FROM APPROXIMATE SYNCHRONIZATION TO IDENTICAL SYNCHRONIZATION IN COUPLED SYSTEMS

CHIH-WEN SHIH*

Department of Applied Mathematics
National Chiao Tung University
Hsinchu, Taiwan 300

JUI-PIN TSENG

Department of Mathematical Sciences
National Chengchi University
Taipei, Taiwan 116

(Communicated by Yuan Lou)

Abstract. We establish a framework to investigate approximate synchronization of coupled systems under general coupling schemes. The units comprising the coupled systems may be nonidentical and the coupling functions are nonlinear with delays. Both delay-dependent and delay-independent criteria for approximate synchronization are derived, based on an approach termed sequential contracting. It is explored and elucidated that the synchronization error, the distance between the asymptotic state and the synchronous set, decreases with decreasing difference between subsystems, difference between the row sums of connection matrix, and difference of coupling time delays between different units. This error vanishes when these factors decay to zero, and approximate synchronization becomes identical synchronization for the coupled system comprising identical subsystems and connection matrix with identical row sums, and with identical coupling delays. The application of the present theory to nonlinearly coupled heterogeneous FitzHugh-Nagumo neurons is illustrated. We extend the analysis to study approximate synchronization and asymptotic synchronization for coupled Lorenz systems and show that for some coupling schemes, the synchronization error decreases as the coupling strength increases, whereas in another case, the error remains at a substantial level for large coupling strength.

1. Introduction. Synchronization has appeared as a common coherent behavior in many biological and physical systems [29, 47, 50, 62]. Complex networks of coupled nodes, in which each node is a dynamical system, have been widely exploited to model many complex behaviors in sciences, engineering, and society, and have attracted much attention in recent decades, see [4, 5, 61] and the references therein. In addition, there is a large number of studies on synchronization of coupled neural networks because of its potential engineering applications such as secret communication [12, 57], pattern recognition [21], and parallel image processing [28].

2010 Mathematics Subject Classification. Primary: 34D06, 92B25; Secondary: 34D05, 34C15.

Key words and phrases. Approximate synchronization, identical synchronization, asymptotic synchronization, coupled system, asymptotic behavior.

The authors are supported in part by the Ministry of Science and Technology of Taiwan.

* Corresponding author: Chih-Wen Shih.
The collective behavior of a coupled system is determined by not only the dynamics of individual subsystems, but also the coupling structure between subsystems. A network of coupled systems can exhibit a variety of interesting behaviors that are qualitatively different from their isolated behaviors. In addition, coupling time delays or communication delays can modify the collective behaviors. Indeed, time delays have been incorporated into neural network modeling and artificial systems.

To investigate synchronization phenomena of coupled systems and complex networks, various models and scenarios have been studied, including complete or identical synchronization, phase synchronization, lag synchronization, partial synchronization, generalized synchronization, and almost synchronization. Among these works, identical synchronization of homogeneous systems consisting of identical units is the simplest form of synchronization, and has been studied extensively.

However, due to various possible deviations, natural imperfection, and wider consideration such as real-world complex networks, heterogeneous systems which consist of nonidentical units are more practical, see for instance and the references therein. For heterogeneous systems, it is natural to relax the notion of identical synchronization to more general sense. The description of synchronous behaviors for heterogeneous systems is more complicated, and the formulations and treatments for identical synchronization in homogeneous systems may no longer be valid. Recently, there were some papers focusing on global dynamics of synchronization for heterogeneous networks. Several notions centering around nonidentical synchronization were considered therein, according to the respective approach employed. They include the so-called approximate synchronization, quasi-synchronization, bounded synchronization, and practical synchronization. In [18], by introducing a virtual target trajectory, some conditions of quasi-synchronization were derived, and synchronization error was estimated for heterogeneous networks under linear diffusive coupling. In [66], bounded synchronization of some dynamical networks was established by introducing the average dynamics of all nodes. A criterion of global synchronization subject to bounded maximum state deviation between nodes was then derived. Bounded synchronization, in a sense different from the one in [66], for asymmetrical coupled networks was investigated in [59]. In [41, 42], an approach was developed to analyze approximate synchronization of a coupled system comprising two nonidentical systems of nonlinear differential equations, where a uniform estimate of the synchronization error with respect to the parameters of the systems was derived. Chaotic approximate synchronization between nonidentical master-slave systems was addressed in [20, 24, 52]. Applications of control schemes to studying approximate synchronization in heterogeneous networks were shown in [19, 25, 48]. The studies on practical synchronization of coupled nonidentical systems with linearly diffusive couplings can be found in [36, 38]. Frequency synchronization and phase synchronization of phase models, such as the Kuramoto model, have been investigated in [16, 22, 53].

To summarize, the equations considered in the studies of approximate synchronization originated from motivations in physics, networks, or engineering, largely carry linear, and/or diffusive (or akin to diffusive) coupling functions. This can be seen from the above-mentioned papers [18, 24, 41, 52, 59, 66] for the ones without delays, and [19, 20] for the ones with delays. On the other hand, depending on the methodologies employed, some of those investigations conclude local dynamics,
whereas others established global dynamics. In particular, the works which analyze the stability of synchronous manifold or near-synchronous manifold via linearization and the ones involving master stability function conclude local dynamics, see also [1, 2, 49, 51]. On the other hand, Lyapunov function theory and ideas developed from this theory are typical tools for concluding global synchronization. In addition, the criteria for synchronization in some of these papers, such as [59, 66], involve certain differential matrix inequalities, and a treatment is needed to examine the criteria.

Another direction of investigations on synchronization comes from biological perspectives, where systems of equations that are used to model the biological processes are highly nonlinear. For example, Hill-type functions are used to describe the repressions of transcriptions in gene regulation. Such functions appear in the coupling terms to describe interactions among cells or proteins in those models [27, 55]. In addition, transcriptional and translational time delays, or transmission delays are often taken into account in these models. This makes complicated the coupled systems modeling these biological processes, and new mathematical treatments are required, see [3, 10].

From the literature survey, it is tempting to develop a framework which can accommodate the basic content of approximate synchronization for nonlinearly coupled systems and heterogeneous complex networks originated form various motivations. And this framework should be able to cover general forms of coupling with delays. It is also important to know what factors contribute to synchronization error and to identify whether the cause of error is resulted from the differences between individual units or coupling structure. The goal of this work is to establish such a general framework.

In this paper, we shall adopt the term approximate synchronization to describe the common nonidentical synchronization of heterogeneous coupled systems and network systems with a bound on the synchronization error. Let us consider $N$ subsystems which may be nonidentical:

$$\dot{x}_i(t) = F_i(x_i^t, t), \quad i \in \mathcal{N} := \{1, \ldots, N\},$$

where the dynamics of the $i$th subsystem is governed by $F_i = (F_{i,1}, \ldots, F_{i,K})$ which is a continuously differentiable function of $x_i^t \in C([-\tau_M, 0]; \mathbb{R}^K)$, and $x_i^t(\theta) = x_i(t + \theta)$ for $\theta \in [-\tau_M, 0]$. Herein, we denote by $C([-\tau_M, 0]; \mathbb{R}^n)$ the set of continuous functions from $[-\tau_M, 0]$ to $\mathbb{R}^n$, for an $n \in \mathbb{N}$, and consider time delays in the range $[0, \tau_M]$, $\tau_M \geq 0$. A general coupled network system comprising these $N$ subsystems can be described as

$$\dot{x}_i(t) = F_i(x_i^t, t) + \sum_{j \in \mathcal{N}} \omega_{ij}(t)G_{ij}(x_i^t, x_j^t, t), \quad t \geq t_0,$$

where $x_i(t) = (x_{i,1}(t), \ldots, x_{i,K}(t)) \in \mathbb{R}^K$. Herein, coupling functions $G_{ij}$ are smooth functions, each $\omega_{ij}(t)$ is a bounded smooth function of $t$, which represents the time-dependent connection weight from node $j$ to node $i$, and $W(t) := [\omega_{ij}(t)]_{N \times N}$ is the connection matrix. We denote by $(x_1^0, \ldots, x_N^0)$ the evolution of system (2) at time $t$ from the initial condition $(x_1^0, \ldots, x_N^0)$ in $C([-\tau_M, 0]; \mathbb{R}^{KN})$ at $t = t_0$, and $(x_1(t), \ldots, x_N(t))$ is the corresponding solution of system (2). System (2) reduces to ordinary differential equations when $\tau_M = 0$.

While our approach can treat system (2), the formulation is very cumbersome. So in this paper, we illustrate our approach by considering system (2) with $F_i(x_i^t, t) =$
\( F_i(x_i(t), t), \omega_{ij}(t) \equiv \omega_{ij} \), and function \( G_{ij} = (G_{ij,1}, \ldots, G_{ij,K}) \) of the following form:

\[
G_{ij,k}(x_i^t, x_j^t, t) = G_k(x_j(t - \tau_{ij})),
\]

for \( k \in K := \{1, \ldots, K\} \), where \( G_k : \mathbb{R} \to \mathbb{R} \) are continuously differentiable and nondecreasing functions, and \( \tau_{ij} \in [0, \tau_M] \). In this situation, the component form (the \( k \)th component in the \( i \)th subsystem) of system (2) becomes

\[
\dot{x}_{i,k}(t) = F_{i,k}(x_i(t), t) + \sum_{j \in \mathcal{N}} \omega_{ij} G_k(x_{j,k}(t - \tau_{ij})), \quad t \geq t_0, \ i \in \mathcal{N}.
\]

System in the form of (4) is still very general, and is commonly considered in the studies of synchronization problems in the literature. For example, the systems that were studied in [18, 30, 33, 36, 41, 54, 58, 59, 63, 65, 66] are in the form of (4), or similar to (4).

Usually, network system (2) or (4) is called homogeneous if it comprises identical subsystems, i.e., \( F_i \equiv F_j \), for all \( i, j \in \mathcal{N} \), and called heterogeneous otherwise. It is important to note that for homogeneous systems, the synchronous set \( \mathcal{S} := \{(\Phi, \ldots, \Phi) \in C([-\tau_M, 0]; \mathbb{R}^{NK}) : \Phi \in C([-\tau_M, 0]; \mathbb{R}^K)\} \) is positively invariant under the evolution of system (4), provided

\[
\delta_r := \max\{|\tau_{ij} - \tau_{rs}| : i, j, r, s \in \mathcal{N}\} = 0,
\]

and

\[
\delta_k := \max\{|\kappa_i - \kappa_j| : i, j \in \mathcal{N}\} = 0,
\]

where \( \kappa_i \) is the \( i \)th row sum of the connection matrix:

\[
\kappa_i := \sum_{j \in \mathcal{N}} \omega_{ij}, \ i \in \mathcal{N}.
\]

Herein, \( \delta_r \) measures the difference of coupling delays between different components, and \( \delta_k \) measures the difference between row sums of connection matrix. Notably, \( \delta_r = 0 \) if system (4) has identical coupling delays or does not admit coupling delays; \( \delta_k = 0 \) if the connection matrix has identical row sums. For homogeneous systems without coupling delays or with identical coupling delay, there have been a large amount of investigations on global identical synchronization, especially in the diffusive coupling case: \( \kappa_i = 0 \) for all \( i \in \mathcal{N} \), see [30, 33, 45, 58, 63, 65] and the references therein. Herein, system (4) is said to attain global identical synchronization if

\[
x_{i,k}(t) - x_{j,k}(t) \to 0 \text{ as } t \to \infty, \text{ for all } i, j \in \mathcal{N}, \ k \in K,
\]

for every solution \((x_1(t), \ldots, x_N(t))\) of system (4). There were also extensive studies on local synchronization, where the stability of synchronous set \( \mathcal{S} \) is discussed, see [23, 33, 39, 58]. Basically, the invariance of synchronous set under the flow generated by system (4) is a prerequisite for identical synchronization of (4). However, for a heterogeneous system (4) or a homogeneous system (4) without condition (5), or without (6), the synchronous set is not invariant in general, and thus the realization of identical synchronization is impractical. We note that for coupled systems with mismatched parameters, some studies on the stability of near-synchronous states have been reported in [1, 2, 49, 51]. Another interesting point is that the asymptotic collective dynamics for (4) is distinct from the ones for individual subsystems when isolated, if some \( \kappa_i \) is nonzero. This leads to the fact that there can be more varieties for the collective behaviors of non-diffusively coupled systems than the ones of diffusively coupled systems.
Let $\delta F_{ij} := F_i - F_j$, and
\[
\|\delta F_{ij}\| = \|F_i - F_j\| := \sup\{\|F_i(\Xi, t) - F_j(\Xi, t)\| : \Xi \in \mathbb{R}^K, t \geq t_0\},
\]
where $\|\cdot\|$ is certain norm on $\mathbb{R}^K$. Then $\delta F := \max\{|\delta F_{ij}| : i, j \in \mathcal{N}\}$ measures the maximal difference between subsystems. In this paper, the coupled system (4) is said to attain global ($\varepsilon$-)approximate synchronization if there exists an $\varepsilon \geq 0$ such that
\[
\limsup_{t \to \infty} \|x_i(t) - x_{i+1}(t)\| \leq \varepsilon, \quad \text{for all } i \in \mathcal{N} - \{\mathcal{N}\},
\]
for every solution $(x_1(t), \ldots, x_N(t))$ of system (4) (i.e., in the sense of global dynamics). The expression in (10) indicates an estimation of the synchronization error, and $\varepsilon$ herein stands for a bound of the error. Moreover, $\varepsilon = 0$ corresponds to identical synchronization. Certainly, this definition with (10) will be meaningful in the sense of approximate synchronization if $\varepsilon$ decreases with decreasing inhomogeneity of the system, i.e., $\varepsilon$ decreases as $\delta F$ decreases, and $\varepsilon \to 0$ as $\delta F \to 0$. We shall show that this is indeed the case if the row sums of connection matrix and coupling delays are identical, i.e., (5) and (6) hold. When the row sums of connection matrix or coupling delays are not identical, network connection may also contribute to the error $\varepsilon$. This is quite natural, but has not been elucidated mathematically, likely because the existing approaches are insufficient for such a task. It will be revealed that $\varepsilon$ decreases as each of $\delta F$, $\delta x$, and $\delta \tau$ decreases, and $\varepsilon \to 0$ as these factors decay to zero. Our results are much more thorough than the related studies in the literature. For instance, in [20, 24, 25], a bound of synchronization error similar to (10) was formulated for master–slave systems with parameter mismatches in subsystems. The estimated bound was shown to depend on the magnitude of parameter mismatches. The present approach can also establish approximate synchronization if (10) is replaced by $\limsup_{t \to \infty} \|x_i(t) - x_j(t)\| \leq \varepsilon$ for all $i, j \in \mathcal{N}$, cf. [59], or $\limsup_{t \to \infty} \|x_i(t) - s(t)\| \leq \varepsilon$ for all $i \in \mathcal{N}$, where $s(t) := \frac{1}{\mathcal{N}} \sum_{j \in \mathcal{N}} x_j(t)$, cf. [66].

If there is a coupling strength $c$ considered in the coupling term, say if the summation term in (4) is modified to
\[
c \sum_{j \in \mathcal{N}} \omega_{ij} G_k(x_{j,k}(t - \tau_{ij}))
\]
one may also discuss asymptotic synchronization. System (4) is said to attain asymptotic synchronization if it attains $\varepsilon$-approximate synchronization and $\varepsilon \to 0$ as $c \to \infty$. With the invariant manifold theory, such asymptotic synchronization has been discussed by Hale [17]. However, it is not true that approximate synchronization always yields asymptotic synchronization. With the present framework of approximate synchronization, we are able to explore this interesting issue.

When system (4) is homogeneous, i.e., $F_i = F_j$ for all $i, j \in \mathcal{N}$, and (5) and (6) hold, the formulation in this investigation reduces to the one for identical synchronization. Recently, an iteration argument named “sequential contracting” has been applied to study identical synchronization to conclude global dynamics in [45, 46, 54]. More precisely, the framework for identical synchronization of coupled systems in a general form was established in [45], via the sequential contracting approach. When considering a coupled system with connection in additive form similar to (4), the condition of circulant coupling in [45] was lifted in [54]. This was completed by performing a suitable transformation on the coupling terms in the difference-differential equations associated with the coupled system. In this paper,
to take into account distinct subsystems, nonidentical row sums of connection matrix, and variant coupling delays between different components, a new formulation on the associated difference-differential equations is developed. In particular, we are able to derive an efficacious decomposition on the difference-differential equations, which allows us to extract and collect the above-mentioned inhomogeneous factors from the decomposition. Combining with such a formulation, we then perform a more general transformation on the coupling terms in the difference-differential equations.

Actually, in the somehow limited publications on approximate synchronization of heterogeneous system (4), the couplings are commonly linear and diffusive (i.e., \( G_k \) is linear for all \( k \) and \( \kappa_i = 0 \) for all \( i \)) and without coupling delays (i.e., \( \tau_M = 0 \)), cf. [18, 36, 41, 59, 66]. To the best of our knowledge, there has not been an analytical result on approximate synchronization for a heterogeneous system (4), with nonlinear and delayed coupling, in the sense of (10) or similar to (10). We remark that basically, a criterion is needed to attain synchronization, even for a coupled system comprising identical subsystems. Thus, deriving a criterion for approximate synchronization is also the key objective in our framework, in addition to indicating how synchronization error depends on the difference between subsystems and on the coupling structure.

Dissipativeness which implies the existence of a globally attracting set, is a basic property often imposed on the systems considered for synchronization and other asymptotic behaviors [17, 30, 45, 54]. It was utilized directly or indirectly in various methodologies for synchronization theory [13, 18, 19, 20, 24, 48, 52]. In those papers, the synchronization criteria involve the ranges or bounds of solutions of the systems, and the dissipative property are actually utilized, although this might not have been spelled out. In fact, dissipative property for coupled systems is a highly nontrivial issue, and was a chief concern in [17, 41, 42]. The idea of sequential contracting [45, 54] is to start with a preliminary estimation on the attracting region, and then make use of the dissipative property inside the coupled systems via an iterative argument. We therefore make the following assumption:

Assumption (D): All solutions of system (4) eventually enter, and then remain in some compact set \( Q^N := Q \times \ldots \times Q \subset \mathbb{R}^{NK} \), where

\[
Q := [\hat{q}_1, \hat{q}_1] \times \ldots \times [\hat{q}_K, \hat{q}_K] \subset \mathbb{R}^K.
\]

Under assumption (D), every solution \((x_1(t), \ldots, x_N(t))\) of system (4) exists on \([t_0, \infty)\), where \(x_i(t) = (x_{i,1}(t), \ldots, x_{i,K}(t)), i \in \mathcal{N}\).

The remainder of this paper is organized as follows. In Section 2, we present the sequential-contracting framework to derive both delay-dependent and delay-independent criteria for approximate synchronization of system (4). In Section 3, with the criteria derived in Section 2, we study the approximate synchronization of nonlinearly and nondiffusively coupled FitzHugh-Nagumo neurons. We discuss approximate synchronization and asymptotic synchronization for coupled Lorenz equations in Section 4. This paper then ends with a conclusion. Some technical parts are arranged in the Appendices.

2. Approximate synchronization via sequential contracting. Our approach to establishing the approximate synchronization of system (4) relies on a new setting of attracting set, an efficacious decomposition of the associated difference-differential system, and the sequential contracting technique.
Recall that the evolution \((x^t_1, \ldots, x^t_N)\) of (4) is an element in \(C([-\tau_M, 0]; \mathbb{R}^{NK})\), and \((x_i(t), \ldots, x_N(t))\) is the corresponding solution of (4), where \(x_i(t) = (x_{i,1}(t), \ldots, x_{i,K}(t)) \in \mathbb{R}^K\). That is, \(x^t_i = (x_{i,1}^t, \ldots, x_{i,K}^t) \in C([-\tau_M, 0]; \mathbb{R}^K)\) and satisfies \(x_i(t + \theta) = x_i^t(\theta)\) for \(\theta \in [-\tau_M, 0]\) and \(t \geq t_0\), where each \(x_{i,k}^t\) is a function in \(C([-\tau_M, 0]; \mathbb{R})\), which satisfies
\[
x_{i,k}^t(\theta) = x_{i,k}(t + \theta)
\]
for \(\theta \in [-\tau_M, 0]\). Let us denote by \((x_{i,k}^t)'\) the derivative of function \(x_{i,k}^t\). It follows from (11) that
\[
(x_{i,k}^t)'(\theta) = (x_{i,k})'(t + \theta) = \dot{x}_{i,k}(t + \theta),
\]
and thus \((x_{i,k})'\) and \(\dot{x}_{i,k}\) both represent the derivative of function \(x_{i,k}\). Notably, \(\tau_M \geq 0\) is the upper bound of delay magnitude, and \(\tau_M = 0\) reduces to the ODE case. In the following expressions, we take \((\Phi_1, \ldots, \Phi_N)\) with \(\Phi_i = (\phi_{i,1}, \ldots, \phi_{i,K}) \in C([-\tau_M, 0]; \mathbb{R}^K)\), \(i \in N\), as a general element in \(C([-\tau_M, 0]; \mathbb{R}^{NK})\), and it indicates the evolution of (4) when \((\Phi_1, \ldots, \Phi_N) = (x^t_1, \ldots, x^t_N)\), i.e., \(\phi_{i,k}(\theta) = x_{i,k}(t + \theta)\) for \(\theta \in [-\tau_M, 0]\). For instance, \(\phi_{i,k}(0)\) and \(\phi_{i,k}(-\tau)\) correspond to \(x_{i,k}(t)\) and \(x_{i,k}(t - \tau)\), respectively, when they represent the components of a solution of (4). In addition, we take \(\Xi = (\xi_1, \ldots, \xi_K)\) as a general element in \(\mathbb{R}^K\).

