Coherence and Concentration in Tightly-Connected Networks

Hancheng Min and Enrique Mallada

***Abstract*—The ability to achieve coordinated behavior—engineered or emergent—on networked systems has attracted widespread interest over several fields. This interest has led to remarkable advances in developing a theoretical understanding of the conditions under which agents within a network can reach an agreement (consensus) or develop coordinated behavior, such as synchronization. However, much less understood is the phenomenon of network coherence. Network coherence generally refers to nodes' ability in a network to have a similar dynamic response despite heterogeneity in their individual behavior. In this paper, we develop a general framework to analyze and quantify the level of network coherence that a system exhibits by relating coherence with a low-rank property of the system. More precisely, for a networked system with linear dynamics and coupling, we show that, as the network connectivity grows, the system transfer matrix converges to a rank-one transfer matrix representing the coherent behavior. Interestingly, the non-zero eigenvalue of such a rank-one matrix is given by the harmonic mean of individual nodal dynamics, and we refer to it as the coherent dynamics. Our analysis unveils the frequency-dependent nature of coherence and a non-trivial interplay between dynamics and network topology. We further show that many networked systems can exhibit similar coherent behavior by establishing a concentration result in a setting with randomly chosen individual nodal dynamics.**

I. Introduction

Coordinated behavior in network systems has been a popular subject of research in many fields, including physics [2], chemistry [3], social sciences [4], and biology [5]. Within engineering, coordination is essential for the proper operation of many networked systems including power networks [6], [7], data and sensor networks [8], [9], and autonomous transportation [10]–[13]. While there exist many expressions of this behavior, two forms of coordination have particularly received thorough attention by the control community: Consensus and synchronization.

Consensus [4], [11]–[16], on one hand, refers to the ability of the network nodes to asymptotically reach a common value over some quantities of interest. Many extensions of this problem include the study of robustness and performance of consensus networks in the presence of noise [12]–[14], time-delay [15], [16], and switching graph topology [16]. Synchronization [5], [8]–[10], [17]–[19], on the other hand, refers to the ability of network nodes to follow a commonly defined trajectory. Although for nonlinear systems synchronization is a structurally stable phenomenon, in the linear case [10], [17]–[19], synchronization requires the existence of a common internal model that acts as a virtual leader [18], [19].

H. Min and E. Mallada are with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218, USA{hanchmin, mallada}@jhu.edu

Preliminary version of this work, covering an alternative version of the results in Section V, was presented in [1].

A closely related notion of coordination emerges when looking at how the network agents collectively respond to disturbances. In this setting, agents with noticeably different input-output responses, can present a similar, i.e., coherent, response when interconnected. A vast body of work, triggered by the seminal paper [13], has quantitatively studied the role of the network topology in the emergence of coherence. Examples include, directed [14] and undirected [20] consensus networks, transportation networks [13], and power networks [7], [21], [22]. The key technical approach amounts to quantify the level of coherence by computing the $\mathcal{H}_2$-norm of the system for appropriately defined nodal disturbance and performance signals. Broadly speaking, the analysis shows a reciprocal dependence between the performance metrics and the non-zero eigenvalues of the network graph Laplacian, validating the fact that strong network coherence (low $\mathcal{H}_2$-norm) results from the high connectivity of the network (large Laplacian eigenvalues). Unfortunately, the analysis strongly relies on a homogeneity [13], [14], [20]–[22] or proportionality [7] assumption of the nodal transfer functions, and thus fails to characterize how individual heterogeneous node dynamics affect the overall coherent network response.

In this paper, we seek to overcome these limitations by formalizing network coherence by the presence of a low-rank structure, of the system transfer matrix, that appears when the network connectivity is high. More precisely, we show that for linear networks, as the network's effective algebraic connectivity (a frequency-dependent connectivity notion) grows, the system transfer matrix converges to a rank-one transfer matrix that exactly represents the coherent behavior. Interestingly, the non-zero eigenvalue of such a rank-one matrix is given by the harmonic mean of individual nodal dynamics, which we refer to it as coherent dynamics. Furthermore, we provide the concentration result of such coherent dynamics in large-scale stochastic networks where the node dynamics are given by random transfer functions.

This frequency domain analysis provides a deeper characterization of the role of both, network topology and node dynamics, on the coherent behavior of the network. In particular, our results make substantial contributions towards the understanding of coordinated and coherent behavior of network systems in many ways:

- We present a general framework in frequency domain to analyze the coherence of heterogeneous networks. We show that network coherence emerges as we increase the effective algebraic connectivity, which is a frequency-varying quantity that depends on the network coupling strength and dynamics;
- Our results suggest that network coherence is a frequency-dependent phenomenon. That is, the ability for nodes to respond coherently depends on the frequency

composition of the input disturbance;

- With the exact characterization of network's coherent dynamics, our analysis is further applied in settings where the composition of the network is unknown and and only distributional information is present. More precisely, we show that the a tightly-connected network where node dynamics are given by random transfer functions has its coherent dynamics well approximated by a deterministic transfer function that only depends on the statistical distribution of node dynamics.
- Our analysis also leads to a general methodology to accurately aggregate coherent nodes. In particular, for power networks, our results theoretically demonstrate the difficulty of aggregating synchronous generators with heterogeneous turbine dynamics.

The paper is organized as follows. In Section III and Section IV we present respectively the point-wise and uniform convergence results of network transfer matrix as the network connectivity increases. In Section V, the dynamics concentration in large-scale networks is discussed. In Section VI, we apply our analysis to consensus networks and synchronous generator networks. Conclusions are presented in Section VII.

Notation: For a vector x, $\|x\| = \sqrt{x^T x}$ denotes the 2-norm of x, and for a matrix A, $\sigma_i(A)$ denotes the ith smallest singular value of A, $\|A\|$ denotes the spectral norm of A. Particularly, if A is real symmetric, we let $\lambda_i(A)$ denote the ith smallest eigenvalue of A. For two sets S_1, S_2, we let $S_1 \backslash S_2$ denote the set difference.

We let I_n denote the identity matrix of order n, $\mathbb{1}$ denote column vector $[1, \cdots, 1]^T$, $[n]$ denote the set $\{1, 2, \cdots, n\}$ and $\mathbb{N}_+$ denote the set of positive integers. Also, we write complex numbers as $a + jb$, where $j = \sqrt{-1}$.

II. Problem Setup

Consider a network consisting of n nodes ($n \geq 2$), indexed by $i \in [n]$ with the block diagram structure in Fig.1. L is the Laplacian matrix of the weighted graph that describes the network interconnection. We further use $f(s)$ to denote the transfer function representing the dynamics of network coupling, and $G(s) = \operatorname{diag}\{g_i(s)\}$ to denote the nodal dynamics, with $g_i(s),\ i \in [n]$, being an SISO transfer function representing the dynamics of node i.

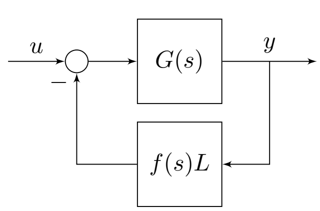

Fig. 1. Block diagram of general networked dynamical systems

Under this setting, we can compactly express the transfer matrix from the input signal vector u to the output signal vector y by

$$\begin{aligned} T(s) =&\ (I_n + G(s)f(s)L)^{-1}G(s) \\ =&\ (I_n + \operatorname{diag}\{g_i(s)\}f(s)L)^{-1}\operatorname{diag}\{g_i(s)\}\,. \end{aligned} \quad (1)$$

Many existing networks can be represented by this structure. For example, for the first-order consensus network [11], [15], $f(s) = 1$, and the node dynamics are given by $g_i(s) = \frac{1}{s}$. For power networks [7], [22], $f(s) = \frac{1}{s}$, $g_i(s)$ are the dynamics of the generators, and L is the Laplacian matrix representing the sensitivity of power injection w.r.t. bus phase angles. Finally, in transportation networks [11], [12], $g_i(s)$ represent the vehicle dynamics whereas $f(s)L$ describes local inter-vehicle information transfer.

Throughout this paper, we make the following assumptions, all of which are mild and commonly satisfied by several models that analyze the above-mentioned applications.

Assumption 1. *The closed-loop system* (1) *satisfies:*

1) *All $g_i(s),\ i = 1, \cdots, n$ and $f(s)$ are rational proper transfer functions;*
2) *Laplacian matrix L is real symmetric;*
3) *Any pole of $f(s)$ is not a zero of any of $g_i(s),\ i = 1, \cdots, n$.*

A straight forward application of the symmetry assumption in L comes from its eigendecomposition

$$L = V\Lambda V^T\,, \quad (2)$$

where $V = \left[\frac{\mathbb{1}}{\sqrt{n}}, V_\perp\right]$, $VV^T = V^TV = I_n$, and $\Lambda = \operatorname{diag}\{\lambda_i(L)\}$ with $0 = \lambda_1(L) \leq \lambda_2(L) \leq \cdots \leq \lambda_n(L)$.

Using now (2) we can rewrite $T(s)$ as

$$\begin{aligned} T(s) =&\ (I_n + \operatorname{diag}\{g_i(s)\}f(s)L)^{-1}\operatorname{diag}\{g_i(s)\} \\ =&\ (\operatorname{diag}\{g_i^{-1}(s)\} + f(s)L)^{-1} \\ =&\ (\operatorname{diag}\{g_i^{-1}(s)\} + f(s)V\Lambda V^T)^{-1} \\ =&\ V(V^T\operatorname{diag}\{g_i^{-1}(s)\}V + f(s)\Lambda)^{-1}V^T\,. \end{aligned} \quad (3)$$

As mentioned before, we are interested in characterizing the behavior of the closed-loop system $T(s)$ of (1) as the connectivity of L, i.e. $\lambda_2(L)$, increases. To gain some insight we first consider the following simplified example.

A. Motivating Example: Homogeneous Node Dynamics

Suppose $g_i(s)$ are homogeneous, i.e. $g_i(s) = g(s)$, and for simplicity $f(s) = 1$. Then using (3) one can decompose $T(s)$ as follows

$$T(s) = \frac{1}{n}g(s)\mathbb{1}\mathbb{1}^T + V_\perp\operatorname{diag}\left\{\frac{1}{g^{-1}(s) + \lambda_i(L)}\right\}_{i=2}^{n} V_\perp^T\,, \quad (4)$$

where the network dynamics decouple into two terms: 1) the dynamics $\frac{1}{n}g(s)\mathbb{1}\mathbb{1}^T$ that is independent of network topology and corresponds to the coherent behavior of the system; 2) the remaining dynamics that are dependent on the network structure via both, the eigenvalues $\lambda_i(L), i = 2, \cdots, n$ and the eigenvectors $V_\perp$. We observe that $\frac{1}{n}g(s)\mathbb{1}\mathbb{1}^T$ is dominant in $T(s)$ as long as $\lambda_2(L)$, which lower bounds all $\lambda_i(L), i = 2, \ldots, n$, is large enough to make the norm of the second term in (4) sufficiently small. Following such observation, we can conclude that for almost every $s_0 \in \mathbb{C}$, except for the poles of $g(s)$, the following holds:

$$\lim_{\lambda_2(L)\to\infty}\left\|T(s_0) - \frac{1}{n}g(s_0)\mathbb{1}\mathbb{1}^T\right\| = 0\,.$$

Furthermore, one can verify that if a compact set $S \subset \mathbb{C}$ contains neither zeros nor poles of $g(s)$, the following holds:

$$\lim_{\lambda_2(L)\to\infty} \sup_{s\in S} \left\| T(s) - \frac{1}{n} g(s)\mathbb{1}\mathbb{1}^T \right\| = 0 \,.$$

Such uniform convergence result suggests that in networks with large algebraic connectivity, the response of the entire system is close to one of $\frac{1}{n}g(s)\mathbb{1}\mathbb{1}^T$. We refer $\frac{1}{n}g(s)\mathbb{1}\mathbb{1}^T$ as the coherent dynamics[1] in the sense that in such system, the inputs are aggregated, and all nodes have exactly the same response to the aggregate input. *Therefore, coherence of the network in Fig.1 corresponds to, in the frequency domain, that the network's transfer matrix approximately has a particular rank-one structure, and such coherence emerges as we increase the network connectivity.*

B. The Goal of This Work

In this paper, we would like to show that even when $g_i(s)$ are heterogeneous, similar convergence result still holds. More precisely, we will, in the following sections, discuss the point-wise and uniform convergence of $T(s)$ to a transfer matrix of the form $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T$, as the algebraic connectivity $\lambda_2(L)$ increases. However, unlike the homogeneous node dynamics case where the coherent behavior is driven by $\bar{g}(s) = g(s)$, we will show that the coherent dynamics $\bar{g}(s)$ are given by the harmonic mean of $g_i(s), i = 1, \cdots, n$, i.e.,

$$\bar{g}(s) = \left(\frac{1}{n} \sum_{i=1}^{n} g_i^{-1}(s) \right)^{-1} . \tag{5}$$

Such asymptotic analyses serve two main purposes. Firstly, using the coherent dynamics $\bar{g}(s)$, one can infer the point-wise convergence of $T(s)$ as $\lambda_2(L)$ increases. In particular, we show that

1) For a point that is neither a pole nor a zero of $\bar{g}(s)$, $T(s)$ converges to $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T$;
2) Poles of $\bar{g}(s)$ are asymptotically poles of $T(s)$;
3) Zeros of $\bar{g}(s)$ are asymptotically zeros of $T(s)$.

Secondly, uniform convergence of $T(s)$ explains the coherent behavior/response of a tightly-connected network subject to disturbances. To see the connection, recall the Inverse Laplace Transform [23, Theorem 3.20] computes the system time-domain response by integration on the line $\{\sigma + j\omega : \omega \in [-\infty, +\infty]\}$ in frequency domain with a suitable σ. Then uniform convergence of $T(s)$ on this line would show that time-domain response of the network convergences to one of the coherent dynamics $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T$ as network connectivity increases. However, we will see that such convergence does not hold for most networks through our analysis and subsequent examples, which suggests that *the coherence we analyze here is a frequency-dependent phenomenon*. On that note, one generally resort to weaker convergence results on a line segment $\{\sigma + j\omega : \omega \in [-\omega_c, +\omega_c]\}$ for some $\omega_c > 0$, which, once established, justifies the coherent behavior of tightly-connected networks subject to low-frequency disturbances. Since the set above is a compact subset of $\mathbb{C}$, we are mostly discussing the uniform convergence results over compact sets.

One additional feature of our analysis is that it can be further applied in settings where the composition of the network is unknown and only distributional information is present. More precisely, we will extend such convergence results by considering a tightly-connected network where node dynamics are given by random transfer functions. As the network size grows, the coherent dynamics $\bar{g}(s)$ converges in probability to a deterministic transfer function. We term such a phenomenon, where a family of uncertain large-scale tightly-connected systems concentrates to a common deterministic system, *dynamics concentration*.

[1] We also refer $g(s)$ as the coherent dynamics since transfer matrix of the form $\frac{1}{n}g(s)\mathbb{1}\mathbb{1}^T$ is uniquely determined by its non-zero eigenvalue $g(s)$.

C. Preliminaries

Before presenting our results, we first state a few facts and preliminary results that are used in later proofs.

1) Basic Results on Vectors and Matrices: The following results can be found in [24].

- *Norm Inequalities:* Let $x, y \in \mathbb{R}^n$ and $A, B \in \mathbb{R}^{m\times n}$, we have

  2-Norm compatibility: $$\|Ax\| \le \|A\|\|x\| \,, \tag{6}$$
  2-Norm sub-multiplicativity: $$\|AB\| \le \|A\|\|B\| \,, \tag{7}$$
  Cauchy-Schwarz: $$|x^T y| \le \|x\|\|y\| \,. \tag{8}$$

- *Inverse of Block Matrix:* For block matrix $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ with A, D being square matrices, its inverse can be written as

$$M^{-1} = \begin{bmatrix} S^{-1} & -S^{-1}BD^{-1} \\ -D^{-1}CS^{-1} & D^{-1} + D^{-1}CS^{-1}BD^{-1} \end{bmatrix} , \tag{9}$$

  where $S = A - BD^{-1}C$, provided that all relevant inverses exist.

2) Inequalities for singular values: We also provide several inequalities for matrix singular values from [24, 7.3.P16] which will be used in our proofs.

Lemma 1 (Weyl's Inequality)**.** *Let A, B be square matrices of order n, the following inequalities hold:*

$$\|AB\| \ge \|A\|\sigma_1(B) \,, \tag{10a}$$
$$\sigma_1(AB) \le \|A\|\sigma_1(B) \,, \tag{10b}$$
$$\sigma_1(A + B) \le \sigma_1(A) + \|B\| \,. \tag{10c}$$

Lemma 1 allows us to obtain a useful bound on the spectral norm of $(A + B)^{-1}$. We state it as another Lemma as it will be repeatedly used in sections III and IV:

Lemma 2. *Let A, B be square matrices of order n. If $\sigma_1(A) \ge \|B\| > 0$, then the following inequality holds:*

$$\|(A + B)^{-1}\| \le \frac{1}{\sigma_1(A) - \|B\|} \,.$$

Proof. By (10c), we have

$$\sigma_1(A) \le \sigma_1(A + B) + \|-B\| \,.$$

Then as long as $\sigma_1(A) \geq \|B\| > 0$, it leads to

$$\frac{1}{\sigma_1(A+B)} \leq \frac{1}{\sigma_1(A) - \|B\|},$$

and the left-hand side is exactly $\|(A+B)^{-1}\|$. $\square$

3) Grounded Laplacian Matrix: For a $n \times n$ Laplacian matrix L, we select an index set $I \subset [n]$. Then the grounded Laplacian $\tilde{L}$ is the principal submatrix of L obtained by removing the rows and columns corresponding to the index set I. The following lemma relates the eigenvalues of $\tilde{L}$ and L.

