose that $w_* = \max_{e\in\mathcal{E}} w(e)$. Then, there are fundamental tradeoffs between normalized performance and graph sparsity measures in the following sense

$$\Pi_\rho(L)\,\|A\|_{\ell_0} \;\geq\; 2\varrho^*(n-1), \tag{38}$$

and

$$\Pi_\rho(L)\,\|A\|_{\mathcal{S}_{0,1}} \;\geq\; 2\varrho^* \tag{39}$$

when $n > 2$, in which A is the adjacency matrix of the coupling graph and $\varrho^* = w_*^{-1}\Pi_\rho(L_{\mathcal{K}_n})$, where $L_{\mathcal{K}_n}$ is Laplacian matrix of the unweighted complete graph.

Proof: Since it is assumed that the coupling graph of the network is connected, the sparsity measure $\|A\|_{\ell_0}$ is always bounded from below by $2(n-1)$; with equality sign if the coupling graph is a tree. Thus, the following inequality holds on the cone of positive semidefinite matrices

$$L \;\preceq\; \frac{\|A\|_{\ell_0}}{2(n-1)}\,L \;\preceq\; \frac{w_*\|A\|_{\ell_0}}{2(n-1)}\,L_{\mathcal{K}_n}.$$

From monotonicity property, it follows that

$$\rho(L)\,\|A\|_{\ell_0}^\alpha \;\geq\; 2^\alpha(n-1)^\alpha w_*^{-\alpha}\,\rho(L_{\mathcal{K}_n}).$$

---

[5]Two graphs are called isospectral if and only if their Laplacian matrices have the same multi-sets of eigenvalues

[6]This means that their adjacency matrices are not permutation-similar.

By taking α'th root from both sides, one can conclude the desired inequality (38). When $n > 2$, the localized sparsity measure $\|A\|_{\mathcal{S}_{0,1}}$ is always greater of equal to 2. Therefore, the following relation holds

$$L \preceq \frac{1}{2}\|A\|_{\mathcal{S}_{0,1}} L \preceq \frac{1}{2} w_* \|A\|_{\mathcal{S}_{0,1}} L_{\mathcal{K}_n}.$$

By utilizing the monotonicity property, we get

$$\rho(L)\|A\|_{\mathcal{S}_{0,1}}^{\alpha} \geq 2^{\alpha} w_*^{-\alpha} \rho(L_{\mathcal{K}_n}).$$

The desired inequality (39) follows from taking α'th root from both sides of the inequality. ∎

The monotonicity property of a systemic performance measure implies that link removal will lead to performance deterioration. Theorem 2 quantifies this inherent interplay by saying that sparsity and performance cannot be improved indefinitely both at the same time. As we will see in the following subsection, this is exactly why we need to perform reweighing after link elimination procedure in order to achieve an approximation that meets (12).

B. Existence and Algorithms

The next theorem enables us to harness the monotonicity property of homogeneous systemic measures in our network approximations.

Theorem 3: Suppose that two linear consensus networks $\mathfrak{N}(L)$ and $\mathfrak{N}(L_s)$ are endowed with a homogeneous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$ of order $-\alpha$. For a given constant $\epsilon \in (0,1)$, the two networks are ϵ-approximation of each other, i.e., property (12) holds, if and only if their state matrices satisfy

$$(1-\epsilon)L \preceq L_s \preceq (1+\epsilon)L. \tag{40}$$

Proof: According to the monotonicity and homogeneity properties of system measures, it follows that if (40) holds then we have

$$(1+\epsilon)^{-\alpha}\rho(L) \leq \rho(L_s) \leq (1-\epsilon)^{-\alpha}\rho(L). \tag{41}$$

Therefore, according to (41) and (12), $\mathfrak{N}(L_s)$ is an ϵ-approximation of $\mathfrak{N}(L)$. Let us consider the following measures

$$\rho_v(L) = v^{\mathrm{T}} L^{\dagger} v \tag{42}$$

for all $v \in \mathbb{R}^n$. This operator is a homogeneous systemic performance measure of order -1. For all $v \notin \mathrm{Span}\{\mathbb{1}\}$, inequality (12) yields

$$-\epsilon \leq \frac{\rho_v(L) - \rho_v(L_s)}{\rho_v(L_s)} \leq \epsilon.$$

Thus, it follows that

$$(1+\epsilon)^{-1} \leq \frac{v^{\mathrm{T}} L_s^{\dagger} v}{v^{\mathrm{T}} L^{\dagger} v} \leq (1-\epsilon)^{-1}. \tag{43}$$

Since $v^{\mathrm{T}} L^{\dagger} v > 0$, inequalities (43) can be rewritten as

$$(1+\epsilon)^{-1} v^{\mathrm{T}} L^{\dagger} v \leq v^{\mathrm{T}} L_s^{\dagger} v \leq (1-\epsilon)^{-1} v^{\mathrm{T}} L^{\dagger} v. \tag{44}$$

We know that L and L_s are Laplacian matrices and (44) holds for all $v \notin \mathrm{Span}\{\mathbb{1}\}$; therefore, we get

$$(1+\epsilon)^{-1} L^{\dagger} \preceq L_s^{\dagger} \preceq (1-\epsilon)^{-1} L^{\dagger}.$$

This inequality can be rewritten to obtain the desired result

$$(1-\epsilon)L \preceq L_s \preceq (1+\epsilon)L.$$

∎

The result of the above theorem is crucial as it enables us to take advantage of monotonicity property of systemic performance measures in our approximations. For two given networks $\mathfrak{N}(L_1)$ and $\mathfrak{N}(L_2)$, inequality $\rho(L_2) \leq \rho(L_1)$ can be realized through several possible scenarios; for example, network $\mathfrak{N}(L_2)$ can be constructed by (i) adding new weighted edges to the coupling graph of network $\mathfrak{N}(L_1)$, (ii) increasing weights of some of the existing links in network $\mathfrak{N}(L_1)$, (iii) rewiring topology of network $\mathfrak{N}(L_1)$ while ensuring $L_1 \preceq L_2$.

The next result proves existence of an abstraction for every given linear consensus network.

Theorem 4: Suppose that a network $\mathfrak{N}(L)$ with coupling graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ endowed with a homogeneous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$ of order $-\alpha$ and a design parameter $d > 2$ are given. Then, it is possible to construct another network $\mathfrak{N}(L_s)$ with coupling graph $\mathcal{G}_s = (\mathcal{V}, \mathcal{E}_s, w_s)$ such that
(i) $\mathfrak{N}(L_s)$ is a $(\frac{\sqrt{8d}}{d+2}, d)$-abstraction of network $\mathfrak{N}(L)$;
(ii) $\mathcal{E}_s \subset \mathcal{E}$.

Proof: It is well known that the characteristic polynomial of $A + vv^{\mathrm{T}}$ can be computed based on the characteristic polynomial[7] of A, its eigenvalues μ_i, and eigenvectors u_i as follows

$$P_{A+vv^{\mathrm{T}}}(x) = P_A(x)\left(1 - \sum_i \frac{\langle v, u_i \rangle^2}{x - \mu_i}\right)$$

in which $\langle v, u_i \rangle := v^{\mathrm{T}} u_i$. Moreover, we know that the eigenvalues of $A + vv^{\mathrm{T}}$ interlace those of A. Reference [17] suggests a framework which provides intuition as to where these new eigenvalues are located. If we start with an empty graph on n nodes, then after $M = \lceil d(n-1)/2 \rceil$ iterations of choosing links from the original graph and adjusting its weight, it is shown that the eigenvalues of the resulted graph are controlled by maintaining two barrier potential functions. According to [17, Th. 1.1], coupling graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ has a weighted subgraph $\hat{\mathcal{G}} = (\mathcal{V}, \hat{\mathcal{E}}, \hat{w})$ with $|\hat{\mathcal{E}}| = \lceil d(n-1)/2 \rceil$ that satisfies

$$L \preceq L_{\hat{\mathcal{G}}} \preceq \left(\frac{1 + \sqrt{d/2}}{1 - \sqrt{d/2}}\right)^2 L \tag{45}$$

where $L_{\hat{\mathcal{G}}}$ is the Laplacian matrix of graph $\hat{\mathcal{G}}$. We define $\mathcal{G}_s = (\mathcal{V}, \hat{\mathcal{E}}, w_s)$ by its Laplacian matrix, which is given by

$$L_s := \frac{(1 - \sqrt{d/2})^2}{1 + d/2} L_{\hat{\mathcal{G}}}. \tag{46}$$

Therefore, according to (45) and (46), it follows that

$$\left(1 - \frac{\sqrt{2d}}{d/2 + 1}\right) L \preceq L_s \preceq \left(1 + \frac{\sqrt{2d}}{d/2 + 1}\right) L. \tag{47}$$

Using (47) and Theorem 3, it yields that $\mathfrak{N}(L_s)$ is a $(\frac{\sqrt{2d}}{d/2+1}, d)$-abstraction of $\mathfrak{N}(L)$. Since $\mathcal{E}_s = \hat{\mathcal{E}}$ and the fact that $\hat{\mathcal{E}}$ is obtained by taking samples from $\mathcal{E}$, one concludes that $\mathcal{E}_s \subset \mathcal{E}$. ∎

In our next result, we show that every consensus network has a sparsification such that: (i) it yields a better systemic performance than the original network, and (ii) the total weight sum of the coupling graph of the spars network is controlled, i.e., it is less than a constant multiple, which is independent of the network size, of the total weight sum of the original network.