Via \(Q\) and \([\tilde{q}_k, \hat{q}_k]\) defined in assumption (D) and the connection matrix \(W = [\omega_{ij}]_{N \times N}\), we define the following quantities successively:
\[
\|F_k\| := \max_{i \in N} \{\sup_{t \geq t_0} \{\|F_{i,k}(\Xi, t)\| : \Xi \in Q, t \geq t_0\}\},
\]
\[
\tilde{G}_k := \sup \{\|G_k(\xi)\| : \xi \in [\tilde{q}_k, \hat{q}_k]\},
\]
\[
\hat{D}_k := \inf \{G_k(\xi) : \xi \in [\tilde{q}_k, \hat{q}_k], \|\xi\| \geq 0\},
\]
\[
\hat{D}_k := \sup \{G_k(\xi) : \xi \in [\tilde{q}_k, \hat{q}_k], \|\xi\| \geq 0\},
\]
\[
\Lambda_k := \|F_k\| + \tilde{G}_k \cdot (\max_{i \in N} \sum_{j \in N} |\omega_{ij}|).
\]

We further define the following subset of \(C([-\tau_M, 0]; \mathbb{R}^{NK})\):
\[
C^N_Q := \{(\Phi_1, \ldots, \Phi_N) : \Phi_i = (\phi_{i,1}, \ldots, \phi_{i,K}) \in C_Q, \text{ for } i \in N\},
\]
where
\[
C_Q := \{\Phi = (\phi_1, \ldots, \phi_K) \in C([-\tau_M, 0]; \mathbb{R}^K) : \phi_k(\theta) \in [\tilde{q}_k, \hat{q}_k] \text{ for } \theta \in [-\tau_M, 0], \text{ and } |\phi_k'(\theta)| \leq \Lambda_k \text{ for } \theta \in (-\tau_M, 0) \text{ and } k \in K\}.
\]

**Remark 1.** Assumption (D) was also imposed to study the identical synchronization of homogeneous coupled systems in [45, 54]. Therein, a notion similar to \(C_Q\) was adopted, and can be described by
\[
\{(\phi_1, \ldots, \phi_K) \in C([-\tau_M, 0]; \mathbb{R}^K) : \phi_k(\theta) \in [\tilde{q}_k, \hat{q}_k] \text{ for } \theta \in [-\tau_M, 0] \text{ and } k \in K\}.
\]

Compared to the set in (20), \(C_Q\) in (19) further considers boundedness of \(\phi_k'\): \(|\phi_k'(\theta)| \leq \Lambda_k\), where \(\Lambda_k\) is determined by not only set \(Q\) but also the individual subsystems \(F_{i,k}\), connection matrix \([\omega_{ij}]\), and coupling function \(G_k\), as shown in (13)–(17). Such a further consideration of \(\phi_k'\) in \(C_Q\) is to measure the synchronization error caused by the heterogeneous coupling delays, see Remark 3 below.
Lemma 2.1. Under assumption (D), starting from any initial point in $\mathcal{C}([-\tau_M, 0]; \mathbb{R}^{NK})$, the evolution $(x^1, \ldots, x^N)$ eventually enters and then remains in $\mathcal{C}_Q^N$. More precisely, there exists a $t_0 + \tau_M$ so that $x^i(t) = (x^i_{1,1}, \ldots, x^i_{1,K}) \in \mathcal{C}_Q$, i.e.,

$$x^i_{1,k}(\theta) \in [\hat{q}_k, \check{q}_k] \text{ if } \theta \in [-\tau_M, 0], \text{ and } |(x^i_{1,k})'(\theta)| \leq \Lambda_k \text{ if } \theta \in (-\tau_M, 0),$$

for all $t \geq \tilde{t}$, $i \in \mathcal{N}$, and $k \in \mathcal{K}$.

Proof. By assumption (D), there exists a $\tilde{t} \geq t_0 + \tau_M$ so that $x^i(t) \in \mathcal{Q}$, for all $t \geq \tilde{t} - \tau_M$ and $i \in \mathcal{N}$; hence, $x^i_{1,k}(\theta) = x^i_{1,k}(t + \theta) \in [\hat{q}_k, \check{q}_k]$ if $\theta \in [-\tau_M, 0]$, for all $t \geq \tilde{t}$, $i \in \mathcal{N}$, and $k \in \mathcal{K}$. Accordingly, as seen from system (4), and (12)–(14), and (17),

$$|x^i_{1,k}(t + \theta)| \leq |\delta F_k| + \sum_{j \in \mathcal{N}} |\omega_{ij}| \tilde{G}_k \leq \Lambda_k,$$

for $t \geq \tilde{t}$, $\theta \in (-\tau_M, 0)$, $i \in \mathcal{N}$, and $k \in \mathcal{K}$. This completes the proof. \hfill \Box

For later use, we denote the index set:

$$\mathcal{A} := \{(i, k) : i \in \mathcal{N} - \{N\}, k \in \mathcal{K}\}. \tag{21}$$

The $i$th individual subsystem of (4) is defined by $F_i = (F_i, \ldots, F_{i,K})$. For each $(i, k) \in \mathcal{A}$, we set

$$\delta F_{i,k}(\Xi, t) := F_{i,k}(\Xi, t) - F_{i+1,k}(\Xi, t), \tag{22}$$

for $\Xi \in \mathbb{R}^K$. Such $(\delta F_{i,1}, \ldots, \delta F_{i,K})$ depicts the difference between two adjacent subsystems.

Let $(x_1(t), \ldots, x_N(t))$ be an arbitrary solution of system (4), with $x_i(t) = (x_{i,1}(t), \ldots, x_{i,K}(t))$, $i \in \mathcal{N}$. Set $z_i(t) = x_i(t) - x_{i+1}(t)$, where $z_i(t) = (z_{i,1}(t), \ldots, z_{i,K}(t))$, for $i \in \mathcal{N} - \{N\}$. From (4), with (22), we see that $(z_1(t), \ldots, z_{N-1}(t))$ satisfies the following difference-differential system associated with system (4):

$$\dot{z}_{i,k}(t) = H_{i,k}(x^i_{1,1}, \ldots, x^i_{1,N}, t), \quad (i, k) \in \mathcal{A}, \quad t \geq t_0, \tag{23}$$

where $z_{i,k}(t) = x_{i,k}(t) - x_{i+1,k}(t)$, and

$$H_{i,k}(x^i_{1,1}, \ldots, x^i_{1,N}, t) := F_{i,k}(x_i(t), t) - F_{i,k}(x_{i+1}(t), t) + \delta F_{i,k}(x_{i+1}(t), t) + \tilde{G}_{i,k}(x^i_{1,1}, \ldots, x^i_{1,N}), \tag{24}$$

with

$$\tilde{G}_{i,k}(x^i_{1,1}, \ldots, x^i_{1,N}) := \sum_{j \in \mathcal{N}} [\omega_{ij} G_k(x_{j,k}(t - \tau_{ij})) - \omega_{(i+1)j} G_k(x_{j,k}(t - \tau_{(i+1)j}))]. \tag{25}$$

We shall establish the $\varepsilon$-approximate synchronization of system (4) by showing

$$\limsup_{t \to \infty} \|z_i(t)\| \leq \varepsilon \text{ for all } i \in \mathcal{N} - \{N\},$$

for some $\varepsilon \geq 0$, under some conditions. The error $\varepsilon$ off the synchronous set $\mathcal{S}$ naturally depends on the difference of units, represented by $\delta F_{i,k}$, and the structure of network which is imbedded in $\tilde{G}_{i,k}$. For the factors from the network structure, the row sums $\kappa_i$ of the connection matrix, defined in (7), and coupling delays $\tau_{ij}$, defined in (3), play important roles. Sequential contracting is an iterative argument which can be used to conclude the asymptotic behavior of $z_{i,k}(t)$ satisfying (23). To analyze the difference-differential equation (23), we shall recompose the terms in $H_{i,k}$ so that (23) can be rewritten as

$$\dot{z}_{i,k}(t) = h_{i,k}(x_{i,k}(t), x_{i+1,k}(t), t) + \tilde{h}_{i,k}(x_{i,k}(t - \tau_{ii}), x_{i+1,k}(t - \tau_{i,i})) + w_{i,k}(t), \tag{26}$$
for \((i,k) \in \mathcal{A}\) and \(t \geq t_0\), where
\[
 w_{i,k}(t) := \tilde{w}_{i,k}(x_1^t, \ldots, x_N^t, t) + e_{i,k}(x_1^t, \ldots, x_N^t, t),
\]
and
\[
\begin{align*}
\tilde{w}_{i,k}(x_1^t, \ldots, x_N^t, t) &= u_{i,k}(x_i(t), x_{i+1}(t), t) + v_{i,k}(x_1^t, \ldots, x_N^t), \\
e_{i,k}(x_1^t, \ldots, x_N^t, t) &= \delta F_{i,k}(x_{i+1}(t), t) + \sigma_{i,k}(x_1^t) + \alpha_{i,k}(x_1^t, \ldots, x_N^t).
\end{align*}
\]

The purpose of recasting system (23) into system (26) is to exploit and utilize the dissipative structure of the coupled systems and the coupling structure contained in \(\tilde{G}_{i,k}\), so that a scheme can be formulated for the attraction of \(z_{i,k}(t)\) satisfying (23) toward a neighborhood of zero, as \(t\) gets large. The dissipative property of coupled systems is certainly associated with the characters imbedded in subsystems depicted by \(F_{i,k}\) and the coupling terms. In particular, we collect the terms involving \(x_{i,k}(t)\) in \(F_{i,k}\) and \(\tilde{G}_{i,k}\), and the terms involving \(x_{i+1,k}(t)\) in \(F_{i+1,k}\) and \(\tilde{G}_{i,k}\) into a function \(h_{i,k}\); we also collect terms involving \(x_{i,k}(t - \tau_{i})\) and \(x_{i+1,k}(t - \tau_{i})\) in \(\tilde{G}_{i,k}\) to form function \(\tilde{h}_{i,k}\). For example, if there is a decay term \(-\mu_k \tilde{e}_{i,k}(t)\) in \(F_{i,k}\), and an analogous \(-\mu_k \tilde{e}_{i+1,k}(t)\) in \(F_{i+1,k}\), then these terms generate in \(h_{i,k}\)
\[
-\mu_k [x_{i,k}(t) - x_{i+1,k}(t)] = -\mu_k z_{i,k}(t)
\]
which contributes to a contraction of \(z_{i,k}(t)\) towards zero, if \(\mu_k > 0\). In addition, \(e_{i,k}\) contains \(\delta F_{i,k}\) which depicts the difference between the \(i\)th and the \((i+1)\)th subsystems, \(\sigma_{i,k}\) which is associated with the difference between row sums of connection matrix \(W\), and \(\alpha_{i,k}\) which is associated with the inhomogeneity of coupling delays, cf. Remark 2. Moreover, \(\tilde{w}_{i,k}\) collects all the remaining terms.

Now, let us present how \(h_{i,k}, \tilde{h}_{i,k}, \tilde{w}_{i,k}, \sigma_{i,k}, \alpha_{i,k}\) are derived, and the process to arrive (26) from (23). For the first two terms in (24), we consider the following decomposition
\[
F_{i,k}(\Xi, t) - F_{i,k}(\bar{\Xi}, t) = f_{i,k}(\xi_k, \bar{\xi}_k, t) + u_{i,k}(\Xi, \bar{\Xi}, t),
\]
for \(\Xi = (\xi_1, \ldots, \xi_K), \bar{\Xi} = (\bar{\xi}_1, \ldots, \bar{\xi}_K) \in \mathbb{R}^K, t \geq t_0,\) and \((i,k) \in \mathcal{A}\). Such a decomposition aims to collect in \(f_{i,k}\) certain terms involving \(\xi_k, \bar{\xi}_k\), and leave the other terms in \(u_{i,k}\). Conditions on \(f_{i,k}\) and \(u_{i,k}\) will be imposed in assumption (F) below. We will illustrate such decomposition in Sections 3 and 4.

For \(\tilde{G}_{i,k}\) in (25), we perform the following manipulation. We compose a matrix whose entries consist of coupling weights \(\omega_{ij}\) and row sums \(\kappa_i\) of the connection matrix:
\[
\hat{W} = [\hat{\omega}_{ij}]_{1 \leq i,j \leq N} \in \mathbb{R}^{N \times N},
\]
where
\[
\hat{\omega}_{ij} := \begin{cases} 
\omega_{ii} - \kappa_i & \text{if } i = j, \\
\omega_{ij} & \text{otherwise.}
\end{cases}
\]
From \(\hat{W}\), we further define matrix \(\tilde{W}\):
\[
\tilde{W} = [\tilde{\omega}_{ij}]_{1 \leq i,j \leq N-1} := C \hat{W} \hat{C}^T (\hat{C} \hat{C}^T)^{-1} \in \mathbb{R}^{(N-1) \times (N-1)},
\]
for \(i = 1, \ldots, N-1\), and leave the other terms in \(u_{i,k}\). Conditions on \(f_{i,k}\) and \(u_{i,k}\) will be imposed in assumption (F) below. We will illustrate such decomposition in Sections 3 and 4.
where $T$ means transpose, and $C$ is an $(N - 1) \times N$ matrix defined by

$$
C := \begin{pmatrix}
1 & -1 & 0 & \cdots & 0 \\
0 & 1 & -1 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & 1 & -1
\end{pmatrix},
$$

Then $\hat{W}$ is well defined and satisfies

$$
\hat{W} = \hat{W} C,
$$

(30)

cf. [35]. In addition, it is shown in Appendix A that

$$
\bar{\omega}_{ii} = -\left( \sum_{j=i+1}^{N} \bar{\omega}_{ij} + \sum_{j=1}^{i} \bar{\omega}_{(i+1)j} \right), \text{ for } i = 1, \ldots, N - 1.
$$

(31)

With (30), we shall decompose $\tilde{G}_{i,k}$ in (25) into four parts in the following lemma. Recall the notation $(\Phi_1, \ldots, \Phi_N), \Phi_j = (\phi_j, 1, \ldots, \phi_j, k) \in C([-\tau_M, 0]; \mathbb{R}^K), j \in N$, introduced in the beginning of this section.

**Lemma 2.2.** For $(i, k) \in A$, $\tilde{G}_{i,k}$ can be decomposed as

$$
\tilde{G}_{i,k}(\Phi_1, \ldots, \Phi_N) = \tilde{g}_{i,k}(\phi_i, k(-\tau_{i}), \phi_{i+1, k}(-\tau_{i})) + v_{i,k}(\Phi_1, \ldots, \Phi_N)
$$

$$
+ \sigma_{i,k}(\Phi_{i+1}) + o_{i,k}(\Phi_1, \ldots, \Phi_N),
$$

where

$$
\tilde{g}_{i,k}(\phi_i, k(-\tau_{i}), \phi_{i+1, k}(-\tau_{i})) := (\kappa_i + \bar{\omega}_{ii})[G_k(\phi_i, k(-\tau_{i})) - G_k(\phi_{i+1, k}(-\tau_{i}))],
$$

$$
v_{i,k}(\Phi_1, \ldots, \Phi_N) := \sum_{j \in N - \{i, N\}} \omega_{ij} [G_k(\phi_j, k(-\tau_{ij})) - G_k(\phi_{j+1, k}(-\tau_{ij}))],
$$

$$
\sigma_{i,k}(\Phi_{i+1}) := (\kappa_i - \kappa_{i+1})[G_k(\phi_{i+1, k}(-\tau_{(i+1)}(i+1)))],
$$

$$
o_{i,k}(\Phi_1, \ldots, \Phi_N) := \kappa_i[G_k(\phi_{i+1, k}(-\tau_{i})) - G_k(\phi_{i+1, k}(-\tau_{(i+1)}))] + \kappa_j[G_k(\phi_{j+1, k}(-\tau_{(j+1)})) - G_k(\phi_{j+1, k}(-\tau_{(j+1)}(i+1)))]
$$

$$
+ \sum_{j \in N - \{i, N\}} \omega_{ij} [G_k(\phi_{j+1, k}(-\tau_{ij})) - G_k(\phi_{j+1, k}(-\tau_{(i+1)})]
$$

$$
+ \sum_{j \in N} \omega_{(i+1)j} [G_k(\phi_{j, k}(-\tau_{ij})) - G_k(\phi_{j, k}(-\tau_{(i+1)}))] = 0
$$

and $\kappa_i$, $\bar{\omega}_{ij}$, and $\omega_{ij}$ are defined in (7), (28), and (29), respectively. Moreover, $o_{i,k}(\Phi_1, \ldots, \Phi_N) = 0$ under condition (5), and $\sigma_{i,k}(\Phi_{i+1}) = 0$ under condition (6).

**Proof.** From (30), we have $C\hat{W} \Xi = \hat{W} C \Xi$, i.e.,

$$
\sum_{j \in N} [\bar{\omega}_{ij} - \bar{\omega}_{(i+1)j}] \xi_j = \sum_{j \in N - \{N\}} [\bar{\omega}_{ij} (\xi_j - \xi_{j+1})],
$$

(32)

for all $i \in N - \{N\}$ and $\Xi = (\xi_1, \ldots, \xi_N)^T \in \mathbb{R}^N$. From (25) and the definition of $\bar{\omega}_{ij}$ in (28), we obtain

$$
\tilde{G}_{i,k}(\Phi_1, \ldots, \Phi_N) = \kappa_i G_k(\phi_{i, k}(-\tau_{i})) - \kappa_{i+1} G_k(\phi_{i+1, k}(-\tau_{(i+1)i}))
$$

$$
+ \sum_{j \in N} [\bar{\omega}_{ij} G_k(\phi_{j, k}(-\tau_{ij})) - \bar{\omega}_{(i+1)j} G_k(\phi_{j, k}(-\tau_{(i+1)j}))].
$$
From this expression, we further decompose
\[ \kappa_i G_k(\phi_{i,k}(-\tau^{i})) - \kappa_{i+1} G_k(\phi_{i+1,k}(-\tau^{i+1})) \]
\[ = \kappa_i[G_k(\phi_{i,k}(-\tau^{i})) - G_k(\phi_{i+1,k}(-\tau^{i}))] + \kappa_i[G_k(\phi_{i+1,k}(-\tau^{i})) - G_k(\phi_{i+1,k}(-\tau^{i+1}))] \]
\[ - G_k(\phi_{i+1,k}(-\tau^{i+1})) + \kappa_i[G_k(\phi_{i+1,k}(-\tau^{i+1})) - G_k(\phi_{i+1,k}(-\tau^{i+1+1}))] \]
\[ + (\kappa_i - \kappa_{i+1}) G_k(\phi_{i+1,k}(-\tau^{i+1+1})), \]
and
\[ \sum_{j \in N} [\tilde{\omega}_{ij} G_k(\phi_{j,k}(-\tau^{ij})) - \tilde{\omega}_{(i+1)j} G_k(\phi_{j,k}(-\tau^{(i+1)j}))] \]
\[ = \sum_{j \in N} [\tilde{\omega}_{ij} - \tilde{\omega}_{(i+1)j}] G_k(\phi_{j,k}(-\tau^{ij})) \]
\[ + \sum_{j \in N} \tilde{\omega}_{(i+1)j} [G_k(\phi_{j,k}(-\tau^{ij})) - G_k(\phi_{j,k}(-\tau^{(i+1)j}))] \]
\[ = \sum_{j \in N - \{N\}} \tilde{\omega}_{ij} [G_k(\phi_{j,k}(-\tau^{ij})) - G_k(\phi_{j+1,k}(-\tau^{i+1j}))] \]
\[ + \sum_{j \in N - \{N\}} \tilde{\omega}_{(i+1)j} [G_k(\phi_{j,k}(-\tau^{ij})) - G_k(\phi_{j,k}(-\tau^{(i+1)j}))] \]
\[ = \sum_{j \in N - \{N\}} \tilde{\omega}_{ij} [G_k(\phi_{j,k}(-\tau^{ij})) - G_k(\phi_{j+1,k}(-\tau^{i+1j}))] \]
\[ + \sum_{j \in N - \{N\}} \tilde{\omega}_{(i+1)j} [G_k(\phi_{j,k}(-\tau^{ij})) - G_k(\phi_{j,k}(-\tau^{(i+1)j}))], \tag{33} \]
where the equality (33) follows from (32). We thus obtain the decomposition of \( \tilde{G}_i,k \) in the assertion. \( \square \)

Now, let us go back to consider \( H_{i,k} \) defined in (24). Recall \( \delta F_{i,k} \), \( f_{i,k} \), and \( u_{i,k} \) defined in (22), (27), and Lemma 2.2. By (22), (27), and Lemma 2.2, system (23) can be rewritten as system (26):
\[ \dot{z}_{i,k}(t) = h_{i,k}(x_{i,k}(t), x_{i+1,k}(t), t) + \tilde{h}_{i,k}(x_{i,k}(t - \tau^{i}), x_{i+1,k}(t - \tau^{i})) + u_{i,k}(t), \tag{34} \]
for \( (i, k) \in \mathcal{A} \) and \( t \geq t_0 \), where
\[ w_{i,k}(t) := \tilde{w}_{i,k}(x_{1}^{i}, \ldots, x_{N}^{i}, t) + e_{i,k}(x_{1}^{i}, \ldots, x_{N}^{i}, t), \tag{35} \]
and
\[ h_{i,k}(x_{i,k}(t), x_{i+1,k}(t), t) := f_{i,k}(x_{i,k}(t), x_{i+1,k}(t), t), \tag{36} \]
\[ \tilde{h}_{i,k}(x_{i,k}(t - \tau^{i}), x_{i+1,k}(t - \tau^{i})) := \tilde{f}_{i,k}(x_{i,k}(t - \tau^{i}), x_{i+1,k}(t - \tau^{i})), \tag{37} \]
\[ w_{i,k}(x_{1}^{i}, \ldots, x_{N}^{i}, t) := u_{i,k}(x_{1}^{i}(t), x_{i+1}(t), t) \]
\[ + v_{i,k}(x_{1}^{i}, \ldots, x_{N}^{i}), \tag{38} \]
\[ e_{i,k}(x_{1}^{i}, \ldots, x_{N}^{i}, t) := \delta F_{i,k}(x_{i+1}(t), t) + \sigma_{i,k}(x_{i+1}) \]
\[ + o_{i,k}(x_{1}^{i}, \ldots, x_{N}^{i}). \tag{39} \]
We will see how these terms are connected to the intrinsic dynamics of the subsystems and the coupling terms in the applications in Sections 3 and 4.
Remark 2. (i) The terms $e_{i,k} := \delta F_{i,k} + \sigma_{i,k} + o_{i,k}$ defined in (39) collect the inhomogeneities in the coupled system (4). Indeed, from the definition in (22), $\delta F_{i,k}$ depicts the difference between two adjacent subsystems. From Lemma 2.2, $\sigma_{i,k}$ is associated with the difference between adjacent row sums of connection matrix, and $o_{i,k}$ collects the terms related to the heterogeneity of coupling delays. Notice that $e_{i,k}$ vanishes if the coupled system (4) is homogeneous, and has identical row sums and identical coupling delays in the connection (i.e., (5) and (6) hold).