Lemma 3. *Given a $n \times n$ symmetric Laplacian matrix L, let $\tilde{L}$ be its grounded Laplacian corresponding to a index set $\mathcal{I}$ with $|\mathcal{I}| = m < n$. Then for the least eigenvalue of $\tilde{L}$, the following inequality holds:*

$$\lambda_1(\tilde{L}) \geq \frac{m}{n}\lambda_2(L).$$

The proof is shown in the Appendix.A. This lower bound shows that for weighted graphs with fixed network size n, as $\lambda_2(L) \to \infty$, we also have $\lambda_1(\tilde{L}) \to \infty$. This result is used in section III and IV when we present the convergence analysis regarding zeros of $\bar{g}(s)$.

Now we are ready for the main results of this paper, and we start with the exact characterization of coherent dynamics of tightly-connected networks by proving the point-wise convergence of $T(s)$.

III. Point-wise Coherence

In this section, we analyze the strength of network coherence at a single point in the frequency domain. We start with an important lemma revealing how such coherence is related to algebraic connectivity $\lambda_2(L)$ and feedback dynamics $f(s)$.

Lemma 4. *Let $T(s)$ and $\bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. Suppose that for $s_0 \in \mathbb{C}$ that is not a pole of $f(s)$, we have*

$$|\bar{g}(s_0)| \leq M_1, \text{and } \max_{1\leq i\leq n} |g_i^{-1}(s_0)| \leq M_2,$$

for some $M_1, M_2 > 0$. Then for large enough $\lambda_2(L)$, the following inequality holds:

$$\left\|T(s_0) - \frac{1}{n}\bar{g}(s_0)\mathbb{1}\mathbb{1}^T\right\| \leq \frac{(M_1M_2+1)^2}{|f(s_0)|\lambda_2(L) - M_2 - M_1M_2^2}. \tag{11}$$

Proof. Let $H = V^T\text{diag}\{g_i^{-1}(s_0)\}V + f(s_0)\Lambda$, such that (3) becomes

$$T(s) = VH^{-1}V^T.$$

Then it is easy to see that

$$\begin{aligned}\left\|T(s_0) - \frac{1}{n}\bar{g}(s_0)\mathbb{1}\mathbb{1}^T\right\| &= \|T(s_0) - \bar{g}(s_0)Ve_1e_1^TV^T\| \\ &= \left\|V\left(H^{-1} - \bar{g}(s_0)e_1e_1^T\right)V^T\right\| \\ &= \left\|H^{-1} - \bar{g}(s_0)e_1e_1^T\right\|, \end{aligned} \tag{12}$$

where e_1 is the first column of identity matrix I_n. The first equality holds by noticing that $\frac{\mathbb{1}}{\sqrt{n}}$ is the first column of V, and the last equality comes from the fact that multiplying by a unitary matrix V preserves the spectral norm.

We now write H in block matrix form:

$$\begin{aligned} H =\ & V^T\text{diag}\{g_i^{-1}(s_0)\}V + f(s_0)\Lambda \\ =\ & \begin{bmatrix}\frac{\mathbb{1}^T}{\sqrt{n}} \\ V_\perp^T\end{bmatrix}\text{diag}\{g_i^{-1}(s_0)\}\begin{bmatrix}\frac{\mathbb{1}}{\sqrt{n}} & V_\perp\end{bmatrix} + f(s_0)\Lambda \\ =\ & \begin{bmatrix}\bar{g}^{-1}(s_0) & \frac{\mathbb{1}^T}{\sqrt{n}}\text{diag}\{g_i^{-1}(s_0)\}V_\perp \\ V_\perp^T\text{diag}\{g_i^{-1}(s_0)\}\frac{\mathbb{1}}{\sqrt{n}} & V_\perp^T\text{diag}\{g_i^{-1}(s_0)\}V_\perp + f(s_0)\tilde{\Lambda}\end{bmatrix} \\ :=\ & \begin{bmatrix}\bar{g}^{-1}(s_0) & h_{21}^T \\ h_{21} & H_{22}\end{bmatrix}, \end{aligned}$$

where $\tilde{\Lambda} = \text{diag}\{\lambda_2(L), \cdots, \lambda_n(L)\}$, and we use the fact that $\lambda_1(\Lambda) = 0$.

Inverting H in its block form as in (9), we have

$$H^{-1} = \begin{bmatrix} a & -ah_{21}^TH_{22}^{-1} \\ -aH_{22}^{-1}h_{21} & H_{22}^{-1} + aH_{22}^{-1}h_{21}h_{21}^TH_{22}^{-1}\end{bmatrix},$$

where $a = \frac{1}{\bar{g}^{-1}(s_0) - h_{21}^TH_{22}^{-1}h_{21}}$.

Notice that $||V_\perp|| = 1$ and $||\mathbb{1}|| = \sqrt{n}$, we have

$$\begin{aligned}\|h_{21}\| =\ & \left\|V_\perp^T\text{diag}\{g_i^{-1}(s_0)\}\frac{\mathbb{1}}{\sqrt{n}}\right\| \\ \leq\ & \|V_\perp\|\|\text{diag}\{g_i^{-1}(s_0)\}\|\frac{\|\mathbb{1}\|}{\sqrt{n}} \leq M_2, \end{aligned} \tag{13}$$

where (13) follows from the norm compatibility (6) and that matrix 2-norm is sub-multiplicative (7).

Also, by Lemma 2, when $|f(s_0)|\lambda_2(L) > M_2$, the following holds:

$$\begin{aligned}\|H_{22}^{-1}\| =\ & \|(f(s_0)\tilde{\Lambda} + V_\perp^T\text{diag}\{g_i^{-1}(s_0)\}V_\perp)^{-1}\| \\ \leq\ & \frac{1}{\sigma_1(f(s_0)\tilde{\Lambda}) - \|V_\perp^T\text{diag}\{g_i^{-1}(s_0)\}V_\perp\|} \\ \leq\ & \frac{1}{\sigma_1(f(s_0)\tilde{\Lambda}) - M_2} \leq \frac{1}{|f(s_0)|\lambda_2(L) - M_2}. \end{aligned} \tag{14}$$

Again (14) uses the fact that $\|V_\perp\| = 1$, the function $\frac{1}{y-x}$ is decreasing in y and increasing in x, and, by our assumption, $\|\text{diag}\{g_i^{-1}(s_0)\}\| = \max_{1\leq i\leq n}|g_i^{-1}(s_0)| \leq M_2$.

Lastly, when $|f(s_0)|\lambda_2(L) > M_2 + M_2^2M_1$, a similar reasoning as above, using (13) (14), and our assumption $|\bar{g}(s_0)| \leq M_1$, gives

$$\begin{aligned}|a| \leq\ & \frac{1}{|\bar{g}^{-1}(s_0)| - |h_{21}^TH_{22}^{-1}h_{21}|} \\ \leq\ & \frac{1}{|\bar{g}^{-1}(s_0)| - \|h_{21}\|^2\|H_{22}^{-1}\|} \\ \leq\ & \frac{1}{\frac{1}{M_1} - \frac{M_2^2}{|f(s_0)|\lambda_2(L) - M_2}} \\ =\ & \frac{(|f(s_0)|\lambda_2(L) - M_2)M_1}{|f(s_0)|\lambda_2(L) - M_2 - M_1M_2^2}, \end{aligned} \tag{15}$$

where in the second inequality, we used norm compatibility and Cauchy-Schwarz inequality (8) to upper-bound $|h_{21}^TH_{22}h_{21}|$.

Now we bound the norm of $H^{-1}-\bar{g}(s_0)e_1e_1^T$ by the sum of norms of all its blocks:

$$
\begin{aligned}
&\|H^{-1}-\bar{g}(s_0)e_1e_1^T\| \\
=&\left\|\begin{bmatrix} a\bar{g}(s_0)h_{21}^TH_{22}^{-1}h_{21} & -ah_{21}^TH_{22}^{-1} \\ -aH_{22}^{-1}h_{21} & H_{22}^{-1}+aH_{22}^{-1}h_{21}h_{21}^TH_{22}^{-1}\end{bmatrix}\right\| \\
\le& |a\bar{g}(s_0)h_{21}^TH_{22}^{-1}h_{21}|+2\|aH_{22}^{-1}h_{21}\| \\
&+\|H_{22}^{-1}+aH_{22}^{-1}h_{21}h_{21}^TH_{22}^{-1}\| \\
\le& |a|\|H_{22}^{-1}\|(|\bar{g}(s_0)|\|h_{21}\|^2+2\|h_{21}\|+\|h_{21}\|^2\|H_{22}^{-1}\|) \\
&+\|H_{22}^{-1}\|\,, \qquad (16)
\end{aligned}
$$

Using (13)(14)(15), we can further upper bound (16) as

$$
\begin{aligned}
&\|H^{-1}-\bar{g}(s_0)e_1e_1^T\| \\
\le& \frac{M_1^2M_2^2+2M_1M_2+\frac{M_1M_2^2}{|f(s_0)|\lambda_2(L)-M_2}}{|f(s_0)|\lambda_2(L)-M_2-M_1M_2^2} \\
&+\frac{1}{|f(s_0)|\lambda_2(L)-M_2} \\
=& \frac{(M_1M_2+1)^2}{|f(s_0)|\lambda_2(L)-M_2-M_1M_2^2}\,. \qquad (17)
\end{aligned}
$$

This bound holds as long as $|f(s_0)|\lambda_2(L)>M_2+M_2^2M_1$. Combining (12) and (17) gives the desired inequality. □

Lemma 4 provides an upper bound for the incoherence measure we are interested in, namely how far apart the system transfer matrix is, at a particular point in the frequency domain, from being rank-one with coherent direction $\frac{1}{n}\mathbb{1}\mathbb{1}^T$. This result can be understood in the following aspects:

First of all, a large value of $|f(s_0)|\lambda_2(L)$ is sufficient to have the incoherence measure small, and we term this quantity as *effective algebraic connectivity*. We see that there are two possible ways to achieve such "point-wise coherence": Either we increase the network algebraic connectivity $\lambda_2(L)$, by adding edges to the network and increasing edge weights, etc., or we move our point of interest s_0 to a pole of $f(s)$.

Secondly, the upper bound is frequency-dependent since it is provided at a single point s_0 in the s-domain. To see such dependence, notice that s_0 near a pole of $f(s)$ has large effective algebraic connectivity, hence the system is more coherent around poles of $f(s)$; On the contrary, s_0 near a pole of $\bar{g}(s)$ requires large M_1 for the condition of Lemma 4 to hold, and readers can check that s_0 near a zero of $\bar{g}(s)$ requires large M_2, therefore for at these points, it is more difficult for us to upper bound the incoherence measure by Lemma 4. Such dependence makes it challenging to understand the network coherence uniformly in the entire frequency domain.

Last but not least, Although Lemma 4 provides a sufficient condition for the network coherence to emerge, i.e. the increasing effective algebraic connectivity, it is still unknown whether such a condition is necessary. In other words, we do not know whether low effective algebraic connectivity means some kind of incoherence. This problem seems trivial for the extreme case: if $|f(s_0)|=0$ or $L=0$, the feedback loop vanishes, and every node responses independently, but certainly not otherwise.

When the condition in Lemma 4 is satisfied, the system is asymptotically coherent, i.e. $T(s_0)$ converges to $\frac{1}{n}\bar{g}(s_0)$ as the effective algebraic connectivity $|f(s_0)|\lambda_2(L)$ increases. As we mentioned above, we can achieve this by increasing either $\lambda_2(L)$ or $|f(s_0)|$, provided that the other value is fixed and non-zero. Subsection III-A considers the former and Subsection III-B the latter.

Before presenting with the results, we define

Definition 1. *For transfer function $g(s)$ and $s_0\in\mathbb{C}$, s_0 is a generic point of $g(s)$ if s_0 is neither a pole nor a zero of $g(s)$.*

As we have seen through the discussions above, we always require some generic point assumptions for either $\bar{g}(s)$, $f(s)$, or both. Those points are of the most interest in this paper but we will provide some results for the cases where the generic assumption fails.

A. Convergence at Generic Points of $f(s)$

In this section we keep s_0 fixed and present the point-wise convergence result of $T(s_0)$ as $\lambda_2(L)$ increases. This requires s_0 to be a generic point of $f(s)$.

Notice that for any s_0 that is also a generic point of $\bar{g}(s)$, we can always find such $M_1, M_2>0$ and large enough $\lambda_2(L)$ for the upper bound in (11) to hold. Furthermore, given fixed M_1 and M_2, one can let the upper bound be arbitrarily small by increasing $\lambda_2(L)$, which leads to the point-wise convergence of $T(s_0)$, as stated in the following theorem.

Theorem 1. *Let $T(s)$ and $\bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. If $s_0\in\mathbb{C}$ is a generic point of both $\bar{g}(s)$ and $f(s)$, then*

$$\lim_{\lambda_2(L)\to+\infty}\left\|T(s_0)-\frac{1}{n}\bar{g}(s_0)\mathbb{1}\mathbb{1}^T\right\|=0\,.$$

Proof. Since s_0 is not a pole of $\bar{g}(s)$, $|\bar{g}(s_0)|$ is trivially upper bounded by some $M_1>0$. Also, it is easy to see that s_0 is not a zero of $\bar{g}(s)$ if and only if s_0 is not a zero of any $g_i(s)$. Then $\max_{1\le i\le n}|\bar{g}^{-1}(s_0)|$ is upper bounded by some $M_2>0$. Therefore the conditions of Lemma 4 are satisfied. We finish the proof by taking $\lambda_2(L)\to+\infty$ on both sides of (11). □

Theorem 1 establishes the emergence of coherence at generic points of $\bar{g}(s)$. This forms the basis of our analysis, yet requires such s_0 satisfying generic conditions. A more careful analysis shows that, as $\lambda_2(L)\to+\infty$, the pole of $\bar{g}(s)$ is asymptotically a pole of $T(s)$, and the zero of $\bar{g}(s)$ is asymptotically a zero of $T(s)$, as stated in the following two theorems.

Theorem 2. *Let $T(s)$ and $\bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. If $s_0\in\mathbb{C}$ is a pole of $\bar{g}(s)$ and a generic point of $f(s)$, then*

$$\lim_{\lambda_2(L)\to+\infty}\|T(s_0)\|=+\infty\,.$$

Proof. Similarly to the proof of Lemma 4, we define $H=V^T\mathrm{diag}\{g_i^{-1}(s_0)\}V+f(s_0)\Lambda$ and now we need to show that $\|T(s_0)\|=\|H^{-1}\|$ grows unbounded as $\lambda_2(L)\to+\infty$.

Write H in block matrix form:

$$H=\begin{bmatrix}\bar{g}^{-1}(s_0) & \frac{\mathbb{1}^T}{\sqrt{n}}\mathrm{diag}\{g_i^{-1}(s_0)\}V_\perp \\ V_\perp^T\mathrm{diag}\{g_i^{-1}(s_0)\}\frac{\mathbb{1}}{\sqrt{n}} & V_\perp^T\mathrm{diag}\{g_i^{-1}(s_0)\}V_\perp+f(s_0)\tilde{\Lambda}\end{bmatrix}$$

$$:= \begin{bmatrix} 0 & h_{21}^T \\ h_{21} & H_{22} \end{bmatrix},$$

by noticing that $\bar{g}^{-1}(s_0) = 0$ because s_0 is a pole of $\bar{g}(s)$.

Inverting H in its block form gives

$$\begin{aligned} H^{-1} &= \begin{bmatrix} a & -ah_{21}^T H_{22}^{-1} \\ -aH_{22}^{-1}h_{21} & H_{22}^{-1} + aH_{22}^{-1}h_{21}h_{21}^T H_{22}^{-1} \end{bmatrix} \\ &= a \begin{bmatrix} 1 \\ -H_{22}^{-1}h_{21} \end{bmatrix} \begin{bmatrix} 1 & -h_{21}^T H_{22}^{-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & H_{22}^{-1} \end{bmatrix}, \end{aligned} \tag{18}$$

where a now is given by $a = -\frac{1}{h_{21}^T H_{22}^{-1} h_{21}}$.

Then from (18), when $\lambda_2(L)$ is large enough, we can lower bound $\|H^{-1}\|$ by

$$\begin{aligned} \|H^{-1}\| &\geq \left\| a \begin{bmatrix} 1 \\ -H_{22}^{-1}h_{21} \end{bmatrix} \begin{bmatrix} 1 \\ -H_{22}^{-1}h_{21} \end{bmatrix}^T \right\| - \left\| \begin{bmatrix} 0 & 0 \\ 0 & H_{22}^{-1} \end{bmatrix} \right\| \\ &= \frac{1}{|h_{21}^T H_{22}^{-1} h_{21}|} \left\| \begin{bmatrix} 1 \\ -H_{22}^{-1}h_{21} \end{bmatrix} \right\|^2 - \|H_{22}^{-1}\| \\ &\geq \frac{1}{|h_{21}^T H_{22}^{-1} h_{21}|} - \|H_{22}^{-1}\| \\ &\geq \frac{1}{\|h_{21}\|^2 \|H_{22}^{-1}\|} - \|H_{22}^{-1}\|, \end{aligned} \tag{19}$$

where in the second inequality, we simply use the fact that the norm of a vector is lower bounded by its first entry.