Proposition 1: For a given consensus network $\mathfrak{N}(L)$ with coupling graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ and every $d > 2$, there exists a consensus network $\mathfrak{N}(L_s)$ with coupling graph $\mathcal{G}_s = (\mathcal{V}, \mathcal{E}_s, w_s)$ that has at most $dn/2$ links and $\mathcal{E}_s \subset \mathcal{E}$. Moreover, we have
(i) the total weight of coupling graph of $\mathfrak{N}(L_s)$ is controlled, i.e.,

$$\sum_{e \in \mathcal{E}_s} w_s(e) \leq \left(\frac{\sqrt{2d} + 2}{\sqrt{2d} - 2}\right)^2 \sum_{e \in \mathcal{E}} w(e); \tag{48}$$

---

[7]The characteristic polynomial of matrix $A \in \mathbb{R}^{n \times n}$ is defined by

$$p_A(t) = \det(tI_n - A).$$

(ii) $\mathfrak{N}(L_s)$ has a superior performance with respect to $\mathfrak{N}(L)$, i.e.,

$$\rho(L_s) \leq \rho(L),$$

for every homogeneous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$.

Proof: According to Theorem 4, $\mathfrak{N}(L)$ has a $(\frac{\sqrt{8d}}{d+2}, d)$-abstraction, $\mathfrak{N}(L_{\hat{s}})$, with $\mathcal{E}_{\hat{s}} \subset \mathcal{E}$, which means that we have

$$\left(1 - \frac{\sqrt{2d}}{d/2+1}\right)^\alpha \leq \frac{\rho(L)}{\rho(L_{\hat{s}})} \leq \left(1 + \frac{\sqrt{2d}}{d/2+1}\right)^\alpha \quad (49)$$

for every homogeneous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$ of order $-\alpha$. Let us define $L_s = \frac{d+2-\sqrt{8d}}{d+2} L_{\hat{s}}$. Then, it follows that

$$\rho(L_{\hat{s}}) = \left(\frac{d+2}{d+2-\sqrt{8d}}\right)^\alpha \rho(L_s). \quad (50)$$

Finally, by substituting $\rho(L_{\hat{s}})$ from (50) in the left-hand side inequality of (49), we obtain

$$\rho(L_s) \leq \rho(L).$$

It is straightforward to verify that operator $\Xi(L) = \left(\sum_{e \in \mathcal{E}} w(e)\right)^{-1}$ is a homogeneous systemic performance measure with $\alpha = 1$. From the left-hand side inequality of (49), we get the desired bound (48). ∎

Remark 4: Proposition 1 demonstrates a tradeoff between the largest possible number of links in an abstraction, i.e., upper bound of $|\mathcal{E}_s|$, and the best achievable ratio of the total weights, i.e., upper bound of $\sum_{e \in \mathcal{E}_s} w_s(e) / \sum_{e \in \mathcal{E}} w(e)$. We have that

$$|\mathcal{E}_s| \leq \frac{1}{2}dn \quad (51)$$

and

$$\frac{\sum_{e \in \mathcal{E}_s} w_s(e)}{\sum_{e \in \mathcal{E}} w(e)} \leq \left(\frac{\sqrt{2d}+2}{\sqrt{2d}-2}\right)^2. \quad (52)$$

Let us fix n. By increasing the value of parameter d, the upper bound in (51) increases, but the upper bound in (52) decreases and gets closer to 1. On the other hand, if $d \to 2$, then the upper bound in (52) tends to infinity.

We next employ a randomized algorithm to compute an (ϵ, d)-abstraction of a given network. A randomized algorithm utilizes a degree of randomness as part of its logic. Randomization allows us to design provably accurate algorithms for problems that are massive and computationally expensive or NP-hard. For this aim based on [16], we sample low-connectivity coupling links with high probability and high-connectivity coupling links with low probability. For a given consensus network $\mathfrak{N}(L)$ with n nodes, we sample links of the coupling graph of this network $M := \lceil dn/2 \rceil$ times in order to produce an (ϵ, d)-abstraction. Let us denote probability of selecting a link $e \in \mathcal{E}$ by $\pi(e)$[8] that is proportional to $w(e)r(e)$, where $w(e)$ and $r(e)$ are the weight and the effective resistance of link e, respectively. In each step of sampling[9], we add the selected link e to the abstraction with weight $w(e)/(M\pi(e))$. All details of our proposed algorithm are explained below. The following result, which is obtained based on a theorem in [16, Thm. 1], provides us with a certificate that the above randomized algorithm is capable of generating a proper abstraction of a given linear consensus network.

---

[8]It is well-know that $\sum_{e \in E} w(e)r(e) = n-1$; therefore, we have $\sum_{e \in \mathcal{E}} \pi(e) = 1$.

[9]A sampling is a discrete probability distribution on a support $\mathcal{E}$ of all possible samples. The probability of selecting link e is denoted by positive number $\pi(e)$ for all $e \in \mathcal{E}$. We also have $\sum_{e \in \mathcal{E}} \pi(e) = 1$, because π is a probability distribution on $\mathcal{E}$.

Algorithm 1: Network Abstraction Algorithm

Input : $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ and $r(e)$ for all $e \in \mathcal{E}$

Output: $\mathcal{G}_s = (\mathcal{V}, \mathcal{E}_s, w_s)$

1 set $\mathcal{G}_s$ to be the empty graph on $\mathcal{V}$ (i.e., $\mathcal{E}_s := \{\}$ and $w_s(.) := 0$)
2 set $\pi(e) = \frac{w(e)r(e)}{n-1}$ for all $e \in \mathcal{E}$
3 for $j = 1$ to $M := \lceil dn/2 \rceil$ do
4   sample a link from $\mathcal{E}$ with probability distribution $\pi \to \{e\}$
5 $\mathcal{E}_s = \mathcal{E}_s \cup \{e\}$
6 $w_s(e) = w_s(e) + \frac{w(e)}{M\pi(e)}$
7 end
8 return $\mathcal{G}_s$

Theorem 5: Suppose that a linear consensus network $\mathfrak{N}(L)$ endowed with a homogeneous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$ of order $-\alpha$ is given and a permissible performance loss parameter $\epsilon \in (1/\sqrt{n}, 1]$ is fixed. Let us pick a real number d that is at the order of $\epsilon^{-2} \log n$. Then, Algorithm 1 produces an (ϵ, d)-abstraction of network $\mathfrak{N}(L)$, whose coupling graph is a subgraph of $\mathfrak{N}(L)$, with probability at least 0.5.

Proof: Let us consider the following projection matrix

$$P = W^{1/2}EL^{\dagger}E^{\mathrm{T}}W^{1/2}, \quad (53)$$

where E is m-by-n incidence matrix and W is a diagonal matrix with link weights on its diagonal such that $L = E^{\mathrm{T}}WE$. The m-by-m matrix P has eigenvalue at 0 with multiplicity $m-n+1$ and eigenvalue at 1 with multiplicity $n-1$ [16, Lemma 3]. Now, we show that the sampling of links in Algorithm 1 corresponds to selecting $M = \mathcal{O}(n \log n/\epsilon^2)$ columns at random from matrix P. Then by a concentration lemma of Rudelson [30, Thm. 3.1] and Markov's inequality, with probability at least 0.5, we get

$$\|P - P\Gamma P\|_2 \leq \epsilon, \quad (54)$$

where Γ is a nonnegative diagonal matrix. Then, it is straightforward to show that for every homogeneous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$, we have

$$\left|\frac{\Pi_\rho(L) - \Pi_\rho(L_s)}{\Pi_\rho(L_s)}\right| \leq \epsilon$$

in which $L_s = EW^{1/2}\Gamma W^{1/2}E^{\mathrm{T}}$. One can show that the inequality (54) is equivalent to

$$\sup_{\substack{x \in \mathbb{R}^m \\ x \neq 0}} \frac{|x^{\mathrm{T}}(P - P\Gamma P)x|}{x^{\mathrm{T}}x} \leq \epsilon. \quad (55)$$

Since $\mathrm{Im}\{W^{1/2}E\} \subset \mathbb{R}^m$, it follows that

$$\sup_{\substack{x \in \mathrm{Im}\{W^{1/2}E\} \\ x \neq 0}} \frac{|x^{\mathrm{T}}(P - P\Gamma P)x|}{x^{\mathrm{T}}x} \leq \sup_{\substack{x \in \mathbb{R}^m \\ x \neq 0}} \frac{|x^{\mathrm{T}}(P - P\Gamma P)x|}{x^{\mathrm{T}}x}$$
$$\leq \epsilon.$$

Let us define $x = W^{1/2}Ex'$. Then, we can rewrite (55) as follows

$$\sup_{\substack{x' \in \mathbb{R}^n \\ x' \notin \ker\{W^{1/2}E\}}} \frac{|x'^{\mathrm{T}}(L - L_s)x'|}{x'^{\mathrm{T}}Lx'} \leq \epsilon. \quad (56)$$

For all $x' \in \ker\{W^{1/2}E\}$, one gets $x'^{\mathrm{T}}Lx' = x'^{\mathrm{T}}L_s x' = 0$. As a result, it follows that

$$\sup_{\substack{x' \in \mathbb{R}^n \\ x' \neq 0}} \frac{|x'^{\mathrm{T}}(L - L_s)x'|}{x'^{\mathrm{T}}Lx'} \leq \epsilon, \tag{57}$$

which implies that

$$(1-\epsilon)L \preceq L_s = EW^{1/2}\Gamma W^{1/2}E^{\mathrm{T}} \preceq (1+\epsilon)L. \tag{58}$$

Finally, using this and Theorem 3, we conclude the desired result. ■

Algorithm 1 produces a network abstraction with $\mathcal{O}(n \log n/\epsilon^2)$ feedback links in expectation and runs in approximately linear time $\tilde{\mathcal{O}}(m)$, where m is the number of links (c.f. [31]). This favorable almost-linear-time complexity is achieved by having access to good approximations of all effective resistances. In [16], the authors show that $\mathcal{O}(\log n)$ calls to a solver for a linear system of equations with symmetric diagonally dominant (SDD) matrix can provide sufficiently good approximations to all effective resistances. Moreover, it is shown in [17] that a spectral sparsification with $\mathcal{O}(n/\epsilon^2)$ links can be computed in $\mathcal{O}(n^3 m/\epsilon^2)$ time by employing a slower deterministic algorithm for link selection. The best known classical algorithm for calculating effective resistances relies on solving a Laplacian linear system and takes $\tilde{\mathcal{O}}(m)$ time [17], [32].

