Cluster Synchronization of Kuramoto Oscillators and Brain Functional Connectivity

Rui Kato and Hideaki Ishii, Fellow, IEEE

Abstract—The recent progress of functional magnetic resonance imaging techniques has unveiled that human brains exhibit clustered correlation patterns of their spontaneous activities. It is important to understand the mechanism of cluster synchronization phenomena since it may reflect the underlying brain functions and brain diseases. In this paper, we investigate cluster synchronization conditions for networks of Kuramoto oscillators. The key analytical tool that we use is the method of averaging, and we provide a unified framework of stability analysis for cluster synchronization. The main results show that cluster synchronization is achieved if (i) the inter-cluster coupling strengths are sufficiently weak and/or (ii) the natural frequencies are largely different among clusters. Moreover, we apply our theoretical findings to empirical brain networks. Discussions on how to understand brain functional connectivity and further directions to investigate neuroscientific questions are provided.

Index Terms—Brain networks, cluster synchronization, Kuramoto model, functional connectivity, methods of averaging.

I. INTRODUCTION

Synchronization in complex networks has attracted interests over the past several decades (see the surveys [1], [2]). For example, synchronous generation in power grids [3] and cooperative control of networked robots [4] are representative applications in technological fields. On the other hand, applications in biological fields include physiological rhythms [5] and neuronal oscillations [6]. The phenomenon where all states are consistent among the whole network is called full synchronization or complete synchronization. In contrast, it is known that human brains generate neural activities with wide-range frequency bands, and their characteristic statistical patterns can be observed [7]. One of the important properties found there is that brains exhibit cluster synchronization, that is, some synchronized clusters coexist in the network.

The recent progress in brain imaging techniques allows us to noninvasively measure the anatomical network structure in a whole brain [8], [9]. There, the connections between selected cortical regions by nervous fibers are extracted by, e.g., diffusion tensor imaging. On the other hand, it is known that brains exhibit spontaneous activities even at resting state (without external stimuli). These activities can be quantified by functional magnetic resonance imaging (fMRI), and a characteristic correlation pattern of fMRI signals has been found [10]. This so-called resting-state functional connectivity is considered to serve an important role in brain functions [11]. Recently, in the neuroscience community, functional connectivity together with anatomical structures has been studied through graph-theoretic tools [12].

Moreover, functional connectivity has also been examined for patients subject to brain diseases, where statistical differences between fMRI signals in healthy brains and those in pathological brains were found. For example, the small-world property observed in healthy brains can be disturbed by Alzheimer’s diseases and schizophrenia [13], [14]. In epilepsy, abnormal functional connectivities are related to the significant increase in synchronous neural activities [15]. Hence, theoretical analysis of cluster synchronization may be helpful for understanding the mechanism of healthy and pathological brains [16]. On the other hand, it is increasingly important to change the pathological state of the brain into the healthy one by external interventions [17]. Today, electrical stimulation techniques are practical methods to manipulate neural activities [18]. Optical stimulation methods are also gaining attentions as a new tool to control brain states [19]. Furthermore, there is a need for developing theoretical foundations to understand and control brain dynamics [20].

This paper considers cluster synchronization of Kuramoto oscillators with heterogeneous natural frequencies. Here, cluster synchronization means that the oscillators may be divided into some groups, and their phases are eventually synchronized within the groups at possibly different frequencies. The Kuramoto model is known as a representative mathematical model of coupled oscillators to study collective synchronization phenomena [21]. Stability analysis for full synchronization of Kuramoto oscillators with general network topologies was investigated in the seminal paper [22]. After that work, many related results have been established with several settings (see [2] and the references therein). However, cluster synchronization of Kuramoto oscillators has not been deeply studied except for [23], [24]. A slightly different definition of cluster synchronization was employed in [25], where phase cohesiveness within clusters was explored. On the other hand, some authors considered cluster synchronization problems for general nonlinear systems. To study cluster synchronization, graph symmetry was employed in [26]–[31], and the notion known as external equitable partitions was addressed in [32], [33]. The latter works are more general than the former ones in the sense that any symmetric network allows an external equitable partition. In [34], [35] sufficient conditions for exponential convergence to a cluster synchronization manifold were derived based on contraction analysis. The effects of delays and antagonistic interactions in cluster synchronization were explored in [36]. Moreover, control methodologies for
cluster synchronization were studied in [37]–[39].

Differently from the aforementioned works, our analysis is based on regarding the intra- and inter-cluster dynamics as slow and fast subsystems, respectively. Thus, we start with formulating the intra- and inter-cluster dynamics in a singular perturbation form, where the smallest difference in natural frequencies between clusters is interpreted as a perturbation parameter. Since we are concerned with partial stability of the intra-cluster dynamics, we then linearize the slow subsystem while the solution of the fast subsystem is considered as a function of time. To analyze stability of the obtained linear time-varying system, we utilize the method of averaging. However, existing results such as one in [40] cannot be applied to our system (see Remark 2). To overcome this difficulty, we employ nonmonotonic Lyapunov functions to extend existing methods. We show that our framework is useful in the study of cluster synchronization and that the time-scale property may be a key factor in the statistical patterns of brain activities.

The contributions of this paper are summarized as follows. As a technical contribution, we provide a stability criterion for a class of linear time-varying systems through the method of averaging. Our result generalizes some existing works and plays a crucial role in the subsequent analysis. The main contribution of this paper is to provide a unified framework for stability analysis of cluster synchronization in networks of Kuramoto oscillators. Specifically, we show that cluster synchronization is achieved if (i) the inter-cluster coupling strengths are sufficiently weak and/or (ii) the natural frequencies are largely different among clusters. Compared with [24], where these two properties are analyzed independently, we provide a tradeoff relation between the inter-cluster connectivity and heterogeneity. Intuitively, to achieve synchronization with clustering, large inter-cluster heterogeneity in the frequencies is required when there are strong couplings between clusters. Also, weak inter-cluster connectivity is required when the oscillators have similar frequencies among clusters. The above results are derived under the assumption that the network is partitioned into clusters such that an external equitable partition is fulfilled. For the case without this assumption, we explore practical stability and show that phase cohesiveness within clusters can be made arbitrarily small by strengthening the intra-cluster couplings.

Furthermore, we apply our theoretical findings to a brain network constructed using empirical data. By numerical simulation, we observe that our results on cluster synchronization are valid for the brain network. Moreover, it is shown that cluster synchronization of Kuramoto oscillators is strongly related to functional connectivity, which is identified by a statistical correlation pattern of region-dependent fMRI signals in human brains. Compared with the graph-theoretic analysis in the neuroscience community, our work is based on dynamic properties of brains and provides a theoretical foundation to understand functional connectivity.

The organization of this paper is described as follows. We start with transforming the Kuramoto model into a singular perturbation form in Section II. Based on the method of averaging, we explore a stability criterion for linear time-varying systems in Section III. Then, we derive stability conditions for cluster synchronization of Kuramoto oscillators in Section IV. The case without external equitable partitions is investigated in Section V. We consider applications to brain networks with empirical data in Section VI, and finally, Section VII provides concluding remarks of this paper. A preliminary version of this paper appeared as the conference paper [41]; in the current work, we provide further extensions and the proofs of all results along with more extensive simulations.

Notations: Let $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{Z}$, and $\mathbb{Z}_+$ respectively denote the set of reals, nonnegative reals, integers, and nonnegative integers. We denote by $S^1 \simeq \mathbb{R}/(2\pi \mathbb{Z})$ the unit circle, and by $\mathbb{T}^n := S^1 \times \cdots \times S^1$ the $n$-torus. The $n$-vector whose entries are all $1$ is denoted by $\mathbf{1}_n$. For a matrix $A \in \mathbb{R}^{n \times m}$, $\text{im} A$ and $\ker A$ stand for the image and the kernel of $A$, respectively. The imaginary unit is denoted by $j$. If $S$ is a subspace, then $S^\perp$ is the orthogonal complement of $S$ with respect to the usual inner product. A function $\alpha : [0, \infty) \to [0, \infty)$ is of class $K$ if it is continuous, strictly increasing, and $\alpha(0) = 0$. A function $\beta : [0, \infty) \to [0, \infty)$ is of class $L$ if it is continuous, nonincreasing, and $\beta(a) \to 0$ as $a \to \infty$. We define the distance from a point $x \in \mathbb{R}^n$ to a subset $S \subset \mathbb{R}^n$ by $\text{dist}(x, S) := \inf_{y \in S} \|x - y\|$.}

II. Problem Setting and Preliminaries

In this section we describe the formal definition of cluster synchronization and basic assumptions. Some preliminaries that will be useful in the subsequent sections are also provided.

A. Problem Setting

Let $G = (\mathcal{V}, \mathcal{E})$ be an undirected weighted graph, where $\mathcal{V} = \{1, \ldots, N\}$ is the set of nodes and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of edges. Throughout this paper, it is assumed that $G$ is connected since cluster synchronization can be trivially achieved if the network is separated. The weighted adjacency matrix of $G$ is denoted by $A = [a_{ij}] \in \mathbb{R}^{N \times N}$, where $a_{ij} > 0$ if $(i, j) \in \mathcal{E}$ and $a_{ij} = 0$ if $(i, j) \notin \mathcal{E}$. After selecting an enumeration and an orientation for all (undirected) edges, we define the oriented incidence matrix of $G$ by $B = [b_{ie}] \in \mathbb{R}^{N \times |\mathcal{E}|}$, where

$$b_{ie} := \begin{cases} 1 & \text{if node } i \text{ is the sink node of edge } e, \\ -1 & \text{if node } i \text{ is the source node of edge } e, \\ 0 & \text{otherwise.} \end{cases}$$

We also define the weight matrix $W := \text{diag}([a_{ij}]_{(i,j) \in \mathcal{E}})$ of $G$ based on the enumeration of edges introduced above.

Throughout this paper, each node represents an oscillator and each edge represents an interconnection of two oscillators. We consider the Kuramoto model governed by

$$\dot{\theta}_i = \omega_i + \sum_{j=1}^{N} a_{ij} \sin(\theta_j - \theta_i), \quad i \in \{1, \ldots, N\},$$

where $\theta_i : \mathbb{R}_+ \to S^1$ is the phase of the $i$th oscillator and $\omega_i \in \mathbb{R}$ is its natural frequency. The equations in (1) can be written in vector form:

$$\dot{\theta} = \omega - BW \sin(B^\top \theta),$$

where $\theta := [\theta_1 \cdots \theta_N]^\top$ and $\omega := [\omega_1 \cdots \omega_N]^\top$. 

If a solution \( \theta(t) \) of (2) satisfies \( \text{dist}(\theta(t), \text{span}(\mathbb{1}_N)) \to 0 \) as \( t \to \infty \), then the solution is said to have achieved full synchronization. Note that this can be achieved only if \( \omega \in \text{span}(\mathbb{1}_N) \), that is, all the natural frequencies are identical. The subject of this paper is different from the situations of full synchronization. We are interested in cases where the natural frequencies are not identical and the synchronization is realized in some groups of oscillators. Next, we introduce the notion of this phenomenon, called cluster synchronization.