(ii) Let us summarize the whole decomposition. We obtained in (24) a preliminary decomposition of $H_{i,k}$ into three parts: $F_{i,k}(x_i(t), t) - F_{i,k}(x_{i+1}(t), t)$, $\delta F_{i,k}$, and $\tilde{G}_{i,k}$. In this step, we extracted from $H_{i,k}$ the term $\delta F_{i,k}$ which captures the distinction between (adjacent) subsystems, as mentioned in (i). For $F_{i,k}(x_i(t), t) - F_{i,k}(x_{i+1}(t), t)$, we further decomposed it into two parts: $f_{i,k}$ and $u_{i,k}$, in (27).

By composing matrices $W$ and $\tilde{W}$ in (28)–(29) from connection matrix $W$, and then applying (30), we further decomposed $\tilde{G}_{i,k}$ into four parts: $\hat{g}_{i,k}$, $v_{i,k}$, $\sigma_{i,k}$, and $o_{i,k}$, in Lemma 2.2. In this process, the heterogeneities caused from the connection matrix and coupling delays were extracted from $\tilde{G}_{i,k}$ into $\sigma_{i,k}$ and $o_{i,k}$, respectively.

The idea of recomposing $H_{i,k}$ in (23) and recasting (23) into (34) originates from extending the formulations in [45, 54], where identical synchronization for a coupled system comprising identical subsystems was studied. In particular, Proposition 2.1 in [54] established a decomposition of $H_{i,k}$, which is in the form of the right hand side of (34) with $e_{i,k}(\cdot) \equiv 0$. By applying sequential contracting technique, this led to the convergence of $z_{i,k}(t)$ to zero as $t \to \infty$, for $z_{i,k}(t)$ satisfying the difference-differential equations, and hence yielded the identical synchronization. The transformation (30) has been applied to diffusively coupled systems in [35]. Its application in [54] was to treat a system similar to (4), but with identical row sums for the connection matrix.

In the present investigation, taking into account distinctions between subsystems, and disparities between row sums and between coupling delays in the connection, the key idea is to develop an ingenious and effective decomposition on the associated difference-differential equations, to suitably extract and collect these inhomogeneous factors, and handle the remaining terms by the sequential contracting arguments. The extension of (30) to treat system (4) with nonidentical row sums becomes possible only after such new decomposition is derived. In the rest of this section, we shall modify and extend the sequential contracting technique to the decomposed difference-differential equations (34) to obtain the convergence of $z_{i,k}(t)$ to a neighborhood of zero as $t \to \infty$, which then yields the approximate synchronization of system (4).

Next, let us introduce two assumptions.

Assumption (E): For each $i, j \in \mathcal{N}$ and $k \in \mathcal{K}$, there exists an $\delta_k \geq 0$ such that

$$|F_{i,k}(\Xi, t) - F_{j,k}(\Xi, t)| \leq \delta_k, \quad \text{for all } \Xi \in \mathcal{Q} \text{ and } t \geq t_0.$$  

Assumption (F): For each $(i, k) \in \mathcal{A}$, there exist $\hat{\mu}_{i,k}, \check{\mu}_{i,k} \in \mathbb{R}$, $\rho_{i,k}^u \geq 0$, and $\hat{\mu}_{i,k}^{(l)}$ for $l \in \mathcal{K} - \{k\}$, such that the following properties hold:

(F-i): $\mu_{i,k} \leq \hat{f}_{i,k}(\xi, \eta, t)\mid (\xi - \eta) \mid \leq \check{f}_{i,k}(\xi, \eta, t) = 0$ if $\xi \neq \eta$, and $\check{f}_{i,k}(\xi, \eta, t) = 0$ if $\xi = \eta$, for $\xi, \eta \in [\check{q}_k, \hat{q}_k]$, and

(F-ii): $|u_{i,k}(\Xi, \hat{\Xi}, t)| \leq \rho_{i,k}^u$ and $|u_{i,k}(\Xi, \check{\Xi}, t)| \leq \sum_{l \in \mathcal{K} - \{k\}} \hat{\mu}_{i,k}^{(l)} |\xi_l - \check{\xi}_l|$, for all $\Xi = (\xi_1, \ldots, \xi_K), \hat{\Xi} = (\hat{\xi}_1, \ldots, \hat{\xi}_K), \check{\Xi} = (\check{\xi}_1, \ldots, \check{\xi}_K) \in \mathcal{Q}$. 


Functions $F_{i,k}$ govern the $i$th subsystem in (4) when there is no coupling. Hence, $\delta_k$ in assumption (E) measures the difference between subsystems when they are isolated. Notably, each $\delta F_{i,k}$ defined in (22) satisfies
\[
|\delta F_{i,k}(\Xi, t)| \leq \tilde{\delta}_k, \text{ for all } \Xi \in \mathcal{Q} \text{ and } t \geq t_0,
\] (40)
under assumption (E). Assumption (F) imposes conditions on functions $f_{i,k}$ and $u_{i,k}$ defined in (27), where $f_{i,k}$ and $u_{i,k}$ are determined by functions $F_{i,k}$. We shall illustrate how to justify these assumptions as we consider coupled FitzHugh–Nagumo equations (69) in Proposition 3.3, and coupled Lorenz equations (99) in Proposition 4.3.

In the following proposition, we derive the properties for functions $h_{i,k}$, $h_{i,k}$, $\bar{w}_{i,k}$, and $c_{i,k}$ in (36)-(39). This proposition is drawn with the notations introduced in (5), (7), (14)-(17), (28), and (29).

**Proposition 2.3.** Under assumptions (D), (E), and (F), for every $(i, k) \in A$ and $(\Phi_1, \ldots, \Phi_N) \in \mathcal{C}_Q^N$ with $\Phi_j = (\phi_{i,j,1}, \ldots, \phi_{i,j,K})$, $j \in \mathcal{N}$, the following hold for $t \geq t_0$:

(i): \[
\hat{\mu}_{i,k} \leq \frac{h_{i,k}(\phi_{i,k}(0), \phi_{i+1,k}(0), t)}{\phi_{i,k}(0) - \phi_{i+1,k}(0)} \leq \hat{\mu}_{i,k} \text{ if } \phi_{i,k}(0) - \phi_{i+1,k}(0) \neq 0,
\]
\[
h_{i,k}(\phi_{i,k}(0), \phi_{i+1,k}(0), t) = 0 \text{ if } \phi_{i,k}(0) - \phi_{i+1,k}(0) = 0,
\]
(ii): \[
\tilde{h}_{i,k} \leq \tilde{h}_{i,k}(\phi_{i,k}(\tau_{ii}), \phi_{i+1,k}(\tau_{ii})) \leq \tilde{h}_{i,k}, \text{ and }
\]
\[
\tilde{h}_{i,k}(\phi_{i,k}(\tau_{ii}), \phi_{i+1,k}(\tau_{ii})) \leq \tilde{h}_{i,k}, \text{ if } \phi_{i,k}(\tau_{ii}) - \phi_{i+1,k}(\tau_{ii}) \neq 0,
\]
\[
h_{i,k}(\phi_{i,k}(\tau_{ii}), \phi_{i+1,k}(\tau_{ii})) = 0 \text{ if } \phi_{i,k}(\tau_{ii}) - \phi_{i+1,k}(\tau_{ii}) = 0,
\]
(iii): \[
|\tilde{w}_{i,k}(\Phi_1, \ldots, \Phi_N, t)| \leq \rho_{i,k}^w, \text{ and }
\]
\[
\tilde{w}_{i,k}(\Phi_1, \ldots, \Phi_N, t) \leq \tilde{c}_{i,k},
\]
(iv): \[
|c_{i,k}(\Phi_1, \ldots, \Phi_N, t)| \leq \tilde{c}_{i,k},
\]

\[
\tilde{\mu}_{i,k} := \frac{(\kappa_i + \tilde{w}_{i}) \hat{D}_k}{(\kappa_i + \tilde{w}_{i}) \hat{D}_k} \text{ if } \kappa_i + \tilde{w}_{i} \geq 0, \quad \tilde{\mu}_{i,k} := \frac{(\kappa_i + \tilde{w}_{i}) \hat{D}_k}{(\kappa_i + \tilde{w}_{i}) \hat{D}_k} \text{ if } \kappa_i + \tilde{w}_{i} < 0,
\]

\[
\tilde{\beta}_{i,k} := \frac{\tilde{\mu}_{i,k}^{(j,l)}}{\tilde{\beta}_{i,k}^{(j,l)}} \text{ if } j = i, l \neq k, \quad \tilde{\beta}_{i,k} := \frac{\tilde{\mu}_{i,k}^{(j,l)}}{\tilde{\beta}_{i,k}^{(j,l)}} \text{ if } j \neq i, l = k, \quad \tilde{\beta}_{i,k} := \frac{\tilde{\mu}_{i,k}^{(j,l)}}{\tilde{\beta}_{i,k}^{(j,l)}} \text{ otherwise},
\]

\[
\tilde{c}_{i,k} := \tilde{G}_k |\kappa_i - \kappa_{i+1}| + (2|\kappa_i| + \sum_{j \in \mathcal{N} \setminus \{N\}} |\tilde{\omega}_{ij}|) \hat{D}_k \Lambda_k \delta_r,
\]

where $\mu_{i,k}$ and $\mu_{i,k}$ are defined in assumption (F), $\tilde{\delta}_k$ in assumption (E), $\mu_{i,k}^{(j,l)}$ in assumption (F), and $\rho_{i,k}^h$ and $\rho_{i,k}^w$ are quantities satisfying
\[
\rho_{i,k}^h \geq 2|\kappa_i| + \bar{w}_{i} \bar{G}_k, \quad \rho_{i,k}^w \geq \rho_{i,k}^h + 2\bar{G}_k \sum_{j \in \mathcal{N} \setminus \{i, N\}} |\tilde{\omega}_{ij}|.
\]

**Proof.** Assertions (i)–(iii) can be justified by applying arguments similar to those for assertions (H-i)–(H-iii) of Proposition 2.1 in [54], respectively. Let us focus on verifying the remaining assertion (iv) to complete the proof. Recall that $(\Phi_1, \ldots, \Phi_N) \in \mathcal{C}_Q^N$ with $\Phi_j = (\phi_{i,j,1}, \ldots, \phi_{i,j,K})$, $j \in \mathcal{N}$, i.e.,
\[
\phi_{j,k}(\theta) \in [\hat{q}_k, \hat{q}_k] \text{ if } \theta \in [-\tau_M, 0], \text{ and } |\phi_{j,k}'(\theta)| \leq \Lambda_k \text{ if } \theta \in (-\tau_M, 0),
\]
(42)
for all $j \in \mathcal{N}$ and $k \in \mathcal{K}$. Consequently,
\begin{equation}
|G_k(\phi_{i,k}(\theta))| \leq G_k,
\end{equation}
for all $i, j \in \mathcal{N}$, $k \in \mathcal{K}$, and $\theta \in [-\tau_M, 0]$, recalling (14). Let us verify that
\begin{equation}
|G_k(\phi_{j,k}(\theta)) - G_k(\phi_{j,k}(\theta))| \leq \tilde{D}_k \Lambda_k \delta_t,
\end{equation}
for all $\theta, \vartheta \in \{-r_s : r, s \in \mathcal{N}\}$, $j \in \mathcal{N}$, and $k \in \mathcal{K}$. Notably, (44) holds if $\phi_{j,k}(\theta) = \phi_{j,k}(\vartheta)$. If $\phi_{j,k}(\theta) \neq \phi_{j,k}(\vartheta)$, then applying the mean-value theorem yields the existence of some $\psi_{j,k}$ between $\phi_{j,k}(\theta)$ and $\phi_{j,k}(\vartheta)$, and some $u_{j,k}$ between $\theta$ and $\vartheta$, with $u_{j,k} \in (-\tau_M, 0)$, such that
\begin{equation}
G_k(\phi_{j,k}(\theta)) - G_k(\phi_{j,k}(\theta)) = G_k'(\psi_{j,k})(u_{j,k})(\theta - \vartheta).
\end{equation}
It is apparent that $\psi_{j,k} \in [q_k, q_k]$, $u_{j,k} \in (-\tau_M, 0)$, and $\theta, \vartheta \in [-\tau_M, 0]$, which imply that $G_k'(\psi_{j,k}) \leq \tilde{D}_k$, $o_{j,k}(u_{j,k}) \leq \Lambda_k$, and $|\theta - \vartheta| \leq \delta_t$, with (5), (16), and (42). This justifies (44).

Recall the definitions of $\sigma_{i,k}$ and $\alpha_{i,k}$ in Lemma 2.2. Applying (43) and (44) then leads to
\begin{equation}
\begin{cases}
|\sigma_{i,k}(\Phi_{i+1})| \leq \tilde{G}_k|\kappa_i - \kappa_{i+1}|, \\
|\alpha_{i,k}(\Phi_1, \ldots, \Phi_N)| \leq (2|\kappa_i| + \sum_{j \in \mathcal{N} - \{N\}} |\tilde{\omega}_{ij}| + \sum_{j \in \mathcal{N}} |\tilde{\omega}_{(i+1)j}|)|\tilde{D}_k \Lambda_k \delta_t.
\end{cases}
\end{equation}
The boundedness of $e_{i,k}$ in assertion (iv) can be verified directly through the definition of $e_{i,k}$ in (39), the definitions of $\sigma_{i,k}$ and $\alpha_{i,k}$ in Lemma 2.2, the boundedness of $\delta F_{i,k}$ in (40), and the boundedness of $\sigma_{i,k}$ and $\alpha_{i,k}$ in (45). This completes the proof.

Remark 3. (i) In Proposition 2.3, assertions (i)-(iii) are based on assumptions (D) and (F). Assertion (iv) relies on the additional assumption (E) which measures the variation of subsystems, and on the new setting of $C_Q$, i.e., $|\phi_k(t)| \leq \Lambda_k$ in (19). Moreover, as seen from Remark 1 and (45), the part $\sum_{j \in \mathcal{N}} |\tilde{\omega}_{(i+1)j}|)|\tilde{D}_k \Lambda_k \delta_t$ of $\tilde{e}_{i,k}$ in assertion (iv) measures the heterogeneity caused from coupling delays between components. This explains why a new setting of $C_Q$ is needed, as compared to the one in [45, 54].

(ii) Recall Remark 2 that $e_{i,k} := \delta F_{i,k} + \sigma_{i,k} + \alpha_{i,k}$ defined in (39) collects the inhomogeneous factors in the network system (4). As seen from (5), (6), and assumption (E), $\delta_{tk}$, $\delta_{\kappa}$, and $\delta_t$ measure the differences between subsystems, the differences between row sums, and the variation of coupling delays, respectively. In addition, $\tilde{e}_{i,k}$ defined in Proposition 2.3 decreases to zero as $\delta_{tk}$, $\delta_{\kappa}$, and $\delta_t$ decrease to zero. This reveals that each $\tilde{e}_{i,k}$ converges to zero as the coupled system (4) tends to being homogeneous, and the row sums and coupling delays in the connection tend to being identical. It will be seen in Theorem 2.6 and Remark 4 how the terms $\tilde{e}_{i,k}$ determine an estimate on the synchronization error.

(iii) The terms $\tilde{\beta}_{i,k}$ and $\tilde{\beta}_{i,k}$ in Proposition 2.3 depend on the sign of $\kappa_i + \tilde{\omega}_{ii}$. The sign of $\kappa_i$ is easily seen from the row sum of connection matrix, defined in (7). On the other hand, determining the sign of $\tilde{\omega}_{ii}$ relies on the expression in (31) which is newly derived in Appendix A of this paper. It can be observed from (7) and (31) that, all $\tilde{\omega}_{ii}$ are apt to be negative if all off-diagonal entries of $W$ are nonnegative, or the magnitudes of the negative off-diagonal entries are relatively smaller than those of the positive off-diagonal entries. Suppose that all $\tilde{\omega}_{ii}$ are negative. If all row sums of $W$ are small enough, then $\kappa_i + \tilde{\omega}_{ii} < 0$ for every $i$. For a
nonzero connection matrix \( W \) with nonnegative off-diagonal entries and satisfying the diffusive condition \( (\kappa_i = 0 \text{ for all } i \in \mathcal{N}) \), which is a common assumption in the literature, one does have \( \kappa_i + \bar{\omega}_{ii} = \bar{\omega}_{ii} < 0 \) for every \( i \).

For later use, with the quantities \( \bar{\mu}_{i,k}, \bar{\beta}_{i,k}, \bar{\beta}_{i,k}^{(j,l)}, \tilde{\mu}_{i,k}^{(j,l)} \), and \( \epsilon_{i,k} \) determined in Proposition 2.3, we define

\[
\begin{align*}
\eta_{i,k} &:= -\bar{\mu}_{i,k} - \bar{\beta}_{i,k} + \bar{\beta}_{i,k} \tau_{ii}(\bar{\mu}_{i,k} + \bar{\beta}_{i,k} + \tilde{\beta}_{i,k}), \\
\bar{\mu}_{i,k}^{(j,l)} &:= \bar{\beta}_{i,k}^{(j,l)} - \bar{\beta}_{i,k}^{(j,l)},
\end{align*}
\]

where

\[
\bar{\beta}_{i,k} := \max\{ |\tilde{\beta}_{i,k}|, |\beta_{i,k}| \}.
\]

Herein, the term \( \eta_{i,k} \) is delay-dependent, since it involves the magnitude of delay \( \tau_{ii} \).