Remark 5: By putting together results of Theorems 4 and 5, we observe an intrinsic tradeoff between the number of feedback links M and the permissible performance loss parameter ϵ. These two design factors move in opposite directions, i.e., if we decrease M, ϵ increases, and vice versa. According to Theorems 4 and 5, one can deduce that the number of feedback links M decreases inversely with the square of ϵ.

C. Guaranteed Performance Bounds

In the following, we show that our proposed abstraction algorithm approximately preserves frequency characteristics of the original (dense) network (see Fig. 3). Our abstraction method shares some common roots with the classical model reduction techniques, where the objective is to find a reduced model that yields small $\mathcal{H}_2$-norm error (c.f., [8]). Our first result gives a tight upper bound on the $\mathcal{H}_2$-norm error of two linear consensus networks in terms of their Laplacian matrices.

Lemma 1: Suppose that $\mathfrak{N}(L)$ and $\mathfrak{N}(\hat{L})$ are two given consensus networks governed by dynamics (7). Then, we have

$$\frac{\|G - \hat{G}\|_{\mathcal{H}_2}^2}{\|G\|_{\mathcal{H}_2}^2} \leq \frac{\mathrm{Tr}\left(\hat{L}^\dagger + L^\dagger - 4(L+\hat{L})^\dagger\right)}{\mathrm{Tr}(L^\dagger)}, \tag{59}$$

where $G(s)$ and $\hat{G}(s)$ are transfer matrices of $\mathfrak{N}(L)$ and $\mathfrak{N}(\hat{L})$ from input ξ to output y, respectively.

Proof: In the first step, we define an augmented dynamical network $\mathfrak{N}^*$ using the two given networks

$$\begin{cases} \dot{z}(t) = -\begin{bmatrix} L + \frac{1}{n}J_n & 0 \\ 0 & L_s + \frac{1}{n}J_n \end{bmatrix} z(t) + \begin{bmatrix} M_n \\ M_n \end{bmatrix} \xi(t) \\ y(t) = \begin{bmatrix} M_n & -M_n \end{bmatrix} z(t) \end{cases} \tag{60}$$

in which $z \in \mathbb{R}^{2n}$, $\xi \in \mathbb{R}^n$, $y \in \mathbb{R}^n$. If we denote the transfer matrix of network (60) from ξ to y by G^*, then one can show that

$$\|G^*\|_{\mathcal{H}_2} = \|G - G_s\|_{\mathcal{H}_2}. \tag{61}$$

It is known that calculating the $\mathcal{H}_2$-norm of a linear time-invariant system reduces to solving an Algebraic Lyapunov Equation (ALE) [33]. Let us form the corresponding ALE to (61) using the state

matrices of the augmented network

$$\mathfrak{A}X + X\mathfrak{A} = \begin{bmatrix} M_n & -M_n \\ -M_n & M_n \end{bmatrix}, \tag{62}$$

where

$$\mathfrak{A} = -\begin{bmatrix} L + \frac{1}{n}J_n & 0_{n\times n} \\ 0_{n\times n} & L_s + \frac{1}{n}J_n \end{bmatrix}$$

and

$$X = \begin{bmatrix} X_1 & X_2 \\ X_2^{\mathrm{T}} & X_3 \end{bmatrix}.$$

The matrix equation (62) can be decomposed into three Sylvester equations as follow. The first equation is

$$\left(L + \frac{1}{n}J_n\right)X_1 + X_1\left(L + \frac{1}{n}J_n\right) = M_n$$

and its solution is given by

$$X_1 = \frac{1}{2}L^\dagger. \tag{63}$$

The second equation is

$$\left(L_s + \frac{1}{n}J_n\right)X_3 + X_3\left(L_s + \frac{1}{n}J_n\right) = M_n$$

and its unique solution is given by

$$X_3 = \frac{1}{2}L_s^\dagger. \tag{64}$$

Finally, the third one is

$$\left(L + \frac{1}{n}J_n\right)X_2 + X_2\left(L_s + \frac{1}{n}J_n\right) = -M_n \tag{65}$$

with unique solution

$$X_2 = -\int_0^\infty e^{-(L+\frac{1}{n}J_n)t}M_n e^{-(L_s+\frac{1}{n}J_n)t}dt,$$

where the integrand can be reformulated as

$$\begin{aligned} e^{-(L+\frac{1}{n}J_n)t}M_n e^{-(L_s+\frac{1}{n}J_n)t} &= e^{-(L+\frac{1}{n}J_n)t}e^{-(L_s+\frac{1}{n}J_n)t} \\ &+ \frac{1}{n}e^{-(L+\frac{1}{n}J_n)t}J_n e^{-(L_s+\frac{1}{n}J_n)t} \\ &= e^{-(L+\frac{1}{n}J_n)t}e^{-(L_s+\frac{1}{n}J_n)t} + \frac{e^{-2t}}{n}J_n. \end{aligned}$$

By utilizing the Golden-Thompson inequality for Hermitian matrices, it follows that

$$\begin{aligned} &\mathrm{Tr}\left(e^{-(L+\frac{1}{n}J_n)t}M_n e^{-(L_s+\frac{1}{n}J_n)t}\right) \\ &= \mathrm{Tr}\left(e^{-(L+\frac{1}{n}J_n)t}e^{-(L_s+\frac{1}{n}J_n)t} + \frac{e^{-2t}}{n}J_n\right) \\ &\geq \mathrm{Tr}\left(e^{-(L+L_s+\frac{2}{n}J_n)t} + \frac{e^{-2t}}{n}J_n\right) \tag{66} \end{aligned}$$

Therefore, the trace of X_2 can be bounded by

$$\begin{aligned} \mathrm{Tr}(X_2) &= -\mathrm{Tr}\left(\int_0^\infty e^{-(L+\frac{1}{n}J_n)t}M_n e^{-(L_s+\frac{1}{n}J_n)t}dt\right) \\ &= -\int_0^\infty \mathrm{Tr}\left(e^{-(L+\frac{1}{n}J_n)t}M_n e^{-(L_s+\frac{1}{n}J_n)t}\right)dt \\ &\leq -\int_0^\infty \mathrm{Tr}\left(M_n e^{-(L_s+L+\frac{2}{n}J_n)t}\right)dt \\ &= -\mathrm{Tr}\left((L+L_s)^\dagger\right). \tag{67} \end{aligned}$$

Putting all these pieces together, the $\mathcal{H}_2$-norm of the augmented network can be written as

$$\|G^*\|_{\mathcal{H}_2}^2 = \mathrm{Tr}\left(\begin{bmatrix} M_n & M_n \end{bmatrix}\begin{bmatrix} X_1 & X_2 \\ X_2 & X_3 \end{bmatrix}\begin{bmatrix} M_n \\ M_n \end{bmatrix}\right)$$
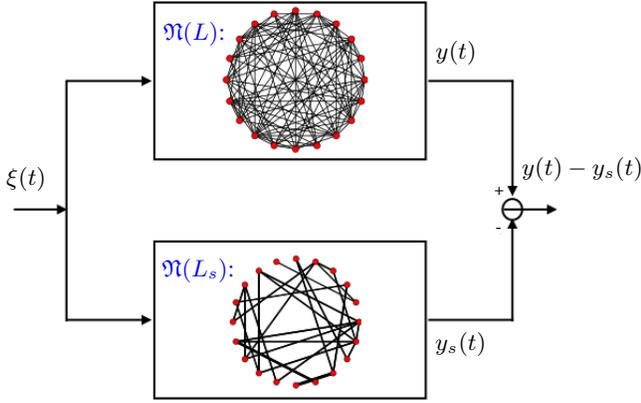
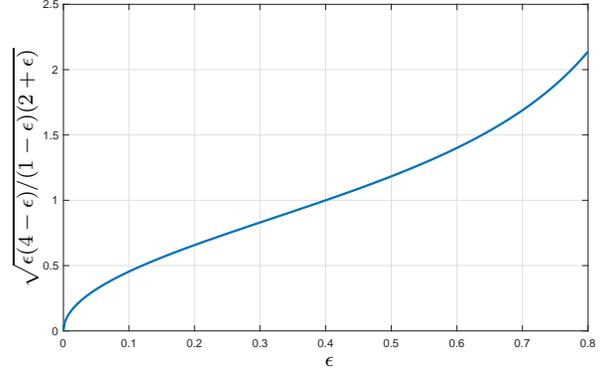Fig. 3: The block diagram of the augmented network (60).