Given an integer \( r \geq 2 \), we denote by \( \Pi = \{C_1, \ldots, C_r\} \) a nontrivial partition of \( \mathcal{V} \), i.e., the following conditions hold: (i) \( C_p \neq \emptyset \) for all \( p \in \{1, \ldots, r\} \), (ii) \( C_p \cap C_q = \emptyset \) if \( p \neq q \), and (iii) \( \bigcup_{p=1}^r C_p = \mathcal{V} \). We call \( C_1, \ldots, C_r \) clusters of the graph \( \mathcal{G} \). It is assumed without loss of generality that every cluster includes at least two nodes. For \( p \in \{1, \ldots, r\} \), let \( \mathcal{G}_p = (\{i,j\} \in \mathcal{E} : i,j \in C_p) \). In this paper, we assume that \( \mathcal{G}_p \) is connected for every \( p \in \{1, \ldots, r\} \). Also, let \( \mathcal{G}_{\text{inter}} = (\mathcal{V}, \mathcal{E}_{\text{inter}}) \) be the subgraph of \( \mathcal{G} \) associated with cluster \( C_p \), where \( \mathcal{E}_p := \{(i,j) \in \mathcal{E} : i,j \in C_p\} \). In this paper, we assume that \( \mathcal{G}_p \) is connected for every \( p \in \{1, \ldots, r\} \).

The cluster synchronization manifold associated with partition \( \Pi = \{C_1, \ldots, C_r\} \) is defined by

\[
\mathcal{S}_\Pi := \{ \theta \in \mathbb{T}^N : \theta_i = \theta_j \text{ for all } i,j \in C_p, p \in \{1, \ldots, r\} \}.
\]

If a solution \( \theta(t) \) of (2) satisfies \( \text{dist}(\theta(t), \mathcal{S}_\Pi) \to 0 \) as \( t \to \infty \), then the solution is said to have achieved cluster synchronization. That is, the phase difference of any two nodes in the same cluster converges to zero, but that of nodes in different clusters does not necessarily converge. Stability of the cluster synchronization manifold is formally defined below.

**Definition 1:** Given a nontrivial partition \( \Pi = \{C_1, \ldots, C_r\} \) of \( \mathcal{V} \), the cluster synchronization manifold \( \mathcal{S}_\Pi \) of (2) is said to be (locally) exponentially stable if there exist \( k \geq 1 \), \( \lambda > 0 \), and \( \delta > 0 \) such that \( \text{dist}(\theta(0), \mathcal{S}_\Pi) < \delta \) implies \( \text{dist}(\theta(t), \mathcal{S}_\Pi) \leq k e^{-\lambda t} \text{dist}(\theta(0), \mathcal{S}_\Pi) \) for all \( t \in \mathbb{R}_+ \).

Clearly, this notion requires that the cluster synchronization manifold \( \mathcal{S}_\Pi \) is positive invariant for (2), that is, \( \theta(0) \in \mathcal{S}_\Pi \) implies \( \theta(t) \in \mathcal{S}_\Pi \) for all \( t \in \mathbb{R}_+ \). It is known that positive invariance of \( \mathcal{S}_\Pi \) is ensured under the following two assumptions [32], [42]:

**Assumption 1:** For every \( p \in \{1, \ldots, r\} \), it holds that \( \omega_i = \omega_j \) for all \( i,j \in C_p \).

**Assumption 2:** For every \( p, q \in \{1, \ldots, r\} \) with \( p \neq q \), it holds that \( \sum_{k \in C_p} a_{ik} = \sum_{k \in C_q} a_{jk} \) for all \( i,j \in C_p \).

Assumption 1 restricts the natural frequencies of oscillators in each cluster to be identical. This is a natural requirement since cluster synchronization implies that the oscillators rotate coherently within clusters. Assumption 2 states that any two nodes in the same cluster have equal weight sums with respect to other clusters. Notice that this assumption depends only on inter-cluster edge weights. It can thus be considered as a generalization of the trivial case where there is no edge between clusters; in such a case, it is clear that under Assumption 1, cluster synchronization can be realized. However, the condition seems somewhat restrictive in that the equality in Assumption 2 must hold strictly. We will relax this aspect in Section V. The following proposition states that these assumptions are necessary for analyzing stability of the cluster synchronization manifold [42].

**Proposition 1:** The cluster synchronization manifold is positive invariant for (2) if Assumptions 1 and 2 hold. These are also necessary when, for every \( p,q \in \{1, \ldots, r\} \) with \( p \neq q \), \( \max_{i \in C_p} \theta_i(t) \neq \max_{j \in C_q} \theta_j(t) \) holds for all \( t \in \mathbb{R}_+ \).

**Remark 1:** A graph partition satisfying the condition in Assumption 2 is called an external equitable partition, which was introduced in [43]. Such a graph partition is known as a sufficient condition for the positive invariance of cluster synchronization manifolds for many classes of systems [32]. In [42], the authors showed that an external equitable partition is necessary and sufficient for the invariance in the Kuramoto model when the oscillators’ frequencies are largely different between clusters as stated in Proposition 1. We also remark that if the network size is large, there are a huge number of partitions that belong to the class of external equitable partitions. The trivial case is where every single node is grouped into a single cluster. For algorithms to find the external equitable partitions with the minimal cardinality, we refer to [44], [45].

**B. Preliminaries**

To analyze synchronization, it is useful to consider relative phases instead of phases themselves. Thus, we transform the Kuramoto model (2) into the dynamics of phase differences. We then show that those dynamics can be written by differential equations in a singular perturbation form.

Recall that we have decomposed the graph \( \mathcal{G} \) as \( \mathcal{G} = \bigcup_{p=1}^r \mathcal{G}_p \cup \mathcal{G}_{\text{inter}} \). With appropriate ordering of the edge indices, we can write the incidence matrix \( \mathbf{B} \) and the weight matrix \( \mathbf{W} \) in the following forms:

\[
\mathbf{B} = \begin{bmatrix} \mathbf{B}_{\text{intra}} & \mathbf{B}_{\text{inter}} \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} \mathbf{W}_{\text{intra}} & 0 \\ 0 & \mathbf{W}_{\text{inter}} \end{bmatrix},
\]

where \( \mathbf{B}_{\text{intra}} \) and \( \mathbf{B}_{\text{inter}} \) are the incidence matrices of \( \bigcup_{p=1}^r \mathcal{G}_p \) and \( \mathcal{G}_{\text{inter}} \), respectively, and \( \mathbf{W}_{\text{intra}} \) and \( \mathbf{W}_{\text{inter}} \) are the weight matrices of \( \bigcup_{p=1}^r \mathcal{G}_p \) and \( \mathcal{G}_{\text{inter}} \), respectively.

As with [24], we consider a spanning tree \( \mathcal{T} = (\mathcal{V}, \mathcal{E}) \) of \( \mathcal{G} \), where \( \mathcal{E} \subset \mathcal{E} \) is the set of selected edges. Let \( \tilde{\mathcal{B}} \in \mathbb{R}^{(|\mathcal{E}| \times (N-1))} \) be the incidence matrix of \( \mathcal{T} \). Without loss of generality, we can write \( \tilde{\mathbf{B}} = \begin{bmatrix} \mathbf{B}_{\text{intra}} & \mathbf{B}_{\text{inter}} \end{bmatrix} \), where \( \mathbf{B}_{\text{intra}} \) and \( \mathbf{B}_{\text{inter}} \) are associated with the edges in \( \bigcup_{p=1}^r \mathcal{E}_p \) and \( \mathcal{E}_{\text{inter}} \), respectively. We denote by \( n \) and \( m \) the dimensions of the range spaces of \( \mathbf{B}_{\text{intra}} \) and \( \mathbf{B}_{\text{inter}} \), respectively. Note that \( n + m = N - 1 \).

Now, we would like to calculate a matrix \( \mathcal{R} \in \mathbb{R}^{(|\mathcal{E}| \times (N-1))} \) such that \( \mathcal{B}^T = \mathcal{R} \tilde{\mathcal{B}}^T \). Under an appropriate partition of \( \mathcal{R} \), we have the relation

\[
\begin{bmatrix} \mathbf{B}_{\text{intra}}^T \\ \mathbf{B}_{\text{inter}}^T \end{bmatrix} = \mathcal{R} \begin{bmatrix} \tilde{\mathbf{B}}_{\text{intra}}^T \\ \tilde{\mathbf{B}}_{\text{inter}}^T \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{R}_3 & \mathbf{R}_4 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{B}}_{\text{intra}}^T \\ \tilde{\mathbf{B}}_{\text{inter}}^T \end{bmatrix}.
\]

Since \( \tilde{\mathcal{B}} \) has full column rank, we have \( \mathcal{R} = \mathbf{B}^T (\tilde{\mathbf{B}}^T) = \mathbf{B}^T (\tilde{\mathbf{B}}^T)^\dagger \). The problem is thus to calculate the Moore-Penrose inverse \( \mathbf{B}^\dagger \) of the column-wise partitioned matrix \( \mathcal{B} \). To this end, we borrow a useful lemma from [46].
Lemma 1: Let \( M_1 \in \mathbb{R}^{m \times n_1} \) and \( M_2 \in \mathbb{R}^{m \times n_2} \). Define the orthogonal projectors \( P_1, P_2 \in \mathbb{R}^{m \times m} \) by \( P_i := I_m - M_i M_i^\dagger \) for \( i \in \{1, 2\} \). Then, we have

\[
[M_1 \quad M_2]^\dagger = \begin{bmatrix} (P_2M_1)^\dagger \\ (P_1M_2)^\dagger \end{bmatrix}
\]

if and only if \( \text{im } M_1 \cap \text{im } M_2 = \{0\} \).

We define the two matrices \( P_{\text{intra}} := I_N - B_{\text{intra}}B_{\text{intra}}^\dagger \) and \( P_{\text{inter}} := I_N - B_{\text{inter}}B_{\text{inter}}^\dagger \). Note that these are orthogonal projectors onto \( \text{ker } B_{\text{intra}} \) and \( \text{ker } B_{\text{inter}} \) respectively. Since \( \text{im } B_{\text{intra}} \cap \text{im } B_{\text{inter}} = \{0\} \), Lemma 1 provides

\[
\tilde{B} = \begin{bmatrix} \tilde{B}_{\text{intra}} \\ \tilde{B}_{\text{inter}} \end{bmatrix}^\dagger = \begin{bmatrix} (P_{\text{inter}}B_{\text{intra}})^\dagger \\ (P_{\text{intra}}B_{\text{inter}})^\dagger \end{bmatrix}.
\]

It thus follows that

\[
R = \begin{bmatrix} B_{\text{intra}}^\dagger \\ B_{\text{inter}}^\dagger \end{bmatrix} \begin{bmatrix} (\tilde{B}_{\text{intra}}P_{\text{inter}})^\dagger \\ (\tilde{B}_{\text{inter}}P_{\text{intra}})^\dagger \end{bmatrix} = B_{\text{intra}}^\dagger B_{\text{inter}}^\dagger (\tilde{B}_{\text{intra}}P_{\text{intra}})^\dagger (\tilde{B}_{\text{inter}}P_{\text{inter}})^\dagger.
\]

We notice that \( \text{ker } B_{\text{intra}}^\dagger = \text{ker } B_{\text{inter}}^\dagger \), which follows from the fact that \( B_{\text{intra}}^\dagger B_{\text{inter}} = 0 \) and \( B_{\text{inter}}^\dagger B_{\text{intra}} = 0 \). Hence, we have \( B_{\text{intra}}(\tilde{B}_{\text{intra}}P_{\text{intra}})^\dagger = 0 \) since \( \text{im } P_{\text{intra}} = \text{ker } B_{\text{intra}} \). The four matrices in (3) are finally given by

\[
R_1 = B_{\text{intra}}^\dagger (\tilde{B}_{\text{intra}}P_{\text{inter}})^\dagger, \quad R_2 = 0,
\]

\[
R_3 = B_{\text{inter}}^\dagger (\tilde{B}_{\text{intra}}P_{\text{inter}})^\dagger, \quad R_4 = B_{\text{inter}}^\dagger (\tilde{B}_{\text{inter}}P_{\text{intra}})^\dagger.
\]