Recall that for an arbitrary solution \( (x_1(t), \ldots, x_N(t)) \) of system (4), where \( x_i(t) = (x_{i,1}(t), \ldots, x_{i,K}(t)) \), \( i \in \mathcal{N} \), by setting \( z_i(t) = x_i(t) - x_{i+1}(t) \) for \( i \in \mathcal{N} - \{N\} \), we see that \( (z_1(t), \ldots, z_{N-1}(t)) \) satisfies system (23) which can be recast into system (26). In the following, let us preview the main process for showing \( \limsup_{t \to \infty} \|z_i(t)\| \leq \varepsilon_\alpha \), for all \( i = 1, \ldots, N-1 \), and some nonnegative \( \varepsilon_\alpha \geq 0 \). This then establishes the approximate synchronization of system (4). Let us first relabel the two-dimensional indices in system (26) into one-dimensional indices, through the bijective mapping \( \ell : \mathcal{A} \to \{1, \ldots, (N-1)K\} \), defined by

\[
\ell(i, k) := (i - 1) \cdot K + k.
\]

The labeling indicates an ordering of \( \ell(i, k) \) by the orders of \( i \) and \( k \) in succession:

\[
\ell(i, k) < \ell(j, l) \text{ if } i < j; \ell(i, k) < \ell(i, l) \text{ if } k < l.
\]

For later use, we define, for \( (i, k) \in \mathcal{A} \),

\[
\begin{align*}
\hat{\mathcal{L}}_{i,k} &:= \{(j, l) \in \mathcal{A} : \ell(j, l) < \ell(i, k)\}, \\
\tilde{\mathcal{L}}_{i,k} &:= \{(j, l) \in \mathcal{A} : \ell(j, l) > \ell(i, k)\}.
\end{align*}
\]

By setting \( z_{\ell(i, k)} := z_{i,k} = x_{i,k} - x_{i+1,k}, \hat{\gamma}_{\ell(i, k)} := \hat{h}_{i,k}, \tilde{\gamma}_{\ell(i, k)} := \tilde{h}_{i,k}, \) and \( w_{\ell(i, k)} \), we can rewrite system (26) as

\[
\begin{align*}
\dot{z}_{\ell(i, k)}(t) &= \hat{h}_{\ell(i, k)}(x_{i,k}(t), x_{i+1,k}(t), t) \\
&\quad + \tilde{h}_{\ell(i, k)}(x_{i,k}(t - \tau_{ii}), x_{i+1,k}(t - \tau_{ii})) + w_{\ell(i, k)}(t),
\end{align*}
\]

for \( (i, k) \in \mathcal{A} \) and \( t \geq t_0 \). For each \( (i, k) \), the equation (53) is analogous to a scalar equation:

\[
\dot{z}(t) = h(x(t), y(t), t) + \bar{h}(x_t, y_t, t) + w(t), \quad t \geq t_0,
\]

where \( z(t) = x(t) - y(t) \). In Appendix B, we introduce precisely this scalar equation (54) (relabelled as (118) therein), and summarize the convergent property of (118) in Proposition 6.1 which reveals that \( z(t) \) converges to an interval \( [-\bar{\mu}, \bar{\mu}] \) as \( t \to \infty \); in addition, \( \bar{\mu} \) can be estimated. Applying such a convergent property to each \( z_{\ell(i, k)}(t) \), \( (i, k) \in \mathcal{A} \), we obtain the convergence of \( z_{\ell(i, k)}(t) \) to an interval \( [-\bar{\mu}_{\ell(i, k)}, \bar{\mu}_{\ell(i, k)}] \), with

\[
0 \leq \bar{\mu}_{\ell(i, k)} \leq \frac{|w_{\ell(i, k)}|_{\text{max}}(\infty)}{\eta_{i,k}},
\]

where \( \eta_{i,k} \) is defined in (46). Herein,

\[
\begin{align*}
|w_{\ell(i, k)}|_{\text{max}}(\infty) &:= \lim_{T \to \infty} |w_{\ell(i, k)}|_{\text{max}}(T), \\
|w_{\ell(i, k)}|_{\text{max}}(T) &:= \sup\{|w_{\ell(i, k)}(t)| : t \geq T\}, \text{ for } T \geq t_0.
\end{align*}
\]
where \( w_{\ell(i,k)} \) is given in (53). We can derive better estimations on \( \tilde{p}_{\ell(i,k)} \) for all \((i,k) \in A\). Notice that for each \((i,k)\), if estimates of \( \tilde{p}_{\ell(j,l)} \) have been obtained for all \((j,l) \in L_{i,k}\), we may use them to bound \( z_{\ell(j,l)}(t) \) which plays the role as \( \phi_{\ell(j,l)}(\theta) - \phi_{\ell(i,k)}(\theta) \) in (41), and hence obtain sharper estimate on \( |w_{\ell(i,k)}|^{\max}(\infty) \), and subsequently sharper estimate on \( \tilde{p}_{\ell(i,k)} \), by (55). We perform such a process to \( z_1(t), \ldots, z_{(N-1)K}(t) \) successively and iteratively and obtain, for each \((i,k) \in A\), a sequence \( \{p_{\ell(i,k)}^{(n)}\}_n^{\infty} \) which satisfies

\[
\tilde{p}_{\ell(i,k)} \leq p_{\ell(i,k)}^{(0)} := (\hat{\mu}_{i,k} + \hat{\beta}_{i,k})/\eta_{i,k},
\]

\[
\tilde{p}_{\ell(i,k)} \leq p_{\ell(i,k)}^{(n)} := \left[ \sum_{(j,l) \in L_{i,k}} m_{i,k}^{(j,l)} p_{\ell(j,l)}^{(n)} + \sum_{(j,l) \in L_{i,k}} \tilde{m}_{i,k}^{(j,l)} p_{\ell(j,l)}^{(n-1)} + \tilde{\epsilon}_{i,k}\right]/\eta_{i,k},
\]

for \( n \geq 1 \). We observe that, with \( p_{\ell(i,k)}^{(n)} \) defined in (56), \( \{(p_1^{(n)}, \ldots, p_{(N-1)K}^{(n)})^T\}_n^{\infty} \) is exactly the Gauss-Seidel iteration for solving the linear system

\[
Mv = e,
\]

where

\[
M = [m_{ij}]_{1 \leq i,j \leq (N-1)K}, m_{ij} := \begin{cases} \eta_{i,k} & \text{if } i = j = \ell(i,k), \\ -\tilde{m}_{i,k}^{(j,l)} & \text{if } i = \ell(i,k) \neq j = \ell(j,l), \end{cases}
\]

\[
e = (\epsilon_1, \ldots, \epsilon_{(N-1)K})^T \in \mathbb{R}^{(N-1)K \times 1}, \quad \epsilon_\ell := \tilde{\epsilon}_{i,k} \text{ if } \ell = \ell(i,k),
\]

with \( \eta_{i,k} \) defined in (46), \( \tilde{m}_{i,k}^{(j,l)} \) in (47), and \( \tilde{\epsilon}_{i,k} \) in Proposition 2.3. Notice that \( M \) is delay-dependent, since the diagonal entries \( \eta_{i,k} \) involve the magnitudes of delay \( \tau_{ii} \), as indicated in (46).

Condition (S1): \( \hat{\mu}_{i,k} + \hat{\beta}_{i,k} < 0 \), and \( \tilde{\beta}_{i,k} \tau_{ii} < 3\hat{\mu}_{i,k}(\hat{\mu}_{i,k} + \hat{\beta}_{i,k})/[(\hat{\mu}_{i,k} + \hat{\beta}_{i,k})(3\hat{\mu}_{i,k} + \hat{\beta}_{i,k} + \tilde{\epsilon}_{i,k})] \), for \((i,k) \in A\).

Notably, condition (S1) is delay-dependent, as it involves \( \tau_{ii} \). Moreover, it implies \( \eta_{i,k} > 0 \) for all \((i,k) \in A\). Now, let us denote

\[
M^{-1} e := (p_1^{*}, \ldots, p_{(N-1)K}^{*})^T \in \mathbb{R}^{(N-1)K \times 1},
\]

\[
\varepsilon_e := \max_{i \in \{1, \ldots, (N-1)K\}} \left\| (p_{(i-1)K+1}^{*}, \ldots, p_{iK}^{*}) \right\|,
\]

\[
M = D_M - L_M - U_M,
\]

where \( D_M, -L_M \) and \(-U_M\) represent the diagonal, strictly lower-triangular and strictly upper-triangular parts of \( M \), respectively.

**Proposition 2.4.** Consider system (23) under assumptions (D), (E), (F) and condition (S1). Then

\[
\limsup_{t \to \infty} \left\| z_i(t) \right\| \leq \varepsilon_e, \text{ for } i = 1, \ldots, N - 1,
\]

if the Gauss-Seidel iteration for linear system (57) converges to \( M^{-1} e \), the unique solution of (57), or equivalently,

\[
\lambda_{\text{syn}} := \max \{|\lambda| : \lambda : \text{eigenvalues of } (D_M - L_M)^{-1}U_M \} < 1.
\]

**Proof.** Recall that under assumptions (D), (E), and (F), system (23) can be recast to system (53), and each \( \ell \)-th component in system (53) takes the form (54), where \( \ell = \ell(i,k) \). In Appendix B, the properties for equation (54) (relabelled as (118)) are summarized in Proposition 6.1, under conditions (H0) and (A1). It can be verified
that conditions (H0) and (A1) hold under assumptions (D), (E), (F) and condition (S1). Applying Proposition 2.3 reveals that $\bar{\mu} = \hat{\mu}_i,k, \bar{\mu} = \hat{\mu}_i,k, \rho^0 = \hat{p}^0_{i,k}, \tilde{\beta} = \hat{\beta}_i,k, \tilde{\sigma} = \tau_{ii}$, for each $(i,k) \in A$, under the formulation in condition (H0). Consequently, for each $(i,k) \in A$, there exists an interval $I_{\ell(i,k)} := [-\bar{\tilde{p}}_{\ell(i,k)}, \bar{\tilde{p}}_{\ell(i,k)}]$ to which $z_{\ell(i,k)}(t)$ converges, as $t \to \infty$. Moreover, the estimation of $\tilde{p}_{\ell(i,k)}$ in (55) holds.

Next, for further estimations on $\tilde{p}_{\ell(i,k)}$, let us show how the sequential estimates $\{\tilde{p}_{\ell(i,k)}^{(n)}\}_{n=0}^\infty$ which satisfy the iterative formula (56) is derived. This can be seen by induction in the following steps (i) and (ii).

(i) First, consider $n = 0$ and all $(i,k) \in A$. By Lemma 2.1, under assumption (D), there exists some $\ell$, with $\ell \geq t_0 \tau_M$, so that $(x_1^i, \ldots, x_N^i) \in C_N$ for all $t \geq \ell$. Accordingly, for each $(i,k) \in A$,

$$|w_{\ell(i,k)}(t)| = |w_{i,k}(t)| \leq \bar{w}_{i,k}(x_1^i, \ldots, x_N^i, t) + |\epsilon_{i,k}(x_1^i, \ldots, x_N^i, t)| \leq \rho^0_{i,k} + \bar{\epsilon}_{i,k},$$

for $t \geq \ell$, recalling (35) and the assertions (iii) and (iv) in Proposition 2.3. Subsequently, $|w_{\ell(i,k)}|^{\max(\infty)} \leq \rho^0_{i,k} + \bar{\epsilon}_{i,k}$, and this yields

$$0 \leq \tilde{p}_{\ell(i,k)} \leq |w_{\ell(i,k)}|^{\max(\infty)}/\eta_{i,k} \leq \tilde{p}^{(0)}_{\ell(i,k)},$$

recalling (55).

(ii) Next, assume that (56) holds and $\tilde{p}^{(n)}_{\ell(i,k)}$ have been defined, and hence $z_{\ell(j,l)}$ converges to $[-\tilde{p}^{(n)}_{\ell(j,l)}, \tilde{p}^{(n)}_{\ell(j,l)}]$ as $t \to \infty$, for $n = 1, \ldots, n_0 - 1, (j,l) \in A$ and $n = n_0$, $(j,l) \in \hat{E}_i$, with $\ell(i,k) < (N - 1)K$. By (35) and the assertions (iii) and (iv) in Proposition 2.3,

$$|w_{\ell(i,k)}(t)| = |w_{i,k}(t)| \leq |\bar{w}_{i,k}(x_1^i, \ldots, x_N^i, t)| + |\epsilon_{i,k}(x_1^i, \ldots, x_N^i, t)| \leq \sum_{(j,l) \in A - \{i(k)} |\tilde{p}_{\ell(j,l)}| + |\tilde{p}_{\ell(j,l)}(t - \tau_{ij})| + \tilde{\epsilon}_{i,k},$$

if $t \geq \hat{\ell}$, where $\hat{\ell}$ was introduced in step (i). Then,

$$|w_{\ell(i,k)}|^{\max(\infty)} \leq \sum_{(j,l) \in \hat{E}_i, k} \tilde{m}_{i,k}(j,l) \tilde{p}^{(n_0)}_{\ell(j,l)} + \sum_{(j,l) \in \hat{E}_i, k} \tilde{m}_{i,k}(j,l) \tilde{p}^{(n_0-1)}_{\ell(j,l)} + \tilde{\epsilon}_{i,k},$$

and hence

$$\tilde{p}_{\ell(i,k)} \leq |w_{\ell(i,k)}|^{\max(\infty)}/\eta_{i,k} \leq \sum_{(j,l) \in \hat{E}_i, k} \tilde{m}_{i,k}(j,l) \tilde{p}^{(n_0)}_{\ell(j,l)} + \sum_{(j,l) \in \hat{E}_i, k} \tilde{m}_{i,k}(j,l) \tilde{p}^{(n_0-1)}_{\ell(j,l)} + \tilde{\epsilon}_{i,k}/\eta_{i,k} = \tilde{p}^{(n_0)}_{\ell(i,k)}.$$

We have thus completed the induction and justified (56).

With $\tilde{p}^{(n)}_{\ell(i,k)}$ defined in (56), we see that $\{(p_1^{(n)}, \ldots, p_{(N-1)K}^{(n)})\}_{n=0}^\infty$ is exactly the Gauss-Seidel iteration for solving the linear system (57). Recall that, for each $(i,k) \in A$, $z_{\ell(i,k)}(t)$ converges to $[-\tilde{p}_{\ell(i,k)}, \tilde{p}_{\ell(i,k)}]$ as $t \to \infty$, and $0 \leq \tilde{p}_{\ell(i,k)} \leq \tilde{p}^{(n)}_{\ell(i,k)}$ for all $n \in \mathbb{N}$. If this Gauss-Seidel iteration converges to the unique solution $(p_1^{*}, \ldots, p_{(N-1)K}^{*}) = \mathbf{M}^{-1}\mathbf{e}$, then $\tilde{p}_{\ell(i,k)} \leq \tilde{p}^{*}_{\ell(i,k)}$ for all $(i,k) \in A$, and

$$\limsup_{t \to \infty} |z_i(t)| = \limsup_{t \to \infty} \|z_{\ell(i,j)}(t), \ldots, z_{\ell(i,K)}(t)\| \leq \|\tilde{p}_{\ell(i,K+1)}^{*}, \ldots, \tilde{p}_{\ell(i,K)}^{*}\| \leq \varepsilon_\text{e},$$

for $i = 1, 2, \ldots, N - 1$. This completes the proof. □
There is an alternative way to estimate the size of intervals \([-\bar{p}_{t(i,k)}, \bar{p}_{t(i,k)}]\). By employing a different sequence of upper-lower bounds in analyzing equation (54) (i.e., (118)), we can obtain a disparate estimate for the interval \([-\bar{p}_{t(i,k)}, \bar{p}_{t(i,k)}]\) to which \(z_{t(i,k)}(t)\) converges, as stated in Proposition 6.2. Subsequently, an alternative iterative formula can be obtained, which represents the Gauss-Seidel iteration for another linear system:

\[
\bar{M}v = e,
\]

where \(e\) is defined in (59), \(\bar{m}_{i,k}\), \(\bar{m}_{i,j}\), and \(\bar{\beta}_{i,k}\) in (46)–(48). Let us denote

\[
\bar{M}^{-1}e := (\tilde{p}_1^*, \ldots, \tilde{p}_{(N-1)K}^*T \in \mathbb{R}^{(N-1)K \times 1},
\]

\[
\bar{\varepsilon}_e := \max_{i \in \{1, \ldots, N-1\}} \| (\tilde{p}_i^* / \cdots, \tilde{p}_{(N-1)K}^* ) \|.
\]

We introduce a delay-independent condition:

Condition (S_2): \(\bar{\beta}_{i,k} < -\bar{\mu}_{i,k}/[1 + (\bar{\rho}_{i,k} + \varepsilon_{i,k})/\bar{\rho}_{i,k}^p], \) for \((i, k) \in \mathcal{A}\).

By arguments similar to those for Proposition 2.4, but using Proposition 6.2 instead of Proposition 6.1, we can derive the following delay-independent result.

**Proposition 2.5.** Consider system (23) under assumptions (D), (E), (F) and condition (S_2). Then

\[
\limsup_{t \to \infty} \| z_i(t) \| \leq \bar{\varepsilon}_e \text{ for } i = 1, \ldots, N-1,
\]

if the Gauss-Seidel iteration for linear system (64) converges to \(\bar{M}^{-1}e\), the unique solution of (64), or equivalently,

\[
\lambda_{\text{syn}} := \max \{ |\lambda| : \lambda : \text{eigenvalues of } (D_{\bar{M}} - L_{\bar{M}})^{-1}U_{\bar{M}} < 1 \}
\]

where \(D_{\bar{M}}, -L_{\bar{M}}, \) and \(-U_{\bar{M}}\) are the diagonal, strictly lower-triangular, and strictly upper-triangular parts of \(\bar{M}\), respectively.

We apply the sequential contracting formulations in Propositions 2.4 and 2.5 to obtain the following main result.

**Theorem 2.6.** Under assumptions (D), (E), (F), and condition (S_1) (resp., (S_2)), system (4) attains \(\varepsilon_e\)-approximate (resp., \(\bar{\varepsilon}_e\)-approximate) synchronization if the Gauss-Seidel iterations for the linear system \(Mv = e\) (resp., \(\bar{M}v = e\)) in (57) (resp., (64)) converges to the unique solution \(M^{-1}e\) (resp., \(\bar{M}^{-1}e\)), or equivalently, (63) (resp., (68)) holds, where \(M, e, \bar{M}, \varepsilon_e, \) and \(\bar{\varepsilon}_e\) are defined in (58), (59), (61), (65), and (67), respectively.

**Remark 4.** Theorem 2.6 established the \(\varepsilon_e\)-approximate (resp., \(\bar{\varepsilon}_e\)-approximate) synchronization of coupled network system (4). It indicates that \(\varepsilon_e \to 0\) (resp., \(\bar{\varepsilon}_e \to 0\)) as \(\|e\| \to 0\), where \(\|e\|\) bounds the effects from all inhomogeneous factors in system (4). In addition, \(\bar{\varepsilon}_{i,k}\) defined in Proposition 2.3, and hence \(\|e\|\) and \(\varepsilon_e\) (resp., \(\bar{\varepsilon}_e\)), decrease to zero as the differences between subsystems, differences between row sums, and differences between coupling delays decrease to zero; restated,

\[
\varepsilon_e (\text{resp., } \bar{\varepsilon}_e) \downarrow 0, \text{ as all } \bar{\delta}_k, \delta_e , \delta_r \downarrow 0.
\]

With the preparation in Lemmas 2.1, 2.2 and Propositions 2.3-2.5, the problem of approximate synchronization is converted to solving a corresponding linear algebraic
system (57) or (64), which is nonhomogeneous if $e \neq 0$, and homogeneous if $e = 0$. If the coupled system (4) has identical subsystems, identical row sums, and identical coupling delays in the connection, expressed in (5) and (6), then $\hat{\epsilon}_{i,k} = 0$ for all $(i, k) \in A$, and thus $e = 0$ and $\varepsilon_a = \hat{\varepsilon}_a = 0$. In this case, the formula in (56) and the assertion in Theorem 2.6 reduce to the ones in [45], where the problem of identical synchronization was converted to solving a corresponding homogeneous linear system of algebraic equations.

3. Approximate synchronization of FitzHugh-Nagumo neurons. In this section, we apply Theorem 2.6 to study approximate synchronization for a coupled system of heterogeneous FitzHugh–Nagumo neurons under delayed sigmoidal coupling:

$$\begin{cases}
\dot{u}_i(t) = U_i(u_i(t), v_i(t)) + c \sum_{j \in \mathcal{N}} \omega_{ij} g(v_j(t - \tau_{ij})), \\
\dot{v}_i(t) = V_i(u_i(t), v_i(t)),
\end{cases} \quad (69)$$

for $t \geq t_0$ and $i \in \mathcal{N} := \{1, \ldots, N\}$, where $c > 0$ is the coupling strength, $\omega_{ij} \in \mathbb{R}$, $\tau_{ij} \in [0, \tau_M]$ for some $\tau_M \geq 0$,

$$U_i(\xi, \eta) := -\xi^3 + (a_i + 1)\xi^2 - a_i \xi - \eta, \quad (70)$$

$$V_i(\xi, \eta) := b_i \xi - \gamma_i \eta, \quad (71)$$

for $\xi, \eta \in \mathbb{R}$, with $a_i, b_i, \gamma_i \geq 0$, and the coupling function $g$ belongs to the following class:

$$\{g \in C^1 : g(0) = 0, \quad L := g'(0) > g'(\xi) > 0, \quad |g(\xi)| < \rho, \quad \text{for } \xi \neq 0\}. \quad (72)$$

In addition, we assume $\sum_{i,j \in \mathcal{N}} |\omega_{ij}| > 0$, so that there are really connections among neurons in system (69). Identical synchronization for system (69) with identical $U_i$ (i.e., $a_i = a$) for all $i$ and identical $V_i$ (i.e., $b_i = b$, $\gamma_i = \gamma$) for all $i$ has been investigated numerically in [6], and analytically in [45, 54].

