Stochastic Synchronization in Nonlinear Network Systems Driven by Intrinsic and Coupling Noise

Zahra Aminzare *   Vaibhav Srivastava †

Abstract

In this paper, we consider a *noisy* network of nonlinear systems in the sense that each system is driven by two sources of state-dependent noise: (1) an intrinsic noise that can be generated by the environment or any internal fluctuations, and (2) a noisy coupling which is generated by interactions with other systems. Our goal is to understand the effect of noise and coupling on synchronization behaviors of such networks. First, we assume that all the systems are driven by a common noise and show how a common noise can be detrimental or beneficial for network synchronization behavior. Then, we assume that the systems are driven by independent noise and study network *approximate* synchronization behavior. We numerically illustrate our results using the example of coupled Van der Pol oscillators.

Keywords. Noisy networks, stochastic synchronization, approximate synchronization, homogeneous networks, heterogeneous networks.

1 Introduction

Coupled nonlinear oscillator models are fundamental in modeling and analyzing the synchronization behavior of systems with rhythmic behavior, including systems in ecology, neuroscience, and engineering [1–7]. Example phenomena that are modeled well by coupled nonlinear oscillators include biological rhythms [8, 9], neural synchrony [2], locomotion gaits [10, 11], collective motion in animal groups [12], fish schools [13], cooperative robotic networks [14], power networks [15], coupled Josephson arrays [16], and sinoatrial pacemakers [17]. These simple models often miss environmental fluctuations as well as internal and external disturbances. In contrast, a stochastic dynamics approach provides a significant compromise in terms of keeping modeling complexity tractable while still capturing important phenomena.

The problem of understanding the influence of stochastic perturbations on the synchronization behavior in networks of nonlinear systems has received some attention in the literature. The influence of noise on the synchronization behavior in a large-scale model of the human brain network is studied in [18] and the authors report that the addition of noise increases the synchronization of global and local dynamics. Noise-induced synchronization in networks of excitable systems

*Zahra Aminzare is with the Department of Mathematics, University of Iowa, IA, USA. zahra-aminzare@uiowa.edu

†Vaibhav Srivastava is with Electrical and Computer Engineering, Michigan State University, East Lansing, MI, USA. vaibhav@egr.msu.edu

is studied in [19], and authors report that for low noise, the solutions remain in the vicinity of the resting state; for large noise, the solutions are asynchronous; and the medium noise, the synchronized periodic responses are obtained. Jafarian et al [20] study stochastic stability of discrete-time phase coupled oscillators and derive sufficient conditions for achieving the phase-cohesiveness. Meng and Riecke [21] study synchronization in networks of multiple coupled oscillator networks and show that for strong inhibitory coupling between networks, the rhythms of each coupled oscillator network synchronize even if the noisy inputs to different oscillator networks are completely uncorrelated.

In addition to the intrinsic noise present at each system in the network, the interconnection noise also plays an important role in synchronization. Experimental studies of cortical areas show that heterogeneity in the connections plays a critical role in their synchronization behaviors [22]. Synchronization behavior has been studied theoretically in large-scale networks of firing-rate and Fitzhugh-Nagumo neurons interconnected with stochastic synapses (see [23] and [24]).

In this paper, we consider a broad network of nonlinear systems (e.g., oscillators) that are coupled through either linear coupling (e.g., gap junction in neuronal populations) or nonlinear coupling (e.g., sinusoidal coupling in coupled Kuramoto oscillators). In addition, we consider two sources of nonlinear stochasticity in the network: one affects the systems, which we will refer to as a common noise (e.g., a common stimulus that drives a population of neurons), and the other perturbs the connection between the systems (e.g., noisy synaptic coupling). The objective is to explore the influence of each network element on the synchronization behavior of the network. In particular, we introduce a synchronization measure that reflects all the network parameters (such as noise intensity and network connectivity) and show how tuning these parameters would alter the synchronization status of the network.

There have been some efforts to find conditions for synchronization in stochastic networks, see for example [25–27], where both the coupling and the common noise intensity are linear functions of the state. There are also some interesting results which guarantee synchronization onset in networks with no coupling but common noise, [28]. In [25], the authors study noise-induced synchronization in a network of nonlinear systems which are coupled through deterministic diffusive coupling. They assume that each system is driven by a common state-dependent noise, where the intensity of the noise is a linear function of the state. In [26], the authors consider a network of nonlinear systems which are coupled through both deterministic and stochastic coupling and characterize the influence of stochastic coupling on the synchronization behavior. However, this work does not consider any common or intrinsic noise.

Motivated by coupled Kuramoto oscillators which are connected through nonlinear coupling and phase equations of coupled noisy oscillators in which the common noise is nonlinear (see e.g., [29, 30]), we first extend the results given in [25, 26] to a network of nonlinear systems which are driven by common noise whose intensity is a nonlinear function of state and are coupled through nonlinear stochastic and deterministic coupling functions. In this scenario, we rigorously characterize the conditions of the nonlinear noise intensity and coupling functions such that they aid synchronization.