Because s_0 is a pole of $\bar{g}(s)$, it cannot be a zero of any $g_i(s)$; otherwise this would lead to the contradiction $\bar{g}(s_0) = 0$. Therefore, $\max_{1\leq i\leq n} |g_i^{-1}(s)|$ is upper bounded by some $M > 0$. Similarly to (13) and (14), we have $\|h_{21}\| \leq M$ and $\|H_{22}^{-1}\| \leq \frac{1}{|f(s_0)|\lambda_2(L) - M}$. Then (19) can be lower bounded by

$$\begin{aligned} \|H^{-1}\| &\geq \frac{1}{\|h_{21}\|^2 \|H_{22}^{-1}\|} - \|H_{22}^{-1}\| \\ &\geq \frac{1}{\frac{M^2}{|f(s_0)|\lambda_2(L) - M}} - \frac{1}{|f(s_0)|\lambda_2(L) - M} \\ &= \frac{(|f(s_0)|\lambda_2(L) - M)^2 - M^2}{M^2(|f(s_0)|\lambda_2(L) - M)}. \end{aligned}$$

This lower bound holds when $|f(s_0)|\lambda_2(L) \geq M$, and it grows unbounded as $\lambda_2(L) \to +\infty$, which finishes the proof. □

Remark 1. *Theorem 2 does not suggest whether the network is asymptotically coherent at poles of $\bar{g}(s)$. Our incoherence measure $\left\|T(s_0) - \frac{1}{n}\bar{g}(s_0)\mathbb{1}\mathbb{1}^T\right\|$ is undefined at such poles. Alternatively, for s_0 the pole of $\bar{g}(s)$, one can prove that when $\tilde{\Lambda}/\lambda_2(L) \to \Lambda_{\lim}$ as $\lambda_2(L) \to +\infty$, we have the limit $\left\|\frac{T(s_0)}{\|T(s_0)\|} - \frac{1}{n}\gamma(\Lambda_{\lim})\mathbb{1}\mathbb{1}^T\right\| \to 0$, for some $\gamma(\Lambda_{\lim}) \in \mathbb{C}$ determined by $\Lambda_{\lim}$ with $|\gamma(\Lambda_{\lim})| = 1$. We leave the formal statement to Appendix.B. This suggests that $T(s_0)$ has the desired rank-one structure for coherence. While the normalized transfer matrix is not discussed in this paper due to the space constraints, such formulation is better for understanding the network coherence at the poles of $\bar{g}(s)$.*

Next, the convergence result regarding the zeros of $\bar{g}(s)$ is stated as

Theorem 3. *Let $T(s)$ and $\bar{g}(s)$ be defined as in* (1) *and* (5), *respectively. If $s_0 \in \mathbb{C}$ is a zero of $\bar{g}(s)$ and a generic point of $f(s)$, then*

$$\lim_{\lambda_2(L)\to+\infty} \|T(s_0)\| = 0.$$

Proof. Since s_0 is a zero of $\bar{g}(s)$, it is the zero of at least one $g_i(s)$. Without loss of generality, suppose $g_i(s_0) = 0$ for $1 \leq i \leq m$ and $g_i(s_0) \neq 0$ for $m+1 \leq i \leq n$.

If $m = n$, then $T(s_0) = 0$. We only consider the non-trivial case when $m < n$. The transfer matrix is now given by

$$\begin{aligned} T(s_0) &= (I_n + G(s_0)f(s_0)L)^{-1}G(s_0) \\ &= \begin{bmatrix} I_m & 0_{m\times(n-m)} \\ 0_{(n-m)\times m} & I_{n-m} + \tilde{G}(s_0)f(s_0)\tilde{L} \end{bmatrix}^{-1} G(s_0) \\ &= \begin{bmatrix} 0_{m\times m} & 0_{m\times(n-m)} \\ 0_{(n-m)\times m} & (I_{n-m} + \tilde{G}(s_0)f(s_0)\tilde{L})^{-1}\tilde{G}(s_0) \end{bmatrix}, \end{aligned} \tag{20}$$

where $\tilde{G}(s) = \mathrm{diag}\{g_{m+1}(s), \cdots, g_n(s)\}$ and $\tilde{L}$ is the *grounded Laplacian* of L by removing the first m rows and columns.

By Lemma 3, when $\lambda_1(\tilde{L})$ is large enough, we have

$$\begin{aligned} \|T(s_0)\| &= \|(I_{n-m} + \tilde{G}(s_0)f(s_0)\tilde{L})^{-1}\tilde{G}(s_0)\| \\ &= \|(\tilde{G}^{-1}(s_0) + f(s_0)\tilde{L})^{-1}\| \\ &\leq \frac{1}{\sigma_1(f(s_0)\tilde{L}) - \|\tilde{G}^{-1}(s_0)\|} \\ &\leq \frac{1}{|f(s_0)|\lambda_1(\tilde{L}) - \|\tilde{G}^{-1}(s_0)\|}. \end{aligned}$$

Since $g_i(s_0) \neq 0$ for $m+1 \leq i \leq n$, $\max_{m+1\leq i\leq n} |g_i^{-1}(s_0)|$ is upper bounded by some $M > 0$. Then we have

$$\|T(s_0)\| \leq \frac{1}{|f(s_0)|\lambda_1(\tilde{L}) - M}. \tag{21}$$

By Lemma 3, we know that $\lambda_1(\tilde{L}) \to +\infty$ as $\lambda_2(L) \to \infty$, then

$$\lim_{\lambda_2(L)\to+\infty} \frac{1}{|f(s_0)|\lambda_1(\tilde{L}) - M} = 0.$$

We finishes the proof by taking $\lambda_2(L) \to +\infty$ on both sides of (21) □

Remark 2. *The limit in Theorem 3 can still be written as $\lim_{\lambda_2(L)\to+\infty} \|T(s_0) - \frac{1}{n}\bar{g}(s_0)\mathbb{1}\mathbb{1}^T\| = 0$, because s_0 is a zero of $\bar{g}(s)$. However, we here emphasize the fact that the system is not coherent at s_0 under normalization because $T(s_0)/\|T(s_0)\|$ does not converge to $\frac{1}{n}\gamma\mathbb{1}\mathbb{1}^T$ for any $\gamma \in \mathbb{C}$ as we can directly see from* (20).

So far, we have shown point-wise convergence of $T(s)$ towards the transfer function $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T$, from which we assess how network coherence emerges as connectivity increases. In Remark 1 and 2 we see that the incoherence measure $\left\|T(s_0) - \frac{1}{n}\bar{g}(s_0)\mathbb{1}\mathbb{1}^T\right\|$ is insufficient for understanding the asymptotic behavior at zeros or poles of $\bar{g}(s)$, and that the alternative measure $\left\|\frac{T(s_0)}{\|T(s_0)\|} - \frac{1}{n}\gamma\mathbb{1}\mathbb{1}^T\right\|$ is a good complement for such purpose.[2] The latter measure, $\left\|\frac{T(s_0)}{\|T(s_0)\|} - \frac{1}{n}\gamma \mathbb{1}\mathbb{1}^T\right\|$, focuses more on the relative scale of eigenvalues of $T(s)$. In this paper, we mostly use the former, $\left\|T(s_0) - \frac{1}{n}\bar{g}(s_0)\mathbb{1}\mathbb{1}^T\right\|$, and particularly when presenting the uniform convergence results; because we are interested in connecting these results to the network time-domain response.

B. Convergence Regarding Poles of $f(s)$

As mentioned before, when s_0 is a pole of $f(s)$, it is a singularity of $T(s)$. Under certain conditions, one can observe that high-gain in $f(s)$ plays a role similar to $\lambda_2(L)$. The result uses Lemma 4 and is stated as follows.

Theorem 4. *Let $T(s)$ and $\bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. Suppose $\lambda_2(L) > 0$. If $s_0 \in \mathbb{C}$ is a generic point of $\bar{g}(s)$ and s_0 is a pole of $f(s)$, then*

$$\lim_{s\to s_0} \left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\| = 0 \,.$$

Proof. Since s_0 is neither a zero nor a pole of $\bar{g}(s)$, $\exists \delta > 0$ such that $\forall s \in U(s_0, \delta) = \{s : |s - s_0| < \delta\}$, we have $|\bar{g}^{-1}(s)| \leq M_1$ and $\max_{1\leq i\leq n} |g_i^{-1}(s)| \leq M_2$ for some $M_1, M_2 > 0$.

By Lemma 4, $\forall s \in U(s_0, \delta)$, the following holds

$$\left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\| \leq \frac{(M_1M_2 + 1)^2}{|f(s)|\lambda_2(L) - M_2 - M_1M_2^2} \,.$$

Taking $s \to s_0$ on both side, notice that $\lim_{s\to s_0} |f(s)| = +\infty$, the limit of right-hand side is 0. □

In other words, at pole of $f(s)$, the network effect is infinitely amplified. The effective algebraic connectivity $|f(s)|\lambda_2(L)$ grows unbounded as s approaching the pole of $f(s)$. As a result, the frequency response of $T(s)$ is exactly the one of $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T$. Hence network coherence naturally arises around poles of $f(s)$.

IV. Uniform Coherence

We now leverage the point-wise convergence results of Section III to characterize conditions for uniform convergence. This will allow us to connect our analysis with time domain implications, as discussed in II-B.

We start by showing uniform convergence of $T(s)$ over compact regions that do not contain any zero or pole of $\bar{g}(s)$. While the uniform convergence does not hold over regions containing such zeros or poles in general, we prove that in some special cases, the uniform convergence around zeros of $\bar{g}(s)$ does hold. Finally, we provide a sufficient condition for uniform convergence of $T(s)$ on the right-half plane, which implies the system converges in $\mathcal{H}_\infty$ norm.

[2] As $\lambda_2(L)$ increases, for pole of $\bar{g}(s)$, the latter measure converges to 0 given suitable conditions but not for the former; for zero of $\bar{g}(s)$, the opposite result holds; for generic point of $\bar{g}(s)$, both incoherence measures converge to 0.

A. Uniform Convergence Around Generic Points of $\bar{g}(s)$

Again, similarly to the point-wise convergence, we discuss uniform convergence of $T(s)$ over set S that satisfies the following assumption

Assumption 2. *$S \subset \mathbb{C}$ satisfies $\sup_{s\in S} |f(s)| < \infty$ and $\inf_{s\in S} |f(s)| > 0$.*

Such an assumption guarantees all points in the closure of S are generic points of $f(s)$. This property prevents any sequence of points in S that asymptotically eliminates or amplifies the network effect on the boundary of S. In subsequent sections, we denote $F_h := \sup_{s\in S} |f(s)|$ and $F_l := \inf_{s\in S} |f(s)| > 0$.

Recall that in Section III, point-wise convergence is proved by choosing $M_1, M_2 > 0$ such that the conditions in Lemma 4 are satisfied at a particular point s_0. Then, finding universal $M_1, M_2 > 0$ that work for every s_0 in a set $S \subset \mathbb{C}$ suffices to show uniform convergence over S. Such a process is straightforward if S is compact:

Theorem 5. *Let $T(s)$ and $\bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. Then given a compact set $S \subset \mathbb{C}$, if S satisfies Assumption 2 and does not contain any zero or pole of $\bar{g}(s)$, we have*

$$\lim_{\lambda_2(L)\to+\infty} \sup_{s\in S} \left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\| = 0 \,.$$

Proof. On the one hand, since S does not contain any pole of $\bar{g}(s)$, $\bar{g}(s)$ is continuous on the compact set S, and hence bounded [25, Theorem 4.15]. On the other hand, because S does not contain any zero of $\bar{g}(s)$, every $g_i^{-1}(s)$ must be continuous on S, and hence bounded as well. It follwos that $\max_{1\leq i\leq n} |g_i^{-1}(s)|$ is bounded on S, and the conditions of Lemma 4 are satisfied for all $s \in S$ with a uniform choice of M_1 and M_2. By (11), we have

$$\sup_{s\in S} \left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\| \leq \frac{(M_1M_2 + 1)^2}{F_l\lambda_2(L) - M_2 - M_1M_2^2} \,,$$

where $F_l = \inf_{s\in S} |f(s)|$. We finish the proof by taking $\lambda_2(L) \to +\infty$ on both sides. □

As we already discussed in Remark 1, if S contains poles of $\bar{g}(s)$, $\sup_{s\in S} \left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\|$ is not a good incoherence measure as it is undefined. The rest of the section mainly discusses the uniform convergence result around zeros of $\bar{g}(s)$.

B. Uniform Convergence Around Zeros of $\bar{g}(s)$

We first define the notion of *Nodal Multiplicity* of a point in complex plane w.r.t. a given network.

Definition 2. *Given $\{g_i(s), i \in [n]\}$, the Nodal Multiplicity of $s_0 \in \mathbb{C}$ is defined as*

$$\mathcal{N}(s_0) := |\{i \in [n] : g_i(s_0) = 0\}| \,,$$

where $|\cdot|$ denotes the set cardinality.

By definition, any zero of $\bar{g}(s)$ must have positive nodal multiplicity. Our finding is that zeros with nodal multiplicity exactly 1 have a special property, which is shown in the following Lemma.

Lemma 5. *Let $T(s), \bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. If $s_0 \in \mathbb{C}$ is a zero of $\bar{g}(s)$ with nodal multiplicity $\mathcal{N}(s_0) = 1$, and we assume that $\exists\delta_0$ such that Assumption 2 holds for $U(s_0, \delta_0) := \{s \in \mathbb{C} : |s - s_0| < \delta_0\}$. Then $\forall\epsilon > 0$, $\exists\delta < \delta_0, \lambda > 0$ such that whenever L satisfies $\lambda_2(L) \geq \lambda$, we have*

$$\sup_{s\in U(s_0,\delta)} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\| < \epsilon\,.$$

The proof is shown in Appendix D. Notice that for given $\epsilon > 0$, the ϵ bound is valid for any $\lambda_2(L) \geq \lambda$, therefore we can prove uniform convergence over compact regions that only contain zeros of $\bar{g}(s)$ with nodal multiplicity 1.

Theorem 6 (Uniform convergence around points with $\mathcal{N}(s_0) = 1$)**.** *Let $T(s), \bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. For a compact set $S \subset \mathbb{C}$ satisfying Assumption 2, if S does not contain any pole of $\bar{g}(s)$, and $\mathcal{N}(s) \leq 1, \forall s \in S$, then we have*

$$\lim_{\lambda_2(L)\to+\infty} \sup_{s\in S} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\| = 0\,.$$

Proof. Firstly, let $\{s_k, 1 \leq k \leq m\}$ be the set of all the zeros of $\bar{g}(s)$ within S. Then $\mathcal{N}(s_k) = 1,\ 1 \leq k \leq m$, and by Lemma 5, $\forall\epsilon > 0$ and every $1 \leq k \leq m$, $\exists\,\delta_{s_k}, \lambda_{s_k} > 0$ such that $\forall L$ satisfying $\lambda_2(L) \geq \lambda_{s_k}$, the following holds:

$$\sup_{s\in U(s_k,\delta_{s_k})} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\| < \epsilon\,.$$

Let $\hat{S} := S \setminus (\bigcup_{k=1}^m U(s_k, \delta_{s_k}))$, then we know that $\hat{S}$ is a compact set that does not contain any pole or zero of $\bar{g}$. By Proposition 5, $\exists\hat{\lambda}$ such that

$$\sup_{s\in\hat{S}} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\| < \epsilon\,.$$

Let $\lambda = \max\left\{\hat{\lambda}, \lambda_{s_1}, \cdots, \lambda_{s_m}\right\}$, then $\forall L$ satisfying $\lambda_2(L) \geq \lambda$, we have:

$$\begin{aligned}
&\sup_{s\in S} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\| \\
&= \max\left\{ \sup_{s\in\hat{S}} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\|, \right. \\
&\qquad \left. \sup_{s\in\bigcup_{k=1}^m U(s_k,\delta_{s_k})} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\| \right\} < \epsilon\,,
\end{aligned}$$

which proves the limit. □

For zeros with nodal multiplicity strictly larger than 1, the analysis is rather complicated. We first look once again at the homogeneous node dynamics setting of Section II to provide some insight.

Example 1. *Consider again a homogeneous network with node dynamics $g(s)$ and $f(s) = 1$, where the transfer matrix is given by*

$$T(s) = \frac{1}{n}g(s)\mathbb{1}\mathbb{1}^T + V_\perp \mathrm{diag}\left\{\frac{1}{g^{-1}(s) + \lambda_i(L)}\right\}_{i=2}^n V_\perp^T\,.$$

The poles of $T(s)$ include 1) the poles of $g(s)$, and 2) any point s_0 that satisfies $g^{-1}(s_0)+\lambda_i(L) = 0$ for a particular i. Notice that if $\lambda_i(L)$ is large, every solution to $g^{-1}(s_0)+\lambda_i(L) = 0$ is close to one of the zeros of $g(s)$. As we increase $\lambda_2(L)$, which effectively increases every $\lambda_i(L), 2 \leq i \leq n$, one can check that at most $n - 1$ poles asymptotically approach each zero of $g(s)$, provided that $\lambda_i(L)$ are distinct. As a result, uniform convergence around any zero of $g(s)$ cannot be obtained due to the presence of poles of $T(s)$ close to them.