System (69) is in the form of (4), with $K = 2$, and $F_{i,k}$ and $G_k$ are now

$$F_{i,1}(\Xi, t) = U_i(\xi, \eta), \quad F_{i,2}(\Xi, t) = V_i(\xi, \eta), \quad (73)$$

$$G_1(x) = cg(x), \quad G_2(x) = 0, \quad (74)$$

where $\Xi = (\xi, \eta) \in \mathbb{R}^2$, $x \in \mathbb{R}$. Set, for $X = a, b, \gamma$,

$$\hat{X} := \max\{X_i : i \in \mathcal{N}\}, \quad \underline{X} := \min\{X_i : i \in \mathcal{N}\}, \quad (75)$$

$$\hat{\omega} := \max\{\sum_{j \in \mathcal{N}} |\omega_{ij}| : i = 1, \ldots, N\}. \quad (76)$$

Now, we show that system (69) satisfies assumption (D) via the following iteratively constructed functions: for $n \in \mathbb{N}$, let

$$P^{(n)}(\xi) := -\xi^3 + (a + 1)\xi^2 - a\xi + c\hat{\omega}\rho^{(n-1)}|\xi|, \quad (77)$$

where $\rho^{(0)} := \rho$, and

$$\rho^{(n)} := \max\{|g(\xi)| : \xi \in [-\sqrt{\hat{\gamma}^2 + b\hat{q}^{(n)}/\hat{\gamma}}, \sqrt{\hat{\gamma}^2 + b\hat{q}^{(n)}/\hat{\gamma}}]\}, \quad (78)$$

$$\hat{q}^{(n)} := \max\{|\xi| : P^{(n)}(\xi) = 0\}. \quad (79)$$

Herein, the terms $a, \hat{a}, b, \hat{\gamma}, \hat{\gamma}$ were introduced in (75), $\hat{\omega}$ in (76), and $\rho$ and function $g$ in (72).
Proposition 3.1. For each \( n \in \mathbb{N} \), all solutions of system (69) eventually enter and remain in \( Q^{(n)} \times \cdots \times Q^{(n)} \subset \mathbb{R}^{2N} \), where
\[
Q^{(n)} := [-\sqrt{\gamma^2 + (\hat{b}^2/b)q^{(n)}/\hat{\gamma}}, \sqrt{\gamma^2 + (\hat{b}^2/b)q^{(n)}/\hat{\gamma}}] \times [-\hat{b}q^{(n)}/\hat{\gamma}, \hat{b}q^{(n)}/\hat{\gamma}] \subset \mathbb{R}^2.
\]

This proposition follows from some energy-like estimates and iterative arguments. We sketch the main justification in Appendix C.

Notice that \( \bar{q}^{(n)} \) are strictly decreasing with respect to \( n \in \mathbb{N} \). Thus, \( Q^{(n)} \) with larger \( n \) provides a smaller attracting region for system (69), and thereby relaxes the condition in the synchronization criterion and yields finer estimation on the synchronization error \( \varepsilon_e \). It will be illustrated in the following Example 1 that \( \varepsilon_e \) becomes smaller if considering \( Q^* = Q^{(5)} \) rather than \( Q^* = Q^{(1)} \).

Throughout this section, thanks to Proposition 3.1, system (69) satisfies assumption (D) with
\[
Q = Q^{(n)} = [-\rho^*_1, \rho^*_1] \times [-\rho^*_2, \rho^*_2] =: Q^*,
\]
where
\[
\rho^*_1 := \sqrt{\gamma^2 + (\hat{b}^2/b)\bar{q}^{(n)}/\hat{\gamma}}, \quad \rho^*_2 := \hat{b}\bar{q}^{(n)}/\hat{\gamma},
\]
for some fixed \( n \in \mathbb{N} \). Applying (70)–(76) then reveals that the quantities defined in (13)–(17) now satisfy
\[
\|F_1\| \leq f^*_1, \|F_2\| \leq f^*_2, G_1 = c\bar{\rho}_g, D_1 = c\bar{L}_g, \bar{D}_1 = cL, \quad (81)
\]
\[
\bar{G}_2 = D_2 = \bar{D}_2 = 0, \Lambda_1 \leq f^*_1 + c\bar{\rho}_g\bar{\omega}, \Lambda_2 \leq f^*_2, \quad (82)
\]
where \( \rho^*_1 \) are introduced in (80), \( L \) in (72), \( \bar{\omega} \) in (76), and
\[
f^*_1 := (\rho^*_1)^3 + (\hat{\alpha} + \hat{\beta})(\rho^*_1)^2 + \hat{\alpha}\rho^*_1 + \rho^*_2, \quad f^*_2 := \hat{b}\rho^*_1 + \hat{\gamma}\rho^*_2, \quad (83)
\]
\[
\bar{\rho}_g := \max\{|g(\xi) : \xi \in [-\rho^*_1, \rho^*_1]|\} = \max\{|g(-\rho^*_1)|, |g(\rho^*_1)|\}, \quad (84)
\]
\[
\bar{L}_g := \min\{g'(\xi) : \xi \in [-\rho^*_1, \rho^*_1]\}. \quad (85)
\]

Proposition 3.2. System (69) satisfies assumption (E) with
\[
\bar{\delta}_k = \begin{cases} (\hat{\alpha} - \hat{a})[(\rho^*_1)^2 + \rho^*_1] & \text{if } k = 1, \\ (\hat{b} - \hat{b})\rho^*_2 + (\hat{\gamma} - \hat{\gamma})\rho^*_2 & \text{if } k = 2, \end{cases}
\]
where \( \rho^*_1 \) and \( \rho^*_2 \) are defined in (80).

Proof. Based on (22), (70), (71), and (73), a direct computation yields
\[
|F_{i,k}(\Xi, t) - F_{j,k}(\Xi, t)| = \begin{cases} (a_i - a_j)\xi^2 - (a_i - a_j)\xi & \text{if } k = 1, \\ (b_i - b_j)\xi - (\gamma_i - \gamma_j)\eta & \text{if } k = 2, \end{cases}
\]
\[
\leq \begin{cases} (\hat{a} - \hat{a})[(\rho^*_1)^2 + \rho^*_1] & \text{if } k = 1, \\ (\hat{b} - \hat{b})\rho^*_2 + (\hat{\gamma} - \hat{\gamma})\rho^*_2 & \text{if } k = 2, \end{cases}
\]
for \( \Xi = (\xi, \eta) \in Q^* \), where \( Q^* \) is defined in (80). The assertion thus follows. \( \square \)

The following proposition confirms assumption (F) for system (69). Its justification is similar to Proposition 3.2 of [54] and is omitted.

Proposition 3.3. System (69) satisfies assumption (F) with \( \hat{\mu}_{i,1} = a^*_i, \hat{\mu}_{i,1} = (a_i^2 - a_i + 1)/3 > 0, \hat{\mu}_{i,2} = \hat{\mu}_{i,2} = -\gamma_i, \rho^*_1 = 2\rho^*_2, \rho^*_2 = 2b_i\rho^*_1, \mu^{(2)}_{i,1} = 1, \mu^{(1)}_{i,2} = b_i, \)
where \( \rho^*_k, k = 1, 2, \) are defined in (80), and
\[
a^*_i := -3(\rho^*_1)^2 - 2(a_i + 1)\rho^*_1 - a_i. \quad (86)
\]
Next, we derive a delay-dependent synchronization criterion for system (69) based on Propositions 3.1–3.3, and Theorem 2.6. Recall δ, defined in (5), κi in (7), w in (28), w in (29), L in (72), and ρ in (80) and ρ in (80), and f, g, L, and a in (83)–(86), respectively.

**Theorem 3.4.** System (69) attains ε-approximate synchronization if

\[(a_i^2 - a_i + 1)/3 + c(\kappa_i + w) \bar{L}_g < 0 \quad \text{and} \quad \tau_i < \bar{\tau}_i, \quad \text{for} \quad i = 1, \ldots, N - 1, \quad (87)\]

and the Gauss-Seidel iterations for the linear system (57) converge to the unique solution \( M^{-1}e \) of (57), where \( e \) is as defined in (61), and the entries \( m_{ij} \) of matrix \( M \) and the components \( \epsilon_{ij} \) of vector \( e \) are

\[
m_{ij} = \begin{cases} \frac{1}{3} - \frac{c(\kappa_i + w)}{\bar{L}_g} & \text{if} \quad i = j = \ell(i,1) \\ \gamma_{i,l} - \tilde{\beta}_{i,k}^{(j,l)} & \text{if} \quad i = j = \ell(i,2) \\ 0 & \text{otherwise,} \end{cases} \]

\[
\epsilon_{ij} = \begin{cases} (\hat{a} - \hat{a})[c(\kappa_i + w) + c(\kappa_i + w) \bar{L}_g] & \text{if} \quad \ell = \ell(i, k) \quad \text{and} \quad k = 1 \\ (\hat{b} - \hat{b})c + (\gamma - \tilde{\gamma})c & \text{if} \quad \ell = \ell(i, k) \quad \text{and} \quad k = 2, \end{cases} \]

where

\[
\tilde{\beta}_{i,k}^{(j,l)} := \begin{cases} 1 & \text{if} \quad i = j, \quad (k, l) = (1, 2) \\ b_{ik} & \text{if} \quad i = j, \quad (k, l) = (2, 1) \\ 0 & \text{otherwise,} \end{cases} \]

\[
\tilde{\rho}_{i,k}^{(j,l)} := \begin{cases} c(\kappa_i + w) \bar{L}_g & \text{if} \quad \ell = \ell(i, k) \quad \text{and} \quad k = 1 \\ 0 & \text{otherwise,} \end{cases} \]

\[
\tilde{\beta}_i := -\tilde{\rho}_{i,k}^{(j,l)}(a_i^2 - a_i + 1)/3 + c(\kappa_i + w) \bar{L}_g, \quad \tilde{\beta}_{i,k}^{(j,l)} := \frac{\tilde{\beta}_i(3\tilde{\rho}_i + \tilde{\rho}_w + \tilde{\epsilon}_i, 1)}{\bar{L}_g}, \quad (92) \]

\[
\tilde{\delta}_i := (2|\kappa_i| + \sum_{j \in N \setminus\{i\}} |\bar{w}_{ij}| + \sum_{j \in N \setminus\{i\}} |\bar{w}_{ij}|)cL(f_i^* + c \bar{p}_g \bar{w})\delta_c, \quad (93) \]

and

\[
\tilde{\rho}_i^w := 2c(|\kappa_i + w|)\bar{p}_g, \quad \tilde{\rho}_i^w := 2c(|\kappa_i + w|)\bar{p}_g, \quad (95) \]

**Proof.** From Propositions 3.1–3.3, system (69) satisfies assumptions (D), (E), and (F). Notably, (87) implies that \( \kappa_i + w < 0 \) for all \( i = 1, \ldots, N - 1 \), because \( (a_i^2 - a_i + 1)/3 > 0 \), \( c > 0 \), and \( \bar{L}_g > 0 \). According to Propositions 3.1–3.3, and (81) and (82), the assertion in Proposition 2.3 holds for \( \mu_{i,1} = a_i^*, \mu_{i,2} = a_i^*/(a_i^2 - a_i + 1)/3, \bar{\mu}_{i,2} = \tilde{\beta}_i, \tilde{\beta}_{i,1} = c(\kappa_i + w)L, \tilde{\beta}_{i,1} = c(\kappa_i + w)L, \rho_{i,1}^w = \tilde{\rho}_i^w, \rho_{i,2}^w = 2b_i, \epsilon_{i,k} = \epsilon_{i,k}, \tilde{\beta}_{i,1}^{(j,l)} := \tilde{\beta}_{i,1}^{(j,l)}, \tilde{\beta}_{i,1}^{(j,l)} := \tilde{\beta}_{i,1}^{(j,l)}, \tilde{\beta}_{i,2} = \tilde{\beta}_{i,2} = 0, \) and arbitrary \( \tilde{\beta}_{i,2} > 0 \). Consequently, we observe that the terms \( \eta_{i,k}, \tilde{m}_{i,k}^{(j,l)}, \) and \( \tilde{\epsilon}_{i,k}, \) defined in (46), (47), and Proposition 2.3, respectively, are now

\[
\eta_{i,k} := \begin{cases} -\frac{1}{3} - \frac{c(\kappa_i + w)}{\bar{L}_g} & \text{if} \quad i = j = \ell(i,1) \\ \gamma_{i,l} - \tilde{\beta}_{i,k}^{(j,l)} & \text{if} \quad i = j = \ell(i,2) \\ 0 & \text{otherwise}, \end{cases} \]

\[
\tilde{m}_{i,k}^{(j,l)} := \begin{cases} \mu_{i,k}^{(j,l)} + \tilde{\beta}_{i,k}^{(j,l)} & \text{if} \quad i = j = \ell(i,1) \\ 0 & \text{if} \quad i = j = \ell(i,2) \end{cases} \]

\[
\tilde{\epsilon}_{i,k} := \tilde{\epsilon}_{i,k}, \]

where
and then confirm that the entries $m_{ij}$ of matrix $M$ and the components $\epsilon_\ell$ of vector $e$, defined in (58) and (59), are now (88) and (89), respectively. Moreover, we can verify that system (4) satisfies condition $(S_1)$ under (87). Consequently, by Theorem 2.6, system (69) attains $\varepsilon_\sigma$-approximate synchronization, if the Gauss-Seidel iterations for the linear system (57) converge to $M^{-1}e$, where matrix $M$ and vector $e$ are defined in (88) and (89). The proof is thus completed. \hfill \Box

**Remark 5.** (i) As seen from Theorem 3.4, the entries $m_{ij}$ in matrix $M$ are determined by both of intrinsic dynamics of the individual FitzHugh–Nagumo neurons and the coupling terms. In particular, the terms $\kappa_i$, $\tilde{\omega}_{ij}$, and $\bar{\omega}_{ij}$ are associated with the connection matrix $W$, as seen in (7), (28), and (29).

(ii) We can extend Remark 3.1 in [54] to make the following observations. In general, the criterion in Theorem 3.4 prefers negative $(a_i^2 - a_i + 1)/3 + c(\kappa_i + \tilde{\omega}_{ii})\tilde{\tau}_i$ of large magnitude and small delay $\tau_{ii}$, so that each $-(a_i^2 - a_i + 1)/3 - c(\kappa_i + \bar{\omega}_{ii})\tilde{\tau}_i\beta_i$ is positive and large. In particular, if the delay in system (69) is nonzero, it requires that the coupling strength $c$ be suitably large, not too small nor too large. When there is no delay, condition (87) reduces to the first inequality therein. If $W$ is the zero matrix, i.e., there is no connection between neurons in system (69), then $\kappa_i = 0$ and $\tilde{\omega}_{ii} = 0$ for every $i$, and condition (87) can not hold.

(iii) The first inequality in (87) implies $\kappa_i + \bar{\omega}_{ii} < 0$ for all $i$. Thus, $\kappa_i + \bar{\omega}_{ii} < 0$ for all $i$ is required in the criterion of Theorem 3.4. Accordingly, as seen in Remark 3(ii), the conditions in Theorem 3.4 are more compatible with the situation that all off-diagonal entries of $W$ are nonnegative, or the magnitudes of the negative off-diagonal entries are relatively smaller than those of the positive off-diagonal entries, and all row-sums $\kappa_i$ of $W$ are small enough.

Let us demonstrate Theorem 3.4 by the following example which is a system comprising three FitzHugh–Nagumo neurons. This example aims at demonstrating the dependence of synchronization error on various heterogeneous factors of the considered system, i.e., from disparities of the subsystems in terms of parameter mismatches, of row sums of the connection matrix, and of the coupling delays. We are not aware of any such example or similar in the literature, which addresses the approximate synchronization of coupled systems with heterogeneity from connection matrix or coupling delays.

**Example 1.** Consider three coupled FitzHugh–Nagumo neurons (69) with $a_i = 0.4 + (i - 1) \cdot \delta_p^* \cdot 10^{-4}$, $b_i = 0.01$, $\gamma_i = 0.02$, $i = 1, 2, 3$, $c = 4.2$, $g(\xi) = 30 \tanh(\xi/30)$, and

$$[w_{ij}]_{1 \leq i, j \leq 3} = \begin{pmatrix}
-0.9 & 0.6 & 0.4 \\
0.4 & -0.7 + \delta_k^* \cdot 10^{-4} & 0.4 \\
-0.1 & 0.6 & -0.4 + \delta_k^* \cdot 10^{-4}
\end{pmatrix},$$

$$\tau_{ij} = \begin{cases}
2 \cdot 10^{-5} & \text{if } i = j \in \{1, 2, 3\}, \\
2 \cdot 10^{-5} + \delta_r^* \cdot 10^{-8} & \text{if } i, j \in \{1, 2, 3\}, i \neq j,
\end{cases}$$

where $\delta_p^*$, $\delta_k^*$, $\delta_r^* \geq 0$. We note that $\delta_p^*$, $\delta_k^*$, and $\delta_r^*$ designate the degrees of inhomogeneity from nonidentical subsystems, row-sums of connection matrix, and coupling delays, respectively. It is apparent that three neurons are identical if $\delta_p^* = 0$, the connection matrix has identical row sums if $\delta_k^* = 0$, and the coupling delays are identical if $\delta_r^* = 0$. We shall consider the system with various combinations of $\delta_p^*$, $\delta_k^*$, and $\delta_r^*$. Note that each neuron has the stable equilibrium at the origin as its only attractor when isolated. Restated, all solutions converge to the origin if we consider
the system with \( c = 0 \), instead of \( c = 4.2 \). We shall show that the system with
\((\delta_F, \delta_g, \delta_r) = (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 1)\), satisfies the condition of our synchronization theory in Theorem 3.4, respectively. Moreover, as illustrated in Figures 1(a)-1(e), if \((\delta_F^*, \delta_g^*, \delta_r^*) = (0, 0, 0)\), the system exhibits global identical synchronization with oscillations, whereas if \((\delta_F^*, \delta_g^*, \delta_r^*) = (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 1)\), the system exhibits convergence to oscillations with some synchronization error. That is, this example indicates that all neurons synchronize perfectly if the system comprises identical neurons, and has identical row sums of the connection matrix, and identical coupling delays \((\delta_F^* = \delta_g^* = \delta_r^* = 0)\). On the other hand, the non-perfect synchronization, with synchronization error, shall occur if the system is heterogeneous, due to parameter mismatch of subsystems/neurons \((\delta_F^* \neq 0)\), or nonidentical row sums of the connection matrix \((\delta_g^* \neq 0)\), or nonidentical coupling delays \((\delta_r^* \neq 0)\). In addition, this coupled system exhibits oscillatory behavior which is disparate from the dynamics as the system is uncoupled \((c = 0)\).