Fig. 4: This plot presents the upper bound given by Theorem 6 on the $\mathcal{H}_2$-norm error of a consensus network and its ϵ-approximation network.

$$= \operatorname{Tr}\left(\begin{bmatrix} X_1 & X_2 \\ X_2 & X_3 \end{bmatrix} \begin{bmatrix} M_n & M_n \\ M_n & M_n \end{bmatrix} \right). \tag{68}$$

From (63), (64) and (68), it follows that

$$\begin{aligned} \|G^*\|_{\mathcal{H}_2}^2 &= \frac{1}{2}\operatorname{Tr}(L^\dagger + L_s^\dagger) - \operatorname{Tr}((X_2 + X_2^{\mathrm{T}})M_n) \\ &\le \frac{1}{2}\operatorname{Tr}(L^\dagger + L_s^\dagger) - 2\operatorname{Tr}\left((L+L_s)^\dagger\right), \end{aligned}$$

where (67) is used in the last inequality. Finally, from this and (61), we conclude the desired result

$$\|G - G_s\|_{\mathcal{H}_2}^2 \le \frac{1}{2}\operatorname{Tr}\left(\hat{L}^\dagger + L^\dagger - 4(L+\hat{L})^\dagger\right).$$

The right-hand side of inequality (59) is always non-negative, i.e.,

$$0 \le \operatorname{Tr}\left(\hat{L}^\dagger + L^\dagger - 4(L+\hat{L})^\dagger\right).$$

This is because of the fact that $\operatorname{Tr}(L^\dagger)$ is convex on $\mathfrak{L}_n$ and the following inequality holds

$$\operatorname{Tr}\left((L+\hat{L})^\dagger\right) \le \frac{1}{4}\operatorname{Tr}\left(L^\dagger\right) + \frac{1}{4}\operatorname{Tr}\left(\hat{L}^\dagger\right).$$

The inequality (40) implies proximity of state matrices of the original and its abstraction on the cone of positive semidefinite matrices. In the following result, it is proven that the frequency specifications of two ϵ-approximations are indeed very similar in $\mathcal{H}_2$ sense.

Theorem 6: If $\mathfrak{N}(L_s)$ is an ϵ-approximation of $\mathfrak{N}(L)$ for some $0 \le \epsilon < 1$, then

$$\frac{\|G - G_s\|_{\mathcal{H}_2}}{\|G\|_{\mathcal{H}_2}} \le \sqrt{\frac{\epsilon(4-\epsilon)}{(1-\epsilon)(2+\epsilon)}}, \tag{69}$$

where $G(s)$ and $G_s(s)$ represent the transfer matrices from input ξ to output y of $\mathfrak{N}(L)$ and $\mathfrak{N}(L_s)$, respectively.

Proof: According to the definition of ϵ-approximation and the fact that $\operatorname{Tr}(L^\dagger)$ is a homogeneous systemic performance measure of order -1, we get

$$\operatorname{Tr}(L_s^\dagger) \le \frac{1}{1-\epsilon}\operatorname{Tr}(L^\dagger). \tag{70}$$

Moreover, it follows from Theorem 3 that

$$(2+\epsilon)^{-1}(L)^\dagger \le (L+L_s)^\dagger. \tag{71}$$

By taking trace from both sides of (71), one obtains

$$\operatorname{Tr}\left((L+L_s)^\dagger\right) \ge \frac{1}{2+\epsilon}\operatorname{Tr}\left(L^\dagger\right). \tag{72}$$

By applying result of Lemma 1, we can conclude that

$$\begin{aligned} \frac{\|G - G_s\|_{\mathcal{H}_2}^2}{\|G\|_{\mathcal{H}_2}^2} &\le \frac{\operatorname{Tr}\left(\hat{L}^\dagger + L^\dagger - 4(L+\hat{L})^\dagger\right)}{\operatorname{Tr}(L^\dagger)} \\ &= \frac{\operatorname{Tr}\left(L_s^\dagger\right) + \operatorname{Tr}\left(L^\dagger\right) - 4\operatorname{Tr}\left((L+L_s)^\dagger\right)}{\operatorname{Tr}(L^\dagger)} \\ &\le \frac{\epsilon(4-\epsilon)}{(1-\epsilon)(2+\epsilon)}, \end{aligned} \tag{73}$$

where the last inequality is obtained after using (70) and (72). ∎

Fig. 4 depicts the upper bound in inequality (69) for the relative $\mathcal{H}_2$-norm error of a linear consensus network and its ϵ-approximation.

Corollary 1: Suppose that y is the output of network $\mathfrak{N}(L)$ and y_s is the output of its corresponding ϵ-approximation $\mathfrak{N}(L_s)$ for some $0 \le \epsilon < 1$. Then, the steady-state expected value of their output error can be bounded by

$$\lim_{t\to\infty} \mathbb{E}\left\{ \|y(t) - y_s(t)\|_2^2 \right\} \le \frac{\epsilon(4-\epsilon)}{2(1-\epsilon)(2+\epsilon)}\operatorname{Tr}\left(L^\dagger\right).$$

The proof of this corollary is based on the proof of Theorem 6.

Remark 6: In [15], the authors consider a similar problem of identifying a sparse representation of a given dense linear consensus network. Their proposed method has two steps. First, an optimal sparse network topology is obtained by adding the $\mathcal{H}_2$-norm error (c.f., Fig. 3) with another penalizing term that accounts for sparsity. Then, the optimal link weights are chosen over the identified topology. For both cases, approximation methods based on the Broyden-Fletcher-Goldfarb-Shanno method are employed in [15]. This method chooses a descent direction based on an approximation of the Hessian matrix. Therefore each update requires $\mathcal{O}(n^6)$ operations [15]. In comparison with Algorithm 1 in Section V, the proposed method in [15] is computationally expensive. Furthermore, it does not provide any guaranteed performance certificates.

VI. LOCALIZED NETWORK ABSTRACTION

Our methodology can be extended further to explore several interesting network design problems, such as partial or localized abstraction of a given large-scale consensus network. In this section, we only look at one of such design problems. Let us consider a slightly modified version of (7) by involving a predesigned state

feedback controller

$$\mathfrak{N}(L_0 + L_1) : \begin{cases} \dot{x}(t) = -L_0 x(t) + u(t) + \xi(t) \\ u(t) = -L_1 x(t) \\ y(t) = M_n x(t) \end{cases} \quad (74)$$

with initial condition $x(0) = x_0$, where L_0 is the Laplacian matrix of the open-loop network and the Laplacian matrix L_1 is the predesigned state feedback gain matrix. Let us represent the corresponding coupling graph to L_1 by $\mathcal{G}_1$. If L_1 is obtained via traditional optimal control methods without incorporating sparsity measures, then one should expect to get a dense interconnection topology for $\mathcal{G}_1$; we refer to [18] for discussions on spatially decaying structure of optimal controllers. Therefore, our design objective is to compute a localized abstraction for the closed-loop network $\mathfrak{N}(L_0 + L_1)$ that only sparsifies $\mathcal{G}_1$. Let us represent such an abstraction by $\hat{\mathcal{G}}_1$ with Laplacian $\hat{L}_1$.

Theorem 7: Suppose that a linear consensus network with structure (74), a homogeneous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$ of order $-\alpha$, and a design parameter $d > 2$ are given. For $\epsilon = \frac{\sqrt{8d}}{d+2}$, there exists a subgraph abstraction $\hat{\mathcal{G}}_1 = (\mathcal{V}, \hat{\mathcal{E}}, \hat{w})$ of $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}, w)$ with at most $dn/2$ links that satisfies $\hat{\mathcal{E}} \subset \mathcal{E}$ and

$$\left| \frac{\Pi_\rho(L_0 + L_1) - \Pi_\rho(L_0 + \hat{L}_1)}{\Pi_\rho(L_0 + \hat{L}_1)} \right| \leq \epsilon. \quad (75)$$

Furthermore, it follows that

$$w_{\text{total}}(\hat{L}_1) \leq (1 + \epsilon) w_{\text{total}}(L_1)$$

in which $w_{\text{total}}(L) = \frac{1}{2} \operatorname{Tr}(L)$.

Proof: According to [17, Th. 1.1], coupling graph $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}, w)$ has a weighted subgraph $\mathcal{H} = (\mathcal{V}, \hat{\mathcal{E}}, \hat{w})$ with $|\hat{\mathcal{E}}| = \lceil d(n-1)/2 \rceil$ that satisfies

$$L_1 \preceq L_{\mathcal{H}} \preceq \frac{d + 2 + \sqrt{8d}}{d + 2 - \sqrt{8d}} L_1 \quad (76)$$

where $L_{\mathcal{H}}$ is the Laplacian matrix of graph $\mathcal{H}$. We define $\hat{\mathcal{G}}_1 = (\mathcal{V}, \hat{\mathcal{E}}, \hat{w})$ based on the following Laplacian matrix

$$\hat{L}_1 := \frac{(1 - \sqrt{d/2})^2}{1 + d/2} L_{\mathcal{H}}. \quad (77)$$

From (76) and (77), it follows that

$$\left(1 - \frac{\sqrt{2d}}{d/2 + 1} \right) L_1 \preceq \hat{L}_1 \preceq \left(1 + \frac{\sqrt{2d}}{d/2 + 1} \right) L_1. \quad (78)$$

Moreover, we know that

$$\left(1 - \frac{\sqrt{2d}}{d/2 + 1} \right) L_0 \preceq L_0 \preceq \left(1 + \frac{\sqrt{2d}}{d/2 + 1} \right) L_0. \quad (79)$$

From (78) and (79), we have

$$\left(1 - \frac{\sqrt{8d}}{d + 2} \right)(L_0 + L_1) \preceq (L_0 + \hat{L}_1)$$
$$\preceq \left(1 + \frac{\sqrt{8d}}{d + 2} \right)(L_0 + L_1). \quad (80)$$

Using (80) and the result of Theorem 3, it yields that $\mathfrak{N}(L_0 + \hat{L}_1)$ is a $(\frac{\sqrt{8d}}{d+2}, d)$-abstraction of $\mathfrak{N}(L_0 + L_1)$. ∎

This result is particularly useful is abstraction of large-scale consensus networks where the control objective is to abstract only desired parts of a network without drastically affecting the global performance.