Such observation also seems to hold in general for networks with heterogeneous node dynamics $g_i(s)$. That is, if a zero of $\bar{g}(s)$ is a zero with nodal multiplicity strictly larger than 1, then we expect it to "attract" poles of $T(s)$. But it is difficult to formally prove it since we cannot exactly locate the poles of $T(s)$ in the absence of homogeneity.[3] Surprisingly, there are certain cases where we can still quantify the effect of those poles of $T(s)$ approaching a zero of $\bar{g}(s)$. This essentially disproves the uniform convergence for such cases.

Theorem 7 (Uniform Convergence Failure)**.** *Let $T(s), \bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. Let $f(s) = 1$. Suppose $z \in \mathbb{R}$ is a real zero of all $g_i(s), i \in [n]$ with multiplicity 1, i.e. $\forall i \in [n]$, $g_i(z) = 0, \lim_{s\to z}\frac{g_i(s)}{s-z} \neq 0$. Then for any set S containing z in its interior, $\exists\lambda, M > 0$ such that, for all Laplacian matrices L satisfying $\lambda_2(L) \geq \lambda$, we have*

$$\sup_{s\in S} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\| \geq M\,.$$

The proof is shown in Appendix C. Although Theorem 7 only disproves the uniform convergence around one particular type of zero of $\bar{g}(s)$, namely, such zero must have nodal multiplicity n and it must be a real zero with multiplicity 1 for all $g_i(s)$, we believe a similar result holds for any zero that is "shared" by multiple $g_i(s)$. However, a complete proof is left for future research.

We now provide another point of view of this phenomenon. Suppose in a network of size 2, z_1 is exclusively zero of $g_1(s)$ and z_2 exclusively zero of $g_2(s)$. If z_1 and z_2 are close enough, there must be p a pole of $\bar{g}(s)$ in the small neighborhood of z_1 or z_2. To be more clear, see the following example.

Example 2. *Let $g_1(s) = \frac{s+a}{s^2}, g_2(s) = \frac{s+a+\epsilon}{s^2}$, then $z_1 = -a$ and $z_2 = -a - \epsilon$ are the zeros respectively. The coherent dynamics is given by*

$$\bar{g}(s) = \frac{2}{g_1^{-1}(s) + g_2^{-1}(s)} = \frac{(s+a)(s+a+\epsilon)}{2s^2(s+a+\epsilon/2)}\,.$$

$\bar{g}(s)$ has a pole $p = -a - \epsilon/2$ that is in both $\epsilon/2$-neighborhoods of z_1 and z_2.

By Theorem 2, we know that p is asymptotically a pole of $T(s)$, in other words, there is a pole of $T(s)$ approaching p, as the network connectivity increases. Moreover, z_1 and z_2 being close enough suggests that p is close to z_1 and z_2, as we see in the example. Consequently, two zeros z_1, z_2 being

[3] We can still exactly locate the poles of $T(s)$ when proportionality is assumed, i.e. $g_i(s) = f_i g(s), i \in [n]$ for some $f_i > 0$ and rational transfer function $g(s)$. Such a case can be regarded as the homogeneous case by considering a scaled version of L.

close introduces a pole of $T(s)$ asymptotically approaching a small neighborhood of z_1, z_2. Consider the limit case where the two zeros collapse into a shared zero of $g_1(s), g_2(s)$, we should expect a pole of $T(s)$ approaching this shared zero.

A similar argument can be made for m zeros of different nodes being close to each other, introducing $m-1$ poles of $\bar{g}(s)$ in the small neighborhood that asymptotically attract poles of $T(s)$. This is by no means a rigorous proof of how uniform convergence fails around a zero "shared" by multiple $g_i(s)$, but rather a discussion providing intuition behind such behavior.

At this point, we have proved uniform convergence of $T(s)$ on a compact set S that does not include 1) zeros of $\bar{g}(s)$ with Nodal Multiplicity larger than 1, or 2) poles of $\bar{g}(s)$.

In particular, we find that zero with Nodal Multiplicity larger than 1, i.e. it is "shared" by multiple $g_i(s)$, attracts pole of $T(s)$ as network connectivity increases, which suggests that uniform convergence of $T(s)$ fails around such point. Although we only provide the proof for special cases as in Theorem 7, we conjecture such a statement is true in general and we left more careful analysis for future research.

C. Uniform Convergence on Right-Half Complex Plane

Aside from uniform convergence on compact sets, uniform convergence over the closed right-half plane $\{s : Re(s) \geq 0\}$ is of great interest as well. If we were to establish uniform convergence over the right-half plane for a certain $T(s)$, then given $\bar{g}(s)$ to be stable, the convergence in $\mathcal{H}_\infty$-norm of $T(s)$ towards $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T$ could be guaranteed, i.e., $T(s)$ converges to $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T$ as a system. One trivial consequence is that we can infer the stability of $T(s)$ with a large enough $\lambda_2(L)$ by the stability of $\bar{g}(s)$. Furthermore, given any $\mathcal{L}_2$ input signal, we can make the $\mathcal{L}_2$ difference between output responses of $T(s)$ and $\frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T$ arbitrarily small by increasing the network connectivity.

Unfortunately, for most networks, we encounter with the same issue we have seen when dealing with zeros of $\bar{g}(s)$: When $g_i(s)$ is strictly proper, $g_i(s) \to 0$ as $|s| \to +\infty$, thus, ∞ can be viewed as a zero of $g_i(s)$ by regarding $g_i(s)$ as functions defined on extended complex plane $\mathbb{C} \cup \{\infty\}$. Then for networks that include more than one node whose transfer functions are strictly proper, there will be poles of $T(s)$ approaching $\{\infty\}$ as $\lambda_2(L)$ increases. Notice that those poles could approach $\{\infty\}$ either from the left-half or right-half plane. Apparently, the uniform convergence on the right-half plane will not hold if the latter happens, but even when the former happens, we still need to quantify the effect of such poles because they are approaching the boundary of our set $\{s : Re(s) \geq 0\}$. A similar argument can be made for any set of the form $\{s : Re(s) = \sigma\} = \{\sigma + j\omega : \omega \in [-\infty, +\infty]\}$, which we mentioned in II-B.

Although proving (or disproving) uniform convergence on the right-half plane for general networks is quite challenging, it is much more straightforward for networks consist of only non-strictly proper nodes, as shown in the following theorem:

Theorem 8 (Sufficient condition for uniform convergence on right-half plane)**.** *Let $T(s), \bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. Suppose $g_i(s), i \in [n]$ are not strictly proper, $\bar{g}(s)$ is stable, and $\mathcal{N}(s) \leq 1, \forall Re(s) \geq 0$, then we have*

$$\lim_{\lambda_2(L)\to+\infty} \sup_{Re(s)\geq 0} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\| = 0 \,.$$

Proof. Given $R > 0$, we define the following sets:

$$S_1 := \{s \in \mathbb{C} : Re(s) \geq 0, |s| \leq R\}\,,$$
$$S_2 := \{s \in \mathbb{C} : Re(s) \geq 0, |s| > R\}\,.$$

Apparently, $S_1 \bigcup S_2 = \{s \in \mathbb{C} : Re(s) \geq 0\}$. Then we can show uniform convergence on right-half plane by proving uniform convergence on S_1, S_2 respectively:

Firstly, because all $g_i(s)$ are not strictly proper, each $g_i(s)$ converges to some non-zero value as $|s| \to +\infty$. Then we can choose R large enough so that $\max_{1\leq i\leq n} |g_i^{-1}(s)| \leq M_2, \forall s \in S_2$ for some $M_2 > 0$. Moreover, $|\bar{g}(s)| < M_1, \forall s \in S_2$ for some $M_1 > 0$ because $\bar{g}(s)$ is stable. Then the conditions in Lemma 4 are satisfied, we have

$$\sup_{s\in S_2} \left\| T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T \right\| \leq \frac{(M_1M_2+1)^2}{F_l\lambda_2(L) - M_2 - M_1M_2^2}\,, \tag{22}$$

where $F_l = \inf_{s\in S} |f(s)|$. Taking $\lambda_2(L) \to +\infty$ on both side of (22), we have the uniform convergence on S_2.

Secondly, notice that S_1 is compact, contains no pole of $\bar{g}(s)$ and has $\mathcal{N}(s) \leq 1, \forall s \in S_1$, the uniform convergence is shown by Theorem 6. □

V. Coherence and Dynamics Concentration in Large-scale Networks

Until now we looked into convergence results of $T(s)$ for networks with fixed size n. However, one could easily see that such coherence does not depend on the network size n. In particular, the right-hand side of (11) only depends on n via $\lambda_2(L)$ as long as the bounds regarding $g_i(s)$, i.e. M_1 and M_2 do not scale with respect to n. This implies that coherence can emerge as the network size increases. This is the topic of this section.

More interestingly, in a stochastic setting where all $g_i(s)$ are unknown transfer functions independently drawn from some distribution, their harmonic mean eventually converges in probability to a deterministic transfer function as the network size increases. Consequently, a large-scale stochastic network concentrates to deterministic a system. We term this phenomenon *dynamics concentration*.

A. Coherence in Large-scale Networks

To start with, we revise the problem settings to account for variable network size: Let $\{g_i(s), i \in \mathbb{N}_+\}$ be a sequence of transfer functions, and $\{L_n, n \in \mathbb{N}_+\}$ be a sequence of real symmetric Laplacian matrices such that L_n has order n, particularly, let $L_1 = 0$. Then we define a sequence of transfer matrix $T_n(s)$ as

$$T_n(s) = (I_n + G_n(s)L_n)^{-1}\, G_n(s)\,, \tag{23}$$

where $G_n(s) = \mathrm{diag}\{g_1(s), \cdots, g_n(s)\}$. This is exactly the same transfer matrix shown in Fig.1 for a network of size n. We can then define the coherent dynamics for every $T_n(s)$ as

$$\bar{g}_n(s) = \left(\frac{1}{n}\sum_{i=1}^{n} g_i^{-1}(s)\right)^{-1} . \tag{24}$$

For certain family $\{L_n, n \in \mathbb{N}_+\}$ of large-scale networks, the network algebraic connectivity $\lambda_2(L_n)$ increases as n grows. For example, when L_n is the Laplacian of a complete graph of size n with all edge weights being 1, we have $\lambda_2(L_n) = n$. As a result, network coherence naturally emerges as the network size grows. Recall that to prove the convergence of $T_n(s)$ to $\frac{1}{n}\bar{g}_n(s)\mathbb{1}\mathbb{1}^T$ for fixed n, we essentially seek for $M_1, M_2 > 0$, such that $|\bar{g}_n(s)| \leq M_1$ and $\max_{1\leq i\leq n} |g_i^{-1}(s)| \leq M_2$ for s in a certain set. If it is possible to find a universal $M_1, M_2 > 0$ for all n, then the convergence results should be extended to arbitrarily large networks, provided that network connectivity increases as n grows. To state this condition formally, we need the notion of uniform boundedness for a family of functions.

Definition 3. *Let $\{g_i(s), i \in I\}$ be a family of complex functions indexed by I. Given $S \subset \mathbb{C}$, $\{g_i(s), i \in I\}$ is uniformly bounded on S if*

$$\exists M > 0 \quad s.t. \quad |g_i(s)| \leq M, \quad \forall i \in I,\ \forall s \in S .$$

Now we are ready to show uniform convergence of $T_n(s)$:

Theorem 9. *Let $T_n(s)$ and $\bar{g}_n(s)$ be defined as in* (23) *and* (24)*, respectively. Suppose $\lambda_2(L_n) \to +\infty$ as $n \to \infty$. If both $\{g_i^{-1}(s), i \in \mathbb{N}_+\}$ and $\{\bar{g}_n(s), n \in \mathbb{N}_+\}$ are uniformly bounded on a set $S \subset \mathbb{C}$. then we have*

$$\lim_{n\to\infty} \sup_{s\in S} \left\| T_n(s) - \frac{1}{n}\bar{g}_n(s)\mathbb{1}\mathbb{1}^T \right\| = 0 .$$

Proof. Since both $\{g_i^{-1}(s), i \in \mathbb{N}_+\}$ and $\{\bar{g}_n(s), n \in \mathbb{N}_+\}$ are uniformly bounded on S, $\exists M_1, M_2 > 0$ s.t. $|\bar{g}_n(s)| \leq M_1$ and $\max_{1\leq i\leq n} |g_i^{-1}(s)| \leq M_2$ for every $n \in \mathbb{N}_+$ and $s \in S$. By Lemma 4, $\forall n \in \mathbb{N}_+$,

$$\sup_{s\in S} \left\| T_n(s) - \frac{1}{n}\bar{g}_n(s)\mathbb{1}\mathbb{1}^T \right\| \leq \frac{(M_1M_2+1)^2}{F_l\lambda_2(L_n) - M_2 - M_1M_2^2} , \tag{25}$$

where $F_l = \inf_{s\in S} |f(s)|$. We already assumed that $\lambda_2(L_n) \to +\infty$ as $n \to +\infty$, therefore the proof is finished by taking $n \to +\infty$ on both sides of (25). ☐

Remark 3. *Similarly to Theorem 5, uniform convergence is achieved on a set away from zeros or poles of $\bar{g}(s)$. The uniform boundedness condition is preventing any point in the closure of S from asymptotically becoming a zero of any $g_i^{-1}(s)$ or a pole of $\bar{g}(s)$ as n increases.*

B. Dynamics Concentration in Large-scale Networks

Now we consider the cases where the node dynamics are unknown (stochastic). For simplicity, we constraint our analysis to the setting where the node dynamics are independently sampled from the same random rational transfer function with all or part of the coefficients are random variables, i.e. the nodal transfer functions are of the form

$$g_i(s) \sim \frac{b_m s^m + \ldots b_1 s + b_0}{a_l s^l + \ldots a_1 s + a_0} , \tag{26}$$

for some $m, l > 0$, where $b_0, \cdots, b_m$, $a_0, \cdots, a_l$ are random variables.

To formalize the setting, we firstly define the random transfer function to be sampled. Let $\Omega = \mathbb{R}^d$ be the sample space, $\mathcal{F}$ the Borel σ-field of Ω, and $\mathbb{P}$ a probability measure on Ω. A sample $w \in \Omega$ thus represents a d-dimensional vector of coefficients. We then define a random rational transfer function $g(s, w)$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that all or part of the coefficients of $g(s, w)$ are random variables. Then for any $w_0 \in \Omega$, $g(s, w_0)$ is a rational transfer function.

Now consider the probability space $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$. Every $\mathbf{w} \in \Omega^\infty$ give an instance of samples drawn from our random transfer function:

$$g_i(s, w_i) := g(s, w_i), i \in \mathbb{N}_+ ,$$

where w_i is the i-th element of $\mathbf{w}$. By construction, $g_i(s, w_i), i \in \mathbb{N}_+$ are i.i.d. random transfer functions. Moreover, for every $s_0 \in \mathbb{C}$, $g_i(s_0, w_i), i \in \mathbb{N}_+$ are i.i.d. random complex variables taking values in the extended complex plane (presumably taking value ∞).

Now given $\{L_n, n \in \mathbb{N}_+\}$ a sequence of $n \times n$ real symmetric Laplacian matrices, consider the random network of size n whose nodes are associated with the dynamics $g_i(s, w_i), i = 1, 2, \cdots, n$ and coupled through L_n. The transfer matrix of such a network is given by

$$T_n(s, \mathbf{w}) = (I_n + G_n(s, \mathbf{w})L_n)^{-1} G_n(s, \mathbf{w}) , \tag{27}$$

where $G_n(s, \mathbf{w}) = \mathrm{diag}\{g_1(s, w_1), \cdots, g_n(s, w_n)\}$.

Then under this setting, the coherent dynamics of the network is given by

$$\bar{g}(s, \mathbf{w}) = \left(\frac{1}{n}\sum_{i=1}^{n} g_i^{-1}(s, w_i)\right)^{-1} . \tag{28}$$

Now given a compact set $S \subset \mathbb{C}$ of interest, and assuming suitable conditions on the distribution of $g(s, w)$, we expect that the random coherent dynamics $\bar{g}(s, \mathbf{w})$ would converge uniformly in probability to its expectation

$$\hat{g}(s) = \left(\mathbb{E}g^{-1}(s, w))\right)^{-1} := \left(\int_\Omega g^{-1}(s, w) d\mathbb{P}(w)\right)^{-1} , \tag{29}$$

for all $s \in S$, as $n \to \infty$. The following Lemma provides a sufficient condition for this to hold.