Define the vectors of intra- and inter-cluster phase differences as \( x := B_{\text{intra}}^\dagger \theta \) and \( z := B_{\text{inter}}^\dagger \theta \), respectively. For the Kuramoto model (2), we choose a perturbation parameter \( \varepsilon \) as the reciprocal of the smallest natural frequency difference between clusters. Without loss of generality, we assume \( \varepsilon \) to be positive, namely,

\[
\varepsilon := \frac{1}{\min_{(i,j) \in E_{\text{inter}}} |\omega_i - \omega_j|}.
\]

Then, under Assumption 1, we obtain the state equations in the following singular perturbation form:

\[
\dot{x} = f(x, z), \quad \varepsilon \dot{z} = \eta + \varepsilon g(x, z),
\]

where \( \eta := \varepsilon B_{\text{inter}}\omega \) is a constant vector representing the ratio of natural frequency differences with respect to \( \varepsilon \), and

\[
f(x, z) := -\tilde{B}_{\text{intra}}^\dagger B_{\text{intra}}W_{\text{intra}} \sin(R_1x - \tilde{B}_{\text{inter}}^\dagger B_{\text{inter}}W_{\text{inter}} \sin(R_3x + R_4z), \quad g(x, z) := -\tilde{B}_{\text{inter}}^\dagger B_{\text{inter}}W_{\text{inter}} \sin(R_1x - \tilde{B}_{\text{inter}}^\dagger B_{\text{intra}}W_{\text{intra}} \sin(R_3x + R_4z).
\]

The equations (5) and (6), respectively, describe the intra- and inter-cluster dynamics.

We now introduce the concept of partial stability defined in Chap. 4 of [47].

Definition 2: Consider the dynamical system described by

\[
\dot{x}_1 = f_1(x_1, x_2), \quad \dot{x}_2 = f_2(x_1, x_2).
\]

Suppose that \( f_1(0, x_2) = 0 \) for all \( x_2 \). Then, the zero solution \( x_1(t) \equiv 0 \) is said to be partially exponentially stable uniformly in \( x_2 \) if there exist \( k \geq 1, \lambda > 0, \) and \( \delta > 0 \) such that

\[
\|x_1(t)\| < \lambda \text{ implies } \|x_1(t)\| \leq \lambda e^{-\delta t}\|x_1(0)\| \quad \text{for all } t \geq 0.
\]

Recall that under Assumptions 1 and 2, the cluster synchronization manifold \( S_{\text{H}} \) is positive invariant for (2). This implies that \( f(0, z) = 0 \) for all \( z \in \mathbb{R}^m \). For the Kuramoto model (2) and its associated state equations (5) and (6), exponential stability of \( S_{\text{H}} \) in the sense of Definition 1 is equivalent to partial exponential stability of \( x(t) \equiv 0 \) in the sense of Definition 2. Therefore, in what follows, we will explore partial stability of the dynamical system described by (5) and (6) to derive conditions for cluster synchronization.

III. Stability Theory via Averaging

In this section we present a sufficient condition for stability of a class of linear time-varying systems. We use an averaging principle to show that there is a nonmonotonic Lyapunov function for the given linear time-varying system characterized by a small perturbation parameter. The result in this section will play a crucial role for deriving stability conditions for cluster synchronization in Section IV.

First, we give a general result involving nonmonotonic Lyapunov functions (see [48] for details). Briefly, different from the requirement of \( V \) to be a Lyapunov function, \( V(x(t)) \) must decrease only at certain sampled instants along solutions \( x(t) \) and time intervals where it increases may be present.

Proposition 2: Consider the nonlinear system described by

\[
\dot{x} = f(t, x), \quad (7)
\]

where \( f: \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n \) is a continuous function. Suppose that \( f(t, 0) = 0 \) for all \( t \in \mathbb{R}_+ \) and that the solutions of (7) for \( t \in \mathbb{R}_+ \) exist and are unique in a neighborhood of \( x = 0 \). Let \( V: \mathbb{R}^n \rightarrow \mathbb{R} \) be a continuous function such that the following properties hold:

1) There exist \( c_1, c_2 > 0 \) such that

\[
c_1\|x\|^2 \leq V(x) \leq c_2\|x\|^2
\]

for all \( x \in \mathbb{R}^n \).

2) For some time sequence \( \{t_k\}_{k \in \mathbb{Z}_+} \) with the properties that \( t_0 = 0, \ t_k < t_{k+1} \) for all \( k \in \mathbb{Z}_+ \), and \( t_k \rightarrow \infty \) as \( k \rightarrow \infty \), there exist \( c_3 > 0 \) and \( \delta > 0 \) such that for any solution \( x: \mathbb{R}_+ \rightarrow \mathbb{R}^n \) of (7) satisfying \( \|x(0)\| < \delta \), it holds

\[
\frac{V(x(t_{k+1})) - V(x(t_k))}{t_{k+1} - t_k} \leq -c_3\|x(t_k)\|^2
\]

for all \( k \in \mathbb{Z}_+ \).

Then, the zero solution \( x(t) \equiv 0 \) is exponentially stable.

Now, consider the system of linear differential equations

\[
\dot{x} = \varepsilon A(t, \varepsilon) x,
\]

where \( A(t, \varepsilon) \) is a time-varying matrix. If \( \varepsilon \) is small enough, then the solutions of the perturbed system converge uniformly to the solutions of the unperturbed system. This result is stated in the following proposition.

Proposition 3: Let \( A(t, \varepsilon) \) be a time-varying matrix such that \( A(t, \varepsilon) \rightarrow A(t) \) as \( \varepsilon \rightarrow 0 \). Suppose that \( A(t) \) satisfies the conditions of Proposition 2. Then, the solutions of the perturbed system converge uniformly to the solutions of the unperturbed system.
where \( A : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times n} \) is continuous and bounded in its arguments and \( \varepsilon > 0 \) is a small parameter. Assume that the time average of \( A(t, \varepsilon) \) exists for \( \varepsilon = 0 \). That is, the matrix

\[
A_{av} := \lim_{T \to \infty} \frac{1}{T} \int_T^{t+T} A(s, 0) \, ds
\]  

(9)
can be defined uniformly in \( t \). Also, it is assumed that there exist a class \( \mathcal{L} \) function \( \beta : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \) and a class \( \mathcal{K} \) function \( \alpha : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \) such that

\[
\left\| \frac{1}{T} \int_T^{t+T} [A(s, \varepsilon) - A_{av}] \, ds \right\| \leq \beta(T) + \alpha(\varepsilon)
\]  

(10)
for all \( t \in \mathbb{R}_+ \).

The following theorem can be established using Proposition 2.

Theorem 1: Suppose that the matrix \( A_{av} \) in (9) is a Hurwitz matrix and that the inequality (10) holds. Then, there exists \( \varepsilon^* > 0 \) such that if \( \varepsilon < \varepsilon^* \), then the zero solution \( x(t) \equiv 0 \) of (8) is exponentially stable.

Proof: See Appendix A. ■

Remark 2: Here, we give some technical comments on Theorem 1 in comparison with existing results. By the method of averaging, a Lyapunov function for the average system can be employed to prove stability of the original system (see Chap. 10 in [40]). To apply the method in [40], it is required that \( A(t, \varepsilon) \) in (8) is Lipschitz with respect to \( \varepsilon \) uniformly in \( t \), i.e., there exists \( L > 0 \) such that for every \( \varepsilon_1, \varepsilon_2 > 0 \) sufficiently small, we have \( \|A(t, \varepsilon_1) - A(t, \varepsilon_2)\| \leq L|\varepsilon_1 - \varepsilon_2| \) uniformly in \( t \). In the present case, however, the matrix \( A(t, \varepsilon) \) in (8) does not satisfy the Lipschitz condition. Thus, we cannot use the conventional method in [40] to prove stability of our system. To overcome this difficulty, we employed nonsmooth Lyapunov functions and applied Proposition 2 to prove stability. The use of nonsmooth Lyapunov functions was also considered for linear time-varying systems in [49], [50]. In those papers, the authors addressed the case where \( A(t, \varepsilon) \) is independent of \( \varepsilon \). Thus, Theorem 1 can be viewed as an extension of the results in [49], [50]. On the other hand, the authors of [51] considered the case where the uniform average in (9) does not exist, and the time-dependent average matrix was used in stability analysis. In contrast, we consider the case where the uniform average in (9) can be defined for \( \varepsilon = 0 \), but it may not necessarily be defined for \( \varepsilon > 0 \).

Remark 3: We note that in Proposition 2, the positive-definite function \( V \) need not be differentiable. That is, one has the advantage that a nonsmooth Lyapunov function candidate can be used in stability analysis. This property may be useful to study synchronization problems. For related results using nonsmooth Lyapunov functions, we refer to, e.g., [52].

IV. CLUSTER SYNCHRONIZATION CONDITIONS FOR KURAMOTO OSCILLATORS

In this section we provide our main results. Through the analytical method developed in the previous section, we derive sufficient conditions for exponential stability of a cluster synchronization manifold associated with the Kuramoto model. This section addresses two problems. First, we consider cluster synchronization with two or more clusters. Then, we restrict our attention to the case of two clusters and carry out further analysis. Numerical examples are also presented to demonstrate our results.

A. Multi-Cluster Case

Recall that the problem under consideration is to analyze partial stability of the nonlinear system described by (5) and (6). First, we linearize the intra-cluster dynamics (5) at \( x = 0 \), which corresponds to the cluster synchronization manifold, while \( z \) is interpreted as a function of time. To do so, let us introduce a new time variable \( \tau := t/\varepsilon \). In this time scale, the equations (5) and (6) are written as

\[
\frac{dx}{d\tau} = \varepsilon f(x, z),
\]

(11)
\[
\frac{dz}{d\tau} = \eta + \varepsilon g(x, z).
\]

(12)
For a fixed \( \varepsilon \), let \( \zeta(\cdot, \varepsilon) \) denote the solution of the initial value problem for (12) under \( x = 0 \):

\[
\frac{dz}{d\tau} = \eta + \varepsilon g(0, z), \quad z(0) = z_0,
\]

(13)
where we have omitted the dependence on \( z_0 \) because it does not affect the results. Note that when \( \varepsilon = 0 \), the solution to (13) is given by \( \zeta(\tau, 0) = z_0 + \eta \tau \). Linearization of \( f(x, z) \) around \( x = 0 \) and \( z = \zeta(\tau, \varepsilon) \) leads us to the Jacobian

\[
J(\tau, \varepsilon) := \frac{\partial f(x, z)}{\partial x} \bigg|_{(x, z) = (0, \zeta(\tau, \varepsilon))}
\]
\[
= -\hat{B}_{intra}^T B_{intra} W_{intra} R_1
\]
\[
- \hat{B}_{inter}^T B_{inter} W_{inter} \text{diag}[R_4 \cos(\zeta(\tau, \varepsilon))] R_3.
\]

(14)
The linearized system is thus described by

\[
\frac{dx}{d\tau} = \varepsilon J(\tau, \varepsilon)x.
\]

(15)
The following proposition motivates us to analyze stability based on the linearization principle.