VII. PARALLEL NETWORK ABSTRACTION

Building upon the results of the previous section, we introduce a distributed and parallel implementation of our proposed localized algorithm. The main advantage of parallel abstraction is that several localized abstraction problems can be solved simultaneously by dividing a large-scale network abstraction problem into smaller in size localized problems.

Definition 4: A base subgraph $\mathcal{G}_0$ of a network is a subgraph which is formed by those feedback links that will stay unchanged throughout the abstraction process.

Let us denote Laplacian matrix of a base subgraph $\mathcal{G}_0$ by L_0. For a given natural number p, the original coupling graph can be partitioned as the union of a base subgraph and p other subgraphs $\mathcal{G}_1, \ldots, \mathcal{G}_p$, i.e.,

$$L = L_0 + \sum_{i=1}^{p} L_i, \quad (81)$$

where L_i is Laplacian matrix of subgraph $\mathcal{G}_i$. Without loss of generality, one may assume that the node set of all subgraphs is $\mathcal{V}$, where $\mathcal{V}$ is the node set of the original graph. This assumption implies that the corresponding Laplacian matrices are compatible n-by-n matrices. The network setup for parallel implementation of abstraction takes the following form

$$\begin{cases} \dot{x}(t) = -L_0 x(t) + u(t) + \xi(t) \\ u(t) = -\sum_{i=1}^{p} L_i x(t) \\ y(t) = M_n x(t) \end{cases} \quad (82)$$

with initial condition $x(0) = x_0$.

Assumption 4: The corresponding subgraphs to L_i for $i = 1, \ldots, p$ are link-disjoint and dense.

Theorem 8: Suppose that a linear consensus network with structure (82), a homogeneous systemic performance measure $\rho : \mathfrak{L}_n \to \mathbb{R}_+$ of order $-\alpha$, and a number $d > 2$ are given. If the coupling graph of the network can be decomposed as (81), then for $\epsilon = \frac{\sqrt{8d}}{d+2}$ there exists a set of subgraph sparsifier $\{\hat{\mathcal{G}}_i\}_{i=1}^{p}$ for $\{\mathcal{G}_i\}_{i=1}^{p}$ where each sparsifier subgraph has at most $dn/2$ links and the global performance index satisfies

$$\left| \frac{\Pi_\rho(L_0 + \sum_{i=1}^{p} L_i) - \Pi_\rho(L_0 + \sum_{i=1}^{p} \hat{L}_i)}{\Pi_\rho(L_0 + \sum_{i=1}^{p} \hat{L}_i)} \right| \leq \epsilon. \quad (83)$$

Furthermore, it follows that

$$w_{\text{total}}(\hat{L}_i) \leq (1 + \epsilon) w_{\text{total}}(L_i),$$

for $i = 1, 2, \cdots, p$.

Proof: For every $i = 1, \cdots, p$, coupling subgraph $\mathcal{G}_i = (\mathcal{V}, \mathcal{E}_i, w_i)$ has a weighted subgraph $\mathcal{H}_i = (\mathcal{V}, \hat{\mathcal{E}}_i, \hat{w}_i)$ with $|\hat{\mathcal{E}}_i| = \lceil d(n-1)/2 \rceil$ whose Laplacian satisfies

$$L_i \preceq L_{\mathcal{H}_i} \preceq \frac{d + 1 + 2\sqrt{d}}{d + 1 - 2\sqrt{d}} L_i \quad (84)$$

where $L_{\mathcal{H}_i}$ is the Laplacian matrix of graph $\mathcal{H}_i$ [17, Th. 1.1]. We define every $\hat{\mathcal{G}}_i = (\mathcal{V}, \hat{\mathcal{E}}_i, \hat{w}_i)$ according to the following Laplacian matrix

$$\hat{L}_i := \frac{(1 - \sqrt{d/2})^2}{1 + d/2} L_{\mathcal{H}_i}. \quad (85)$$

By putting (84) and (85) together, it follows that

$$\left(1 - \frac{\sqrt{8d}}{d + 2} \right) L_i \preceq \hat{L}_i \preceq \left(1 + \frac{\sqrt{8d}}{d + 2} \right) L_i. \quad (86)$$

Moreover, we know that

$$\left(1 - \frac{\sqrt{8d}}{d+2}\right) L_0 \preceq L_0 \preceq \left(1 + \frac{\sqrt{8d}}{d+2}\right) L_0. \qquad (87)$$

From (86) and (87), we get

$$\left(1 - \frac{\sqrt{2d}}{d/2+1}\right)\left(L_0 + \sum_{i=1}^{p} L_i\right) \preceq L_0 + \sum_{i=1}^{p} \hat{L}_i$$
$$\preceq \left(1 + \frac{\sqrt{2d}}{d/2+1}\right)\left(L_0 + \sum_{i=1}^{p} L_i\right). \quad (88)$$

From (88) and the result of Theorem 3, one can conclude that $\mathfrak{N}(L_0 + \sum_{i=1}^{p} \hat{L}_i)$ is a $(\frac{\sqrt{8d}}{d+2}, d)$-abstraction of $\mathfrak{N}(L_0 + \sum_{i=1}^{p} L_i)$. ∎

Algorithm 1 can be implemented on several parallel localized processing units to abstract every $\mathcal{G}_i$ for $i = 1, \cdots, p$. This parallelization scheme cuts the time complexity of solving an abstraction problem down to $\tilde{\mathcal{O}}(m_{\max})$, where $m_{\max}$ is the number of links of the densest graph among $\mathcal{G}_1, \ldots, \mathcal{G}_p$.

VIII. Shortcomings of ℓ_0/ℓ_1-Regularized Sparsification Methods

In order to put our proposed methodology into perspective, we discuss some of the shortcomings of the ℓ_0/ℓ_1-regularization based sparsification methods. The common approach is to formulate an optimal control problem that is augmented by a penalty term to promote sparsity. The resulting optimal control problem can be usually cast as a bilinear matrix optimization problem and convexified using alternating methods [5], [34]. While ℓ_0/ℓ_1-regularization based methods generally do not scale with network size and suffer from high time complexities, we would like to accentuate a more important issue by means of an example that shows ℓ_0/ℓ_1-regularized methods do not always return sparse solutions. Let us consider a linear consensus network that is governed by (7) whose feedback structure is represented by a complete graph with identical link weights w_0, Laplacian matrix L_0, link set $\mathcal{E}_0$, and incidence matrix E_0. The control objective is to eliminate or reweigh some links of the network while minimizing $\mathcal{H}_2$-norm of the network from ξ to y. This problem can be formulated as the following ℓ_0-problem

$$\underset{w(e)}{\text{Minimize}} \quad \|G_F\|_{\mathcal{H}_2}^2 + \frac{\gamma}{2} \|A_F\|_{\ell_0} \qquad (89)$$

subject to:

$$F = \sum_{e \in \mathcal{E}_0} w(e) b_e b_e^{\mathsf{T}} \qquad (90)$$

$$w(e) \leq w_0 \qquad (91)$$

$$L_0 - F \in \mathfrak{L}_n \qquad (92)$$

in which $G_F(s)$ is the transfer matrix of network $\mathfrak{N}(L_0 - F)$ from ξ to y, A_F is the adjacency matrix of $\mathfrak{N}(L_0 - F)$, and b_e is the column of incidence matrix E_0 that corresponds to link e. Boundedness of the cost function as well as constraint (92) ensures connectivity of the coupling graph and positive semi-definiteness of the Laplacian matrix of $\mathfrak{N}(L_0 - F)$. In the penalty term, $\gamma > 0$ is a design parameter and ℓ_0 sparsity measure of matrix A_F is equal to the total number of nonzero elements of A_F, which is defined by (2). For an undirected graph, the value of this sparsity measure is equal to twice the number of the edges. The value of cost function (89) is greater than or equal to $\gamma(n-1)$. One can get arbitrarily close to this lower bound through the following steps. Let us select a spanning tree $\mathcal{T}_0$ from link set $\mathcal{E}_0$ and set elements of F in a way that $L_0 - F$ becomes the Laplacian matrix of $\mathcal{T}_0$ with link weights $w_0 - w(e)$. It is known that quantity $\|G_F\|_{\mathcal{H}_2}$ is bounded if and only if the corresponding coupling graph to $L_0 - F$ has at least $n-1$ links [12], which in our case $\mathcal{T}_0$ does have

$n-1$ links. Therefore, the value of $\|G_F\|_{\mathcal{H}_2}$ can be made arbitrarily close to 0 by letting $w(e)$ get closer to $-\infty$.