Lemma 6. *Consider the probability space $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$. Let $\bar{g}_n(s, \mathbf{w})$ and $\hat{g}(s)$ be defined as in* (28) *and* (29)*, respectively, and given a compact set $S \subset \mathbb{C}$, let the following conditions hold:*

1) *$g^{-1}(s, w)$ is uniformly bounded on $S \times \Omega$;*
2) *$\{\bar{g}_n(s, \mathbf{w}), n \in \mathbb{N}_+\}$ are uniformly bounded on $S \times \Omega^\infty$;*
3) *$\exists L > 0$ s.t. $|g_1^{-1}(s_1, w) - g_1^{-1}(s_2, w)| \leq L|s_1 - s_2|$, $\forall w \in \Omega, \forall s_1, s_2 \in S$;*
4) *$\hat{g}(s)$ is uniformly continuous.*

Then, $\forall \epsilon > 0$*, we have*

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{s\in S}\left\|\frac{1}{n}\bar{g}_n(s,\mathbf{w})\mathbb{1}\mathbb{1}^T - \frac{1}{n}\hat{g}(s)\mathbb{1}\mathbb{1}^T\right\| \geq \epsilon\right) = 0\,.$$

The proof is shown in Appendix E. This lemma suggests that our coherent dynamics $\bar{g}_n(s,\mathbf{w})$, as n increases, converges uniformly on S to its expected version $\hat{g}(s)$. Then provided that the coherence is obtained as the network size grows, we would expect that the random transfer matrix $T_n(s,\mathbf{w})$ to concentrate to a deterministic one $\frac{1}{n}\hat{g}(s)\mathbb{1}\mathbb{1}^T$, as the following theorem shows.

Theorem 10. *Given probability space* $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$*. Let* $T_n(s,\mathbf{w})$ *and* $\hat{g}(s)$ *be defined as in* (27) *and* (29)*, respectively. Suppose* $\lambda_2(L_n) \to +\infty$ *as* $n \to +\infty$*. Given a compact set* $S \subset \mathbb{C}$*, if all the conditions in Lemma 6 hold, then* $\forall \epsilon > 0$*, we have*

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{s\in S}\left\|T_n(s,\mathbf{w}) - \frac{1}{n}\hat{g}(s)\mathbb{1}\mathbb{1}^T\right\| \geq \epsilon\right) = 0\,.$$

Proof. Firstly, notice that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{s\in S}\left\|T_n(s,\mathbf{w}) - \frac{1}{n}\hat{g}(s)\mathbb{1}\mathbb{1}^T\right\| \geq \epsilon\right)\\
\leq\, &\mathbb{P}\left(\sup_{s\in S}\left\|T_n(s,\mathbf{w}) - \frac{1}{n}\bar{g}_n(s)\mathbb{1}\mathbb{1}^T\right\| + \right.\\
&\qquad\left.\sup_{s\in S}\left\|\frac{1}{n}\bar{g}_n(s,\mathbf{w})\mathbb{1}\mathbb{1}^T - \frac{1}{n}\hat{g}(s)\mathbb{1}\mathbb{1}^T\right\| \geq \epsilon\right)\\
\leq\, &\mathbb{P}\left(\sup_{s\in S}\left\|T_n(s,\mathbf{w}) - \frac{1}{n}\bar{g}_n(s,\mathbf{w})\mathbb{1}\mathbb{1}^T\right\| \geq \frac{\epsilon}{2}\right) + \\
&\qquad\mathbb{P}\left(\sup_{s\in S}\left\|\frac{1}{n}\bar{g}_n(s,\mathbf{w})\mathbb{1}\mathbb{1}^T - \frac{1}{n}\hat{g}(s)\mathbb{1}\mathbb{1}^T\right\| \geq \frac{\epsilon}{2}\right)\,.
\end{aligned}$$

The second term converges to 0 as $n \to +\infty$ by Lemma 6. For the first term, we show below that it becomes exactly 0 for large enough n. Still, we assume $\{\bar{g}_n(s,\mathbf{w})\}$ and $\{g_i^{-1}(s,\mathbf{w})\}$ are uniformly bounded on S by $M_1, M_2 > 0$ respectively. By Lemma 4, choosing large enough n s.t.

$$\begin{aligned}
&\mathbb{P}\left(\sup_{s\in S}\left\|T_n(s,\mathbf{w}) - \frac{1}{n}\bar{g}_n(s,\mathbf{w})\mathbb{1}\mathbb{1}^T\right\| \geq \frac{\epsilon}{2}\right)\\
\leq\, &\mathbb{P}\left(\frac{(M_1M_2+1)^2}{F_l\lambda_2(L_n) - M_2 - M_1M_2^2} \geq \frac{\epsilon}{2}\right)\,,
\end{aligned}$$

then we can choose even larger n such that the probability on the right-hand side is 0 because $\lambda_2(L_n) \to +\infty$ as $n \to \infty$. ☐

Remark 4. *Lemma 6 requires* $g^{-1}(s,w)$ *to be uniformly bounded on* $S \times \Omega$*. That is, for* $s_0 \in S$*,* $g^{-1}(s_0,w)$ *is a bounded complex random variable. In [1], a weaker condition, that* $g^{-1}(s_0,w)$ *is a sub-Gaussian complex random variable, is considered. This allows to show that point-wise convergence in probability can be achieved whenever* $\lambda_2(L_n)$ *grows polynomially in* n*.*

In summary, because the coherent dynamics of the tightly-connected network is given by the harmonic mean of all node dynamics $g_i(s)$, it concentrates to its harmonic expectation $\hat{g}(s)$ as the network size grows. As a result, in practice, the coherent behavior of large-scale tightly-connected networks depends on the empirical distribution of $g_i(s)$, i.e. a collective effect of all node dynamics rather than every individual node dynamics. For example, two different realizations of large-scale network with dynamics $T_n(s,\mathbf{w})$ exhibit similar coherent behavior with high probability, in spite of the possible substantial differences in individual node dynamics.

VI. Examples and Applications

In this section, we give two examples of coherent networks. Firstly, we look into the first order continuous time consensus networks. As opposed to the standard analysis in state-space, we can view the dynamic behavior of reaching consensus as a result of network dynamics getting close to its coherent dynamics $\bar{g}(s)$. This example also gives an illustration of dynamics concentration in large-scale consensus networks.

Secondly, we apply our analysis to investigate coherence in power networks. For coherent generator groups, we find that $\frac{1}{n}\hat{g}(s)$ generalized typical aggregate generator models which are often used for model reduction in power networks [26]. Moreover, we show that heterogeneity in generator dynamics usually leads to high-order aggregate dynamics, making it challenging to find a reasonably low-order approximation.

A. Consensus Networks: A Frequency Domain Viewpoint

Consider the network in Fig. 1 with the dynamics of each node given by

$$g_i(s) = \frac{k_i}{s},$$

and $f(s) = 1$. This gives us the standard continuous time consensus network [11]. When $g_i(s)$ is subject to a unit impulse input $\delta(t)$, it is equivalent to setting initial condition at $t = 0$ of node i to be k_i.

We sample k_i from $Unif(1,5)$, then $\frac{s}{k_i}$ is a bounded random variable for any fixed s. The graph of the network is a k-regular ring, i.e., every node is connected to its $2k$ nearest neighbors. We set the weight to be 1 for all edges and $k \approx 0.15n$, then each node is connected to roughly $1/3$ of other nodes in the network. It can be shown that the algebraic connectivity of the graph Laplacian $|\lambda_2|$ is $\Omega(n)$ [27], then this network should exhibit dynamics concentration, according to the convergence result in Section V. The coherent dynamics is given by

$$\bar{g}(s) = \frac{1}{s}\left(\sum_{i=1}^n k_i^{-1}\right)^{-1},$$

and the expected dynamics is given by

$$\hat{g}(s) = \frac{1}{s}\left(\mathbb{E}k_i^{-1}\right)^{-1} = \frac{4}{\ln 5}\frac{1}{s}.$$

Suppose the network is subject to an impulse input $u(t) = \delta(t)\mathbb{1}$. Then, the impulse response of $\hat{g}(s)$ subject to $\frac{1}{n}\mathbb{1}^T u(t) = \delta(t)$ can represent the coherent response of all nodes, which is given by $\hat{g}(t) = \frac{4}{\ln 5}\chi_{\geq 0}(t)$. Here $\chi_{\geq 0}(t)$ is the unit step function.

We plot the impulse responses for networks with different size $n = 20, 50, 100, 500$ in Fig. 2. As network size grows, the

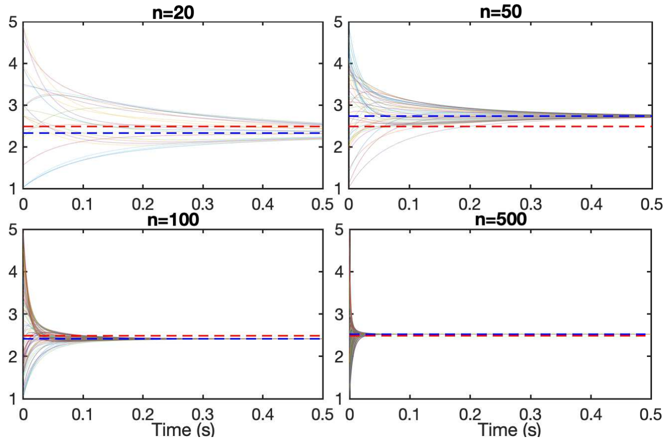

Fig. 2. Coherence and dynamics concentration in the consensus network. Response from individual nodes are plotted in solid lines, and we also plot coherent response $\bar{g}(t)$ (blue) and expected coherent response $\hat{g}(t)$ (red) in solid lines.

network exhibits a more coherent response and the consensus value gets close to the expected average $4/\ln 5$ as suggested by $\hat{g}(s)$. To be specific, we effectively increase the network algebraic connectivity $\lambda_2(L)$ by letting the network size grow. As a result, the time-domain response of every node in the network is getting close to the impulse response of $\bar{g}(s)$, and eventually the one of $\hat{g}(s)$.

B. Aggregate Dynamics of Synchronous Generator Networks

We now look at the case of power networks and leverage our analysis to provide a general framework to aggregate coherent generator groups. Consider the transfer matrix of power generator networks [7] linearized around its steady-state point, given by the following block diagram: This is exactly

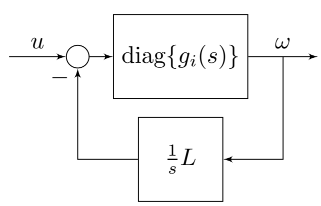

Fig. 3. Block Diagram of Linearized Power Networks

the block structure shown in Fig. 1 with $f(s) = \frac{1}{s}$. Here, the network output, i.e., the frequency deviation of each generator, is denoted by ω. Generally, the $g_i(s)$ are modeled as strictly positive real transfer functions and we assume L is connected.

1) Coherence Analysis: We utilize the convergence results from previous sections to characterize the coherent behavior of such networks. We still denote the coherent dynamics as $\bar{g}(s)$ defined in (5).

Firstly, Notice that $s = 0$ is a pole of $f(s) = \frac{1}{s}$, then by Theorem 4, $T(0)$ is exactly $\frac{1}{n}\bar{g}(0)\mathbb{1}\mathbb{1}^T$, which suggests that in steady state, frequency outputs are the same for all generators. Moreover, another consequence of Theorem 4 is as follows.

Claim 1. *Given fixed L, $\forall \epsilon > 0$, $\exists \eta_c > 0$ such that*

$$\sup_{\eta\in[-\eta_c,\eta_c]} \left\| T(j\eta) - \frac{1}{n}\bar{g}(j\eta)\mathbb{1}\mathbb{1}^T \right\| < \epsilon\,.$$

Proof. This is the direct application of Theorem 4 according to the definition of the limit. □

The claim suggests that the network is naturally coherent in the low-frequency range. In other words, for any fixed network L, there is a low-frequency range such that the network responds coherently to disturbances in that frequency range. Additionally, we know that such frequency range can be arbitrarily wide, given sufficiently large network connectivity, suggested by the following claim.

Claim 2. *$\forall \eta_c > 0$, we have*

$$\lim_{\lambda_2(L)\to\infty} \sup_{\eta\in[-\eta_c,\eta_c]} \left\| T(j\eta) - \frac{1}{n}\bar{g}(j\eta)\mathbb{1}\mathbb{1}^T \right\| = 0\,.$$

Proof. Provided that $g_i(s)$ are strictly positive real [28], this is a direct application of Theorem 6. □

Loosely speaking, the generator network is coherent for certain low-frequency range and the width of such frequency range increases when the network is more connected ($\lambda_2(L)$ increases).

Furthermore, notice that ∞ can be regarded as a zero of $f(s)$, around which the network effect diminishes, i.e. for sufficiently large η, the effective algebraic connectivity $|f(j\eta)\lambda_2(L)| = \frac{\lambda_2(L)}{|\eta|}$ can be arbitrarily small. Hence given any fixed L, there is a high-frequency range such that the network does not exhibit coherence under disturbances within such range.

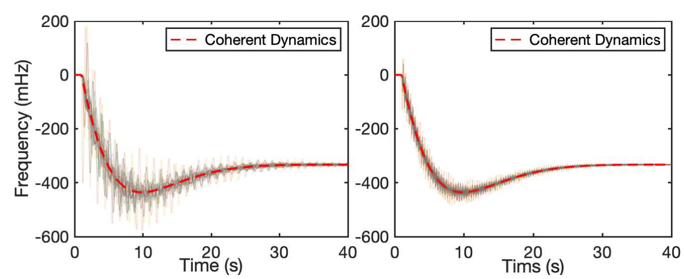

Fig. 4. Step responses of Icelandic Grid without (Left) and with (Right) connectivity $\lambda_2(L)$ scaled up. The responses of coherent dynamics $\bar{g}(s)$ are shown in red dashed lines.

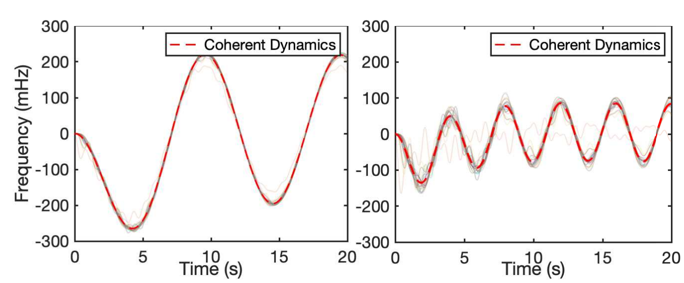

Fig. 5. Responses of Icelandic Grid under sinusoidal disturbances of frequency $w_{\text{low}} = 0.1\text{rad/s}$ (Left) and $w_{\text{high}} = 0.25\text{rad/s}$ (Right). The responses of coherent dynamics $\bar{g}(s)$ are shown in red dashed lines.

We verify our analysis with simulations on the Icelandic power grid [29]. As shown in Fig. 4, the network step response is more coherent, i.e. response of every single node (generator) is getting closer to the one of the coherent dynamics $\bar{g}(s)$, when the network connectivity is scaled up. For the plot on the right, the connectivity $\lambda_2(L)$ is scaled up to 10 times the original. Then Fig. 5 shows such coherence is also frequency-dependent, where the generators respond less coherently under disturbances of higher frequency.

The discussions and simulations above suggest that the coherent dynamics $\bar{g}(s)$ characterize well the overall response/behavior of generators. This leads to a general methodology to analyze the aggregate dynamics of such networks, that we describe next.

2) Aggregate Dynamics of Generator Networks: Let

$$g_{\mathrm{aggr}}(s) := \frac{1}{n}\bar{g}(s) = \left(\sum_{i=1}^{n} g_i^{-1}(s)\right).$$

Our analysis suggests that the transfer function $T(s)$ representing a network of generators is close $g_{\mathrm{aggr}}(s)\mathbb{1}\mathbb{1}^T$ within the low-frequency range, for sufficiently high network connectivity $\lambda_2(L)$. We can also view $g_{\mathrm{aggr}}(s)$ as the aggregate generator dynamics, in the sense that it takes the sum of disturbances $\mathbb{1}^T u = \sum_{i=1}^n u_i$ as its input, and its output represents the coherent response of all generators.

Such a notion of aggregate dynamics is important in modeling large-scale power networks [26]. Generally speaking, one seeks to find an aggregate dynamic model for a group of generators using the same structure (transfer function) as individual generator dynamics, i.e. when generator dynamics are modeled as $g_i(s) = g(s;\theta_i)$, where θ_i is a vector of parameters representing physical properties of each generator, existing works [30]–[32] propose methods to find aggregate dynamics of the form $g(s;\theta_{\mathrm{aggr}})$ for certain structures of $g(s;\theta)$. Our $g_{\mathrm{aggr}}(s)$ justifies their choices of θ_{aggr}, as shown in the following example.

Example 3. *For generators given by the swing model* $g_i(s) = \frac{1}{m_i s + d_i}$ *, where* m_i, d_i *are the inertia and damping of generator* i*, respectively. The aggregate dynamics are*

$$g_{\mathrm{aggr}}(s) = \frac{1}{m_{\mathrm{aggr}} s + d_{\mathrm{aggr}}}\,, \tag{30}$$

where $m_{\mathrm{aggr}} = \sum_{i=1}^n m_i$ *and* $d_{\mathrm{aggr}} = \sum_{i=1}^n d_i$*.*

Here the parameters are $\theta = \{m, d\}$. The aggregate model given by (30) is consistent with the existing approach of choosing inertia m and damping d as the respective sums over all the coherent generators.

However, as we show in the next example, when one considers more involved models, it is challenging to find parameters θ_{aggr} that accurately fit the aggregate dynamics.