Proposition 3: Consider the nonlinear system (5) and (6). The zero solution \( x(t) \equiv 0 \) is partially exponentially stable uniformly in \( z \) if and only if the linear time-varying system (15) is (uniformly) exponentially stable.

Proof: Interpreting \( z \) as a function of time, we can use the standard stability analysis techniques (e.g., Chap. 4 of [40]) to conclude the proof. Specifically, the sufficiency part follows from Lyapunov’s indirect method and the necessity part follows from the converse Lyapunov theorem. ■

We are ready to derive cluster synchronization conditions. Particularly, we analyze stability of (15) based on Theorem 1. From (4), we note that the smaller \( \varepsilon \) is, the larger the inter-cluster natural frequency differences are. The following result provides sufficient conditions for cluster synchronization. Its proof is given in the next subsection.

Theorem 2: Let \( \Sigma_k \) be the cluster synchronization manifold associated with the Kuramoto model (2), where \( \Pi = \{C_1, \ldots, C_r\} \) is a nontrivial partition of \( \mathcal{V} \) such that the
subgraphs $G_1, \ldots, G_c$ are all connected. Suppose that Assumptions 1 and 2 hold. Then, the following statements hold:

1) There exists $a^* > 0$ such that if
   \[ a_{ij} < a^*, \quad (i, j) \in E_{\text{inter}}, \]
   then $S_{\Pi}$ is exponentially stable.

2) There exists $\omega^* > 0$ such that if
   \[ |\omega_i - \omega_j| > \omega^*, \quad (i, j) \in E_{\text{inter}}, \]
   then $S_{\Pi}$ is exponentially stable.

Theorem 2 indicates that cluster synchronization can be achieved if (i) the inter-cluster coupling strengths are sufficiently weak and/or (ii) the natural frequencies are sufficiently different between clusters. Our approach provides a unified framework for stability analysis in that the two conditions (1) and 2) in the above theorem are simultaneously derived. We must highlight that this is achieved based on the use of a single nonmonotonic Lyapunov function. This allows us to characterize a certain tradeoff between the two conditions. We must note that these two stability conditions are similar to those derived in [24] (Theorems 3.2 and 3.3). There, different methods are applied to obtain the two conditions separately regarding the inter-cluster edge weights and the natural frequency differences.

Furthermore, our result shows that there is a critical value of the natural frequency differences between clusters, above which the cluster synchronization manifold is exponentially stable. This has an advantage since Theorem 3.3 in [24] requires that $|\omega_i - \omega_j| \rightarrow \infty$ for all $(i, j) \in E_{\text{inter}}$. The main difference in the analysis lies in the fact that our result does not depend on a particular Lyapunov function. Instead, we found a nonmonotonic Lyapunov function for the cluster synchronization manifold to prove the exponential stability via Proposition 2. We will see in Section IV-B that a specific value of the natural frequency difference for which cluster synchronization is realized can be found.

Remark 4: The analysis in this section is based on the so-called two-timescale averaging, that is, averaging is applied to a singularly perturbed system [53]. However, our goal was to show partial stability of the slow dynamics and stability of the fast dynamics is not involved. A similar setting was considered in [54], where the fast subsystem is restricted to a scalar system and the fast state is interpreted as a time variable. In contrast, the fast subsystem given in (6) has vector states, and hence, the result in [54] is not applicable.

Remark 5: Synchronization in networks with periodically switching topologies was considered in [55], and the nonperiodic case was explored in [56]. In those papers, it is allowed that the graph is disconnected at certain time instants, but it must be connected in some sense. By contrast, our setting is more general since the network dynamics becomes unstable in some time intervals, that is, the pointwise Jacobian in (15) may have unstable eigenvalues. We have utilized the inequality (10) to show that fast oscillations generated from the inter-cluster dynamics support cluster synchronization. Moreover, we note that our results can also address chimera states considered in [57], [58], where synchronous and asynchronous groups coexist in the network. This is because each oscillator in the asynchronous group can be regarded as an isolated cluster.

B. Proof of Theorem 2

Consider the linearized system (15). We define the uniform average of $J(\tau, \varepsilon)$ by

\[ J_{\text{av}} := \lim_{T \to \infty} \frac{1}{T} \int_{\tau}^{\tau+T} J(s, 0) \, ds \]

\[ = -\tilde{B}_{\text{intra}}^T B_{\text{intra}} W_{\text{intra}} R_1. \tag{16} \]

The above matrix is well defined because each element of $\cos(\zeta(\tau, 0))$ in (14) is periodic in $\tau$. If these elements have a common period, then the vector-valued function $\cos(\zeta(\tau, 0))$ is periodic in $\tau$, and otherwise it is almost periodic in $\tau$. Clearly, $J_{\text{av}}$ is determined by only the intra-cluster edge weights, and thus, it can be written as $J_{\text{av}} = \text{blk-diag}(J_1, \ldots, J_c)$. Note that $J_i$ is a Hurwitz matrix if and only if $G_i$ is connected (see Lemma 3.1 in [24]). Therefore, the average Jacobian $J_{\text{av}}$ is a Hurwitz matrix.

The following lemma characterizes the relation between the original and the average Jacobians.

Lemma 2: There exists $\gamma > 0$ such that

(i) $\gamma = O(\|W_{\text{intra}}\|_{\infty})$

(ii) for all $\varepsilon > 0$ and all $T > 0$, the following inequality holds uniformly in $\tau \in \mathbb{R}_+$:

\[ \left\| \frac{1}{T} \int_{\tau}^{\tau+T} [J(s, \varepsilon) - J_{\text{av}}] \, ds \right\| \leq \gamma \left( \frac{1}{T} + \varepsilon \right). \]

Proof: See Appendix B.

It can be observed from Lemma 2 that the time average of the difference between $J(\tau, \varepsilon)$ and $J_{\text{av}}$ is small enough if one of the following holds:

1) $\gamma$ is small;
2) $T$ is large and $\varepsilon$ is small.

To satisfy the condition in (10), we can choose a class $\mathcal{L}$ function $\beta(T) = \gamma/T$ and a class $\mathcal{K}$ function $\alpha(\varepsilon) = \gamma \varepsilon$.

Let $\Phi(\tau + T, \tau)$ be the state-transition matrix of (15), and let $\Phi_{\text{av}}(\tau + T, \tau) := e^{T J_{\text{av}}}$. Then, we consider $H(\tau + T, \tau) := \Phi(\tau + T, \tau) - \Phi_{\text{av}}(\tau + T, \tau)$. By the Peano–Baker series of the transition matrices, we have

\[ H(\tau + T, \tau) \]

\[ = \varepsilon \int_{\tau}^{\tau+T} [J(s, \varepsilon) - J_{\text{av}}] \, ds \]

\[ + \sum_{i=2}^{\infty} \varepsilon_i \int_{\tau}^{\tau+T} J(s_1, \varepsilon) \cdots J(s_{i-1}, \varepsilon) \, ds_1 \cdots ds_1 \]

\[ - \sum_{i=2}^{\infty} (\varepsilon T)^i J_{\text{av}}^i. \]

Thus, it can be obtained that

\[ \|H(\tau + T, \tau)\| \leq \varepsilon \left\| \int_{\tau}^{\tau+T} [J(s, \varepsilon) - J_{\text{av}}] \, ds \right\| \]

\[ + 2 \sum_{i=2}^{\infty} \frac{(\varepsilon T)^i}{i!} H_{\text{av}}. \]
where \( \rho := \sup_{\tau \in \mathbb{R}, \varepsilon > 0} \| J(\tau, \varepsilon) \| \). Notice from (14) that \( \rho \) is finite. Then, Lemma 2 implies that

\[
\| H(\tau + T, \tau) \| \leq \gamma \varepsilon T \left( \frac{1}{T} + \varepsilon \right) + 2(e^{\rho T} - 1 - \rho \varepsilon T).
\]

We note that \( J_{\text{av}} \) has only the negative of the nonzero eigenvalues of the graph Laplacians associated with the subgraphs \( G_1, \ldots, G_r \). This implies that \( \| \Phi_{\text{av}}(\tau + T, \tau) \| \leq e^{-\lambda_2 \varepsilon T} \), where \( \lambda_2 \) is the smallest algebraic connectivity of \( G_1, \ldots, G_r \).

We thus obtain

\[
\| \Phi(\tau + T, \tau) \| \leq \| \Phi_{\text{av}}(\tau + T, \tau) \| + \| H(\tau + T, \tau) \|
\leq e^{-\lambda_2 \varepsilon T} + \gamma \varepsilon T \left( \frac{1}{T} + \varepsilon \right) + 2(e^{\rho T} - 1 - \rho \varepsilon T)
=: \kappa(\gamma, \varepsilon, T).
\]

The problem is to show \( \kappa(\gamma, \varepsilon, T) < 1 \) for some \( T > 0 \) if \( \gamma \) or \( \varepsilon \) is sufficiently small. To find such a constant \( T \), we first let \( \varepsilon \) be fixed. At \( \gamma = 0 \), we have \( \kappa(0, \varepsilon, T) = e^{-\lambda_2 \varepsilon T} + 2(e^{\rho T} - 1 - \rho \varepsilon T) \).

Because \( \rho < \infty \), there exists \( T > 0 \) such that \( \kappa(0, \varepsilon, T) < 1 \). This means by continuity that \( \kappa(\gamma, \varepsilon, T) < 1 \) is satisfied for sufficiently small \( \gamma > 0 \), that is, small \( \| W_{\text{inter}}(\tau) \|_{\infty} \) by Lemma 2 (i); by the definition of \( W_{\text{inter}} \), this corresponds to taking the edge weights of \( W_{\text{inter}} \) to be small. Similarly, if we let \( \gamma \) be fixed, the same conclusion can be obtained for sufficiently small \( \varepsilon > 0 \), whose procedure has been shown in the proof of Theorem 1. Notice that a specific value of \( \varepsilon \) such that \( \kappa(\gamma, \varepsilon, T) < 1 \) for a suitably chosen \( T \) can be found numerically. By definition of \( \varepsilon \) in (4), we conclude that the natural frequencies are required to be sufficiently different among clusters. The proof is now complete.

C. Two-Cluster Case

In the previous subsections we have not restricted the number of clusters. However, the obtained stability conditions are somewhat conservative. We here consider cluster synchronization with two clusters \( (\tau = 2) \) and derive a more specific stability condition, which may be less conservative.

For \( i \in C_1 \) and \( j \in C_2 \), define

\[
\bar{\omega} := |\omega_i - \omega_j|, \quad \bar{a} := \sum_{k \in C_2} a_{ik} + \sum_{k \in C_1} a_{jk}.
\]

We restrict our attention to the case where \( \bar{\omega} > \bar{a} \). In that case, the solutions to (12) converge to a periodic orbit with the period

\[
T_2 = \frac{2\pi \bar{\omega}}{\sqrt{\bar{\omega}^2 - \bar{a}^2}} \tag{18}
\]
(see Lemma 3.4 in [24]). If \( \bar{\omega} < \bar{a} \), the solutions to (12) converge to constant values. This corresponds to frequency synchronization as considered in [22].

Given a nontrivial partition \( \Pi = (C_1, C_2) \), one can assume without loss of generality that \( R_4 = 1_{\text{intra}} \). We note that \( \varepsilon = 1/\bar{\omega} \). Similarly to the previous subsection, we obtain the linearized system

\[
\frac{dx}{d\tau} = \varepsilon J_2(\tau, \varepsilon)x, \tag{19}
\]

where \( J_2(\tau, \varepsilon) := J_{\text{intra}} + J_{\text{inter}} \cos(\gamma(\tau, \varepsilon)) \). Here, we have defined \( J_{\text{intra}} := -\bar{B}_{\text{intra}}^T B_{\text{intra}} W_{\text{intra}} R_1 \) and \( J_{\text{inter}} := -\bar{B}_{\text{inter}}^T B_{\text{inter}} W_{\text{inter}} R_3 \). We remark that the Jacobian cannot be transformed into the above form in the multi-cluster case.