After relaxing ℓ_0-measure by ℓ_1-norm and representing the $\mathcal{H}_2$-norm in terms of Laplacian spectrum, the objective function (89) can be rewritten as

$$\frac{1}{2} \sum_{i=2}^{n} \lambda_i (L_0 - F)^{-1} + \frac{\gamma}{2} \|A_F\|_{\ell_1} \qquad (93)$$

where $\|A_F\|_{\ell_1}$ is equal to the sum of the absolute values of all elements of A_F and

$$\|A_F\|_{\ell_1} = 2 \sum_{e \in \mathcal{E}_0} (w_0 - w(e)) = \text{Tr}(L_0 - F).$$

The cost function (93) can be simplified further to obtain

$$\frac{1}{2} \sum_{i=2}^{n} \lambda_i (L_0 - F)^{-1} + \frac{\gamma}{2} \sum_{i=2}^{n} \lambda_i (L_0 - F). \qquad (94)$$

Thus, the ℓ_1-problem is the minimization problem with cost function (94) and constraints (90)-(92). This problem is convex and has a unique solution. This follows from the following facts that $\mathcal{H}_2$-norm is a systemic measure and convex, trace is a linear operator and convex, and sum of these two convex functions results in a convex cost function (94). The inequality (91) also represents a convex set. Moreover, positive semi-definiteness of $L_0 - F$ and connectivity of the resulting network $\mathfrak{N}(L_0 - F)$ are both convex constraints; therefore, constraint (92) is convex. By applying the arithmetic and geometric means inequality, one can show that the spectral function (94) is lower bounded by constant $(n-1)\sqrt{2\gamma}$. By respecting all the constraints and for all design parameters $\gamma > \frac{1}{8n^2 w_0^2}$, this lower bound can be achieved by a complete graph with identical link weights $w(e)^* = \frac{1}{n\sqrt{2\gamma}} - w_0$. An interesting scenario happens when $\gamma = 2$ and $w_0 > \frac{1}{4}n^{-1}$ where the lower bounds on the least achievable cost values for both ℓ_0-problem and its relaxed ℓ_1-problem coincide and become $2(n-1)$. In this case, the optimal networks from solving the ℓ_0-problem and ℓ_1-problem have very different sparsity patterns: the ℓ_0-problem solution is a network with a spanning tree topology and arbitrarily large link weights, while ℓ_1-problem identifies a network whose coupling graph is complete with identical link weights. This shows that ℓ_1-relaxation of a ℓ_0-regularized optimal control does not always provide sparse solutions.

IX. Illustrative Examples

In this section, we present several numerical examples to illustrate our theoretical findings.

Example 1: We first consider a consensus network with 40 agents defined over an unweighted coupling graph with two dense components which are connected by a single link, i.e., a cut edge. Each of the components is obtained by adding 100 uniformly and randomly selected links to an empty graph with 20 nodes; see Fig. 5(a). Based on Algorithm 1, we sample low-connectivity coupling links (i.e., feedback gains) with high probability and high-connectivity coupling links with low probability. The probability of selecting a link from the coupling graph is depicted in Fig. 6. One observes that the probability of selecting the cut edge as an important link is much higher than the probability of choosing other links. Fig. 5(b) shows a (0.5, 3.05)-abstraction of the network after applying Algorithm 1 that has 61 links and meets all requirements of Definition 3. The coupling graph of the abstraction is weighted and has about 70% fewer links than the original network. Although, we set $\epsilon = 0.5$ when running Algorithm 1, the performance loss of the resulting abstraction is less than 24% according to Table II.
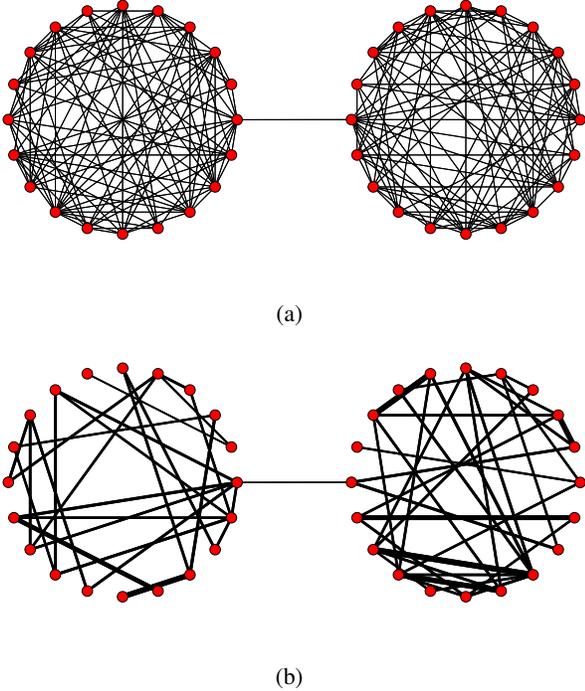
(a)

(b)

Fig. 5: (a) An unweighted coupling graph with 40 nodes, 201 links, $\|G\|_{\mathcal{H}_2} = 2.7837$, and $w_{\text{total}}(L) = 201$. (b) An abstraction of Fig. 5(a) with 40 nodes, 61 links, $\|G_s\|_{\mathcal{H}_2} = 3.0805$, and $w_{\text{total}}(L_s) = 199.88$.
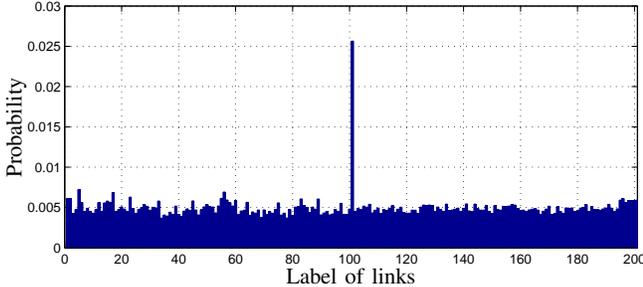
Systemic Performance Measure	$\frac{|\Pi_\rho(L_s) - \Pi_\rho(L)|}{\Pi_\rho(L_s)} \times 100$
System Hankel Norm: $\frac{1}{2}\lambda_2^{-1}$	19.65 %
Squared $\mathcal{H}_2$ Norm: $\frac{1}{2}\sum_{i=2}^{n}\lambda_i^{-1}$	18.34 %
Zeta Spectral Norm: $\left(\sum_{i=2}^{n}\lambda_i^{-2}\right)^{\frac{1}{2}}$	15.26 %
Local Deviation Error: $\frac{1}{2}\sum_{i\in\mathcal{V}}d_i^{-1}$	23.16 %

TABLE II: Relative performance loss percentage of network Fig. 5(a) with respect to its abstraction Fig. 5(b).

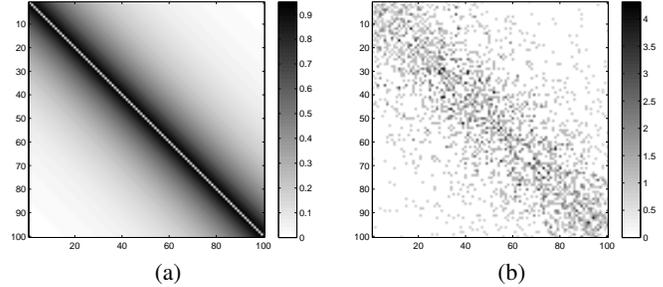(a)   (b)

Fig. 7: (a) This plot demonstrates the sparsity pattern of adjacency matrix of network in Example 2 with coupling parameters $c = 1$ and $\gamma = 0.05$. This network has 100 agents and $4,950$ links and the color intensity at each pixel of this plot shows magnitude of the corresponding element in the adjacency matrix. (b) This plot depicts the sparsity pattern of adjacency matrix of an abstraction that has 1114 links. The relative $\mathcal{H}_2$ error of these two networks is $\|G - G_s\|_{\mathcal{H}_2}/\|G\|_{\mathcal{H}_2} = 0.18$ and the ratio of their total weights is $w_{\text{total}}(L_s)/w_{\text{total}}(L) = 1.0028$.

Fig. 6: The probability of selecting a link of the coupling graph given by Fig. 5(a) as an important link. A cut edge is the most important link with the highest probability: if we throw out the cut edge, the coupling graph of the resulting abstraction will be disconnected for sure.

Example 2: Let us consider a consensus network with 100 agents and exponentially decaying couplings that are defined by

$$w(\{i,j\}) = \begin{cases} c\,e^{-\gamma|i-j|} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad (95)$$

where c and γ are positive numbers and $i, j \in \mathcal{V}$. This class of networks arises in various applications where there is a notion of spatial distance between the subsystems; we refer to [18] for more details. Fig. 7(a) shows the adjacency matrix of the coupling graph of this network. According to Theorem 4, this network has a $(0.5, 27.85)$-abstraction. Fig. 7(b) illustrates a $(0.5, 22.28)$-abstraction of the network after applying Algorithm 1, where the design parameter ϵ is set to 0.5. The original network has $4,950$ links, while its abstraction has $1,114$ coupling links and meets all requirements of Definition 3. As a result, abstraction achieves 77.49 % sparsification. Although, in

this example, we allow 50% performance loss, numerous simulation examples assert that the resulting bounds for performance loss can be comparably smaller. As it is seen from Table III, the relative performance loss percentage is less than 11%. Fig. 8 depicts the probability distribution of sampling process to select important links in Algorithm 1, where the color intensity of pixels show how important that link is. According to Algorithm 1, low-connectivity coupling links are sampled with higher probability than high-connectivity coupling links.

Example 3: Let us consider a dynamical network consists of 100 agents which are randomly distributed in a 30×30 square-shape area in space and are coupled over a proximity graph. Every agent is connected to all of its spatial neighbors within a closed ball of radius $r = 10$. Fig. 9(a) shows the resulting coupling graph of this dynamical network that has 100 nodes and $1,291$ links and Fig. 9(b) depicts an $(0.5, 16.62)$-abstraction of this network, which is obtained using Algorithm 1 with $\epsilon = 0.5$. The number of coupling links in this abstraction is 831, which is 35.63% sparsification, and meets all requirements of Definition 3. Table IV summarizes the percentage of the relative performance loss with respect to some systemic performance measures.