Example 4. *For generators given by the swing model with turbine droop*

$$g_i(s) = \frac{1}{m_i s + d_i + \frac{r_i^{-1}}{\tau_i s + 1}}\,, \tag{31}$$

where r_i^{-1} *and* τ_i *are the droop coefficient and turbine time constant of generator* i*, respectively. The aggregate dynamics are given by*

$$g_{\mathrm{aggr}}(s) = \frac{1}{m_{\mathrm{aggr}} s + d_{\mathrm{aggr}} + \sum_{i=1}^n \frac{r_i^{-1}}{\tau_i s + 1}}\,. \tag{32}$$

Here the parameters are $\theta = \{m, d, r^{-1}, \tau\}$. This example illustrates, in particular, the difficulty in aggregating generators with heterogeneous turbine time constants. When all generators have the same turbine time constant $\tau_i = \tau$, then $g_{\mathrm{aggr}}(s)$ in (32) reduces to the typical effective machine model

$$g_{\mathrm{aggr}}(s) = \frac{1}{m_{\mathrm{aggr}} s + d_{\mathrm{aggr}} + \frac{r_{\mathrm{aggr}}^{-1}}{\tau s + 1}}\,,$$

where $r_{\mathrm{aggr}}^{-1} = \sum_{i=1}^n r_i^{-1}$, i.e. the aggregation model is still obtained by choosing parameters $\{m, d, r^{-1}\}$ as the respective sums of their individual values.

If the τ_i are heterogeneous, then $g_{\mathrm{aggr}}(s)$ is a high-order transfer function and cannot be accurately represented by a single generator model. The aggregation of generators essentially asks for a low-order approximation of $g_{\mathrm{aggr}}(s)$. Our analysis reveals the fundamental limitation of using conventional approaches seeking aggregate dynamics with the same structure of individual generators. Furthermore, by characterizing the aggregate dynamics in the explicit form $g_{\mathrm{aggr}}(s)$, one can develop more accurate low-order approximation [33]. Lastly, we emphasize that our analysis does not depend on a specific model of generator dynamics $g_i(s)$, hence it provides a general methodology to aggregate coherent generator networks.

VII. Conclusions

This paper provides various convergence results of the transfer matrix $T(s)$ for a tightly-connected network. The analysis leads to useful characterizations of coordinated behavior and justifies the relation between network coherence and network effective algebraic connectivity. Our results suggest that network coherence is also a frequency-dependent phenomenon, which is numerically illustrated in generator networks. Lastly, concentration results for large-scale tightly-connected networks are presented, revealing the exclusive role of the statistical distribution of node dynamics in determining the coherent dynamics of such networks.

As we already see from the numerical example of synchronous generator networks, when the network is coherent, the entire network can be represented by its aggregate dynamics, because the network has all non-zero eigenvalue being sufficiently large to make the dynamics associated with these eigenvalues negligible. One could extend such observation to less coherent networks which has large eigenvalues except for a few small eigenvalues. We can utilize the same frequency domain analysis to these networks, but instead of only keeping the coherent dynamics, we retain the dynamic associated with all the small eigenvalues (These dynamics are generally coupled due to the heterogeneity in node dynamics). In this way, we can accurately capture the important modes in the network dynamics while significantly reduce the complexity of our model.

Furthermore, for large-scale networks with multiple coherent groups, or communities in a more general sense, one could model the inter-community interactions by replacing the dynamics of each community with its coherent one, or more generally, a reduced one. Although clustering, i.e. finding communities, for homogeneous networks can be efficiently done by various methods such as spectral clustering [34] [35], it is still open for research to find multiple coherent groups in heterogeneous dynamical networks.

Appendix

A. Proof of Lemma 3

Proof. To avoid confusion, here we let $\mathbb{1}_n$ denote all one vector $\mathbb{1}$ of length n.

Without loss of generality, assume $\mathcal{I} = \{1, \cdots, m\}$. Then $\tilde{L}$ is the principal submatrix of L by removing first m rows and columns.

By [36, 3.5], the matrix

$$L - \lambda_2(L)\left(I - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^T\right),$$

is positive semi-definite. Now let $v = \begin{bmatrix} 0 \\ \tilde{v}_1 \end{bmatrix}$ where $\tilde{v}_1$ is the unit eigenvector correspoding to $\lambda_1(\tilde{L})$, we have

$$\begin{aligned} & v^T\left(L - \lambda_2(L)\left(I - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^T\right)\right)v \\ &= \tilde{v}_1^T\tilde{L}\tilde{v}_1 - \lambda_2(L)\left(1 - \frac{1}{n}(\mathbb{1}_{n-m}^T\tilde{v}_1)^2\right) \\ &= \lambda_1(\tilde{L}) - \lambda_2(L)\left(1 - \frac{1}{n}(\mathbb{1}_{n-m}^T\tilde{v}_1)^2\right) \geq 0\,. \end{aligned}$$

We also have $\mathbb{1}_{n-m}^T\tilde{v}_1 \leq \|\mathbb{1}_{n-m}\|\|\tilde{v}_1\| = \sqrt{n-m}$, then the desired lower bound on $\lambda_1(\tilde{L})$ is obtained:

$$\lambda_1(\tilde{L}) \geq \lambda_2(L)\left(1 - \frac{1}{n}(\mathbb{1}_{n-m}^T\tilde{v}_1)^2\right) \geq \frac{m}{n}\lambda_2(L)\,.$$

□

B. Convergence of Normalized Transfer Function at Poles of $\bar{g}(s)$

We present here the formal statement regarding the Remark 1 on the convergence of normalized transfer function at poles of $\bar{g}(s)$.

Proposition 11. *Let $T(s)$ and $\bar{g}(s)$ be defined as in* (1) *and* (5)*, respectively. If $s_0 \in \mathbb{C}$ is a pole of $\bar{g}(s)$ and a generic point of $f(s)$, and additionally, we assume $\|\tilde{\Lambda}/\lambda_2(L) - \Lambda_{\lim}\| \to 0$ as $\lambda_2(L) \to +\infty$, then*

$$\lim_{\lambda_2(L)\to+\infty}\left\|\frac{T(s_0)}{\|T(s_0)\|} - \frac{1}{n}\gamma(\Lambda_{\lim})\mathbb{1}\mathbb{1}^T\right\| = 0\,,$$

where $\gamma(\Lambda_{\lim}) = \frac{h_{12}^T\Lambda_{\lim}^{-1}h_{12}}{|h_{12}^T\Lambda_{\lim}^{-1}h_{12}|}$.

Here $\gamma(\Lambda_{\lim})$ is well-defined since $\Lambda_{\lim}$ must be positive definite, otherwise it contradicts the fact that $\tilde{\Lambda}/\lambda_2(L) - I \succeq 0$.

Proof. Since

$$\frac{T(s_0)}{\|T(s_0)\|} = \frac{T(s_0)}{a}\cdot\frac{|a|}{\|T(s_0)\|}\cdot\frac{a}{|a|}\,,$$

where a is defined as in (18). The desired result follows once we show that $\left\|\frac{T(s_0)}{a} - \frac{1}{n}\mathbb{1}\mathbb{1}^T\right\| \to 0$, $\frac{|a|}{\|T(s_0)\|} \to 1$ and $\frac{a}{|a|} \to \gamma(\Lambda_{\lim})$. Recall that $T(s_0) = VH^{-1}V^T$ and

$$H^{-1} = a\begin{bmatrix} 1 \\ -H_{22}^{-1}h_{21} \end{bmatrix}\begin{bmatrix} 1 & -h_{21}^TH_{22}^{-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & H_{22}^{-1} \end{bmatrix}.$$

For the first limit, we have

$$\begin{aligned} & \left\|\frac{T(s_0)}{a} - \frac{1}{n}\mathbb{1}\mathbb{1}^T\right\| \\ &= \left\|\frac{H^{-1}}{a} - e_1e_1^T\right\| \\ &= \left\|\begin{bmatrix} 0 \\ -H_{22}^{-1}h_{21} \end{bmatrix}\begin{bmatrix} 0 & -h_{21}^TH_{22}^{-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & H_{22}^{-1} \end{bmatrix}/a\right\| \\ &= \|h_{21}^TH_{22}^{-1}\|^2 + \|h_{21}^TH_{22}^{-1}h_{21}H_{22}^{-1}\| \\ &\leq 2\|h_{21}\|^2\|H_{22}^{-1}\|^2 \leq \frac{2M^2}{(|f(s_0)|\lambda_2(L) - M)^2} \to 0\,. \end{aligned}$$

The second limit comes from

$$\begin{aligned} & \|T(s_0)\|/|a| \\ &= \left\|a\begin{bmatrix} 1 \\ -H_{22}^{-1}h_{21} \end{bmatrix}\begin{bmatrix} 1 & -h_{21}^TH_{22}^{-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & H_{22}^{-1} \end{bmatrix}\right\|/|a| \\ &\leq \left\|\begin{bmatrix} 1 \\ -H_{22}^{-1}h_{21} \end{bmatrix}\begin{bmatrix} 1 & -h_{21}^TH_{22}^{-1} \end{bmatrix}\right\| + \|H_{22}^{-1}\|/|a| \\ &\leq 1 + \|H_{22}^{-1}h_{21}\|^2 + \|H_{22}^{-1}\|^2\|h_{21}\|^2\,, \end{aligned}$$

and

$$\begin{aligned} & \|T(s_0)\|/|a| \\ &\geq \left\|\begin{bmatrix} 1 \\ -H_{22}^{-1}h_{21} \end{bmatrix}\begin{bmatrix} 1 & -h_{21}^TH_{22}^{-1} \end{bmatrix}\right\| - \|H_{22}^{-1}\|/|a| \\ &\geq 1 + \|H_{22}^{-1}h_{21}\|^2 - \|H_{22}^{-1}\|^2\|h_{21}\|^2\,, \end{aligned}$$

along with the fact that both the upper and the lower bound converge to 1. To see the limit, notice that

$$\begin{aligned} \left|\|H_{22}^{-1}h_{21}\|^2 + \|H_{22}^{-1}\|^2\|h_{21}\|^2\right| &\leq 2\|H_{22}^{-1}\|^2\|h_{21}\|^2 \\ \left|\|H_{22}^{-1}h_{21}\|^2 - \|H_{22}^{-1}\|^2\|h_{21}\|^2\right| &\leq \|H_{22}^{-1}h_{21}\|^2 + \|H_{22}^{-1}\|^2\|h_{21}\|^2 \\ &\leq 2\|H_{22}^{-1}\|^2\|h_{21}\|^2\,, \end{aligned}$$

and

$$2\|H_{22}^{-1}\|^2\|h_{21}\|^2 \leq \frac{2M^2}{(|f(s_0)|\lambda_2(L) - M)^2} \to 0\,.$$

The last limit is obtained from

$$\begin{aligned} \frac{a}{|a|} &= \frac{h_{12}^TH_{22}^{-1}h_{12}}{|h_{12}^TH_{22}^{-1}h_{12}|} \\ &= \gamma(\Lambda_{\lim})\frac{h_{12}^TH_{22}^{-1}h_{12}}{f^{-1}(s_0)\lambda_2^{-1}(L)h_{12}^T\Lambda_{\lim}^{-1}h_{12}}\frac{|f^{-1}(s_0)\lambda_2^{-1}(L)h_{12}^T\Lambda_{\lim}^{-1}h_{12}|}{|h_{12}^TH_{22}^{-1}h_{12}|}\,, \end{aligned}$$

and the fact that $\frac{h_{12}^TH_{22}^{-1}h_{12}}{f^{-1}(s_0)\lambda_2^{-1}(L)h_{12}^T\Lambda_{\lim}^{-1}h_{12}} \to 1$.

To show the limit of the ratio in the end, notice that

$$
\begin{aligned}
&\left|\frac{h_{12}^T H_{22}^{-1} h_{12}}{f^{-1}(s_0)\lambda_2^{-1}(L) h_{12}^T \Lambda_{\lim}^{-1} h_{12}} - 1\right| \\
=& \left|\frac{h_{12}^T (V_\perp^T \tilde{G}(s_0) V_\perp + f(s_0)\tilde{\Lambda})^{-1} h_{12}}{f^{-1}(s_0)\lambda_2^{-1}(L) h_{12}^T \Lambda_{\lim}^{-1} h_{12}} - 1\right| \\
=& \left|\frac{h_{12}^T \left(\frac{V_\perp^T \tilde{G}(s_0) V_\perp}{f(s_0)\lambda_2(L)} + \frac{\tilde{\Lambda}}{\lambda_2(L)}\right)^{-1} h_{12}}{h_{12}^T \Lambda_{\lim}^{-1} h_{12}} - 1\right| \\
=& \left|\frac{h_{12}^T \left[\left(\frac{V_\perp^T \tilde{G}(s_0) V_\perp}{f(s_0)\lambda_2(L)} + \frac{\tilde{\Lambda}}{\lambda_2(L)}\right)^{-1} - \Lambda_{\lim}^{-1}\right] h_{12}}{h_{12}^T \Lambda_{\lim}^{-1} h_{12}}\right| \\
\leq& \frac{\|h_{12}\|^2}{|h_1 2^T \Lambda_{\lim} h_{12}|}\left\|\left(\frac{V_\perp^T \tilde{G}(s_0) V_\perp}{f(s_0)\lambda_2(L)} + \frac{\tilde{\Lambda}}{\lambda_2(L)}\right)^{-1} - \Lambda_{\lim}^{-1}\right\| \\
\leq& \frac{\|h_{12}\|^2}{|h_1 2^T \Lambda_{\lim} h_{12}|} \frac{\|\Lambda_{\lim}^{-1}\|^2 \|\Delta\|}{1 - \|\Lambda_{\lim}^{-1}\|\|\Delta\|},
\end{aligned}
$$

where

$$
\Delta = \frac{V_\perp^T \tilde{G}(s_0) V_\perp}{f(s_0)\lambda_2(L)} + \frac{\tilde{\Lambda}}{\lambda_2(L)} - \Lambda_{\lim}.
$$

Here we use the fact that for some invertible square matrices A, B, we have

$$
\begin{aligned}
\|(A+B)^{-1} - A^{-1}\| =& \ \|(A+B)^{-1}(A+B-A)A^{-1}\| \\
\leq& \ \|(A+B)^{-1}\|\|B\|\|A^{-1}\| \\
\text{(Lemma 2)} \leq& \ \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|B\|}\|B\|\|A^{-1}\| \\
=& \ \frac{\|A^{-1}\|^2\|B\|}{1 - \|A^{-1}\|\|B\|},
\end{aligned}
$$

provided that $1 - \|A^{-1}\|\|B\| > 0$. To finish the proof, we see that

$$
\begin{aligned}
\|\Delta\| \leq& \left\|\frac{V_\perp^T \tilde{G}(s_0) V_\perp}{f(s_0)\lambda_2(L)}\right\| + \left\|\frac{\tilde{\Lambda}}{\lambda_2(L)} - \Lambda_{\lim}\right\| \\
\leq& \ \frac{M}{f(s_0)\lambda_2(L)} + \left\|\frac{\tilde{\Lambda}}{\lambda_2(L)} - \Lambda_{\lim}\right\| \to 0.
\end{aligned}
$$

This shows that

$$
\begin{aligned}
&\left|\frac{h_{12}^T H_{22}^{-1} h_{12}}{f^{-1}(s_0)\lambda_2^{-1}(L) h_{12}^T \Lambda_{\lim}^{-1} h_{12}} - 1\right| \\
\leq& \frac{\|h_{12}\|^2}{|h_1 2^T \Lambda_{\lim} h_{12}|} \frac{\|\Lambda_{\lim}^{-1}\|^2\|\Delta\|}{1 - \|\Lambda_{\lim}^{-1}\|\|\Delta\|} \to 0,
\end{aligned}
$$

which is the desired limit. □

C. Proof of Theorem 7

Proof. Let $U(s_0, R) := \{s \in \mathbb{C} : |s_0 - s| < R\}$ be the open ball centered at s_0 with radius R.

Notice that z is a zero of $\bar{g}(s)$ and it is in the interior of S, then given $M > 0$, $\exists R_1 > 0$, s.t. $U(z, R_1) \subset S$, and $\sup_{s \in U(z,R_1)} |\bar{g}(s)| \leq M$.

Suppose the above is true for some $M > 0$, then

$$
\begin{aligned}
&\sup_{s\in S}\left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\| \\
\geq& \sup_{s\in U(z,R_1)}\left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\| \\
\geq& \sup_{s\in U(z,R_1)}\left(\|T(s)\| - \sup_{s'\in U(z,R_1)}\left\|\frac{1}{n}\bar{g}(s')\mathbb{1}\mathbb{1}^T\right\|\right) \\
=& \sup_{s\in U(z,R_1)}\left(\|T(s)\| - \sup_{s'\in U(z,R_1)}|\bar{g}(s')|\right) \\
\geq& \sup_{s\in U(z,R_1)}\|T(s)\| - M.
\end{aligned}
$$

Apparently we only need to find such M and also show that

$$
\exists s^* \in U(z, R_1), \text{ s.t. } \sup_{s\in U(z,R_1)}\|T(s)\| \geq \|T(s^*)\| \geq 2M.
$$

By the assumption, every $g_i(s)$, $i = 1, \cdots, n$ can be written as

$$
g_i(s) = (s - z)h_i(s).
$$

Here, $h_i(s)$ is rational and we denote

$$
h_i(z) = h_{i0} \quad i = 1, \cdots, n,
$$

where $h_{i0} \in \mathbb{R}$ and $h_{i0} \neq 0$, $i = 1, \cdots, n$. We let $h_{\max} = \max_{i\in[n]}|h_{i0}|$ and $h_{\min} = \min_{i\in[n]}|h_{i0}| > 0$.