We then generalize the results to approximate synchronization for heterogeneous noisy networks, i.e., networks of nonlinear systems in which the local noise is not the same for each system. The approximate synchronization behavior is similar to the practical synchronization [31] and quasi-synchronization [32] behavior studied in the context of heterogeneous deterministic perturbation to network of nonlinear systems. Similar to the practical synchronization behavior, we show that by making the coupling strength strong enough, the steady state behavior of the system can be driven arbitrarily close to synchronization.

Our main goal is to find conditions that foster synchronization in networks of coupled stochastically perturbed systems, in which the systems are subject to a common perturbation or perturbations through their interactions with other systems in the network. Here, we model both stochastic perturbations by *nonlinear multiplicative (state-dependent) Itô* terms. We introduce a synchronization metric that depends on the intrinsic dynamics of each system, the coupling function and the underlying network topology, the common noise which drives the systems, and the noise which affects the connections. We will analyze each factor's conductive or destructive effects on the network's synchronization. First, we show that, in general, adding multiplicative state-dependent noise to a synchronized deterministic network is detrimental to synchronization. The network may synchronize with small common noise but desynchronize with large common multiplicative noise. We then show that adding multiplicative noise can aid synchronization if a linear function of the state lower bounds the common multiplicative noise.

Our main contributions are twofold: (1) to generalize the (complete) synchronization conditions in [25] and [26] to networks with common *nonlinear* state-dependent noise and *skew-symmetric nonlinear* coupling functions, which are a generalization of diffusive coupling functions; and (2) to allow heterogeneous intrinsic noise and provide conditions that guarantee approximate synchronization.

The remainder of the paper is organized as follows. In Section 2, we study the "complete" synchronization behavior of homogeneous noisy networks where the individuals are driven by a common noise. We provide conditions that guarantee stochastic synchronization in such networks and introduce a class of homogeneous noisy networks that take advantage of noise to foster synchronization. In contrast, in Section 3, we study the "approximate" synchronization behavior of heterogeneous noisy networks where the individuals are driven by independent noises and show how coupling fosters synchronization in these heterogeneous networks. In Section 4, we numerically illustrate the theoretical results using the example of coupled Van der Pol oscillators. We conclude in Section 5. All the proofs are given in Section 6.

2 Stochastic synchronization in homogeneous noisy networks

In this section, we consider a network of N coupled identical systems with two sources of state-dependent noise: (1) an intrinsic noise which is common among all systems and can be generated by the environment, and (2) a coupling noise which is generated by interactions with other systems. For $i = 1, \ldots, N$, let the stochastic differential equation (SDE)

$$d\phi_i = \underbrace{\mathcal{F}(\phi_i, t)dt + \sigma\mathcal{K}(\phi_i, t)dW}_{\text{homogeneous intrinsic dynamics}} + \underbrace{\sum_{j=1}^{N} c_{ij}\left(\epsilon\mathcal{H}(\phi_j, \phi_i)dt + \delta\mathcal{C}(\phi_j, \phi_i)dW_{ij}(t)\right)}_{\text{coupling dynamics}} \tag{1}$$

describe the dynamics of system i with state $\phi_i \in \mathbb{R}^n$. The intrinsic and coupling dynamics of system i are described as below.

Intrinsic dynamics. The systems are identical and governed by an $n-$dimensional vector of nonlinear functions, $\mathcal{F}$. There is a source of noise in (1) which is common among all the systems in the network and described by $\sigma\mathcal{K}(\phi_i, t)dW$. The constant $\sigma \geq 0$ is the common noise intensity, $\mathcal{K} : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}^{n\times n}$, and W is an $n-$dimensional vector of independent standard Wiener processes. Since the intrinsic dynamics are common among all the systems, we refer to this network as a *homogeneous* noisy network.

Coupling dynamics. Denote the underlying network graph by $\mathcal{G}$ and assume that it is an undirected and weighted graph with weight c_{ij}, i.e., $c_{ij} = c_{ji} \geq 0$, with $c_{ij} > 0$ if i and j are connected; and $c_{ij} = 0$ if i and j are not connected. The interaction between system i and another system, say j, influences the dynamics of i through a deterministic term $c_{ij}\epsilon\mathcal{H}(\phi_j, \phi_i)dt$ and a stochastic term $c_{ij}\delta\mathcal{C}(\phi_j, \phi_i)dW_{ij}$, where $\mathcal{H}, \mathcal{C} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$, and $\mathbf{W}_i = (W_{i1}, \dots, W_{iN})^\top$ is a vector of independent standard Wiener processes. The processes W and $\mathbf{W}_i$'s are assumed to be mutually independent. The constants $\epsilon \geq 0$ and $\delta \geq 0$ respectively describe the coupling strength and interaction noise intensity of the overall network while ϵc_{ij} and δc_{ij} respectively specify the coupling strength and noise intensity of each connection.