The following result provides an explicit stability condition. In particular, this condition includes that in Theorem 3.6 in [24], where \( J_{\text{intra}} \) is assumed to be of the form \( J_{\text{intra}} = \alpha I \) for some \( \alpha < 0 \).

Theorem 3: Let \( S_\Pi \) be the cluster synchronization manifold associated with the Kuramoto model (2), where \( \Pi = (C_1, C_2) \) is a nontrivial partition of \( V \) such that the subgraphs \( \gamma_1 \) and \( \gamma_2 \) are connected. Suppose that Assumptions 1 and 2 hold and that \( \bar{\omega} > \bar{a} \). If the matrices \( J_{\text{intra}} \) and \( J_{\text{inter}} \) commute, i.e., if \( J_{\text{intra}} J_{\text{inter}} = J_{\text{inter}} J_{\text{intra}} \), then \( S_\Pi \) is exponentially stable.

Proof: The inter-cluster dynamics under \( x = 0 \) is governed by

\[
\frac{dz}{d\tau} = 1 - \varepsilon \bar{a} \sin(z).
\]

Since \( \varepsilon \bar{a} < 1 \), the solution is periodic with the period \( T_2 \) in (18) and has zero average. Notice that the uniform average is given as

\[
J_{2,\text{av}} := \frac{1}{T_2} \int_0^{T_2} J_2(s, 0) \, ds = J_{\text{intra}}.
\]

Notice that for the two cluster case, the inequality in Lemma 2 is translated into the following equality:

\[
\frac{1}{T_2} \int_0^{T_2} [J_2(s, \varepsilon) - J_{2,\text{av}}] \, ds = 0 \tag{20}
\]

Since \( J_{\text{intra}} \) and \( J_{\text{inter}} \) commute, we have

\[
J_2(\tau, \varepsilon) \int_\sigma^\tau J_2(s, \varepsilon) \, ds = \int_\sigma^\tau J_2(s, \varepsilon) \, ds J_2(\tau, \varepsilon)
\]

for all \( \sigma, \tau \in \mathbb{R}_+ \). In the proof of Theorem 2, we employed the Peano–Baker series of state transition matrices. This takes a complex form in general, leading us to a conservative stability condition. Here, by the above property, the Peano–Baker series becomes simpler in its form (see Chap. 4 of [59]).

Let \( \Phi_2(\tau + T, \tau) \) denote the state transition matrix of (19), and let \( \Phi_{2,\text{av}}(\tau + T, \tau) := e^{T J_{2,\text{av}}} \).

Then, we have

\[
\Phi_2(\tau + T, \tau) - \Phi_{2,\text{av}}(\tau + T, \tau) = \sum_{i=2}^{\infty} \frac{\varepsilon^i}{i!} \left( \int_{\tau}^{\tau+T} [J_2(s, \varepsilon) - J_{2,\text{av}}] \, ds \right)^i.
\]

Due to (20), it follows that \( \Phi_2(\tau + T, \tau) - \Phi_{2,\text{av}}(\tau + T, \tau) = 0 \), which implies

\[
\Phi_2(\tau + T, \tau) = \Phi_{2,\text{av}}(\tau + T, \tau) = e^{T J_{2,\text{av}}}.
\]

Since \( J_{2,\text{av}} \) is a Hurwitz matrix, \( |\Phi_2(\tau + T, \tau)| < 1 \) can be established. Therefore, we conclude the proof.

Remark 6: Two-cluster synchronization can be considered as a nonlinear version of bipartite consensus, which is studied in the presence of antagonistic interactions [60], [61]. The Kuramoto model with antagonistic interactions is considered in [62], where the underlying network is represented by a signed graph. In this paper we do not consider signed graphs.
explicitly, but our results can be extended to that setting by utilizing a signed external equitable partition condition [32] in place of Assumption 2.

D. Numerical Examples

We present two numerical examples. The first one is related to Theorem 2 and the second one demonstrates Theorem 3.

**Example 1:** Consider the network shown in Fig. 1, which is partitioned into three clusters as indicated. The adjacency matrix of this network is given by

\[
A = \begin{bmatrix}
0 & a_1 & b_1 & b_{1/2} & 0 & 0 & 0 \\
-a_1 & 0 & b_{1/2} & b_1 & 0 & 0 & 0 \\
0 & 0 & a_2 & 0 & 0 & 0 & 0 \\
a_{1/2} & 0 & a_2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & a_3 & a_3 & 0 \\
0 & 0 & 0 & 0 & 0 & a_3 & a_3 \\
0 & 0 & 0 & 0 & 0 & 0 & a_3
\end{bmatrix}.
\]

Notice that in this case, Assumption 2 is satisfied for any parameter value. We set all the edge weights to be 1, that is, \(a_1 = a_2 = a_3 = b_1 = b_2 = 1\). The natural frequencies are set as \(\omega_1 = 0\) for \(i \in C_1\), \(\omega_i = 5\) for \(i \in C_2\), and \(\omega_i = 10\) for \(i \in C_3\). We present in Fig. 2 the phase evolutions of the oscillators, where the oscillators are synchronized within clusters and the inter-cluster oscillations can be observed. Theoretically, we obtain the tradeoff curve for which \(\kappa(\gamma, \varepsilon, T)\) in the proof of Theorem 2 is equal to one as shown in Fig. 3. The cluster synchronization manifold is guaranteed to be exponentially stable for any pair \((\gamma, \varepsilon)\) in the left-bottom area.

**Example 2:** Consider the network shown in Fig. 4, where the network is partitioned into two clusters. The adjacency matrix is given by

\[
A = \begin{bmatrix}
0 & a_1 & 0 & 0 & b \\
a_1 & 0 & a_1 & 0 & b \\
0 & a_1 & 0 & b & 0 \\
0 & 0 & b & 0 & a_2 \\
0 & b & 0 & a_2 & 0 \\
b & 0 & 0 & a_2 & 0
\end{bmatrix}.
\]

In this case we have the linearized system (19) with

\[
J_{\text{intra}} = \begin{bmatrix}
-2a_1 & a_1 & 0 & 0 \\
a_1 & -2a_1 & 0 & 0 \\
0 & 0 & -2a_2 & a_2 \\
0 & 0 & a_2 & -2a_2
\end{bmatrix},
\]

which is a Hurwitz matrix, and

\[
J_{\text{inter}} = \begin{bmatrix}
-b & 0 & 0 & -b \\
0 & -b & -b & 0 \\
0 & -b & -b & 0 \\
-b & 0 & 0 & -b
\end{bmatrix}.
\]

One can observe that \(J_{\text{intra}}\) and \(J_{\text{inter}}\) commute if \(a_1 = a_2\) or \(b = 0\). Thus, we set \(a_1 = a_2 = 1\) and \(b = 1\). The simulation result with \(\omega_i = 5\) for \(i \in C_1\) and \(\omega_i = 1\) for \(i \in C_2\) is shown in Fig. 5 (top). In this setting, Theorem 3 can provide a less conservative stability condition compared with Theorem 2, which requires that \(\bar{\omega} > 125\) for achieving cluster synchronization. Next, we would like to present an example where cluster synchronization does not occur. We change the edge weights within cluster \(C_1\) to a much smaller value \(a_1 = 0.01\). Then, \(J_{\text{intra}}\) and \(J_{\text{inter}}\) do not commute, and the simulation result with the same natural frequencies is presented in Fig. 5 (bottom). This indicates that the cluster synchronization manifold is not necessarily stable.

V. PRACTICAL STABILITY WITHOUT EXTERNAL EQUITABLE PARTITIONS

In Section IV we have studied cluster synchronization under the assumption of external equitable partitions. This hypothesis may be restrictive to apply to real-world networks. Here, we remove Assumption 2 and explore the question of whether it is possible to achieve arbitrarily small partial phase cohesiveness within clusters by tuning network parameters. To do so, we
consider practical stability under perturbation from an external equitable partition. In particular, we are concerned with the stability concept defined below.

**Definition 3:** The cluster synchronization manifold $\mathcal{S}_\Pi$ of (2) is $(\delta, \rho)$-practically stable if for given $\delta > 0$ and $\rho > 0$, there exists $T > 0$ such that $\text{dist}(\theta(0), \mathcal{S}_\Pi) < \delta$ implies $\text{dist}(\theta(t), \mathcal{S}_\Pi) \leq \rho$ for all $t \geq T$.

Similar definitions of practical stability under perturbations can be found in [63], [64]. To proceed, we impose the following assumption instead of Assumption 2.

**Assumption 3:** There exists $K \geq 0$ such that for every $p, q \in \{1, \ldots, r\}$ with $p \neq q$, it holds that $|\sum_{k \in C_q} (a_{ik} - a_{jk})| \leq K$ for all $i, j \in C_p$. We remark that as long as the edge weights are finite, there always exists $K$ such that the condition in Assumption 3 is satisfied. Also, we note that Assumption 2 corresponds to the case where $K = 0$, and hence, $K$ represents how large the variation from an external equitable partition is.

The following result states that the distance between the solutions of (2) and the cluster synchronization manifold can be made arbitrarily close by strengthening the connectivity within clusters.

**Theorem 4:** Let $\mathcal{S}_\Pi$ be the cluster synchronization manifold associated with the Kuramoto model (2), where $\Pi$ is a non-trivial partition of $V$ such that the subgraphs $G_1, \ldots, G_r$ are all connected. Suppose that Assumptions 1 and 3 hold. Then, for any $\rho > 0$, there exist $\delta > 0$ and $\alpha' > 0$ depending on $K$ such that if

$$a_{ij} > \alpha' \quad (i, j) \in \mathcal{E}_p, \quad p \in \{1, \ldots, r\},$$

then $\mathcal{S}_\Pi$ is $(\delta, \rho)$-practically stable.