Since for $i \in [n]$, $h_i(s)$ is rational and z is not a pole of $h_i(s)$, $\exists R_2 > 0$ s.t. $h_i(s)$,$i = 1, \cdots, n$ are holomorphic on $U(z, R_2)$, i.e. every $h_i(s)$ has expansion

$$
h_i(s) = h_{i0} + \sum_{k=1}^{\infty}\frac{h_i^{(k)}(z)}{k!}(s-z)^k := h_{i0} + r_i(s),
$$

where $h_i^{(k)}(\cdot)$ is the k-th derivative of $h_i(\cdot)$. For every $i \in [n]$, expand $r_i(s)$ in Taylor series, then using Cauchy's estimation formula [37], we have $\forall s \in U(z, \frac{R_2}{2})$,

$$
|r_i(s)| \leq \sum_{k=1}^{\infty}\frac{M_{h_i}}{R_2^k}|s-z|^k = \frac{M_{h_i}\frac{|s-z|}{R_2}}{1 - \frac{|s-z|}{R_2}} \leq \frac{2M_{h_i}}{R_2}|s-z|,
$$

where $M_{h_i} = \max_{|s-z|=R_2}|h_i(s)|$.

For $s \in U(z, \frac{R_2}{2})$, $\text{diag}\{g_i(s)\}$ can be expanded as

$$
\text{diag}\{g_i(s)\} = (s-z)\text{diag}\{h_i(s)\} = (s-z)(H_0 + R(s)),
$$

where $H_0 = \text{diag}\{h_{i0}\}, R(s) = \text{diag}\{r_i(s)\}$. And we also have

$$
\|R(s)\| \leq \max_{i\in[n]}|r_i(s)| \leq M_h|s-z|,
$$

where $M_h = \max_{i\in[n]}\frac{2M_{h_i}}{R_2}$.

Now for $T(s)$, we start from

$$
\begin{aligned}
&\|T(s)\| \\
&= \|(I + \text{diag}\{g_i(s)\}L)^{-1}\text{diag}\{g_i(s)\}\| \\
&= |s-z|\|(I + \text{diag}\{g_i(s)\}L)^{-1}\text{diag}\{h_i(s)\}\| \\
\text{(10a)} &\geq |s-z|\|(I + \text{diag}\{g_i(s)\}L)^{-1}\|\sigma_1(\text{diag}\{h_i(s)\}) \\
&= |s-z|\frac{\sigma_1(\text{diag}\{h_i(s)\})}{\sigma_1(I + \text{diag}\{g_i(s)\}L)}. \qquad \text{(C.1)}
\end{aligned}
$$

For $s \in U(z, \frac{R_2}{2})$, the numerator can be lower bounded by

$$\begin{aligned}\sigma_1(\text{diag}\{h_i(s)\}) = &\ \sigma_1(H_0 + R(s)) \\ \text{(10c)} \geq &\ h_{\min} - \|R(s)\| \\ \geq &\ h_{\min} - M_h|s-z|\,. \end{aligned} \tag{C.2}$$

Then, let

$$L_H = H_0^{1/2} L H_0^{1/2}\,,$$

which is semi-positive definite. the denominator can be upper bounded by

$$\begin{aligned}&\sigma_1(I + \text{diag}\{g_i(s)\}L) \\ &= \sigma_1(I + (s-z)H_0L + (s-z)R(s)L) \\ &= \sigma_1(H_0^{1/2}[I + (s-z)L_H \\ &\qquad + (s-z)H_0^{-1/2}R(s)H_0^{-1/2}L_H]H_0^{-1/2}) \\ &\text{(10b)} \\ &\leq \frac{h_{\max}^{1/2}}{h_{\min}^{1/2}}\sigma_1\left(I + (s-z)L_H + (s-z)H_0^{-1/2}R(s)H_0^{-1/2}L_H\right) \\ &\text{(10c)} \\ &\leq \frac{h_{\max}^{1/2}}{h_{\min}^{1/2}}\left(\sigma_1\left(I + (s-z)L_H\right) + |s-z|\|L_H\|\frac{\|R(s)\|}{h_{\min}}\right)\,. \end{aligned} \tag{C.3}$$

Suppose $\lambda_n(L_H) = \|L_H\| \geq \max\{\frac{1}{R_1}, \frac{2}{R_2}\}$, then for $s^* = z - \frac{1}{\lambda_n(L_H)} \in U(z, \min\{R_1, \frac{R_2}{2}\})$, we have

1) $I + (s^*-z)L_H = I - \frac{L_H}{\lambda_n(L_H)}$ is symmetric and has eigenvalue 0. Then $\sigma_1(I + (s^*-z)L_H) = 0$;
2) $|s^*-z|\|L_H\| = 1$.

From (C.3) and (10c), we have

$$\begin{aligned}&\sigma_1(I + \text{diag}\{g_i(s^*)\}L) \\ &\leq \frac{h_{\max}^{1/2}}{h_{\min}^{1/2}}\left(\sigma_1\left(I + (s^*-z)L_H\right) + |s^*-z|\|L_H\|\frac{\|R(s^*)\|}{h_{\min}}\right) \\ &\leq |s^*-z|\|L_H\|\frac{h_{\max}^{1/2}}{h_{\min}^{3/2}}\|R(s^*)\| \\ &\leq \frac{h_{\max}^{1/2}}{h_{\min}^{3/2}}\|R(s^*)\| \leq \frac{h_{\max}^{1/2}}{h_{\min}^{3/2}}M_h|s^*-z|\,. \end{aligned} \tag{C.4}$$

Combing (C.1)(C.2)(C.4), let $\lambda_n(L_H)$ large enough such that $\lambda_n(L_H) \geq \max\{\frac{2M_h}{h_{\min}}, \frac{1}{R_1}, \frac{2}{R_2}\}$, we have

$$\begin{aligned}\|T(s^*)\| \geq &\ |s^*-z|\frac{\sigma_1(\text{diag}\{h_i(s^*)\})}{\sigma_1(I + \text{diag}\{g_i(s^*)\}L)} \\ \geq &\ \frac{h_{\min} - \frac{M_h}{\lambda_n(L_H)}}{\frac{h_{\max}^{1/2}}{h_{\min}^{3/2}}M_h} \geq \frac{h_{\min}^{5/2}}{2h_{\max}^{1/2}M_h}\,. \end{aligned}$$

Since

$$\lambda_n(L_H) \geq \lambda_n(L)h_{\min} \geq \lambda_2(L)h_{\min}\,,$$

by (10a), we can have $\lambda_n(L_H)$ arbitrarily large by increasing $\lambda_2(L)$.

In summary, we can pick R_2 for the expansion of $h_i(s)$ to exist, then pick $M = \frac{h_{min}^{5/2}}{4h_{\max}^{1/2}M_h}$, which also determines R_1, and let

$$\lambda = \frac{1}{h_{\min}}\max\left\{\frac{2M_h}{h_{\min}}, \frac{1}{R_1}, \frac{2}{R_2}\right\}\,.$$

Then we conclude that $\forall L$ s.t. $\lambda_2(L) \geq \lambda$,

$$\exists s^* = z - \frac{1}{\lambda_n(L_H)} \in U(z, R_1) \subset S,\ \text{s.t. } \|T(s^*)\| \geq 2M\,,$$

which implies

$$\sup_{s\in S}\left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\| \geq \sup_{s\in U(z,R_1)}\|T(s)\| - M \geq M\,.$$

□

D. Proof of Lemma 5

Proof. We denote $U(s_0, R) := \{s \in \mathbb{C} : |s_0 - s| < R\}$ the open ball centered at s_0 with radius R.

By assumptions, we have $F_l = \inf_{s\in U(s_0,\delta_0)} f(s) > 0$ and $F_h = \sup_{s\in U(s_0,\delta_0)} f(s) < \infty$. Without loss of generality, we assume $F_l = 1$. For $s \in U(s_0, \delta_0)$, we have

$$1 \leq |f(s)| \leq F_h\,. \tag{D.1}$$

Consider any set S, we have

$$\sup_{s\in S}\left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\| \leq \sup_{s\in S}\|T(s)\| + \sup_{s\in S}|\bar{g}(s)|\,.$$

For the second term, since $\bar{g}(s)$ is rational and s_0 is a zero of $\bar{g}(s)$, by continuity of $\bar{g}(s)$, $\exists\delta_1 > 0$ such that

$$\sup_{s\in U(s_0,\delta_1)}|\bar{g}(s)| \leq \frac{\epsilon}{2}\,. \tag{D.2}$$

Now we bound the first term. To start with, notice that $\mathcal{N}(s_0) = 1$. Without loss of generality, assume $g_1(s_0) = 0$ and $g_i(s_0) \neq 0, 2 \leq i \leq n$. Then again by continuity of $g_i^{-1}(s), 2 \leq i \leq n$, $\exists\delta_2, M > 0$ such that

$$\sup_{s\in U(s_0,\delta_2)}\max_{2\leq i\leq n}|g_i^{-1}(s)| \leq M\,. \tag{D.3}$$

On the other hand, since s_0 is a zero of $g_1(s)$, one can pick $\delta_3 > 0$ such that

$$\sup_{s\in U(s_0,\delta_3)}|g_1(s)| \leq \frac{\epsilon}{36n + 2\epsilon nMF_h^2}\,. \tag{D.4}$$

We let $\delta = \min\{\delta_0, \delta_1, \delta_2, \delta_3\}$ and we would like to bound $\sup_{s\in U(s_0,\delta)}\|T(s)\|$ by $\frac{\epsilon}{2}$ under sufficiently large $\lambda_2(L)$. This would imply, together with (D.2), that

$$\sup_{s\in U(s_0,\delta)}\left\|T(s) - \frac{1}{n}\bar{g}(s)\mathbb{1}\mathbb{1}^T\right\| \leq \epsilon\,.$$

The remaining of this proof is to bound $\sup_{s\in U(s_0,\delta)}\|T(s)\|$ by $\frac{\epsilon}{2}$.

We write L in block form as $\begin{bmatrix} l_{11} & L_{21}^T \\ L_{21} & \tilde{L}\end{bmatrix}$, by separating its first row and column from the rest. Here $\tilde{L}$ is a grounded Laplacian of L, and $\tilde{L}$ is invertible as long as $\lambda_2(L) \neq 0$ according to Lemma 3.

Since $L\mathbb{1}_n = 0$, we have $L_{21} + \tilde{L}\mathbb{1}_{n-1} = 0$, which gives

$$\tilde{L}^{-1}L_{21} = -\mathbb{1}_{n-1}\,, \tag{D.5a}$$

$$L_{21}^T\tilde{L}^{-1}L_{21} = -L_{21}^T\mathbb{1}_{n-1} = l_{11}\,. \tag{D.5b}$$

We use these equalities later.

We define $\tilde{G}(s) = \mathrm{diag}\{g_2(s), \cdots, g_n(s)\}$ and $\tilde{T}(s) = (I_{n-1} + \tilde{G}(s)f(s)\tilde{L})^{-1}\tilde{G}(s)$. Through some computation, we can write $T(s)$ in the following block form

$$\begin{bmatrix} A(s) & -A(s)f(s)L_{21}^T\tilde{T}(s) \\ -A(s)f(s)\tilde{T}(s)L_{21} & \tilde{T}(s) + A(s)f^2(s)\tilde{T}(s)L_{21}L_{21}^T\tilde{T}(s)\end{bmatrix},$$

where

$$A(s) = \frac{g_1(s)}{1 + g_1(s)\left(f(s)l_{11} - f^2(s)L_{21}^T\tilde{T}(s)L_{21}\right)}\,.$$

.

Then an upper bound of $\|T(s)\|$ is given by

$$\begin{aligned}
&\|T(s)\| \\
\le\;& \left\|A(s)\begin{bmatrix} 1 & f(s)L_{21}^T\tilde{T}(s) \\ f(s)\tilde{T}(s)L_{21} & f^2(s)\tilde{T}(s)L_{21}L_{21}^T\tilde{T}(s)\end{bmatrix}\right\| \\
&\quad + \left\|\begin{bmatrix} 0 & 0 \\ 0 & \tilde{T}(s)\end{bmatrix}\right\| \\
=\;& |A(s)|\left\|\begin{bmatrix} 1 \\ f(s)\tilde{T}(s)L_{21}\end{bmatrix}\begin{bmatrix} 1 \\ f(s)\tilde{T}(s)L_{21}\end{bmatrix}^T\right\| + \|\tilde{T}(s)\| \\
\le\;& |A(s)|\left(1 + \|f(s)\tilde{T}(s)L_{21}\|\right)^2 + \|\tilde{T}(s)\|\,.
\end{aligned} \tag{D.6}$$

Now we bound every term in (D.6) for $s \in U(s_0, \delta)$ through following steps:

1) We firstly bound $|A(s)|$. By Woodbury matrix identity [24, 0.7.4], we have

$$\begin{aligned}
f(s)\tilde{T}(s) =\;& f(s)\left(I_{n-1} + \tilde{G}(s)f(s)\tilde{L}\right)^{-1}\tilde{G}(s) \\
=\;& \left(\tilde{L} + f^{-1}(s)\tilde{G}^{-1}(s)\right)^{-1} \\
=\;& \tilde{L}^{-1} - \tilde{L}^{-1}\left(f^{-1}(s)\tilde{G}(s) + \tilde{L}^{-1}\right)^{-1}\tilde{L}^{-1}\,.
\end{aligned} \tag{D.7}$$

By (D.7) and (D.5), we have

$$\begin{aligned}
& l_{11} - f(s)L_{21}^T\tilde{T}(s)L_{21} \\
=\;& l_{11} - L_{21}^T\left(\tilde{L}^{-1} - \tilde{L}^{-1}\left(f^{-1}(s)\tilde{G}(s) + \tilde{L}^{-1}\right)^{-1}\tilde{L}^{-1}\right)L_{21} \\
=\;& (l_{11} - L_{21}^T\tilde{L}^{-1}L_{21}) \\
&\quad + L_{21}^T\tilde{L}^{-1}\left(f^{-1}(s)\tilde{G}(s) + \tilde{L}^{-1}\right)^{-1}\tilde{L}^{-1}L_{21} \\
&\text{(D.5b)} \\
=\;& L_{21}^T\tilde{L}^{-1}\left(f^{-1}(s)\tilde{G}(s) + \tilde{L}^{-1}\right)^{-1}\tilde{L}^{-1}L_{21} \\
&\text{(D.5a)} \\
=\;& \mathbb{1}_{n-1}^T\left(f^{-1}(s)\tilde{G}(s) + \tilde{L}^{-1}\right)^{-1}\mathbb{1}_{n-1}\,.
\end{aligned} \tag{D.8}$$

When $\lambda_1(\tilde{L}) \ge 2MF_h$, the following holds:

$$\begin{aligned}
&\left|\mathbb{1}_{n-1}^T\left(f^{-1}(s)\tilde{G}(s) + \tilde{L}^{-1}\right)^{-1}\mathbb{1}_{n-1}\right| \\
&\le (n-1)\left\|\left(f^{-1}(s)\tilde{G}(s) + \tilde{L}^{-1}\right)^{-1}\right\| \\
\text{(Lemma 2)} &\le \frac{n-1}{\sigma_1(\tilde{G}(s))|f(s)|^{-1} - \|\tilde{L}^{-1}\|} \\
\text{(D.1)(D.3)} &\le \frac{n-1}{1/(MF_h) - 1/\lambda_1(\tilde{L})} \le 2nMF_h\,,
\end{aligned} \tag{D.9}$$

which when combined with (D.4) gives the following bound on $|A(s)|$:

$$\begin{aligned}
& |A(s)| \\
=\;& \frac{|g_1(s)|}{\left|1 + g_1(s)f(s)\left(l_{11} - L_{21}^Tf(s)\tilde{T}(s)L_{21}\right)\right|} \\
&\text{(D.8)} \\
=\;& \frac{|g_1(s)|}{\left|1 + g_1(s)f(s)\mathbb{1}_{n-1}^T\left(f^{-1}(s)\tilde{G}(s) + \tilde{L}^{-1}\right)^{-1}\mathbb{1}_{n-1}\right|} \\
\le\;& \frac{|g_1(s)|}{1 - |g_1(s)||f(s)|\left|\mathbb{1}_{n-1}^T\left(f^{-1}(s)\tilde{G}(s) + \tilde{L}^{-1}\right)^{-1}\mathbb{1}_{n-1}\right|} \\
&\text{(D.4)(D.9)(D.1)} \\
\le\;& \frac{\frac{\epsilon}{36n + 2n\epsilon MF_h^2}}{1 - \frac{2\epsilon nMF_h^2}{36n + 2n\epsilon MF_h^2}} = \frac{\epsilon}{36n}\,.
\end{aligned} \tag{D.10}$$