For now, we only assume that $\mathcal{F}$, $\mathcal{H}$, $\mathcal{K}$, and $\mathcal{C}$ are nonlinear functions and they are nice enough so that (1) has a unique solution, for example, they are Lipschitz and satisfy a linear growth condition. See [33, Section 2.3] for more details. Later in Theorems 1 and 2 below, we will discuss appropriate conditions of these functions.

In what follows we review definitions of *stochastic stability* and *stochastic synchronization*. Subsequently, in Theorem 1, we will provide a sufficient condition that guarantees stochastic synchronization in (1). In the following section, Theorem 2 will discuss more conditions that foster synchronization in such networks.

Definition 1 (***Stochastic stability***)**.** *Let $x(t)$ be a solution of an SDE. Then,*

Moment exponential stability. *$x(t)$ is $p-$th ($p > 0$) moment exponentially stable if there are a pair of positive constants C and c and a neighborhood Ω_0 of $x(0)$ such that for any solution y with $y(0) \in \Omega_0$*

$$\mathbb{E}\|y(t) - x(t)\|^p < C\, \mathbb{E}\|y(0) - x(0)\|^p e^{-ct}, \quad \forall t > 0,$$

where $\mathbb{E}$ denotes the expected value and $\|\cdot\|$ denotes the Euclidean norm. When $p = 2$, it is said to be exponentially stable in mean square.

Almost sure exponential stability. *$x(t)$ is almost sure exponentially stable if there is a neighborhood Ω_0 of $x(0)$ such that for any solution y with $y(0) \in \Omega_0$*

$$\limsup_{t\to\infty} \frac{1}{t}\, \log\|y(t) - x(t)\| < 0, \quad \textit{almost surely (a.s.)},$$

which means $\mathbb{P}\left\{\limsup_{t\to\infty} \frac{1}{t}\, \log\|y(t) - x(t)\| < 0\right\} = 1$.

Clearly, the $p-$th moment exponentially stability means that the solution y tends to x exponentially fast, and the so called $p-$th moment Lyapunov exponent of y is negative:

$$\limsup_{t\to\infty} \frac{1}{t}\, \log\|y(t) - x(t)\|^p < 0.$$

Also, the left hand side of (1) is called sample Lyapunov exponent of y. In general $p-$th moment stability and almost sure exponential stability are not equivalent and additional conditions are required to deduce one from the other [33, Section 4.4].

Definition 2 (***Stochastic invariance***)**.** *A set* $\mathcal{S}$ *is called an invariant set for an SDE, if for any* $x_0 \in \mathcal{S}$, $\mathbb{P}\{x(t) \in \mathcal{S},\ \forall t \geq 0\} = 1$, *where* $x(t)$ *is a solution of the SDE starting from* x_0 *at* $t = 0$.

Definition 3 (***Stochastic synchronization***)**.** *Let* $\mathcal{S}$ *be the set of states defined by* $\mathcal{S} := \{x = (x_1, \ldots, x_N)^\top \mid x_1 = \cdots = x_N\}$. *We say that a network stochastically synchronizes if* $\mathcal{S}$ *is stochastically invariant and for any solution* $x(t)$ *there exists* $s(t) \in \mathcal{S}$ *such that* $x(t)$ *converges to* $s(t)$ *exponentially fast, that is*

$$\mathbb{E}\|x(t) - s(t)\|^p < C\ \mathbb{E}\|x(0) - s(0)\|^p e^{-ct}, \quad \forall t > 0, \quad \text{and some } c, C > 0, \tag{2}$$

or

$$\limsup_{t\to\infty} \frac{1}{t}\ \log\|x(t) - s(t)\| < 0, \quad a.s. \tag{3}$$

Although the systems in (1) can be of any arbitrary dimension, in the following theorems, for the ease of notation, we assume that the state variables are 1-dimensional, $n = 1$.

We denote the Laplacian matrix of the underlying network graph $\mathcal{G}$ by $L_{[c]}$ (where the subscript $[c]$ represents the weights c_{ij}) and its eigenvalues by $0 = \lambda_{1,[c]} \leq \lambda_{2,[c]} \leq \cdots \leq \lambda_{N,[c]}$.