Example 4: Let us consider a simple model for the formation control of a group of autonomous vehicles, which is given by

$$\begin{bmatrix} \dot{x}(t) \\ \dot{v}(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ -L & -\beta L \end{bmatrix} \begin{bmatrix} x(t) \\ v(t) \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} \xi(t) \quad (96)$$

$$y(t) = M_n v(t) \quad (97)$$

where $\beta > 0$ is a design parameter. Each vehicle has a position and a velocity variable. The state variable of the entire network is denoted by $\begin{bmatrix} x(t) & v(t) \end{bmatrix}^{\mathrm{T}}$ and is measured relative to a pre-specified desired trajectory $x_d(t)$ and velocity $v_d(t)$. Without loss of generality, we may assume that the position and velocity of each

Systemic Performance Measure	$\frac{\lvert\Pi_\rho(L_s)-\Pi_\rho(L)\rvert}{\Pi_\rho(L_s)}\times 100$
System Hankel Norm: $\frac{1}{2}\lambda_2^{-1}$	10.72 %
Squared $\mathcal{H}_2$ Norm: $\frac{1}{2}\sum_{i=2}^{n}\lambda_i^{-1}$	6.44 %
Zeta Spectral Norm: $\left(\sum_{i=2}^{n}\lambda_i^{-2}\right)^{\frac{1}{2}}$	9.69 %
Local Deviation Error: $\frac{1}{2}\sum_{i\in\mathcal{V}}d_i^{-1}$	3.07 %

TABLE III: Relative performance loss percentage of network Fig. 7(a) with respect to its abstraction Fig. 7(b) that has 77.49 % fewer links.
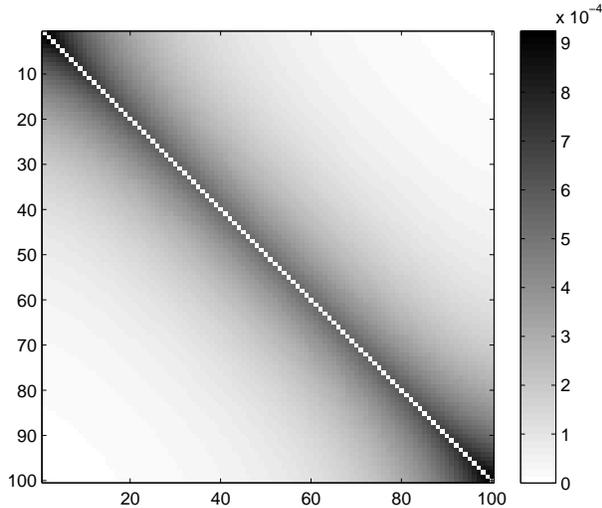

Fig. 8: This plot presents the probability distribution of the sampling process to choose important links in Example 2. The color intensity of each pixel shows importance of that link for sampling.
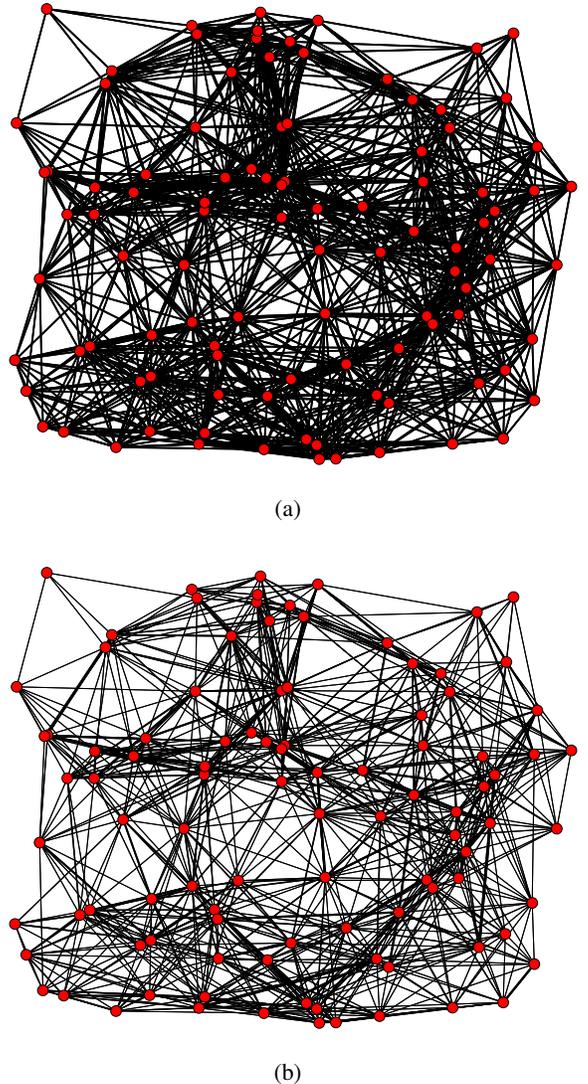

(a)

(b)

Fig. 9: (a) An unweighted coupling (proximity) graph of a consensus network with 100 agents is presented. Every agent is connected to all of its spatial neighbors within a closed ball of radius $r = 10$. This graph has $1,291$ links and $w_{\text{total}}(L) = 1,291$. (b) This graph shows a $(0.5, 16.62)$-abstraction of the network. Our abstraction algorithm results in a network with a weighted coupling graph that has 831 links and $w_{\text{total}}(L_s) = 1293.4$. The relative $\mathcal{H}_2$ error of these two networks is $\lVert G - G_s\rVert_{\mathcal{H}_2}/\lVert G\rVert_{\mathcal{H}_2} = 0.17$.

vehicle are scalar variables. The reason is that one can decouple higher D-dimensional models into D decoupled (96)-(97) models. The overall objective is for the network to reach a desired formation pattern, where each autonomous vehicle travels at the constant desired velocity v_d while preserving a pre-specified distance between itself and each of its neighbors. In this model, the state feedback controller uses both position and velocity measurements and L is, in fact, the corresponding feedback gain, which represents the coupling topology in the controller array, and constant β is a design parameter [28], [35]. We consider the steady-state variance of the performance output of this network as the performance measure. This quantity is indeed equivalent to the square of the $\mathcal{H}_2$-norm of the system from the exogenous disturbance input to the performance output [28], [35]–[39]. The squared $\mathcal{H}_2$-norm of (96)-(97) can be characterized in terms of Laplacian eigenvalues of the coupling graph as follows

$$\Theta_2^2(L) := \lim_{t\to\infty}\mathbb{E}\left[y^{\mathrm{T}}(t)y(t)\right] = \frac{1}{2\beta}\sum_{i=2}^{n}\lambda_i^{-2} = \frac{1}{2\beta}\zeta_2^2(L). \quad (98)$$

We refer to [11] for more details. This quantity is a homogeneous systemic performance measure; therefore, we can apply our abstraction algorithm. Suppose that the coupling graph of network (96)-(97) is given by Fig. 9(a). Then, as we mentioned in Example 3, Fig. 9(b) illustrates one example of $(0.5, 16.62)$-abstraction of this network that is obtained using Algorithm 1. As shown in Example 3, this network has fewer coupling links (831 links) compare to the original network. Table V presents the percentage of the relative performance loss, where both systemic performance measures are homogeneous of order -2. This example shows that our proposed abstraction algorithm can

be successfully applied to second-order linear consensus networks as well.

X. DISCUSSION

We have introduced a notion of abstraction for a class of linear consensus networks based on notions of spectral sparsification. There have been several close-in-spirit notions of graph sparsifications in the context of theoretical computer science. While these other notions are interesting for their own sake from a combinatorial standpoint, their connections to performance analysis and synthesis of dynamical networks are not trivial and require further scrutiny. In this context, for a given graph there are several ways to define sparse subgraphs, namely, distance sparsifiers that approximate all pairwise distances

Systemic Performance Measure	$\frac{|\Pi_\rho(L_s)-\Pi_\rho(L)|}{\Pi_\rho(L_s)} \times 100$
Hankel Norm: $\frac{1}{2}\lambda_2^{-1}$	19.65 %
Squared $\mathcal{H}_2$-Norm: $\frac{1}{2}\sum_{i=2}^n \lambda_i^{-1}$	18.34 %
Zeta Spectral Norm: $\left(\sum_{i=2}^n \lambda_i^{-2}\right)^{\frac{1}{2}}$	15.26 %
Local Deviation Error: $\frac{1}{2}\sum_{i\in\mathcal{V}} d_i^{-1}$	23.16 %

TABLE IV: Relative performance loss percentage of network Fig. 9(a) with respect to its abstraction Fig. 9(b) that has 35.63% fewer links.

Systemic Performance Measure	$\frac{|\Pi_\rho(L)-\Pi_\rho(L_s)|}{\Pi_\rho(L_s)} \times 100$
Squared $\mathcal{H}_2$-Norm: $\frac{1}{2\beta}\sum_{i=2}^n \lambda_i^{-2}$	17.58 %
Local Deviation Error: $\frac{1}{2\beta}\sum_{i=1}^n d_i^{-2}$	11.38 %

TABLE V: Relative performance loss percentage of network (96)-(97) with coupling graph Fig. 9(a) with respect to its abstraction Fig. 9(b).

up to a multiplicative and/or additive error (see [40] and subsequent research on spanners), cut sparsifiers that approximate every cut to an arbitrarily small multiplicative error [41], spectral sparsifier that approximate every eigenvalue to an arbitrarily small multiplicative error [31], and many more. It is shown that sparsifiers can be constructed by sampling links according to their strength, effective resistance [16], edge connectivity [32], or by sampling random spanning trees [42]. Benczúr and Karger propose a randomized algorithm to construct a cut sparsifier in $\mathcal{O}(m\log^2 n)$ time for unweighted graphs and $\mathcal{O}(m\log^3 n)$ time for weighted graphs [32], [41]. The notion of spectral sparsifier is stronger than cut sparsifier, which implies spectral sparsifiers are also cut sparsifiers. One of our current research directions is to investigate all these methods of sparsification and explore their connections to dynamical networks.

In our analysis, we assumed that the input matrix of the network is an identity matrix. Further extension of the current work includes linear consensus networks over directed coupling graphs with arbitrary input matrices. However, this generalization is challenging and even the problem formulation demands some appropriate conditions to guarantee the boundedness of performance measures.