**Proof:** Instead of (5), we consider the nonlinear system $\dot{x} = h(t, x)$, where $h$ is a function such that $h(t, x) = f(x, z(t))$ with $z(t)$ satisfying (6). This can be considered as the perturbed system

$$\dot{x} = h_0(t, x) + p(t, x),$$

where $h_0(t, x)$ is a vector field associated with the external equitable partition condition and $p(t, x)$ is a bounded perturbation term. In particular, these are given as

$$h_0(t, x) := -\hat{B}_{\text{intra}}^T \hat{B}_{\text{inter}} W_{\text{intra}} \sin(R_1 x)$$

$$- \hat{B}_{\text{intra}}^T \hat{B}_{\text{inter}} W_{\text{inter}}^\prime \sin(R_3 x + R_4 z(t)),$$

$$p(t, x) := -\hat{B}_{\text{intra}}^T \hat{B}_{\text{inter}} (W_{\text{inter}} - W_{\text{inter}}^\prime) \times \sin(R_3 x + R_4 z(t)),$$

where $W_{\text{inter}}^\prime$ is a certain weight matrix with respect to inter-cluster edges for which the condition in Assumption 3 holds with $K = 0$. Also, for all $t \in \mathbb{R}_+$ and all $x \in \mathbb{R}^n$, we have

$$\|p(t, x)\| \leq \|\hat{B}_{\text{intra}}^T \hat{B}_{\text{inter}}\| \|W_{\text{inter}} - W_{\text{inter}}^\prime\| \|x\| \leq cK,$$

where $c := \|\hat{B}_{\text{intra}}^T \hat{B}_{\text{inter}}\|$. Assume that the zero solution $x'(t) \equiv 0$ to the unperturbed system $\dot{x}' = h_0(t, x')$ is exponentially stable, i.e., there exists $k > 0$, $\lambda > 0$, and $\delta' > 0$ such that $\|x'(t)\| < \delta'$ implies $\|x(t)\| \leq k e^{-\lambda(t-t_0)} \|x(t_0)\|$ for all $t \geq t_0$. We note that $\lambda$ can be made arbitrarily large by strengthening the algebraic connectivities of the subgraphs $G_1, \ldots, G_r$. This can be seen from the result in Section IV-A. According to the converse Lyapunov theorem, the unperturbed system allows a Lyapunov function $V : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$. Let $c_1, c_2, c_3, c_4 > 0$ be constants such that for all $t \in \mathbb{R}_+$ and all $x \in \mathbb{R}^n$,

$$c_1 \|x\|^2 \leq V(t, x) \leq c_2 \|x\|^2,$$

$$\frac{\partial V(t, x)}{\partial t} + \frac{\partial V(t, x)}{\partial x} h_0(t, x) \|x\| \leq -c_3 \|x\|^2,$$

whenever $\|x\| < d$ for some $d > 0$. In what follows, we show that the solutions to the perturbed system (21) converge to a certain small residual set.

Assume that $cK < (c_3/c_4) \sqrt{c_1/c_2} d$. From Lemma 9.2 in [40], any solution to (21) such that $\|x(0)\| < \sqrt{c_1/c_2} d$ satisfies the following inequality with some $T > 0$:

$$\|x(t)\| < \frac{c_4}{c_3} \sqrt{\frac{c_2}{c_1} cK}, \quad t \geq T.$$  

By choosing $c_1, c_2, c_3, c_4$ as in Theorem 4.14 in [40], we can observe that the larger $\lambda$ is, the smaller $c_4/c_3$ is, while $c_2/c_1$ remains the same. Thus, by strengthening the algebraic connectivities of $G_1, \ldots, G_r$, we can make the right-hand side in (22) small arbitrarily. For the same reason, the hypothesis that $cK < (c_3/c_4) \sqrt{c_1/c_2} d$ is allowed. Finally, we verify that for any $\rho > 0$, $(\delta, \rho)$-practical stability is realizable for some $\delta > 0$. Given $\rho > 0$, we make $\lambda$ large such that $(c_4/c_3) \sqrt{c_2/c_1} cK = \rho$. Then, by defining $\delta := \sqrt{c_1/c_2} d$, we conclude that if $\|x(0)\| < \delta$, then $x(t)$ eventually converges.
to the set \( \{ x \in \mathbb{R}^n : \| x \| \leq \rho \} \). This immediately results in the \((\delta, \rho)\)-practical stability of the cluster synchronization manifold \( S_{\Pi} \). Therefore, the proof is now complete. \( \blacksquare \)

Theorem 4 indicates that even if Assumption 2 is not satisfied, clustered patterns of synchronization can arise. Furthermore, we can design the network parameters \( a_{ij} \) to realize a desired partial phase cohesiveness within clusters, which is characterized by \( \rho \) in the above theorem. We note that this theorem is helpful to proceed from analysis to control of synchronization patterns.

Remark 7: Here, we discuss differences between our result and those in [25], where partial phase cohesiveness within clusters is considered with a slightly different setting. In this paper we consider the perturbation of the edge weights from an external equitable partition. This can be seen as a generalization of the case of completely separated clusters, i.e., the case where \( a_{ij} = 0 \) for all \( (i, j) \in \mathcal{E}_{\text{inter}} \). Therefore, combining the result in this section with those in Section IV, we can conclude that if the natural frequencies are largely different among clusters, the inter-cluster couplings do not necessarily have to be small for realizing partial synchronization patterns. This fact cannot be obtained from the results in [25].

VI. APPLICATIONS TO BRAIN NETWORKS

This section focuses on applications of the cluster synchronization conditions derived so far to brain networks. We first explain the construction of human brain networks, followed by the exploration of synchronization patterns in neural activities. However, neural activities of whole brains cannot be measured directly, and fMRI measures blood flow-dependent signals. Thus, we then consider hemodynamic responses [65] and the relation between synchronization patterns of neural activities and functional connectivity obtained from fMRI signals.

A. CONSTRUCTION OF BRAIN NETWORKS FROM EMPIRICAL DATA

Mapping of human brain networks with noninvasive methods has been an active research area in the neuroscience community. Here, we consider the brain networks identified in [9], whose data including their weighted adjacency matrices can be found at the USC Multimodal Connectivity Database [66]. In [9], the authors used diffusion spectrum imaging to extract the network structure with 998 cortical subregions of equal sizes for 5 human subjects. Those subregions are classified into anatomically determined 66 cortical regions.

Following [67], [68], we process the empirical data to construct a representative network for computer simulation. We first binarize all the 5 weighted adjacency matrices and then average them into the 66 cortical regions, that is, sum up all the weights in each region and divide it by the number of the target regions. After that, we average the 66-region networks of 5 human subjects and normalize it so that the weights belong to the range \([0, 10]\). The resulting representative network is shown in Fig. 6, where the image was generated by BrainNet Viewer [69] and only 30% of edges are visualized. We also present in Fig. 7 the obtained adjacency matrix. Note that this network is nonsymmetric but is almost symmetric.

B. CLUSTER SYNCHRONIZATION OF NEURAL ACTIVITIES

In general, functional connectivity is similar to the underlying structural connectivity. However, they are not consistent and there are some regions between which functional connectivity is strong but structural connectivity is weak [67]. This fact can be explained by cluster synchronization phenomena as explored in Section IV.

We briefly introduce a mathematical model for neural activities in brain networks. Because of the complexity of human brains, the model is required to represent dynamical behaviors at an appropriate spatial scale. Here, we are interested in macroscopic behaviors of the populations of neurons. The macroscopic dynamics of brains were investigated in [70], [71], and the so-called Wilson–Cowan model was proposed. This model consists of the populations of neurons with excitatory and inhibitory types, and can represent oscillations in the phase space of the mean numbers of activated excitatory and inhibitory neurons. Such a model was used to analyze functional connectivity in [67], [72].

Here, we adopt the simpler Kuramoto model, which was employed in [68], [73]. The derivation of a multiplexed Kuramoto model from the Wilson–Cowan model was investigated...
in [74]. In functional connectivity, it is known that delays in the communication among oscillators also play a certain role [72]. We, however, ignore the time delays and focus on the effect of heterogeneity of the natural frequencies. Thus, the model considered here is the one in (2). Using the adjacency matrix obtained above, we simulate the phase evolution of the Kuramoto oscillators with various choices of parameters.

We consider three clusters in the brain network, each consisting of 22 nodes as represented in Figs. 6 and 7. To measure the synchronization level in each cluster, we introduce the order parameters defined by $r_p(t) = \frac{1}{|C_p|} \sum_{i \in C_p} e^{\theta_i(t)}$ for $p \in \{1, 2, 3\}$. In geometric interpretations, the order parameters represent the distance from the origin to the mean position of the oscillators on the unit circle. That is, $r_p(t) = 1$ when all the oscillators in the cluster $C_p$ are located at the same point, and $r_p(t) = 0$ when they are well spread with equal distances. We also define the global order parameter to quantify the synchronization level among the whole network as $r(t) = \frac{1}{N} \sum_{i=1}^{N} e^{\theta_i(t)}$. Furthermore, we calculate the Pearson correlation of the neural activities among cortical regions, where the neural signal in the $i$th region is written as $z_i(t) = \sin(\theta_i(t))$.

We simulate the neural activities with the Kuramoto model for 100 seconds. Then, after removing the data of the first 40 seconds, we compute the Pearson correlation coefficients among all pairs of regions. In the simulations, we produce the oscillators’ frequencies $\omega_i/2\pi$ by Gaussian distributions with mean $\mu$ and standard deviation $\sigma$. For the first simulation, we set $(\mu, \sigma) = (60, 0.5)$, and the order parameters as well as the correlation matrix are presented in Fig. 8. One can observe in the left figures that the synchronization levels of the individual clusters (in the top plot) are high but the global synchronization level (in the bottom plot) is also high. This indicates that specific correlation patterns of neural activities do not appear as we can also confirm from the right figure.

Next, we demonstrate that the heterogeneity in the natural frequencies plays a key role for the appearance of synchronization patterns. We produce the natural frequencies in clusters $C_1$, $C_2$, and $C_3$ by Gaussian distributions with $(\mu, \sigma) = (50, 0.5)$, $(\mu, \sigma) = (60, 0.5)$, and $(\mu, \sigma) = (70, 0.5)$, respectively. The simulation result is presented in Fig. 9, and it can be seen that the order parameters within clusters are moderately high while the global order parameter is averagely low. Hence, a correlation pattern can be observed. This result exhibits that the heterogeneity in the natural frequencies is important for the functional patterns observed in human brains and is also consistent with the conclusions discussed in [75].

Finally, we aim to realize stronger synchronization levels within clusters. From Theorem 4, this can be realized by strengthening the intra-cluster coupling strengths. Thus, we double the intra-cluster coupling strengths while the natural frequencies are kept the same as in the second simulation. The obtained simulation result is shown in Fig. 10. In this case, the order parameters within clusters are very high while the global order parameter is averagely low.

C. Hemodynamic Responses and Functional Connectivity

As already stated in the introduction, the fMRI techniques allow us to measure the underlying brain activities. These are detected as slow fluctuations (about 0.1 Hz) of the blood oxygen level-dependent (BOLD) signals, which are related to the blood volume and the deoxyhemoglobin content of the cortical regions. This is based on the fact that the brain activity at a region consumes oxygen in the blood flowing around that region. Functional connectivity is then defined as the correlation level of the BOLD signals among different cortical regions. In the last part of this section, we discuss the relation of cluster synchronization with functional connectivity.

Following [65], we describe a mathematical model of hemodynamic responses. The model consists of two parts. The first part represents a linear relation between the neural activity and the regional cerebral blood flow (rCBF). In particular, the neural activity $z_i$ causes the increase of the vasodilatory signal $s_i$, which is integrated to obtain the blood inflow $f_i$. The signal $s_i$ is regulated by the self feedback and by the blood flow feedback. The second part links rCBF and fMRI BOLD signals with nonlinear relation and is called the Balloon-Windkessel model. In this model, the difference between the inflow $f_i$ and the outflow $f_i^{\text{out}}$ results in the increase of the regional blood volume $v_i$. The deoxyhemoglobin content $q_i$ changes depending on the blood flow delivery and the concentration itself. The BOLD signal $y_i$ is then obtained, depending nonlinearly on $v_i$ and $q_i$. In particular, the hemodynamic model is described by

\begin{align}
\dot{s}_i &= z_i - \kappa s_i - \gamma (f_i - 1), \\
\dot{f}_i &= s_i, \\
\tau \dot{v}_i &= f_i - f_i^{\text{out}}(v_i), \\
\tau \dot{q}_i &= f_i E(f_i) - f_i^{\text{out}} q_i v_i,
\end{align}

where $f_i^{\text{out}}(v_i) = v_i^{1/\alpha}$ and $E(f_i) = 1 - (1 - E_0)^{1/f_i}$. The BOLD signal is calculated as

\begin{equation}
y_i = V_0(k_1(1 - q_i) + k_2(1 - q_i/v_i) + k_3(1 - v_i)),
\end{equation}

where $V_0 = 0.02$, $k_1 = 7E_0$, $k_2 = 2$, and $k_3 = 2E_0 - 0.2$. We show in Table I the meanings of the parameters appearing in (23)–(26) and their values estimated in [65].