Let us explain how we implement our approach and verify that this system satisfies the criterion in Theorem 3.4. We first consider the case \((\delta_F^*, \delta_g^*, \delta_r^*) = (1, 1, 1)\). From (5), (7), (28), (29), (72), and (76), we compute to obtain \(\rho = 30, L = 1, \bar{a} = 0.4002, \bar{a} = 0.4, \bar{b} = \bar{b} = 0.01, \hat{\gamma} = \hat{\gamma} = 0.02, \hat{\omega} = 1.9, \delta_r = 10^{-6}, \kappa_1 = 0.1, \kappa_2 = \kappa_3 = 0.1001\), and

\[
\tilde{W} = [\tilde{w}_{ij}]_{1 \leq i, j \leq 3} = \begin{pmatrix}
-1 & 0.6 & 0.4 \\
0.4 & -0.8 & 0.4 \\
-0.1 & 0.6 & -0.5
\end{pmatrix},
\]

(96)

\[
\bar{W} = [\bar{w}_{ij}]_{1 \leq i, j \leq 2} = \begin{pmatrix}
-1.4 & 0 \\
0.5 & -0.9
\end{pmatrix}.
\]

(97)

By Propositions 3.1, (79), and (80), we obtain \(\bar{q} = 6.6897686713\), and then confirm that the system satisfies assumption (D) with \(Q^* = Q^{(1)} = [-\rho_1^*, \rho_1^{*}] \times [-\rho_2^*, \rho_2^{*}]\), where \(\rho_1^* := \sqrt{x^2 + (\tilde{b}^2/\bar{b})\bar{q}^{(1)})/\hat{\gamma} \approx 34.1112609962, \rho_2^* := b\bar{q}^{(1)}/\hat{\gamma} \approx 3.3448843356\). Subsequently, we compute \(f_1^* \approx 4.1337355475 \cdot 10^4, f_2^* \approx 0.4080102967, \tilde{\rho}_q \approx 24.4024696862\) and \(\tilde{L}_g \approx 0.3383549702\) by (83), (84) and (85), and then obtain \(\bar{\delta}_1 \approx 0.2395378775, \bar{\delta}_2 = 0, \bar{a}_1^* \approx -3.586645911 \cdot 10^3, \bar{a}_2^* \approx -3.586652833 \cdot 10^3\), by Propositions 3.2 and 3.3, respectively. We further compute \(\tilde{\rho}_1^w = 2.6647496897 \cdot 10^2, \tilde{\rho}_2^w = 1.6396409821 \cdot 10^2, \tilde{\rho}_1^w = 6.6897686712, \tilde{\rho}_2^w \approx 1.0918014135 \cdot 10^2, \tilde{\omega}_1 \approx 0.0055819125, \tilde{\omega}_2 \approx 0.048845223, \tilde{\beta}_1 \approx 1.96216019771 \cdot 10^4, \tilde{\beta}_2 \approx 1.2063901770 \cdot 10^4, \tilde{\epsilon}_{1,1} \approx 0.2553688273, \tilde{\epsilon}_{2,1} \approx 0.2444223999, \tilde{\epsilon}_{1,2} = \tilde{\epsilon}_{2,2} = 0, \tilde{\tau}_1 \approx 8.0541599304 \cdot 10^{-5}, \tilde{\tau}_2 \approx 5.9901565126 \cdot 10^{-5}\), from (89) and (92)–(95). Consequently, we can verify that condition (87) holds, and compute \(M\) and \(e\) defined in (58) and (59), expressed by (88) and (89) for system (69), as

\[
M \approx \begin{pmatrix}
1.2016527646 & -1 & 0 & 0 \\
-0.01 & 0.02 & 0 & 0 \\
-2.1 & 0 & 0.6421258855 & -1 \\
0 & 0 & 0 & -0.01 & 0.02
\end{pmatrix},
\]

\(e \approx \begin{pmatrix}
0.2553688273 & 0 \\
0 & 0.2444223999 \\
0 & 0
\end{pmatrix}.
\)

(98)

For \(M\) and \(e\) in (98), we can compute to obtain \(\lambda_{\text{syn}} \approx 0.7786635164\) defined in (63), and \(\varepsilon_e \approx 7.0974002444\) defined in (61), using the sup-norm. Accordingly, the system attains \(\varepsilon_e\)-approximate synchronization, with \(\varepsilon_e \approx 7.0974002444\), by Theorems 2.6 and 3.4.
Similarly, if considering \((\delta^*_p, \delta^*_c, \delta^*_e) = (0, 0, 0)\) (resp., \((1, 0, 0), (0, 1, 0), \) and \((0, 0, 1)\)), we can also verify that the system attains \(\varepsilon_e\)-approximate synchronization, where \(\varepsilon_e = 0\) (resp., \(\varepsilon_e \approx 0.6243646822\), \(\varepsilon_e \approx 0.2157757652\), and \(\varepsilon_e \approx 0.1517520741\), due to \(\lambda_{\text{syn}} \approx 0.7784992773\) (resp., \(\lambda_{\text{syn}} \approx 0.7785278369, \lambda_{\text{syn}} \approx 0.7786349514, \) and \(\lambda_{\text{syn}} \approx 0.7784992773\)).

Recall that, in Proposition 3.1, \(\tilde{q}^{(n)}\), \(n \in \mathbb{N}\), are strictly decreasing with respect to \(n\). Thus, for larger \(n\), \(Q^{(n)}\) provides a smaller attracting region for the dynamics of system \((69)\), which actually yields finer estimation on the synchronization error. For example, if we take \(Q^* = Q^{(5)}\) so that assumption (D) is satisfied, then \(\lambda_{\text{syn}}\) and \(\varepsilon_e\) can be computed as

\[
\begin{align*}
(\delta^*_p, \delta^*_c, \delta^*_e) &= (1, 1, 1) : \lambda_{\text{syn}} \approx 0.5991110279, \varepsilon_e \approx 1.9968422690, \\
(\delta^*_p, \delta^*_c, \delta^*_e) &= (1, 0, 0) : \lambda_{\text{syn}} \approx 0.5990134389, \varepsilon_e \approx 1.8955660626, \\
(\delta^*_p, \delta^*_c, \delta^*_e) &= (0, 1, 0) : \lambda_{\text{syn}} \approx 0.5990880707, \varepsilon_e \approx 0.0610916471, \\
(\delta^*_p, \delta^*_c, \delta^*_e) &= (0, 0, 1) : \lambda_{\text{syn}} \approx 0.5989904858, \varepsilon_e \approx 0.0393798903.
\end{align*}
\]

The real synchronization errors can be seen in Figures 1(a)-1(d), which appear to be much smaller than the theoretical estimates. It is also interesting to see from these computations that the inhomogeneity from the row sums of connection matrix contributes the largest synchronization error. We note that these errors decrease as the magnitudes of the inhomogeneities diminish, according to our theory.

**Remark 6.** (i) As shown in Example 1, theoretical estimation on the synchronization error depends on the estimate of the global attracting set in the phase space, and thus may be coarse. The main goal of this study is to establish approximate synchronization criterion for coupled oscillators, such as the one in Theorem 3.4. The error estimate comes with the theory to confirm that the error decreases as the inhomogeneities of the coupled system decrease. Once the criterion is met for certain parameter and delay values, the actual deviation of the asymptotical state from the synchronization set \(S\) can be seen from numerical simulations.

(ii) As mentioned in the Introduction, the existing works on approximate synchronization commonly considered linear coupling without delays \([18, 36, 59, 66]\). The approximate synchronization of two nonlinearly coupled systems without delay and linearly coupled systems without delay were reported in \([41]\) and \([42]\), respectively. In addition, it is commonly required in those works that the connection matrix satisfies diffusion condition \([18, 36, 41, 59, 66]\), and its off-diagonal entries have the same sign \([18, 36, 59]\). Notably, the off-diagonal entries of connection matrix in our Example 1 have mixed signs and the connection matrix does not satisfy diffusion condition. To the best of our knowledge, synchronization of the system in this example cannot be concluded by the methods employed in those papers, even for the case \(\delta^*_c = \delta^*_e = 0\).

4. **Asymptotic synchronization in coupled Lorenz equations.** We employ the framework of approximate synchronization established in Section 2 to investigate asymptotic synchronization of coupled systems. It will be shown that whether if approximate synchronization yields asymptotic synchronization depends on the coupling structure. We shall illustrate the scenario via a system comprising two diffusively coupled Lorenz equations. Herein, the Lorenz equation we consider was proposed by Lorenz in 1963, cf. \([32]\). First, we show that the synchronization error decreases as the coupling strength increases in the following system
Approximate synchronization in coupled systems

Figure 1. Plot of \((u_1(t), v_1(t))\) for the solution \((u_i(t), v_i(t))_{i=1,2,3}\) of system (69), starting from \((1.3, -1.2, 1.5, -1.5, 1.4, -1.3)\) at \(t_0 = 0\), and evolution of \(Err(t) := \max_{1 \leq i \leq 2} \{\max \{|u_i(t) - u_{i+1}(t)|, |v_i(t) - v_{i+1}(t)|\}\},\) for various \((\delta^p, \delta^s, \delta^\star)\) in Example 1.

\[
\begin{align*}
\dot{x}_{1,1}(t) &= \sigma x_{1,2}(t) - x_{1,1}(t) + c[x_{2,1}(t) - x_{1,1}(t)] \\
\dot{x}_{1,2}(t) &= x_{1,1}(t)[\rho - x_{1,3}(t)] - x_{1,2}(t) + c[x_{2,2}(t) - x_{1,2}(t)] \\
\dot{x}_{1,3}(t) &= x_{1,1}(t)x_{1,2}(t) - \beta_1 x_{1,3}(t) + c[x_{2,3}(t) - x_{1,3}(t)] \\
\dot{x}_{2,1}(t) &= \sigma x_{2,2}(t) - x_{2,1}(t) + c[x_{1,1}(t) - x_{2,1}(t)] \\
\dot{x}_{2,2}(t) &= x_{2,1}(t)[\rho - x_{2,3}(t)] - x_{2,2}(t) + c[x_{1,2}(t) - x_{2,2}(t)] \\
\dot{x}_{2,3}(t) &= x_{2,1}(t)x_{2,2}(t) - \beta_2 x_{2,3}(t) + c[x_{1,3}(t) - x_{2,3}(t)],
\end{align*}
\] (99)

where \(\sigma = 10, \beta_1 = \beta = 8/3, \rho = 28, \beta_2 = \beta + \vartheta\) for some \(\vartheta \geq 0\), and \(c \geq 0\) is the coupling strength. Herein, the magnitude of \(\vartheta\) can be regarded as the degree of heterogeneity in the coupled system (99), and the system is homogeneous if \(\vartheta = 0\), and heterogeneous if \(\vartheta \neq 0\). Each Lorenz equation in (99) admits the well-known
chaotic attractor when isolated (c = 0). System (99) is in the form of (4), with $N = 2$, $K = 3$, $\tau_{ij} = 0$, the connection matrix

$$
W = [\omega_{ij}]_{2 \times 2} = \begin{pmatrix} -c & c \\ c & -c \end{pmatrix},
$$

(100)

and functions $F_{i,k}$, $G_k$ are now

$$
F_{i,1}(\Xi, t) = \sigma(\xi_2 - \xi_1),
$$

(101)

$$
F_{i,2}(\Xi, t) = \xi_1(\rho - \xi_3) - \xi_2,
$$

(102)

$$
F_{i,3}(\Xi, t) = \xi_1\xi_2 - \beta_3\xi_3,
$$

(103)

$$
G_k(\eta) = \eta,
$$

(104)

for $i = 1, 2$, $k = 1, 2, 3$, $\Xi = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3, \eta \in \mathbb{R}$. We can find a globally attracting set for system (99) which is independent of the coupling strength $c$:

**Proposition 4.1.** For any $c > 0$, all solutions of system (99) eventually enter, and then remain in the compact set $\hat{Q}^{(c)} \times \check{Q}^{(c)}$, where

$$
\hat{Q}^{(c)} := [\tilde{R}_1^{(c)}, \hat{R}_1^{(c)}] \times [\tilde{R}_2^{(c)}, \hat{R}_2^{(c)}] \times [\tilde{R}_3^{(c)}, \hat{R}_3^{(c)}] \subset \mathbb{R}^3,
$$

with $\tilde{R}_k^{(c)} := \hat{R}_k - \varsigma$, $k = 1, 2, \hat{R}_3^{(c)} := \hat{R}_3 - \varsigma$, and for $k = 1, 2$,

$$
\tilde{R}_k := -\hat{R}_k := \sqrt{2[(\hat{r}_1)^2 + (\hat{r}_2)^2 + (\max\{|\hat{r}_3 - (\sigma + \rho)|, |\hat{r}_3 - (\sigma + \rho)|\})^2]},
$$

$$
\hat{R}_3 := \sqrt{2[(\hat{r}_1)^2 + (\hat{r}_2)^2 + (\max\{|\hat{r}_3 - (\sigma + \rho)|, |\hat{r}_3 - (\sigma + \rho)|\})^2] + \sigma + \rho},
$$

with

$$
\hat{r}_1 := (\sigma + \rho)(\beta + \vartheta)\sqrt{1/(2\beta\sigma)}, \hat{r}_2 := (\sigma + \rho)(\beta + \vartheta)\sqrt{1/(2\beta)},
$$

$$
\hat{r}_3 := (1 + \sqrt{2})(\beta + \vartheta)(\sigma + \rho)/(2\beta), \check{r}_3 := |\beta - \sqrt{2}(\beta + \vartheta)|(\sigma + \rho)/(2\beta).
$$

The proof of this proposition is arranged in Appendix D. In what follows, we consider system (99) which satisfies assumption (D), with $N = 2$, $K = 3$, and $Q = \hat{Q}^{(c)}$, for an arbitrarily fixed $c > 0$. With (104), $\hat{G}_k$, $\hat{D}_k$, and $\check{D}_k$ defined in (14)-(16) are now

$$
\hat{G}_k = \hat{R}_1^{(c)} - \hat{G}_k, \hat{D}_k = \check{D}_k = 1, k = 1, 2, 3.
$$

(105)

**Proposition 4.2.** For any $c > 0$, system (99) satisfies assumption (E) with $\tilde{\delta}_1 = \tilde{\delta}_2 = 0$, and $\tilde{\delta}_3 = \vartheta \cdot \tilde{R}_3^{(c)}$, where $\tilde{R}_3^{(c)}$ is defined in Proposition 4.1.

**Proof.** Based on (22) and (101)-(103), and recalling $\beta_1 - \beta_2 = -\vartheta$, a computation shows $|F_{1,k}(\Xi, t) - F_{2,k}(\Xi, t)| = 0$ for $k = 1, 2$, and

$$
|F_{1,3}(\Xi, t) - F_{2,3}(\Xi, t)| = |(\beta_1 - \beta_2) \cdot \xi_3| \leq \vartheta \cdot \max\{|\tilde{R}_3^{(c)}|, |\check{R}_3^{(c)}|\} = \vartheta \cdot \check{R}_3^{(c)},
$$

for $\Xi = (\xi_1, \xi_2, \xi_3) \in \hat{Q}^{(c)}$, where $\hat{Q}^{(c)}$ is defined in Proposition 4.1. The assertion thus follows.

**Proposition 4.3.** For any $c > 0$, system (99) satisfies assumption (F) with $\tilde{\mu}_{1,k} = \tilde{\mu}_{1,k} = -\sigma, -1, -\beta$ for $k = 1, 2, 3$, respectively, and

$$
\rho_{\mu}^{(c)} = \begin{cases} 
2\sigma \hat{R}_1^{(c)} & \text{if } k = 1, \\
2\hat{\rho}^{(c)} \hat{R}_1^{(c)} + \hat{R}_1^{(c)}[\check{R}_3^{(c)} - \tilde{R}_3^{(c)}] & \text{if } k = 2, \\
4\hat{R}_1^{(c)} \hat{R}_2^{(c)} & \text{if } k = 3,
\end{cases}
$$
where \( \bar{\rho}^{(c)} := \max\{|\rho - \hat{R}_3^{(c)}|, |\rho - \tilde{R}_3^{(c)}|, \} \), and \( \hat{R}_1^{(c)}, \hat{R}_2^{(c)}, \tilde{R}_3^{(c)} \), and \( \hat{R}_3^{(c)} \) are defined in Proposition 4.1.

Proof. Based on (101)-(103), and recalling \( \beta_1 = \beta \), a direct computation yields

\[
F_{1,k}(\Xi, t) = F_{1,k}(\tilde{\Xi}, t) = f_{1,k}(\xi_k, \psi_k, t) + u_{1,k}(\Xi, \tilde{\Xi}, t), \quad k = 1, 2, 3,
\]

for \( \Xi = (\xi_1, \xi_2, \xi_3), \tilde{\Xi} = (\psi_1, \psi_2, \psi_3) \in \mathbb{R}^3 \), where

\[
f_{1,k}(\xi, \eta, t) = \begin{cases} -\sigma(\xi - \eta) & \text{if } k = 1, \\ -(\xi - \eta) & \text{if } k = 2, \\ -\beta(\xi - \eta) & \text{if } k = 3, 
\end{cases}
\]

\[
u_{1,k}(\Xi, \tilde{\Xi}, t) = \begin{cases} \sigma(\xi_2 - \eta_2) & \text{if } k = 1, \\ (\rho - \eta_3)(\xi_1 - \eta_1) - \xi_1(\xi_3 - \eta_3) & \text{if } k = 2, \\ \xi_1(\xi_2 - \eta_2) + \eta_2(\xi_1 - \eta_1) & \text{if } k = 3. 
\end{cases}
\]

Consider \( \Xi = (\xi_1, \xi_2, \xi_3), \tilde{\Xi} = (\eta_1, \eta_2, \eta_3) \in \tilde{Q}^{(c)}, \) where \( \tilde{Q}^{(c)} \) is defined in Proposition 4.1. Then

\[
|\rho - \eta_3| \leq \bar{\rho}^{(c)} \text{ and } \hat{R}_k^{(c)} \leq \xi_k, \eta_k \leq \tilde{R}_k^{(c)}, \quad k = 1, 2, 3.
\]

Recall \( \hat{R}_k^{(c)} = -\tilde{R}_k^{(c)} < 0, \quad k = 1, 2. \) The assertion follows from (106) and (107).

Let

\[
M^* = \begin{pmatrix} \sigma + 2c & -\sigma & 0 \\ -\bar{\rho} & 1 + 2c & -\hat{R}_1 \\ -\hat{R}_2 & -\hat{R}_1 & \beta + 2c \end{pmatrix}, \quad e^* = \vartheta \cdot \begin{pmatrix} 0 \\ 0 \\ \hat{R}_3 \end{pmatrix},
\]

and

\[
q_1(c) := 2(\hat{R}_1 + \bar{\rho}c),
\]

\[
q_2(c) := 32\sigma \hat{R}_1 \hat{R}_2 c^3 + 4[\sigma^2 \bar{\rho}^2 + \hat{R}_1] + 2\sigma \bar{\rho} \hat{R}_1^2 + 4\sigma(\beta + 1 + \sigma)\hat{R}_1 \hat{R}_2 c^2
+ 4\sigma(\bar{\rho}^2 \beta + \hat{R}_1^2 + \beta + 2)\hat{R}_1^2 + 2(\beta + \sigma + \sigma)\hat{R}_1 \hat{R}_2 c^2
+ 16\bar{\rho}^2 \beta^2 + 2\bar{\rho} \hat{R}_2^2 + 4\beta \hat{R}_1 \hat{R}_2 \hat{R}_1^2)^{1/2},
\]

\[
q_3(c) := 16c^3 + 8(1 + \beta + \sigma)c^2 + 4(\beta + 2)\sigma^2 c + 2\beta\sigma c,
\]

where

\[
\bar{\rho} := \max\{|\rho - \hat{R}_3|, |\rho - \tilde{R}_3|\}
\]

and \( \hat{R}_1, \hat{R}_2, \) and \( \tilde{R}_3 \) are defined in Proposition 4.1.