Acknowledgment

References

[1] M. Siami and N. Motee, "Systemic measures for performance and robustness of large–scale interconnected dynamical networks," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, Dec. 2014, pp. 5119–5124.

[2] ——, "Schur–convex robustness measures in dynamical networks," in *Proceedings of the 2014 American Control Conference*, 2014, pp. 5198–5203.

[3] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, Sept 2004.

[4] M. Siami and N. Motee, "Robustness and performance analysis of cyclic interconnected dynamical networks," in *Proceedings of the 2013 SIAM Conference on Control and Its Application*, Jan. 2013, pp. 137–143.

[5] M. Fardad, F. Lin, and M. R. Jovanović, "Design of optimal sparse interconnection graphs for synchronization of oscillator networks," *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2457–2462, 2014.

[6] T. H. Summers, F. L. Cortesi, and J. Lygeros, "On submodularity and controllability in complex dynamical networks," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 91–101, March 2016.

[7] A. Girard and G. J. Pappas, "Approximation metrics for discrete and continuous systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 5, pp. 782–798, May 2007.

[8] S. Gugercin, A. C. Antoulas, and C. Beattie, "h_2 model reduction for large-scale linear dynamical systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 2, pp. 609–638, 2008.

[9] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, no. 1, pp. 17–32, Feb 1981.

[10] M. Johansson and A. Rantzer, "Computation of piecewise quadratic lyapunov functions for hybrid systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 555–559, Apr 1998.

[11] M. Siami and N. Motee, "Fundamental limits on robustness measures in networks of interconnected systems," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, Dec. 2013, pp. 67–72.

[12] ——, "Fundamental limits and tradeoffs on disturbance propagation in linear dynamical networks," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 4055–4062, 2016.

[13] X. Wu and M. R. Jovanović, "Sparsity-promoting optimal control of consensus and synchronization networks," in *Proceedings of the 2014 American Control Conference*, 2014, pp. 2948–2953.

[14] S. Hassan-Moghaddam and M. R. Jovanović, "An interior point method for growing connected resistive networks," in *Proceedings of the 2015 American Control Conference*, Chicago, IL, 2015, pp. 1223–1228.

[15] N. Dhingra, F. Lin, M. Fardad, and M. R. Jovanović, "On identifying sparse representations of consensus networks," *IFAC Proceedings Volumes*, vol. 45, no. 26, pp. 305 – 310, 2012.

[16] D. A. Spielman and N. Srivastava, "Graph sparsification by effective resistances," in *Proceedings of the 14th Annual ACM Symposium on Theory of Computing*, ser. STOC '08. New York, NY, USA: ACM, 2008, pp. 563–568.

[17] J. Batson, D. Spielman, and N. Srivastava, "Twice-ramanujan sparsifiers," *SIAM Review*, vol. 56, no. 2, pp. 315–334, 2014.

[18] N. Motee and Q. Sun, "Sparsity and spatial localization measures for spatially distributed systems," *SIAM Jounal on Control and Optimization*, vol. 55, no. 1, pp. 200–235, 2017.

[19] C. Chen, G. Iyengar, and C. C. Moallemi, "An axiomatic approach to systemic risk," *Management Science*, vol. 59, no. 6, pp. 1373–1388, 2013.

[20] M. Siami and N. Motee, "Growing linear consensus networks via systemic performance measures," *IEEE Transactions on Automatic Control, Submitted*, 2016.

[21] S. Hawking, "Zeta function regularization of path integrals in curved space-time," *Communications in Mathematical Physics*, vol. 133, no. 55, 1977.

[22] V. D. Blondel, E. D. Sontag, M. Vidyasagar, and J. C. Willems, *Open Problems in Mathematical Systems and Control Theory*. Springer Verlag, 1999.

[23] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial Mathematics, 1997.

[24] M. H. de Badyn, A. Chapman, and M. Mesbahi, "Network entropy: A system-theoretic perspective," in *Proceedings of 54th IEEE Conference Decision and Control*, Dec 2015, pp. 5512–5517.

[25] K. Glover, "All optimal hankel-norm approximations of linear multivariable systems and their l, -error bounds," *International Journal of Control*, vol. 39, no. 6, pp. 1115–1193, 1984.

[26] J. Borwein and A. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples*, ser. CMS Books in Mathematics. Springer New York, 2005.

[27] S. Boyd., "Convex optimization of graph laplacian eigenvalues," in *Proceedings International Congress of Mathematicians*, 2006, pp. 1311–1319.

[28] B. Bamieh, M. Jovanović, P. Mitra, and S. Patterson, "Coherence in large-scale networks: Dimension-dependent limitations of local feedback," *IEEE Transactions on Automatic Control*, vol. 57, no. 9, pp. 2235 –2249, sept. 2012.

[29] M. Newman, *The Laplacian Spectrum of Graphs*. University of Manitoba, 2001.

[30] M. Rudelson and R. Vershynin, "Sampling from large matrices: An approach through geometric functional analysis," *Journal of the ACM*, vol. 54, no. 4, Jul. 2007. [Online]. Available: http://doi.acm.org/10.1145/1255443.1255449

[31] D. A. Spielman and S.-H. Teng, "Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems," in *Proceedings of the Thirty-sixth Annual ACM Symposium on Theory of Computing*, ser. STOC '04. New York, NY, USA: ACM, 2004, pp. 81–90.

[32] W. S. Fung, R. Hariharan, N. J. Harvey, and D. Panigrahi, "A general framework for graph sparsification," in *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*, ser. STOC '11. New York, NY, USA: ACM, 2011, pp. 71–80. [Online]. Available: http://doi.acm.org/10.1145/1993636.1993647

[33] J. Doyle, K. Glover, P. Khargonekar, and B. Francis, "State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems," *IEEE Transactions on Automatic Control*, vol. 34, no. 8, pp. 831–847, Aug 1989.

[34] R. Arastoo, M. Bahavarnia, M. V. Kothare, and N. Motee, "Closed-loop feedback sparsification under parametric uncertainties," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec 2016, pp. 123–128.

[35] S. Patterson and B. Bamieh, "Network coherence in fractal graphs," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, Dec. 2011, pp. 6445–6450.

[36] G. F. Young, L. Scardovi, and N. E. Leonard, "Robustness of noisy consensus dynamics with directed communication," in *American Control Conference (ACC), 2010*, July 2010, pp. 6312–6317.

[37] A. Jadbabaie and A. Olshevsky, "Combinatorial bounds and scaling laws for noise amplification in networks," in *European Control Conference (ECC)*, July 2013, pp. 596–601.

[38] B. Bamieh and D. Gayme, "The price of synchrony: Resistive losses due to phase synchronization in power networks," in *American Control Conference (ACC), 2013*, September 2012, pp. 5815–5820.

[39] E. Lovisari, F. Garin, and S. Zampieri, "Resistance-based performance analysis of the consensus algorithm over geometric graphs," *SIAM Journal on Control and Optimization*, vol. 51, no. 5, pp. 3918–3945, 2013.

[40] D. Peleg and A. A. Schäffer, "Graph spanners," *Journal of Graph Theory*, vol. 13, no. 1, pp. 609–638, 1989.

[41] A. A. Benczúr and D. R. Karger, "Approximating s-t minimum cuts in $\tilde{O}(n2)$ time," in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, ser. STOC '96. New York, NY, USA: ACM, 1996, pp. 47–55. [Online]. Available: http://doi.acm.org/10.1145/237814.237827

[42] N. Goyal, L. Rademacher, and S. Vempala, "Expanders via random spanning trees," in *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA '09. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2009, pp. 576–585.

PLACE PHOTO HERE

Milad Siami received his dual B.Sc. degrees in electrical engineering and pure mathematics from Sharif University of Technology in 2009, M.Sc. degree in electrical engineering from Sharif University of Technology in 2011, and M.Sc. and Ph.D. degrees in mechanical engineering from Lehigh University in 2014 and 2017 respectively. From 2009 to 2010, he was a research student at the Department of Mechanical and Environmental Informatics at the Tokyo Institute of Technology, Tokyo, Japan. He is currently a postdoctoral associate in the Institute for Data, Systems, and Society at MIT. His research interests include distributed control systems, distributed optimization, and applications of fractional calculus in engineering. He received a Gold Medal of National Mathematics Olympiad, Iran (2003) and the Best Student Paper Award at the 5th IFAC Workshop on Distributed Estimation and Control in Networked Systems (2015). Moreover, he was awarded RCEAS Fellowship (2012), Byllesby Fellowship (2013), Rossin College Doctoral Fellowship (2015), and Graduate Student Merit Award (2016) at Lehigh University.

PLACE PHOTO HERE

Nader Motee (S'99-M'08-SM'13) received his B.Sc. degree in Electrical Engineering from Sharif University of Technology in 2000, M.Sc. and Ph.D. degrees from University of Pennsylvania in Electrical and Systems Engineering in 2006 and 2007 respectively. From 2008 to 2011, he was a postdoctoral scholar in the Control and Dynamical Systems Department at Caltech. He is currently an Associate Professor in the Department of Mechanical Engineering and Mechanics at Lehigh University. His current research area is distributed dynamical and control systems with particular focus on issues related to sparsity, performance, and robustness. He is a past recipient of several awards including the 2008 AACC Hugo Schuck best paper award, the 2007 ACC best student paper award, the 2008 Joseph and Rosaline Wolf best thesis award, a 2013 Air Force Office of Scientific Research Young Investigator Program award (AFOSR YIP), a 2015 NSF Faculty Early Career Development (CAREER) award, and a 2016 Office of Naval Research Young Investigator Program award (ONR YIP).