It is known that a characteristic correlation pattern of the BOLD signals appears in the low frequency range less than 0.1 Hz [76]. This low-frequency synchronization can be explained as follows. The state equations of the hemodynamic model in (23)–(26) have an equilibrium point at $[s_i, f_i, v_i, q_i]^T = [0, 1, 1]^T$. Since each signal fluctuates around the equilibrium with small magnitudes, we linearize the hemodynamic model. In what follows, we use $s_i$, $f_i$, $v_i$, and $q_i$ to represent the

| Parameter | Meaning | Value |
|-----------|---------|-------|
| $\kappa$ | Decay rate of signals | 0.65 |
| $\gamma$ | Elimination rate of signals | 0.41 |
| $\tau$ | Time constant for blood flow | 0.98 |
| $\alpha$ | Stiffness exponent | 0.33 |
| $E_0$ | Resting oxygen extraction fraction | 0.34 |
Fig. 8. Synchronization levels for natural frequencies under Gaussian distributions with $(\mu, \sigma) = (60, 0.5)$. Left: Order parameters for individual clusters (top) and the whole system (bottom). Right: Correlation matrix.

Fig. 9. Synchronization levels for natural frequencies under Gaussian distributions with $(\mu, \sigma) = (50, 0.5)$ for cluster $C_1$, $(\mu, \sigma) = (60, 0.5)$ for cluster $C_2$, and $(\mu, \sigma) = (70, 0.5)$ for cluster $C_3$. Left: Order parameters for individual clusters (top) and the whole system (bottom). Right: Correlation matrix.

Fig. 10. Synchronization levels with doubled intra-cluster coupling strengths. Left: Order parameters for individual clusters (top) and the whole system (bottom). Right: Correlation matrix.
deviation from this point with abuse of notation. The linearized model is described by the series connection of the following two systems:

\[
\begin{bmatrix}
    \dot{s}_i \\
    \dot{f}_i
\end{bmatrix} = \begin{bmatrix}
    -\kappa & -\gamma \\
    1 & 0
\end{bmatrix} \begin{bmatrix}
    s_i \\
    f_i
\end{bmatrix} + \begin{bmatrix}
    1 \\
    0
\end{bmatrix} z_i,
\]

\[\eta_i = \begin{bmatrix}
    0 & 1
\end{bmatrix} \begin{bmatrix}
    s_i \\
    f_i
\end{bmatrix},\]

and

\[
\begin{bmatrix}
    \dot{v}_i \\
    \dot{q}_i
\end{bmatrix} = \begin{bmatrix}
    -1/(\tau \alpha) & 0 \\
    -(1-\alpha)/(\tau \alpha) & -1/\tau
\end{bmatrix} \begin{bmatrix}
    v_i \\
    q_i
\end{bmatrix} + \begin{bmatrix}
    1/\tau \\
    \beta(E_0)
\end{bmatrix} \eta_i,
\]

\[y_i = V_0 \begin{bmatrix}
    k_2 - k_3 & -k_1 - k_2
\end{bmatrix} \begin{bmatrix}
    v_i \\
    q_i
\end{bmatrix},\]

where \(\beta(E_0) := [E_0 + (1-E_0) \ln(1-E_0)]/(\tau E_0)\). Both are stable systems, and the hemodynamic model can be interpreted as a low-pass filter [68]. This implies that synchronization at relatively high frequency can be detected as the correlation of the BOLD signals at low frequency. In particular, by a frequency analysis, it can be observed that a neural signal \(z_i(t)\) with frequency around 60 Hz is identified as a BOLD signal \(y_i(t)\) with peak frequency around 0.1 Hz.

We compute the BOLD signals with the hemodynamic model for 100 seconds. The natural frequencies of the Kuramoto oscillators in each cluster are respectively generated by Gaussian distributions with means 50, 60, and 70, and standard deviation 0.5, which are the same setting as the second simulation in the previous subsection. We then compute the Pearson correlation coefficients after removing the data of the first 40 seconds. We present in Fig. 11 the obtained functional connectivity map. From this figure, we see that a pattern of functional connectivity can be observed similarly to the correlation pattern of the neural activities examined in the previous subsection.

We here give some discussions. This paper has been concerned with synchronization patterns in oscillator networks. In particular, weak inter-cluster connections and large heterogeneity in the natural frequencies among clusters are key factors of cluster synchronization. These factors also play certain roles in human brain networks as demonstrated above. Thus, it would be of interest to analyze functional brain networks from these viewpoints and moreover see if healthy and pathological states of brains may be characterized based on this criterion. Compared with the graph-theoretic analysis [12] and numerical studies [68] in the neuroscience community, this paper has provided a theoretical analysis based on dynamical models of oscillators to explain the robust patterns in functional connectivity of brains. There are some possible future directions of this work. One is the brain intervention for cluster synchronization as considered in [16]. Furthermore, the analysis of remote synchronization is important since it is known that some brain regions located remotely can be synchronized [77].

VII. Concluding Remarks

In this paper we have considered cluster synchronization of Kuramoto oscillators. Through a stability criterion based on averaging, we have shown that cluster synchronization is achieved if (i) the inter-cluster couplings are sufficiently weak and/or (ii) the natural frequencies are sufficiently different among clusters. Practical stability has been investigated to remove the hypothesis on network partitions. We have also examined applications to empirical brain networks and have discussed the relation between the cluster synchronization problem for the Kuramoto model and the problem of functional connectivity in brain networks.

Appendix

A. Proof of Theorem 1

The average system associated with (8) is given by

\[\dot{x} = \varepsilon A_{\text{av}} x,\]

where \(A_{\text{av}}\) is defined in (9). Since \(A_{\text{av}}\) is a Hurwitz matrix, there exists a positive-definite matrix \(P \in \mathbb{R}^{n \times n}\) such that \(A_{\text{av}}^T P + P A_{\text{av}} < 0\). We define the positive-definite function \(V(x) := x^T P x\). This is clearly a Lyapunov function for (28), and thus, there exist \(c_1, c_2, c_3 > 0\) such that for all \(x \in \mathbb{R}^n\),

\[c_1 ||x||^2 \leq V(x) \leq c_2 ||x||^2,\]

\[\frac{\partial V(x)}{\partial x} A_{\text{av}} x \leq -c_3 ||x||^2.\]

For a fixed \(t\), we consider the time interval \([t, t+T]\), where \(T > 0\) is determined later. Let \(\Phi(t+T, t)\) denote the state transition matrix of (8), i.e., \(x(t+T) = \Phi(t+T, t)x(t)\). The Peano–Baker series of \(\Phi(t+T, t)\) is given by

\[\Phi(t+T, t) = I + \varepsilon \int_t^{t+T} A(s_1, \varepsilon) ds_1 + \varepsilon^2 \int_t^{t+T} A(s_1, \varepsilon) \int_t^{s_1} A(s_2, \varepsilon) ds_2 ds_1 + \cdots\]

\[= I + \varepsilon \int_t^{t+T} A(s_1, \varepsilon) ds_1 + \sum_{i=2}^{\infty} \varepsilon^i \int_t^{t+T} A(s_1, \varepsilon) \cdots \int_t^{s_{i-1}} A(s_i, \varepsilon) ds_i \cdots ds_1.\]
Let $\Phi_{av}(t + T, t)$ denote the state transition matrix of (28), i.e., $\xi(t + T) = \Phi_{av}(t + T, t)\xi(t)$. This can be calculated as

$$\Phi_{av}(t + T, t) = e^{eT A_{av}} = I + eT A_{av} + \sum_{i=2}^{\infty} \frac{(eT)^i A_{av}^i}{i!}. \quad (32)$$

We then consider the difference $H(t + T, t) := \Phi(t + T, t) - \Phi_{av}(t + T, t)$ between these two transition matrices. From (31) and (32), we have

$$H(t + T, t) = \varepsilon \int_t^{t+T} [A(s, \varepsilon) - A_{av}] \, ds + \sum_{i=2}^{\infty} \varepsilon^i \int_t^{t+T} A(s_1, \varepsilon) \cdots \int_t^{s_{i-1}} A(s_i, \varepsilon) \, ds_i \cdots ds_1 - \sum_{i=2}^{\infty} \frac{(\varepsilon T)^i A_{av}^i}{i!}.$$

It follows that

$$\|H(t + T, t)\| \leq \varepsilon \left\| \int_t^{t+T} [A(s, \varepsilon) - A_{av}] \, ds \right\| + \sum_{i=2}^{\infty} \varepsilon^i \int_t^{t+T} \left\| A(s_1, \varepsilon) \right\| \cdots \int_t^{s_{i-1}} \left\| A(s_i, \varepsilon) \right\| \, ds_i \cdots ds_1 - \sum_{i=2}^{\infty} \frac{(\varepsilon T)^i \|A_{av}\|^i}{i!} \leq \varepsilon T \beta(T) + \alpha(\varepsilon) + 2(e^{eT} - 1 - \rho eT),$$. \]

where $\rho := \sup_{t \in \mathbb{R}_+, \varepsilon > 0} \|A(t, \varepsilon)\|$ and we have used the inequality (10) in the last line. Let us define

$$h(\varepsilon, T) := \varepsilon T \beta(T) + \alpha(\varepsilon) + 2(e^{eT} - 1 - \rho eT). \quad (33)$$

Again, we consider the Lyapunov function candidate $V(x)$. Along the solutions to (8), we have

$$V(x(t + T)) - V(x(t)) = x^T(t) \Phi^T_{av}(t + T, t) P \Phi_{av}(t + T, t) - P x(t) = x^T(t) \Phi^T_{av}(t + T, t) P \Phi_{av}(t + T, t) - P + \Phi^T_{av}(t + T, t) PH(t + T, t) + H^T(t + T, t) P \Phi_{av}(t + T, t) + H^T(t + T, t) PH(t + T, t) x(t).$$

For the first two terms in the far right-hand side, it can be guaranteed from (29) and (30) that

$$x^T(t) \Phi^T_{av}(t + T, t) P \Phi_{av}(t + T, t) - P x(t) \leq c_2 e^{-\frac{\beta}{2} eT} - 1 \|x(t)\|^2.$$\]

Note that $\|P\| \leq c_2$ and $\|\Phi_{av}(t + T, t)\| \leq \sqrt{c_2/c_1} e^{-\frac{\beta}{2} eT}$. Hence, by using $h(\varepsilon, T)$ defined in (33), we obtain

$$V(x(t + T)) - V(x(t)) \leq c_2 \left( e^{-\frac{\beta}{2} eT} - 1 \right) + 2 \sqrt{c_2/c_1} e^{-\frac{\beta}{2} eT} h(\varepsilon, T) + h^2(\varepsilon, T) \times |x(t)|^2 = \tilde{h}(\varepsilon, T) |x(t)|^2. \quad (34)$$

It can be verified that $\tilde{h}(0, T) = 0$. Moreover, we have

$$\limsup_{\varepsilon \to 0^+} \frac{\tilde{h}(\varepsilon, T) - \tilde{h}(0, T)}{\varepsilon} = -T \left( c_3 - 2c_2 \sqrt{c_2/c_1} \beta(T) \right).$$

Because $\beta$ is a strictly decreasing function that converges to zero, the positive constant $T$ can be chosen such that $c_3 - 2c_2 \sqrt{c_2/c_1} \beta(T) > 0$. This implies that for sufficiently small $\varepsilon > 0$, we have $h(\varepsilon, T) < 0$. It follows from (34) that $V(x(t + T)) - V(x(t)) < 0$ whenever $x(t) \neq 0$.