**Theorem 4.4.** System (99) attains \( \varepsilon_{\vartheta,c} \)-approximate synchronization if \( c > c_{\vartheta} := \max\{c_1^+, c_2^+\} \), where \( \varepsilon_{\vartheta,c} := \|M^*\|^{-1}e^*\|, \) \( c_1^+, c_2^+ \) are the largest solutions to \( |\lambda_{\pm}(c)| = 1 \), and functions \( \lambda_{\pm} \) are defined by

\[
\lambda_{\pm}(c) := |\sigma(\hat{R}_2^2 + \bar{\rho} \beta) + q_1(c) \pm q_2(c)|/q_3(c).
\]

In addition, system (99) attains asymptotic synchronization, i.e., \( \varepsilon_{\vartheta,c} \to 0 \), as \( c \to \infty \).
Proof. For matrix $W$ in (100), the corresponding $\kappa_i$ defined in (7) and matrix $\check{W}$ defined in (29) are now

$$\kappa_i = 0, \ i = 1, 2, \text{ and } \check{W} = [\bar{G}_{11}] = [-2c].$$  \hspace{1cm} (111)

In addition, $\delta_r = 0$ since $\tau_{ij} \equiv 0$. According to Propositions 4.1, 4.2, and 4.3, the system satisfies assumptions (D), (E), and (F). Accordingly, by (105) and (111), and $\delta_r = 0$, we can choose $\hat{\mu}_{1,k}, \check{\mu}_{1,k}, \hat{\beta}_{1,k}, \check{\beta}_{1,k}, \check{\epsilon}_{1,k}$ introduced in Proposition 2.3 as follows:

$$\hat{\mu}_{1,k} = \check{\mu}_{1,k}, \hat{\beta}_{1,k} = \check{\beta}_{1,k} = -2c, \check{\epsilon}_{1,k} = \check{b}_k,$$

where $\check{b}_k$ and $\check{\mu}_{1,k}$ are determined in Propositions 4.2 and 4.3, respectively. Recall that matrix $M$ and vector $e$, defined in (58) and (59), are defined via $\eta_{1,k}, \bar{m}_{1,k}^{(1,l)}$, and $\check{\epsilon}_{1,k}$. In addition, we can determine $\eta_{1,k}$ and $\bar{m}_{1,k}^{(1,l)}$ which are defined in (46) and (47) respectively via quantities determined in Proposition 2.3. Consequently, with (112) as well as $\check{\mu}_{1,k}$ determined in Propositions 4.3, and $\tau_{ij} \equiv 0$, we can determine $\eta_{1,k}$ and $\bar{m}_{1,k}^{(1,l)}$, and hence matrix $M$ and vector $e$:

$$M = M_c := \begin{pmatrix}
\eta_{1,1} & -\bar{m}_{1,1}^{(1,2)} & -\bar{m}_{1,1}^{(1,3)} \\
-\bar{m}_{1,1}^{(1,2)} & \eta_{1,2} & -\bar{m}_{1,2}^{(1,3)} \\
-\bar{m}_{1,1}^{(1,3)} & -\bar{m}_{1,2}^{(1,3)} & \eta_{1,3}
\end{pmatrix} = \begin{pmatrix}
\sigma + 2c & -\sigma & 0 \\
-\bar{\rho}^{(c)} & 1 + 2c & -\bar{R}_{1}^{(c)} \\
-\bar{R}_{2}^{(c)} & -\bar{R}_{1}^{(c)} & \beta + 2c
\end{pmatrix},$$

$$e = e_c = \begin{pmatrix}
\check{\epsilon}_{1,1} \\
\check{\epsilon}_{1,2} \\
\check{\epsilon}_{1,3}
\end{pmatrix} = \partial \cdot \begin{pmatrix}
0 \\
0 \\
\check{R}_{3}^{(c)}
\end{pmatrix}.$$  \hspace{1cm} (113)

In addition, condition (S1) holds, since $\hat{\mu}_{1,k} + \check{\beta}_{1,k} = -\eta_{1,k} < 0$ for all $k = 1, 2, 3$, and $\tau_{ij} \equiv 0$. Consequently, by Theorem 2.6, system (99) achieves $\varepsilon_{\hat{\rho},c}^{(c)}$–approximate synchronization if the Gauss-Seidel iteration for solving

$$M_c v = e_c,$$

converges to the unique solution of (113), where $\varepsilon_{\hat{\rho},c}^{(c)} := \|M_c^{-1}e_c\|$. For system (113), the eigenvalues of the corresponding matrix $(P_{M_c} - L_{M_c})^{-1}U_{M_c}$, cf. (63), are

$$\lambda_1(c) = 0, \ \lambda_{\pm}^{(c)}(c) = \{\sigma([\bar{R}_{1}^{(c)}])^2 + \bar{\rho}^{(c)} \beta) + q_1^{(c)}(c) \pm q_2^{(c)}(c)/q_3(c),$$

where

$$q_1^{(c)}(c) = 2([\bar{R}_{1}^{(c)}])^2 + \bar{\rho}^{(c)} \sigma |c|,$$

$$q_2^{(c)}(c) = \{32\sigma \bar{R}_{1}^{(c)} \bar{R}_{2}^{(c)} c^3 + 4[\sigma^2(\bar{\rho}^{(c)})^2] + (\bar{R}_{1}^{(c)})^4 + 2\sigma \bar{\rho}^{(c)}(\bar{R}_{1}^{(c)})^2 + 4\sigma(\bar{\rho}^{(c)})^2 \beta + (\bar{R}_{1}^{(c)})^4 + \bar{\rho}^{(c)}(\beta + \sigma)(\bar{R}_{1}^{(c)})^2 + 2(\beta + \sigma \beta + \sigma)\bar{R}_{1}^{(c)} \bar{R}_{2}^{(c)} c + \sigma^2(\bar{\rho}^{(c)})^2 \beta^2 + 2\bar{\rho}^{(c)} \beta (\bar{R}_{1}^{(c)})^2 + 4\beta \bar{R}_{1}^{(c)} \bar{R}_{2}^{(c)} + (\bar{R}_{1}^{(c)})^4\}^{1/2}.$$  \hspace{1cm} (116)

Since $|\lambda_{\pm}^{(c)}(0)| > 1$ and $|\lambda_{\pm}^{(c)}(c)| \to 0$, as $c \to \infty$, there exist largest solutions, say $c_{\hat{\rho},c}^{\pm} > 0$, to $|\lambda_{\pm}^{(c)}(c)| = 1$. Hence, for any $c > 0$, system (99) attains $\varepsilon_{\hat{\rho},c}^{(c)}$–approximate synchronization if $c > c_{\hat{\rho},c}^{\pm} := \max\{c_{\hat{\rho},c}^{+(c)}, c_{\hat{\rho},c}^{-(c)}\}$. Notice that $c_{\hat{\rho},c}^{\pm} \to c_{\hat{\rho},c}^{\pm} \to c_{\hat{\rho},c}$ as $c \downarrow 0$, where $c_{\hat{\rho},c}^{\pm}$, $c_{\hat{\rho},c}$, and $\varepsilon_{\hat{\rho},c}$ are well defined since $|\lambda_{\pm}(0)| > 1$ and $|\lambda_{\pm}(c)| \to 0$, as $c \to \infty$. Accordingly, system (99) attains $\varepsilon_{\hat{\rho},c}$–approximate synchronization if $c > c_{\hat{\rho},c}^{\pm}$. It can be seen that $M^* = \lim_{c \to 0} M_c$, 

\hspace{1cm} (117)
that demonstrates the evolution of function $\vartheta$ and $c$.

Obviously, $\vartheta$ is computed through $\vartheta \approx \rho_{c}$, which yields $\vartheta_{c} := \|M^{*}\|^{-1} \epsilon^{*} \| \rightarrow 0$, as $c \rightarrow \infty$. Accordingly, system (99) attains asymptotic synchronization.

**Example 2.** Consider system (99) with $\vartheta = 1$. By Proposition 4.1, (109), and Theorem 4.4, it can be computed that $R_{1} \approx R_{2} \approx 119.4494$, $R_{3} \approx 157.4494$, and $\tilde{\rho} \approx 129.4494$, and functions $\lambda_{\pm}(c)$ are then determined. Subsequently, we can compute to find $c_{\vartheta} \approx 63.3819$. Herein, we stress that $c_{\vartheta}$ is computed through the analytical result in Theorem 4.4, not through a numerical computation. Thus system (99) attains approximate synchronization when $c > c_{\vartheta}$ and asymptotic synchronization, by Theorem 4.4. Figure 2 demonstrates the evolution of function $\text{Err}(t) := \max_{1 \leq k \leq 3} \{|x_{1,k}(t) - x_{2,k}(t)|\}$, for solutions $(x_{i,j}(t))$ of (99) with $\vartheta = 1$, and $c = 64$, $c = 640$, and $c = 6400$, respectively. Herein, $\text{Err}(t)$ is referred to as the synchronization error for the corresponding solution. It appears in Figure 2 that the synchronization error decreases as the coupling strength increases, and thus strong coupling strength suppresses the synchronization error. We note that the system appears to exhibit chaotic behavior in all cases in Figure 2.

However, approximate synchronization does not always yield asymptotic synchronization, as shown in the following example.

\[\text{Err}(t) := \max_{1 \leq k \leq 3} \{|x_{1,k}(t) - x_{2,k}(t)|\}\]
Example 3. Consider two Lorenz equations which are coupled through their first two components:

\[
\begin{aligned}
\dot{x}_{1,1}(t) &= \sigma[x_{1,2}(t) - x_{1,1}(t)] + c[x_{2,1}(t) - x_{1,1}(t)] \\
\dot{x}_{1,2}(t) &= x_{1,1}(t)[\rho - x_{1,3}(t)] - x_{1,2}(t) + c[x_{2,2}(t) - x_{1,2}(t)] \\
\dot{x}_{1,3}(t) &= x_{1,1}(t)x_{1,2}(t) - \beta x_{1,3}(t) \\
\dot{x}_{2,1}(t) &= \sigma[x_{2,2}(t) - x_{2,1}(t)] + c[x_{1,1}(t) - x_{2,1}(t)] \\
\dot{x}_{2,2}(t) &= x_{2,1}(t)[\rho - x_{2,3}(t)] - x_{2,2}(t) + c[x_{1,2}(t) - x_{2,2}(t)] \\
\dot{x}_{2,3}(t) &= x_{2,1}(t)x_{2,2}(t) - \beta x_{2,3}(t),
\end{aligned}
\]  

(114)

where \(\sigma = 10, \beta = 8/3, \rho = 28, \hat{\sigma} = \sigma + \vartheta, \hat{\beta} = \beta + \vartheta, \hat{\rho} = \rho + \vartheta\) for \(\vartheta \geq 0\) and \(c \geq 0\). Obviously, (114) is homogeneous if \(\vartheta = 0\), and is heterogeneous if \(\vartheta \neq 0\).

Via arguments similar to Theorem 4.4 as considering system (99), we can show that system (114) achieves \(\tilde{\varepsilon}_{\vartheta,c}\)-approximate synchronization if \(c > \tilde{c}_\vartheta\), where

\[
\tilde{\varepsilon}_{\vartheta,c} := \|\tilde{M}^{-1}\tilde{e}\|, \quad \text{with} \quad \tilde{M} = \begin{pmatrix} \sigma + 2c & -\sigma & 0 \\ -\hat{\rho} & 1 + 2c & -\hat{R}_1 \\ -\hat{R}_2 & -\hat{R}_1 & \beta \end{pmatrix}, \quad \tilde{e} = \vartheta \cdot \begin{pmatrix} \hat{R}_1 + \hat{R}_2 \\ \hat{R}_1 \\ \hat{R}_3 \end{pmatrix}.
\]

(115)

Herein, \(\tilde{c}_\vartheta, \tilde{\rho}, \text{ and } \hat{R}_i > 0, i = 1, 2, 3\), can be derived similarly as those for system (99) (not shown herein). By a direct computation, we obtain

\[
\tilde{\varepsilon}_{\vartheta,c} \to \vartheta \hat{R}_3/\beta =: \tilde{\varepsilon}_{\vartheta}^* \Rightarrow \varepsilon_{\vartheta,c} \to \vartheta \hat{R}_3/\beta =: \tilde{\varepsilon}_{\vartheta}^*, \quad \text{as } c \to \infty,
\]

(116)

where \(\tilde{\varepsilon}_{\vartheta}^* > 0\) if \(\vartheta > 0\). This indicates that the estimated bound of synchronization error \(\tilde{\varepsilon}_{\vartheta,c}\) converges to a positive number, and hence does not drop to zero by increasing coupling strength \(c\). If \(\vartheta = 1\), then a computation shows \(\tilde{c}_\vartheta \approx 3013.4427\).

Figure 3 demonstrates the evolution of \(\text{Err}(t)\) for some solutions of (114) with \(\vartheta = 1\) and \(c = 3014, c = 6028, \text{ and } c = 12056\), respectively. It can be seen that the synchronization errors remain at substantial levels for large \(c\).

As seen from the approach in [17], an important and necessary concern for the study of asymptotic synchronization in coupled systems with parameter mismatch is finding a uniform attracting region which is independent of the coupling strength. Finding an estimate of the attracting region for coupled Lorenz equations with parameter variation is actually a challenging task, especially for those coupled through all components, such as system (99). The asymptotic synchronization of two nearly identical Lorenz equations coupled through their first components was discussed in [17], and the work concerning the existence of a uniform attracting region for Lorenz equations with parameter variation coupled through their first components can be found in [41, 42]. In this section, we studied the approximate synchronization of system (99) which is coupled in three components and system (114) which is coupled in two components, and discussed the possibility for them to achieve asymptotic synchronization. We observe that a main difference between systems (99) and (114), which results in whether the asymptotic synchronization holds or not, is about the coupling configuration. The coupling is full in system (99), but there is no coupling between variables \(x_{1,3}\) and \(x_{2,3}\) in system (114). The consequence is that the diagonal entry in the third row of \(M^*\) in (108) is \(\beta + 2c\), whereas the diagonal entry in the third row of \(\tilde{M}\) in (115) is \(\beta\), where \(c\) is the coupling strength. It can be observed that \(\tilde{\varepsilon}_{\vartheta,c}\) in (116) converges to zero if this entry \(\beta\) is replaced by \(\beta + 2c\).
5. Discussion and conclusion. We have established a framework to study approximate synchronization of coupled systems. The framework accommodates a large class of coupled systems and network systems under general coupling. The units which constitute the coupled systems can be identical or nonidentical. The nonidentical ones include possible cases of slightly different units, units with parameter mismatch, or completely disparate units. Communication delay was also taken into account. Under this setting, there naturally exists a synchronization error when the synchronization criterion is met. We have shown that the synchronization error depends on and decreases with decreasing difference between subsystems, difference between row sums of connection matrix, and difference of communication delays between components, as summarized in Remark 4. When these differences decrease to zero, approximate synchronization reduces to identical synchronization for coupled systems comprising identical subsystems, with identical row sums of connection matrix and identical communication delays. Among these inhomogeneous factors, the one from row sums of connection matrix is less obvious. In fact, identical row sums of connection matrix carries a sense of balance for total coupling weights for each unit from all units, and thus plays an important role in the coupling structure.

The present theory provides both delay-dependent and delay-independent criteria for approximate synchronization. The sufficient conditions depend on estimations of the globally attracting region for a coupled system. Finer estimates on the attracting region lead to sharper estimates on the synchronization error and weaker criterion for synchronization. In addition, in order to study asymptotic synchronization, one would expect to obtain an estimate on the attracting set, which is independent of the coupling strength, such as the coupled Lorenz systems in Example 2. However, rigorous confirmation of dissipativeness and estimation of attracting set in nonlinear systems are nontrivial mathematical tasks in general. Many of the existing results on synchronization and/or asymptotic behaviors of dynamical systems, via various mathematical treatments, also rely on such a property. As analytic estimations on globally attracting sets are difficult to obtain in general, examinations of the
synchronization criteria largely rely on numerical computations. Those numerical computations are generally executed after assigning parameter values and choosing some (finitely many) initial points, and hence are not mathematically sufficient. In this paper, we were able to confirm the dissipative property, and obtained estimations of the globally attracting sets for coupled systems comprising nonidentical FitzHugh-Nagumo neurons under sigmoidal (nonlinear) coupling and coupled systems comprising parameter-mismatched Lorenz equations under linearly diffusive coupling.

The approach in this paper can be applied to establish approximate synchronization for coupled systems under quite general coupling, such as (2), and even the ones not in the form of (2) or (4). For instance, we can also establish approximate synchronization for the kinetic models on gene regulatory networks of zebrafish segmentation clocks, which are under very nonlinear coupling with transcriptional and translational delays, cf. [10, 31]. The key point for such an application relies on suitable manipulation of the associated difference-differential equations. This has to be formulated according to the coupling configuration. More precisely, it hinges upon recomposing the difference-differential equations, such as (23) in this paper, into a form, such as (34), for which assertions analogous to the ones in Proposition 2.3 can be formulated. However, such formulation is quite cumbersome.

As mentioned in the Introduction, in previous studies of nonidentical synchronization, the coupled systems considered were under linear diffusive coupling, and without delays or with identical delays. The factors contributed to synchronization error were only partially uncovered. In the present investigation, via a new approach “sequential contracting”, we were able to establish a framework to tackle the problem on nonidentical synchronization for coupled systems in general form. In a systematic way, we have exploited the structure behind the synchronization, including the causes of synchronization error. With this approach, some common restrictions in the literature have been lifted, as mentioned in Remark 6. Furthermore, for practical applications, all conditions in the present theory can be examined in a straightforward fashion or by numerical algebraic computation. Significantly, with this framework, the general theory for synchronization has attained a completeness and the scope on the study of synchronization is expected to be further expanded. For example, one can see how various factors, including the intrinsic terms of the individual subsystems, the coupling structure, coupling strength, and coupling delays play their roles in synchronization, such as the ones shown for coupled FitzHugh-Nagumo neurons in Theorem 3.4, Remark 5, and Example 1. Based on the present framework, we also extended our analysis to investigate asymptotic synchronization for coupled Lorenz equations and illustrated that for some coupling schemes, the synchronization error decreases as the coupling strength increases, whereas for some other coupling schemes, synchronization error can not be suppressed by increasing the coupling strength, as shown in Examples 2 and 3. This has clarified an issue which was addressed, but not settled, in the literature. From the view point of application, the factors that we found can be used to control the synchronization error in network systems. All in all, the exploration in this work has enhanced and expanded the understanding of synchronization and the dynamics of coupled systems.
6. **Appendices.** We outline the derivation for the entries of the transformed connection matrix in Appendix A, and summarize the properties for the scalar equation (54) in Appendix B. We estimate the globally attracting sets for the coupled FitzHugh-Nagumo system and coupled Lorenz equations comprising subsystems with mismatched parameters, in Appendices C and D, respectively.

**Appendix A.** Computation for (31): Let $\hat{W}_i = (\hat{w}_{i1}, \ldots, \hat{w}_{iN}) \in \mathbb{R}^{1 \times N}$, $i \in \mathcal{N}$, and

$$
T = \begin{pmatrix}
1 & 1 & \cdots & 1 \\
0 & 1 & \cdots & \cdots \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & 0 & 1
\end{pmatrix} \in \mathbb{R}^{N \times N}.
$$

Then,

$$
C\hat{W} = \begin{pmatrix} \hat{W}_1 - \hat{W}_2 \\ \vdots \\ \hat{W}_{N-1} - \hat{W}_N \end{pmatrix} \quad \text{and} \quad CT = \begin{pmatrix} I_{N-1} & 0 \end{pmatrix},
$$

(117)

where $I_{N-1}$ is the $(N-1) \times (N-1)$ identity matrix and $0 = (0, \ldots, 0)^T \in \mathbb{R}^{(N-1) \times 1}$. From (30) and (117), it follows that $C\hat{W}T = \hat{W}CT$, and hence

$$
\begin{pmatrix}
\hat{W}_1 - \hat{W}_2 \\
\vdots \\
\hat{W}_{N-1} - \hat{W}_N
\end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\
0 & 1 & \cdots & \cdots \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & 0 & 1
\end{pmatrix} \begin{pmatrix} \hat{W} \\
0 \end{pmatrix},
$$

which implies that for $i = 1, \ldots, N - 1$,

$$
\hat{w}_{ii} = \sum_{j=1}^{i} \hat{w}_{ij} - \sum_{j=1}^{i} \hat{w}_{i(1+j)} = -\left( \sum_{j=i+1}^{N} \hat{w}_{ij} + \sum_{j=1}^{i} \hat{w}_{i(1+j)} \right),
$$

recalling $\sum_{j=1}^{N} \hat{w}_{ij} = 0$.

**Appendix B.** Properties for scalar equation (54): Let $w(t)$ be a bounded continuous function for $t \geq t_0$, and let $h : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\hat{h} : \mathcal{C}([-\tau_M, 0]; \mathbb{R}) \times \mathcal{C}([-\tau_M, 0]; \mathbb{R}) \times \mathbb{R} \to \mathbb{R}$ be continuous functions. Let $x_t, y_t \in \mathcal{C}([-\tau_M, 0]; \mathbb{R})$ for $t \geq t_0$, and set $x(t + \theta) = x_t(\theta)$, $y(t + \theta) = y_t(\theta)$ for $\theta \in [-\tau_M, 0]$. We assume that $x(t)$ and $y(t)$ eventually enter and then remain in some closed and bounded interval $[\bar{q}, \tilde{q}]$; namely, $x(t)$ and $y(t)$ lie in $[\bar{q}, \tilde{q}]$ for all $t \geq \tilde{t}_0$, for some $\tilde{t}_0 \geq t_0$. We suppose that $z(t) = x(t) - y(t)$ satisfies the following scalar equation:

$$
\dot{z}(t) = h(x(t), y(t), t) + \hat{h}(x_t, y_t, t) + w(t), \quad t \geq t_0.
$$

(118)

Set

$$
|w|^{\text{max}}(T) := \sup\{|w(t)| : t \geq T\}, \quad |w|^{\text{max}}(\infty) := \lim_{T \to \infty} |w|^{\text{max}}(T).
$$

We introduce the following conditions:

Condition (H0): There exist $\hat{\mu}, \hat{\mu}, \hat{\beta}, \hat{\beta} \in \mathbb{R}$, $\rho^h > 0$, and $0 \leq \bar{\tau} \leq \tau_M$ such that the following properties hold for all $t \geq t_0$:

\begin{align*}
\text{(H0-i):} & \quad \begin{cases} 
\hat{\mu} \leq \hat{h}(\xi, \eta, t)(\xi, \eta) \leq \hat{\mu} & \text{if } \xi \neq \eta \neq \bar{q},
\hat{h}(\xi, \eta, t) = 0 & \text{if } \xi = \eta = \bar{q}.
\end{cases}
\end{align*}

For $\xi, \eta \in [\bar{q}, \tilde{q}]$,
Proposition 6.2. If \( t(1) \) converges to \( \Omega \)

\[ \Omega(1) \]

Appendix C.

Proposition 6.1. If \( z(t) \) satisfies \((118)\), then \( z(t) \) converges to interval \([-\bar{p}, \bar{p}]\) as \( t \to \infty \), under conditions \((H_0)\) and \((A_1)\). Moreover,

\[ 0 \leq \bar{p} \leq |u|_{\infty}/[\bar{\mu} - \bar{\beta}] \].