2) For $\|f(s)\tilde{T}(s)L_{21}\|$, when $\lambda_1(\tilde{L}) \ge 2M$, we have:

$$\begin{aligned}
\left\|\tilde{T}(s)L_{21}\right\| &= \left\|\left(\tilde{G}^{-1}(s) + f(s)\tilde{L}\right)^{-1}L_{21}\right\| \\
&= \left\|\left(f(s)I + \tilde{L}^{-1}\tilde{G}^{-1}(s)\right)^{-1}\tilde{L}^{-1}L_{21}\right\| \\
\text{(D.5a)} &\le \sqrt{n-1}\left\|\left(f(s)I + \tilde{L}^{-1}\tilde{G}^{-1}(s)\right)^{-1}\right\| \\
\text{(Lemma 2)} &\le \sqrt{n}\left(|f(s)| - \|\tilde{L}^{-1}\|\|\tilde{G}^{-1}(s)\|\right)^{-1} \\
\text{(D.1)(D.3)} &\le \sqrt{n}\left(1 - M/\lambda_1(\tilde{L})\right)^{-1} \le 2\sqrt{n} \le 3\sqrt{n} - 1\,.
\end{aligned} \tag{D.11}$$

3) Lastly, for $\|\tilde{T}(s)\|$, when $\lambda_1(\tilde{L}) \ge M + \frac{4}{\epsilon}$, we have

$$\begin{aligned}
\|\tilde{T}(s)\| &= \left\|\left(\tilde{G}^{-1}(s) + f(s)\tilde{L}\right)^{-1}\right\| \\
\text{(Lemma 2)} &\le \left(|f(s)|\lambda_1(\tilde{L}) - \|\tilde{G}^{-1}(s)\|\right)^{-1} \\
\text{(D.1)(D.3)} &\le \left(\lambda_1(\tilde{L}) - M\right)^{-1} \le \frac{\epsilon}{4}\,.
\end{aligned} \tag{D.12}$$

Apply three bounds obtained from (D.10)(D.11)(D.12) to (D.6), we have:

$$\sup_{s\in U(s_0,\delta)}\|T(s)\| \le \frac{\epsilon}{36n}\cdot 9n + \frac{\epsilon}{4} = \frac{\epsilon}{2}\,,$$

which holds when $\lambda_1(\tilde{L}) \ge \max\{2MF_h, M + \frac{4}{\epsilon}\}$. According to Lemma 3, We can guarantee it by letting $\lambda_2(L) \ge \lambda = n\max\{2MF_h, M + \frac{4}{\epsilon}\}$. We therefore found the desired δ and λ, which finish the proof. □

E. Proof of Lemma 6

Proof. It suffices to show that $\forall \epsilon > 0$,

$$\lim_{n\to+\infty} \mathbb{P}\left(\sup_{s\in S} |\bar{g}_n(s,\mathbf{w}) - \hat{g}(s)| \geq \epsilon\right) = 0\,, \tag{E.1}$$

since $|\bar{g}_n(s,\mathbf{w}) - \hat{g}(s)| = \left\|\frac{1}{n}\bar{g}_n(s,\mathbf{w})\mathbb{1}\mathbb{1}^T - \frac{1}{n}\hat{g}(s)\mathbb{1}\mathbb{1}^T\right\|$.

By the assumptions, $\{\bar{g}_n(s,\mathbf{w}), n \in \mathbb{N}_+, \mathbf{w} \in \Omega^\infty\}$, and $\{g_i^{-1}(s,w), i \in \mathbb{N}_+, w \in \Omega\}$ are uniformly bounded by $M_1 > 0$ and $M_2 > 0$, respectively on S. Then, at any $s \in S$, both $Re\left(g_i^{-1}(s,w)\right)$ and $Im\left(g_i^{-1}(s,w)\right)$ are random variables bounded within $[-M_2, M_2]$. We can simply bound their variances by

$$\begin{aligned}\mathrm{Var}\left(Re\left(g_i^{-1}(s,w)\right)\right) \leq (2M_2)^2 = 4M_2^2\,,\\ \mathrm{Var}\left(Im\left(g_i^{-1}(s,w)\right)\right) \leq (2M_2)^2 = 4M_2^2\,.\end{aligned}$$

Then it follows that

$$\begin{aligned}&\mathrm{Var}\left(Re\left(\bar{g}_n^{-1}(s,\mathbf{w})\right)\right)\\ &= \mathrm{Var}\left(Re\left(n^{-1}\sum_{i=1}^n g_i^{-1}(s,w)\right)\right) \leq 4M_2^2/n\,,\end{aligned}$$

and

$$\begin{aligned}&\mathrm{Var}\left(Im\left(\bar{g}_n^{-1}(s,\mathbf{w})\right)\right)\\ &= \mathrm{Var}\left(Im\left(n^{-1}\sum_{i=1}^n g_i^{-1}(s,w)\right)\right) \leq 4M_2^2/n\,.\end{aligned}$$

By definition of $\hat{g}(s)$ in (28), we have $\mathbb{E}Re\left(\bar{g}_n^{-1}(s,\mathbf{w})\right) = Re\left(\hat{g}(s)\right)$ and $\mathbb{E}Im\left(\bar{g}_n^{-1}(s,\mathbf{w})\right) = Im\left(\hat{g}(s)\right)$, then by Chebyshev's inequality, for $\epsilon > 0$, we have

$$\begin{aligned}&\mathbb{P}\left(\left|\bar{g}_n^{-1}(s,\mathbf{w}) - \hat{g}^{-1}(s)\right| \geq \epsilon\right)\\ &\leq \mathbb{P}\left(\left|Re\left(\bar{g}_n^{-1}(s,\mathbf{w})\right) - Re\left(\hat{g}^{-1}(s)\right)\right| + \right.\\ &\qquad\left.\left|Im\left(\bar{g}_n^{-1}(s,\mathbf{w})\right) - Im\left(\hat{g}^{-1}(s)\right)\right| \geq \epsilon\right)\\ &\leq \mathbb{P}\left(\left|Re\left(\bar{g}_n^{-1}(s,\mathbf{w})\right) - Re\left(\hat{g}^{-1}(s)\right)\right| \geq \epsilon/2\right) +\\ &\qquad \mathbb{P}\left(\left|Im\left(\bar{g}_n^{-1}(s,\mathbf{w})\right) - Im\left(\hat{g}^{-1}(s)\right)\right| \geq \epsilon/2\right)\\ &\leq \frac{4\mathrm{Var}\left(Re\left(\bar{g}_n^{-1}(s,\mathbf{w})\right)\right)}{\epsilon^2} + \frac{4\mathrm{Var}\left(Im\left(\bar{g}_n^{-1}(s,\mathbf{w})\right)\right)}{\epsilon^2}\\ &\leq \frac{32M_2^2}{\epsilon^2 n}\,.\end{aligned} \tag{E.2}$$

On the other hand, we have

$$\begin{aligned}&|\bar{g}_n(s,\mathbf{w})| \leq M_1\\ &\Rightarrow \left|\bar{g}_n^{-1}(s,\mathbf{w})\right| \geq \frac{1}{M_1}\\ &\Rightarrow \left|\bar{g}_n^{-1}(s,\mathbf{w}) - \hat{g}^{-1}(s) + \hat{g}^{-1}(s)\right| \geq \frac{1}{M_1}\\ &\Rightarrow \left|\hat{g}^{-1}(s)\right| \geq \frac{1}{M_1} - \left|\bar{g}_n^{-1}(s,\mathbf{w}) - \hat{g}^{-1}(s)\right|\,.\end{aligned} \tag{E.3}$$

Then given $\epsilon > 0$, $\forall n \in \mathbb{N}_+, \forall s \in S$, the following holds:

$$\begin{aligned}&\mathbb{P}\left(|\hat{g}(s) - \bar{g}_n(s,\mathbf{w})| \geq \epsilon\right)\\ &= \mathbb{P}\left(\left|\bar{g}_n(s,\mathbf{w})\hat{g}(s)\left(\bar{g}_n^{-1}(s,\mathbf{w}) - \hat{g}^{-1}(s)\right)\right| \geq \epsilon\right)\\ &\leq \mathbb{P}\left(|\bar{g}_n(s,\mathbf{w})|\,|\hat{g}(s)|\left|\bar{g}_n^{-1}(s,\mathbf{w}) - \hat{g}^{-1}(s)\right| \geq \epsilon\right)\\ &\leq \mathbb{P}\left(M_1\left|\bar{g}_n^{-1}(s,\mathbf{w}) - \hat{g}^{-1}(s)\right| \geq \epsilon|\hat{g}^{-1}(s)|\right)\\ &\overset{(E.3)}{\leq} \mathbb{P}\left(M_1\left|\bar{g}_n^{-1}(s,\mathbf{w}) - \hat{g}^{-1}(s)\right| \geq \right.\\ &\qquad\left.\frac{\epsilon}{M_1} - \epsilon\left|\bar{g}_n^{-1}(s,\mathbf{w}) - \hat{g}^{-1}(s)\right|\right)\\ &= \mathbb{P}\left(\left|\bar{g}_n^{-1}(s,\mathbf{w}) - \hat{g}^{-1}(s)\right| \geq \frac{\epsilon}{M_1(M_1+\epsilon)}\right)\\ &\overset{(E.2)}{\leq} \frac{32M_2^2M_1^2(M_1+\epsilon)^2}{\epsilon^2 n}\,.\end{aligned}$$

By taking $n \to +\infty$ on both sides, we prove that $\bar{g}_n(s,\mathbf{w})$ converges point-wise to $\hat{g}(s)$ on S.

We now show that $\bar{g}_n(s,\mathbf{w})$ is also stochastic equicontinuous on S. For the definition of stochastic equicontinuity, please refer to [38]. We already assumed that $\bar{g}_n(s,\mathbf{w}) \leq M_1$, $\forall \mathbf{w} \in \Omega^\infty, s \in S$. Then $\forall \mathbf{w} \in \Omega^\infty, \forall s_1, s_2 \in S$, we have

$$\begin{aligned}&|\bar{g}_n(s_1,\mathbf{w}) - \bar{g}_n(s_2,\mathbf{w})|\\ &\leq \left|\bar{g}_n(s_1,\mathbf{w})||\bar{g}_n(s_2,\mathbf{w})||\bar{g}_n^{-1}(s_1,\mathbf{w}) - \bar{g}_n^{-1}(s_2,\mathbf{w})\right|\\ &\leq M_1^2\left|\sum_{i=1}^n\left(g_i^{-1}(s_1,w_i) - g_i^{-1}(s_2,w_1)\right)\right|\\ &\leq M_1^2\sum_{i=1}^n\left|g_i^{-1}(s_1,w_i) - g_i^{-1}(s_2,w_i)\right| \leq nM_1^2L|s_1 - s_2|\,,\end{aligned}$$

where the last inequality is from our third assumption and also the fact that $g_i^{-1}(s,w) = g_1^{-1}(s,w)$ (identically distributed as random functions). By [38, Corollary 2.2], the inequality above is sufficient to establish stochatic equicontinuity of $\bar{g}_n(s,\mathbf{w})$ on S, and combining point-wise convergence and the fourth assumption that $\hat{g}(s)$ is uniform continuous, we get the uniform convergence of $\bar{g}_n(s,\mathbf{w})$ to $\hat{g}(s)$ on S, which gives (E.1). □

References

[1] H. Min and E. Mallada, "Dynamics concentration of large-scale tightly-connected networks," in *IEEE 58th Conf. on Decision and Control*, 2019, pp. 758–763.

[2] P. C. Bressloff and S. Coombes, "Travelling waves in chains of pulse-coupled integrate-and-fire oscillators with distributed delays," *Physica D: Nonlinear Phenomena*, vol. 130, no. 3-4, pp. 232–254, 1999.

[3] I. Z. Kiss, Y. Zhai, and J. L. Hudson, "Emerging coherence in a population of chemical oscillators," *Science*, vol. 296, no. 5573, pp. 1676–1678, 2002.

[4] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[5] R. E. Mirollo and S. H. Strogatz, "Synchronization of pulse-coupled biological oscillators," *SIAM Journal on Applied Mathematics*, vol. 50, no. 6, pp. 1645–1662, 1990.

[6] Y. Jiang, R. Pates, and E. Mallada, "Performance tradeoffs of dynamically controlled grid-connected inverters in low inertia power systems," in *56th IEEE Conf. on Decision and Control*, 12 2017, pp. 5098–5105.

[7] F. Paganini and E. Mallada, "Global analysis of synchronization performance for power systems: Bridging the theory-practice gap," *IEEE Trans. Automat. Contr.*, vol. 65, no. 7, pp. 3007–3022, 2020.

[8] E. Mallada, X. Meng, M. Hack, L. Zhang, and A. Tang, "Skewless network clock synchronization without discontinuity: Convergence and performance," *IEEE/ACM Transactions on Networking (TON)*, vol. 23, no. 5, pp. 1619–1633, 10 2015.

[9] E. Mallada, "Distributed synchronization in engineering networks: The Internet and electric power girds," Ph.D. dissertation, Electrical and Computer Engineering, Cornell University, 01 2014.

[10] R. Sepulchre, D. Paley, and N. Leonard, "Stabilization of planar collective motion with limited communication," *IEEE Trans. Automat. Contr.*, vol. 53, no. 3, pp. 706–719, 2008.

[11] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[12] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Automat. Contr.*, vol. 48, no. 6, pp. 988–1001, 2003.
[13] B. Bamieh, M. R. Jovanovic, P. Mitra, and S. Patterson, "Coherence in large-scale networks: Dimension-dependent limitations of local feedback," *IEEE Trans. Automat. Contr.*, vol. 57, no. 9, pp. 2235–2249, 2012.
[14] E. Tegling, B. Bamieh, and H. Sandberg, "Localized high-order consensus destabilizes large-scale networks," in *2019 American Control Conference (ACC)*, July 2019, pp. 760–765.
[15] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1520–1533, 2004.
[16] Y. Ghaedsharaf, M. Siami, C. Somarakis, and N. Motee, "Centrality in time-delay consensus networks with structured uncertainties," *arXiv preprint arXiv:1902.08514*, 2019.
[17] S. Nair and N. Leonard, "Stable synchronization of mechanical system networks," *SIAM Journal on Control and Optimization*, vol. 47, no. 2, pp. 661–683, 2008.
[18] H. Kim, H. Shim, and J. Seo, "Output consensus of heterogeneous uncertain linear multi-agent systems," *IEEE Trans. Automat. Contr.*, vol. 56, no. 1, pp. 200–206, 2011.
[19] P. Wieland, R. Sepulchre, and F. Allgöwer, "An internal model principle is necessary and sufficient for linear output synchronization," *Automatica*, vol. 47, no. 5, pp. 1068–1074, 2011.
[20] H. G. Oral, E. Mallada, and D. F. Gayme, "Performance of first and second order linear networked systems over digraphs," in *IEEE 56th Annu. Conf. on Decision and Control*, Dec 2017, pp. 1688–1694.
[21] B. Bamieh and D. F. Gayme, "The price of synchrony: Resistive losses due to phase synchronization in power networks," in *2013 American Control Conference*, 2013, pp. 5815–5820.
[22] M. Andreasson, E. Tegling, H. Sandberg, and K. H. Johansson, "Coherence in synchronizing power networks with distributed integral control," in *IEEE 56th Annu. Conf. on Decision and Control*, Dec 2017, pp. 6327–6333.
[23] G. E. Dullerud and F. Paganini, *A course in robust control theory: a convex approach*. Springer Science & Business Media, 2013, vol. 36.
[24] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, NY, USA: Cambridge University Press, 2012.
[25] W. Rudin *et al.*, *Principles of mathematical analysis*. McGraw-hill New York, 1964, vol. 3.
[26] J. H. Chow, *Power system coherency and model reduction*. New York, NY, USA: Springer, 2013.
[27] R. Olfati-Saber, "Algebraic connectivity ratio of ramanujan graphs," in *2007 American Control Conference*. IEEE, 2007, pp. 4619–4624.
[28] H. K. Khalil and J. W. Grizzle, *Nonlinear systems*. Prentice hall Upper Saddle River, NJ, 2002, vol. 3.
[29] U. of Edinburgh. Power systems test case archive. Mar. 2003. [Online]. Available: https://www.maths.ed.ac.uk/optenergy/NetworkData/icelandDyn/
[30] A. J. Germond and R. Podmore, "Dynamic aggregation of generating unit models," *IEEE Trans. Power App. Syst.*, vol. PAS-97, no. 4, pp. 1060–1069, July 1978.
[31] S. S. Guggilam, C. Zhao, E. Dall'Anese, Y. C. Chen, and S. V. Dhople, "Optimizing DER participation in inertial and primary-frequency response," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5194–5205, Sep. 2018.
[32] D. Apostolopoulou, P. W. Sauer, and A. D. Domínguez-García, "Balancing authority area model and its application to the design of adaptive AGC systems," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3756–3764, Sep. 2016.
[33] H. Min, F. Paganini, and E. Mallada, "Accurate reduced-order models for heterogeneous coherent generators," *IEEE Contr. Syst. Lett.*, vol. 5, no. 5, pp. 1741–1746, 2021.
[34] F. R. Bach and M. I. Jordan, "Learning spectral clustering," in *Advances in Neural Information Processing Systems*, 2004, pp. 305–312.
[35] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Advances in Neural Information Processing Systems*, 2001, p. 585–591.
[36] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak Mathematical Journal*, vol. 23, no. 2, pp. 298–305, 1973.
[37] E. Freitag, *Complex analysis*. Krishna Prakashan Media, 1977.
[38] W. K. Newey, "Uniform convergence in probability and stochastic equicontinuity," *Econometrica: Journal of the Econometric Society*, pp. 1161–1167, 1991.