Notice that the above procedure can be performed for every $t \in \mathbb{R}_+$. Hence, there always exists a time sequence $\{t_k\}_{k \in \mathbb{Z}_+}$ satisfying the hypotheses in Proposition 2. Therefore, we conclude that there exists $\varepsilon^* > 0$ such that if $\varepsilon < \varepsilon^*$, the zero solution $x(t) \equiv 0$ of (8) is exponentially stable.

\section*{B. Proof of Lemma 2}

From (14) and (16) we have

$$\int_\tau^{\tau+T} [J(s, \varepsilon) - J_{av}] \, ds = -\tilde{B}^T_{\text{inter}} B_{\text{inter}} W_{\text{inter}} \int_\tau^{\tau+T} \text{diag}(R_4 \cos(\zeta(s, \varepsilon))) \, ds R_3.$$\]

Because each row of $R_4$ contains single 1 or $-1$ and other elements are zeros, it holds that $\|R_4 x\|_\infty = \|x\|_\infty$ for all $x \in \mathbb{R}^m$. It can be verified that

$$\left\| \int_\tau^{\tau+T} [J(s, \varepsilon) - J_{av}] \, ds \right\| \leq \|\tilde{B}^T_{\text{inter}} B_{\text{inter}} W_{\text{inter}}\| \|R_3\| \left\| \int_\tau^{\tau+T} \cos(\zeta(s, \varepsilon)) \, ds \right\|_\infty,$$\]

where we have used the fact that the spectrum norm of the diagonal matrix in the integral part is equal to its induced $\infty$-norm. Note that the solution $\zeta$ to (13) can be written as

$$\zeta(\tau, \varepsilon) = \zeta_0 + \eta \tau - \varepsilon \tilde{B}^T_{\text{inter}} B_{\text{inter}} W_{\text{inter}} R_4 \int_0^\tau \sin(\zeta(s, \varepsilon)) \, ds,$$\]

where we note that $R_4$ can be moved outside of the argument of $\sin(\cdot)$ by the same reason as mentioned above. For notational simplicity, let $\Psi := \tilde{B}^T_{\text{inter}} B_{\text{inter}} W_{\text{inter}} R_4$. Then, we consider

$$e^{\tilde{\zeta}(s, \varepsilon)} = e^{\tilde{\zeta}(s_0 + \eta s)} \odot e^{-\tilde{\zeta} s \Psi f_0 \sin(\tilde{\zeta}(r, \varepsilon)) \, dr},$$\]

where the exponential function of a vector is to be taken elementwise such as $e^x = [e^{x_1} \cdots e^{x_n}]^T$ and $\odot$ denotes
the Hadamard (elementwise) product. By defining \( \psi(s) := [\text{diag}(q)]^{-1} j(\xi(s,\tau)) \), we have
\[
\int_{T} e^{j(\xi(s,\tau))} + e^{-j(\xi(s,\tau))} \, ds \leq 2 + \varepsilon T \| \Psi \|_\infty.
\]

Similarly, we have
\[
\int_{T} e^{-j(\xi(s,\tau))} \, ds \leq 2 + \varepsilon T \| \Psi \|_\infty.
\]

Combining the above two inequalities with the equality
\[
\cos(\xi(s,\tau)) = \frac{e^{j(\xi(s,\tau))} + e^{-j(\xi(s,\tau))}}{2}
\]

yields
\[
\frac{1}{T} \int_{T} \cos(\xi(s,\tau)) \, ds \leq \frac{2}{T} + \varepsilon \| \Psi \|_\infty.
\]

Therefore, we obtain
\[
\left\| B_{\text{intra}}^{T} B_{\text{inter}} W_{\text{inter}} \right\| \| R_3 \| \left( \frac{2}{T} + \varepsilon \| \Psi \|_\infty \right)
\]

\[
\leq \gamma \left( \frac{1}{T} + \varepsilon \right),
\]

where \( \gamma := \left\| B_{\text{intra}}^{T} B_{\text{inter}} W_{\text{inter}} \right\| \| R_3 \| \max\{2, \| \Psi \|_\infty\} \). It is clear that \( \gamma = O\left(\{W_{\text{inter}}\|_\infty\}\right) \). The proof is now complete. 

**REFERENCES**

[1] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. Synchronization in complex networks. *Phys. Rep.*, 469(3):93–153, 2008.

[2] F. Dörfler and F. Bullo. Synchronization in complex networks of phase oscillators: A survey. *Automatica*, 50(6):1539–1564, 2014.

[3] A. E. Motter, S. A. Myers, M. Anghel, and T. Nishikawa. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.*, 8:700–711, 2007.

[4] B. A. Francis and M. Maggiore. Flocking and Rendezvous in Distributed Robotics. Springer, 2016.

[5] L. Glass. Synchronization and rhythmic processes in physiology. *Nature*, 410:277–284, 2001.

[6] D. Zuzik and A. Drago. Neuronal oscillations in cortical networks. *Science*, 304(5679):1926–1929, 2004.

[7] C. D. Smyser and J. S. Shimony. Resting-state fMRI: A review of methods and clinical applications. *Am. J. Neuroradiol.*, 34(10):1866–1872, 2013.

[8] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. Synchronization in complex networks. *Phys. Rep.*, 469(3):93–153, 2008.

[9] B. A. Francis and M. Maggiore. Flocking and Rendezvous in Distributed Robotics. Springer, 2016.

[10] L. Glass. Synchronization and rhythmic processes in physiology. *Nature*, 410:277–284, 2001.

[11] D. Zuzik and A. Drago. Neuronal oscillations in cortical networks. *Science*, 304(5679):1926–1929, 2004.

[12] C. D. Smyser and J. S. Shimony. Resting-state fMRI: A review of methods and clinical applications. *Am. J. Neuroradiol.*, 34(10):1866–1872, 2013.

[13] M. Fox and M. Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.*, 8:700–711, 2007.

[14] M.-E. Lynall, D. S. Bassett, R. Kerwin, P. J. McKenna, M. Kitzbichler, U. Muller, and E. Bullmore. Functional connectivity and brain networks in schizophrenia. *J. Neurosci.*, 30(28):9477–9480, 2010.

[15] S. C. Ponten, F. Bartolomei, and J. C. Stam. Small-world networks and epilepsy: Graph theoretical analysis of intracerebrally recorded mesial temporal lobe seizures. *Clin. Neurophysiol.*, 118(4):918–927, 2007.

[16] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti. A framework to control functional connectivity in the human brain. In *Proc. IEEE Conf. on Decision and Control*, pages 4697–4704, 2019.

[17] S. F. Muldoon, F. Pasqualetti, S. Gu, M. Cieslak, S. T. Grafton, J. M. Vettel, and D. S. Bassett. Simulation-based control of dynamic brain networks. *PLoS Comput. Biol.*, 12(9):e1–23, 2016.

[18] S. F. Cogan. Neural stimulation and recording electrodes. *Annu. Rev. Biomed. Eng.*, 1(1):275–309, 2008.

[19] O. Yizhar, L. E. Fenno, T. J. Davidson, M. Mogri, and K. Deisseroth. Optogenetic in neural systems. *Neuron*, 71(1):9–34, 2011.

[20] E. Tang and D. S. Bassett. Colloquium: Control of dynamics in brain networks. *Rev. Mod. Phys.*, 90(3):031003, 2018.

[21] Y. Kuramoto. Chemical Oscillations, Waves, and Turbulence. Springer, 1984.

[22] A. Jadbabaie, N. Motee, and M. Barahona. On the stability of the Kuramoto model of coupled nonlinear oscillators. *In Proc. American Control Conf.*, pages 4296–4301, 2004.

[23] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti. Exact and approximate stability conditions for cluster synchronization of Kuramoto oscillators. In *Proc. American Control Conf.*, pages 205–210, 2019.

[24] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti. Stability conditions for cluster synchronization in networks of heterogeneous Kuramoto oscillators. *IEEE Trans. Control Netw. Syst.*, 7(1):302–314, 2020.

[25] Y. Qin, Y. Kawano, O. Portoles, and M. Cao. Partial phase cohesiveness in networks of Kuramoto oscillators. *IEEE Trans. Autom. Control*, 2021, to appear.

[26] A. Pogromsky, G. Santoboni, and H. Nijmeijer. Partial synchronization: From symmetry towards stability. *Physica D*, 172(1–4):65–87, 2002.

[27] I. Stewart, M. Golubitsky, and M. Pivato. Symmetry groupoids and patterns of synchrony in coupled cell networks. *SIAM J. Appl. Dyn. Syst.*, 2(4):609–646, 2003.

[28] V. N. Belykh, G. V. Osipov, V. S. Petrov, J. A. K. Suykens, and J. Vandewalle. Cluster synchronization in oscillatory networks. *Chaos*, 18(3):037106, 2008.

[29] W. Lu, B. Liu, and T. Chen. Cluster synchronization in networks of coupled nonidentical dynamical systems. *Chaos*, 20(1):013120, 2010.

[30] M. T. Schaub, N. O’Clery, Y. N. Billeh, J.-C. Delvenne, R. Lambiotte, and M. Barahona. Graph partitions and cluster synchronization in networks of oscillators. *IEEE Trans. Control Syst. Technol.*, 26(9):994821, 2016.

[31] L. V. Gambuzza and M. Frasca. A criterion for stability of cluster synchronization in networks with external equitable partitions. *Automatica*, 50(2):412–421, 2014.

[32] Z. Aminzare, B. Dey, E. N. Davison, and N. E. Leonard. Clustering of diffusively coupled nonlinear systems: A contraction-based approach. *J. Nonlinear Sci.*, 2018.

[33] W. Xia and M. Cao. Clustering in diffusively coupled networks. *Chaos*, 27(4):043114, 2017.

[34] W. Wu, W. Zhou, and T. Chen. Cluster synchronization of linearly coupled complex networks under pinning control. *IEEE Trans. Circuits Syst. I, Reg. Papers*, 56(4):829–839, 2009.

[35] C.-Q. Pham and J.-I. E. Slotine. Stable concurrent synchronization in dynamic system networks. *Neural Netw.*, 20(1):62–77, 2007.

[36] Z. Aminzare, B. Dey, E. N. Davison, and N. E. Leonard. Cluster synchronization of diffusively coupled nonlinear systems: A contraction-based approach. *J. Nonlinear Sci.*, 2018.

[37] L. V. Gambuzza and M. Frasca. A criterion for stability of cluster synchronization in networks with external equitable partitions. *Automatica*, 50(2):412–421, 2014.

[38] Z. Aminzare, B. Dey, E. N. Davison, and N. E. Leonard. Cluster synchronization of diffusively coupled nonlinear systems: A contraction-based approach. *J. Nonlinear Sci.*, 2018.

[39] W. Xia and M. Cao. Clustering in diffusively coupled networks. *Chaos*, 27(4):043114, 2017.

[40] W. Wu, W. Zhou, and T. Chen. Cluster synchronization of linearly coupled complex networks under pinning control. *IEEE Trans. Circuits Syst. I, Reg. Papers*, 56(4):829–839, 2009.