Proposition 6.2. If \( z(t) \) satisfies \((118)\), then \( z(t) \) converges to interval \([-\bar{p}, \bar{p}]\) as \( t \to \infty \), under conditions \((H_0)\) and \((A_2)\). Moreover,

\[ 0 \leq \bar{p} \leq |u|_{\infty}/[-\bar{\mu} - \bar{\beta}] \].

Appendix C. Proof of Proposition 3.1: The derivation is modified from the arguments in Theorem 3.1 of [54], where a coupled system comprising identical FitzHugh–Nagumo neurons was considered. Let \((u_1(t), v_1(t), \ldots, u_N(t), v_N(t))\) be an arbitrary solution of \((69)\). We first show that \((u_i(t), v_i(t))\) converges to \(Q^{(1)}\), for each \(i \in \mathcal{N}\). For \(i \in \mathcal{N}\), by setting \(V_i(t) = V_i(u_i(t), v_i(t)) := u_i^2(t)/2 + v_i^2(t)/(2b_i)\), we obtain

\[ V_i(t) = -u_i^4 + (\tilde{a} + 1)u_i^3 - a_iu_i^2 + c(\sum_{j \in \mathcal{N}} \omega_{ij}g(u_{ij}^\tau))u_i - (\gamma_i/b_i)v_i^2 \]

\[ \leq P(1)(u_i) - (\gamma_i/b_i)v_i^2, \tag{119} \]

for all \(t \geq t_0\), recalling \(|g(u_{ij}^\tau)| < \rho\), where \(u_i = u_i(t), v_i = v_i(t)\), and \(u_{ij} = u_j(t-\tau_{ij})\), for \(i, j \in \mathcal{N}\), and \(P(1)\) is defined in \((77)\). We define

\[ \Omega_i^{(1)} := \{(\xi, \zeta) \in \mathbb{R}^2 : \xi^2 + \zeta^2/b_i \leq [1 + \tilde{b}^2/(b_i\gamma^2)](q^{(1)})^2 \text{ if } |\xi| > q^{(1)} \} \]

\[ \Omega^{(1)} := \{(\xi, \zeta) \in \mathbb{R}^2 : \xi^2 + \zeta^2/b \leq [1 + \tilde{b}^2/(b\gamma^2)](q^{(1)})^2 \text{ if } |\xi| > q^{(1)} \} \]

Notably, \(\Omega_i^{(1)} \subseteq \Omega^{(1)} \subseteq Q^{(1)}\), for all \(i \in \mathcal{N}\).

With \((119)\), applying the arguments similar to those for the proof of Theorem 3.1 in [54] yields that \(\Omega_i^{(1)}\) is positively invariant for \((u_i(t), v_i(t))\), and \((u_i(t), v_i(t))\) converges to \(\Omega_i^{(1)}\), and hence \(Q^{(1)}\), as \(t \to \infty\), for all \(i \in \mathcal{N}\). Furthermore, each \((u_i(t), v_i(t))\) converges to \(\hat{\Omega}^{(1)}\), where

\[ \hat{\Omega}^{(1)} := \{(\xi, \zeta) \in \mathbb{R}^2 : \xi^2 + \zeta^2/b \leq [1 + \tilde{b}^2/(b\gamma^2)](q^{(1)})^2 \text{ if } |\xi| > q^{(1)} \} \]

with \(q^{(1)} := \max\{|\xi| : \hat{P}^{(1)}(\xi) = 0\}\), and \(\hat{P}^{(1)}(\xi) := -\xi^4 + (\tilde{a} + 1)|\xi|^3 -\tilde{a}\xi^2 + c\omega\beta|\xi|\).

Notably, \(\hat{\Omega}^{(1)} \subseteq \Omega^{(1)}\) and \(q^{(1)} < q^{(1)}\) because \(P^{(1)}(0) = P^{(1)}(0) , P^{(1)}(\xi) < P^{(1)}(\xi)\) if \(\xi \neq 0\). Consequently, \((u_i(t), v_i(t))\) eventually enters and then remains in \(\Omega^{(1)}\), and
Synchronization of nearly identical dynamical systems: Size

Appendix D. Proof of Proposition 4.1: For $X = (X_1, X_2) \in \mathbb{R}^6$, with $x_i = (x_{i,1}, x_{i,2}, x_{i,3}) \in \mathbb{R}^3$, $i = 1, 2$, and $\mathcal{O} := (0, 0, \sigma + \rho, 0, 0, \sigma + \rho) \in \mathbb{R}^6$, we define

$$V(X) := \|X - \mathcal{O}\|^2/2 = \sum_{i=1}^{2} \{(x_{i,1})^2 + (x_{i,2})^2 + [x_{i,3} - (\sigma + \rho)]^2\}/2. \quad (120)$$

The derivative of $V$ along the solution of (99) is

$$\dot{V}(X) = -\sum_{i=1}^{2} \{\sigma(x_{i,1})^2 + (x_{i,2})^2 + \beta_i x_{i,3} [x_{i,3} - (\sigma + \rho)]\} + c \sum_{k=1}^{3} \sum_{i=1}^{2} \sum_{j=1}^{2} \omega_{ij} x_{i,k} x_{j,k} - c(\sigma + \rho) \sum_{j=1}^{2} \sum_{i=1}^{2} \omega_{ij} x_{j,3} \leq -\sum_{i=1}^{2} \{\sigma(x_{i,1})^2 + (x_{i,2})^2 + \beta(x_{i,3})^2\} + (\sigma + \rho) \sum_{i=1}^{2} \beta_i x_{i,3},$$

due to that $[\omega_{ij}]_{1 \leq i,j \leq 2}$ is negative definite, and $\beta_i \geq \beta$ and $\sum_{i=1}^{2} \omega_{ij} = 0$, $j = 1, 2$, where $x_{i,k} = x_{i,k}(t)$. Therefore, $\dot{V}(X) \leq \dot{\nu}(X) := \max \{\nu_{(t_1, t_2)}(X) : t_1, t_2 = 1, 2\}$, where

$$\nu_{(t_1, t_2)}(X) := -\sum_{i=1}^{2} \{\sigma(x_{i,1})^2 + (x_{i,2})^2 + \beta(x_{i,3})^2\} + (\sigma + \rho) \sum_{i=1}^{2} \beta_i x_{i,3}$$

$$= -\sum_{i=1}^{2} \{\sigma(x_{i,1})^2 + (x_{i,2})^2 + \beta x_{i,3} - (\sigma + \rho) \beta_i x_{i,3} - (\sigma + \rho)^2 \beta_i^2/4\beta\}. $$

Observe that $\nu_{(t_1, t_2)}(X) < 0$ if $X \in \mathbb{R}^6 - B_{(t_1, t_2)}$, where $B_{(t_1, t_2)} := \{X : \nu_{(t_1, t_2)}(X) \geq 0\}$. It is not difficult to verify that $\mathcal{O} \in \Theta \times \Theta$ and $\bigcup \{B_{(t_1, t_2)} : t_1, t_2 = 1, 2\} \subseteq \Theta \times \Theta$, where $\Theta := [-\hat{r}_1, \hat{r}_1] \times [-\hat{r}_2, \hat{r}_2] \times [-\hat{r}_3, \hat{r}_3] \subseteq \mathbb{R}^3$. We further consider the set

$$\mathcal{B} := \{X : \|X - \mathcal{O}\|^2 \leq 2[(\hat{r}_1)^2 + (\hat{r}_2)^2 + (\max \{|\hat{r}_3 - (\sigma + \rho)|, |\hat{r}_3 - (\sigma + \rho)| \}|)^2]\}. $$

It is not difficult to verify that $\Theta \times \Theta \subseteq \mathcal{B} \subseteq \hat{Q} \times \hat{Q}$, where $\hat{Q} = [\hat{R}_1, \hat{R}_1] \times [\hat{R}_2, \hat{R}_2] \times [\hat{R}_3, \hat{R}_3]$. Obviously, $\dot{\nu}(X) < 0$ if $X \in \mathbb{R}^6 - \Theta \times \Theta$. We hence conclude that

$$\dot{V}(X) \leq \dot{\nu}(X) < 0,$$

for all $X \in \mathbb{R}^6 - \mathcal{B} \subseteq \mathbb{R}^6 - \Theta \times \Theta$. As seen form the definition of function $V$ and set $\mathcal{B}$, an arbitrary solution $X(t)$ of (99) gets closer to $\mathcal{O}$ should $X(t)$ remain in $\mathbb{R}^6 - \mathcal{B}$. Moreover, $X(t)$ eventually converges to $\mathcal{B}$, which yields that $X(t)$ converges to $\hat{Q} \times \hat{Q}$ since $\mathcal{B} \subseteq \hat{Q} \times \hat{Q}$. It follows that $X(t)$ eventually enters, and then remains in $\hat{Q}(\dot{c}) \times \hat{Q}(\dot{c})$, for any $c > 0$.

REFERENCES

[1] S. Acharyya and R. E. Amritkar, Synchronization of coupled nonidentical dynamical systems, EPL, 99 (2012), 40005.
[2] S. Acharyya and R. E. Amritkar, Synchronization of nearly identical dynamical systems: Size instability, Phys. Rev. E, 92 (2015), 052902, 10 pp.
[3] A. Ay, J. Holland, A. Sperlea, G. S. Devakannalal, S. Knieer, S. Sangervasi, A. Stevenson and E. M. Ozbudak, Spatial gradients of protein-level time delays set the pace of the traveling segmentation clock waves, *Development*, **141** (2014), 4158–4167.

[4] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science*, **286** (1999), 509–512.

[5] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D.-U. Hwang, Complex networks: Structure and dynamics, *Phys. Rep.*, **424** (2006), 175–308.

[6] N. Burić and D. Todorović, Dynamics of FitzHugh-Nagumo excitable systems with delayed coupling, *Phys. Rev. E*, **67** (2003), 066222, 15 pp.

[7] N. Burić, K. Todorović and N. Vasović, Synchronization of bursting neurons with delayed chemical synapses, *Phys. Rev. E*, **78** (2008), 036211.

[8] S. A. Campbell, Time delays in neural systems, in *Handbook of Brain Connectivity* (eds. A. McIntosh and V. K. Jirsa), Springer, Berlin, Heidelberg, (2007), 65–90.

[9] J. Cao, Z. Wang and Y. Sun, Synchronization in an array of linearly stochastically coupled networks with time delays, *Phys. A*, **385** (2007), 718–728.

[10] K.-W. Chen, K.-L. Liao and C.-W. Shih, The kinetics in mathematical models on segmentation clock genes in zebrafish, *J. Math. Biol.*, **76** (2018), 97–150.

[11] S. M. Crook, G. B. Ermentrout, M. C. Vanier and J. M. Bower, The role of axonal delay in the synchronization of networks of coupled cortical oscillators, *J. Comput. Neurosci.*, **4** (1997), 161–172.

[12] K. M. Cuomo and A. V. Oppenheim, Circuit implementation of synchronized chaos with applications to communications, *Phys. Rev. Lett.*, **71** (1993), 65–68.

[13] Z. Duan and G. Chen, Global robust stability and synchronization of networks with Lorenz-type nodes, *IEEE Trans. Circuits Syst. II*, **56** (2009), 679–683.

[14] G. B. Ermentrout and N. Kopell, Fine structure of neural spiking and synchronization in the presence of conduction delays, *Proc. Natl. Acad. Sci.*, **95** (1998), 1259–1264.

[15] R. Femat and G. Solis-Peralles, On the chaos synchronization phenomena, *Phys. Lett. A*, **262** (1999), 50–60.

[16] S.-Y. Ha, S. E. Noh and J. Park, Practical synchronization of generalized Kuramoto systems with an intrinsic dynamics, *Netw. Heterog. Media*, **10** (2015), 787–807.

[17] J. K. Hale, Diffusive coupling, dissipation, and synchronization, *J. Dynam. Differential Equations*, **9** (1997), 1–52.

[18] W. He, W. Du, F. Qian and J. Cao, Synchronization analysis of heterogeneous dynamical networks, *Neurocomputing*, **104** (2013), 146–154.

[19] W. He, F. Qian, J. Cao and Q.-L. Han, Impulsive synchronization of two nonidentical chaotic systems with time-varying delay, *Phys. Lett. A*, **375** (2011), 498–504.

[20] W. He, F. Qian, Q.-L. Han and J. Cao, Synchronization error estimation and controller design for delayed Lur’e systems with parameter mismatches, *IEEE Trans. Neural Netw. Learn. Syst.*, **23** (2012), 1551–1563.

[21] F. C. Hoppensteadt and E. M. Izhikevich, Pattern recognition via synchronization in phase-locked loop neural networks, *IEEE Trans. Neural Netw.*, **11** (2000), 734–738.

[22] C.-H. Hsia, C.-Y. Jung and B. Kwon, On the global convergence of frequency synchronization for Kuramoto and Winfree oscillators, *Discrete Contin. Dyn. Syst. Ser. B*, **24** (2019), 3319–3334.

[23] L. Huang, Q. Chen, Y.-C. Lai and L. M. Pecora, Generic behavior of master-stability functions in coupled nonlinear dynamical systems, *Phys. Rev. E*, **80** (2009), 036204.

[24] T. Huang, C. Li and X. Liao, Synchronization of a class of coupled chaotic delayed systems with parameter mismatch, *Chaos*, **17** (2007), 033121, 5 pp.

[25] T. Huang, C. Li, W. Yu and G. Chen, Synchronization of delayed chaotic systems with parameter mismatches by using intermittent linear state feedback, *Nonlinearity*, **22** (2009), 569–584.

[26] X. Huang and J. Cao, Generalized synchronization for delayed chaotic neural networks: A novel coupling scheme, *Nonlinearity*, **19** (2006), 2797–2811.

[27] J. K. Kim and D. B. Forger, A mechanism for robust circadian timekeeping via stoichiometric balance, *Mol. Syst. Biol.*, **8** (2012), 1–14.

[28] V. I. Krinsky, V. N. Biktashev and I. R. Efimov, Autowave principles for parallel image processing, *Phys. D*, **49** (1991), 247–253.
[29] R. Levy, W. D. Hutchison, A. M. Lozano and J. O. Dostrovsky, High-frequency synchronization of neuronal activity in the subthalamic nucleus of Parkinsonian patients with limb tremor, J. Neurosci., 20 (2000), 7766–7775.

[30] C.-H. Li and S.-Y. Yang, Eventual dissipativeness and synchronization of nonlinearly coupled dynamical network of Hindmarsh-Rose neurons, Appl. Math. Model., 39 (2015), 6631–6644.

[31] K.-L. Liao, C.-W. Shih and J.-P. Tseng, Synchronized oscillations in a mathematical model of segmentation in zebrafish, Nonlinearity, 25 (2012), 869–904.

[32] E. N. Lorenz, Deterministic nonperiodic flow, J. Atmospheric Sci., 20 (1963), 130–141.

[33] W. Lu, T. Chen and G. Chen, Synchronization analysis of linearly coupled systems described by differential equations with a coupling delay, Phys. D, 221 (2006), 118–134.

[34] A. Margheri and R. Martins, Generalized synchronization in linearly coupled time periodic systems, J. Differential Equations, 249 (2010), 3215–3232.

[35] G. S. Medvedev, Electrical coupling promotes fidelity of responses in the networks of model neurons, Neural Comput., 21 (2009), 3057–3078.

[36] J. M. Montenbruck, M. Bürger and F. Allgöwer, Practical synchronization with diffusive couplings, Automatica J. IFAC, 53 (2015), 235–243.

[37] M. E. J. Newman, The structure and function of complex networks, SIAM Rev., 45 (2003), 167–256.

[38] E. Panteley, A. Loría and L. Conteville, On practical synchronization of heterogeneous networks of nonlinear systems: Application to chaotic systems, 2015 American Control Conference (ACC), (2015), 5359–5364.

[39] L. M. Pecora and T. L. Carroll, Master stability functions for synchronized coupled systems, Phys. Rev. Lett., 80 (1998), 2109–2112.

[40] C. Pradeep, Y. Cao, R. Murugesu and R. Rakkiyappan, An event-triggered synchronization of semi-Markov jump neural networks with time-varying delays based on generalized free-weighting-matrix approach, Math. Comput. Simulation, 155 (2019), 41–56.

[41] H. M. Rodrigues, Abstract methods for synchronization and applications, Appl. Anal., 62 (1996), 263–296.

[42] H. M. Rodrigues, L. F. C. Alberto and N. G. Bretas, Uniform invariance principle and synchronization. Robustness with respect to parameter variation, J. Differential Equations, 169 (2001), 228–254.

[43] M. G. Rosenblum, A. S. Pikovsky and J. Kurths, From phase to lag synchronization in coupled chaotic oscillators, Phys. Rev. Lett., 78 (1997), 4193–4196.

[44] M. G. Rosenblum, A. S. Pikovsky and J. Kurths, Phase synchronization of chaotic oscillators, Phys. Rev. Lett., 76 (1996), 1804–1807.

[45] C.-W. Shih and J.-P. Tseng, A general approach to synchronization of coupled cells, SIAM J. Appl. Dyn. Syst., 12 (2013), 1354–1393.

[46] C.-W. Shih and J.-P. Tseng, Global synchronization and asymptotic phases for a ring of identical cells with delayed coupling, SIAM J. Math. Anal., 43 (2011), 1667–1697.

[47] W. Singer, Synchronization of cortical activity and its putative role in information processing and learning, Annu. Rev. Physiol., 55 (1993), 349–374.

[48] Q. Song, J. Cao and F. Liu, Synchronization of complex dynamical networks with nonidentical nodes, Phys. Lett. A, 374 (2010), 544–551.

[49] F. Sorrentino and M. Porfiri, Analysis of parameter mismatches in the master stability function for network synchronization, EPL, 93 (2011), 50002.

[50] S. H. Strogatz and I. Stewart, Coupled oscillators and biological synchronization, Scientific American, 269 (1993), 102–109.

[51] J. Sun, E. M. Bollt and T. Nishikawa, Master stability functions for coupled nearly identical dynamical systems, EPL, 85 (2009), 60011.

[52] J. A. K. Suykens, P. F. Curran and L. O. Chua, Robust synthesis for master-slave synchronization of Lur’e systems, IEEE Trans. Circuits Syst. I, 46 (1999), 841–850.

[53] D. Taylor, P. S. Skardal and J. Sun, Synchronization of heterogeneous oscillators under network modifications: Perturbation and optimization of the synchrony alignment function, SIAM J. Appl. Math., 76 (2016), 1984–2008.

[54] J.-P. Tseng, A novel approach to synchronization of nonlinearly coupled network systems with delays, Phys. A, 452 (2016), 266–280.

[55] K. Urie, Y. Morishita and Y. Iwasa, Synchronized oscillation of the segmentation clock gene in vertebrate development, J. Math. Biol., 61 (2010), 207–229.
[56] C. van Vreeswijk, Partial synchronization in populations of pulse-coupled oscillators, *Phys. Rev. E*, 54 (1996), 5522–5537.

[57] G. D. VanWiggeren and R. Roy, Communication with chaotic lasers, *Science*, 279 (1998), 1198–1200.

[58] J.-L. Wang, Z.-C. Yang, T. Huang and M. Xiao, Local and global exponential synchronization of complex delayed dynamical networks with general topology, *Discrete Contin. Dyn. Syst. Ser. B*, 16 (2011), 393–408.

[59] L. Wang, W. Qian and Q.-G. Wang, Bounded synchronization of a time-varying dynamical network with nonidentical nodes, *Internat. J. Systems Sci.*, 46 (2015), 1234–1245.

[60] Y. Wang and J. Cao, Cluster synchronization in non-linearly coupled delayed networks of non-identical dynamic systems, *Nonlinear Anal. Real World Appl.*, 14 (2013), 842–851.

[61] D. J. Watts and S. H. Strogatz, Collective dynamics of ‘small–world’ networks, *Nature*, 393 (1998), 440–442.

[62] J. A. White, C. C. Chow, J. Rit, C. Soto-Treviño and N. Kopell, Synchronization and oscillatory dynamics in heterogeneous, mutually inhibited neurons, *J. Comput. Neurosci.*, 5 (1998), 5–16.

[63] C. W. Wu, Synchronization in networks of nonlinear dynamical systems coupled via a directed graph, *Nonlinearity*, 18 (2005), 1057–1064.

[64] X. Yang, J. Cao and Z. Yang, Synchronization of coupled reaction-diffusion neural networks with time-varying delays via pinning-impulsive controller, *SIAM J. Control Optim.*, 51 (2013), 3486–3510.

[65] W. Yu, J. Cao and J. Lü, Global synchronization of linearly hybrid coupled networks with time-varying delay, *SIAM J. Appl. Dyn. Syst.*, 7 (2008), 108–133.

[66] J. Zhao, D. J. Hill and T. Liu, Global bounded synchronization of general dynamical networks with nonidentical nodes, *IEEE Trans. Automat. Control*, 57 (2012), 2656–2662.

Received March 2019; revised August 2019.

E-mail address: cwshih@math.nctu.edu.tw
E-mail address: jptseng@nccu.edu.tw