Theorem 1 (***Stochastic synchronization: exponential stability in mean square***)**.** *Fix* $\Omega_1 \subset \mathbb{R}$ *and let* $\Omega_2 := \{x - y \mid x, y \in \Omega_1\}$. *Consider* (1) *and assume that:*

i. there exists a constant $\bar{c}_{\mathcal{F}}$ *such that for all* $x, y \in \Omega_1$ *and* $t \geq 0$,

$$(x - y)(\mathcal{F}(x,t) - \mathcal{F}(y,t)) \leq \bar{c}_{\mathcal{F}}(x - y)^2; \tag{4}$$

ii. $\mathcal{H} : \Omega_1 \times \Omega_1 \to \mathbb{R}$ *satisfies* $\mathcal{H}(x,y) = -\mathcal{H}(y,x)$ *and there exists a constant* $\underline{c}_{\mathcal{H}}$ *such that for all* $x, y \in \Omega_1$, $\underline{c}_{\mathcal{H}}(x-y)^2 \leq (x-y)\mathcal{H}(x,y)$;

iii. there exists a non-negative constant $\bar{c}_{\mathcal{C}}$ *such that for all* $x, y \in \Omega_1$, $|\mathcal{C}(x,y)| \leq \bar{c}_{\mathcal{C}}|x - y|$; *and*

iv. there exists a non-negative constant $\bar{c}_{\mathcal{K}}$ *such that for all* $x, y \in \Omega_1$ *and* $t \geq 0$,

$$|\mathcal{K}(x,t) - \mathcal{K}(y,t)| \leq \bar{c}_{\mathcal{K}}|x - y|.$$

Then for any solution $(\phi_1, \ldots, \phi_N)^\top$, *there exists a solution on* $\mathcal{S} := \{x = (x_1, \ldots, x_N)^\top \mid x_1 = \cdots = x_N\}$, *namely* $(\psi(t), \ldots, \psi(t))^\top$, *where* $\psi(t) = \frac{1}{N}\sum_{i=1}^N \phi_i(t)$, *such that*

$$\mathbb{E}\sum_{i=1}^N |\phi_i(t) - \psi(t)|^2 < \mathbb{E}\sum_{i=1}^N |\phi_i(0) - \psi(0)|^2 e^{-ct}, \quad \forall t > 0,$$

and

$$c := -2\bar{c}_{\mathcal{F}} + 2\epsilon\underline{c}_{\mathcal{H}}\lambda - 2\delta^2\bar{c}_{\mathcal{C}}^2\Big(1 - \frac{1}{N}\Big)\lambda_{N,[c^2]} - \Big(1 - \frac{1}{N}\Big)\sigma^2\bar{c}_{\mathcal{K}}^2. \tag{5}$$

In (5), *if* $\underline{c}_{\mathcal{H}} > 0$, *then* $\lambda = \lambda_{2,[c]}$, *otherwise,* $\lambda = \lambda_{N,[c]}$. $\lambda_{N,[c^2]}$ *denotes the largest eigenvalue of the Laplacian matrix of network graph* $\mathcal{G}$ *with weights* c_{ij}^2.

Therefore, the network stochastically synchronizes (in the sense of (2) *with* $p = 2$*) when* $c > 0$.

On the synchronization manifold $\mathcal{S} := \{x = (x_1, \ldots, x_N)^\top \mid x_i = \psi\}$, the dynamics of the network (1) becomes $d\psi = \mathcal{F}(\psi, t)dt + \sigma\mathcal{K}(\psi, t)dW$, in which, a unique solution exists if $\mathcal{F}$ and $\mathcal{K}$ satisfy the Lipschitz and growth conditions: $\exists K_1, K_2 > 0$ such that $\forall x, y$: $\|\mathcal{F}(x) - \mathcal{F}(y)\| + \|\mathcal{K}(x) - \mathcal{K}(y)\| \leq K_1\|x - y\|$, and $\|\mathcal{F}(x)\|^2 + \|\mathcal{K}(x)\|^2 \leq K_2(1 + \|x\|^2)$, where $\|\cdot\|$ denotes the Euclidean norm.

Condition (i) is a one-sided Lipschitz condition for $\mathcal{F}$ (bounded above). Unlike a Lipschitz constant which must be positive, $\bar{c}_{\mathcal{F}}$ could take any values. Although the Lipschitz condition implies one-sided Lipschitz condition for $\mathcal{F}$ with a non-negative $\bar{c}_{\mathcal{F}}$, we assume Condition (i) to allow one-sided Lipschitz condition with any $\bar{c}_{\mathcal{F}}$. This condition is also called QUAD condition since the left-hand side of (4) is bounded by a quadratic term. For $\bar{c}_{\mathcal{F}} < 0$, the condition is equivalent to the vector field $\mathcal{F}$ being *contractive* in L^2 norm. This means that the distance between any two flows decreases and the flows converge to each other exponentially. For more details see [34]. This is an easy condition to check. The best one-sided Lipschitz constant for a differential $\mathcal{F}$ is $\sup_x \lambda_x$ where λ_x is the largest eigenvalue of $\frac{1}{2}(D\mathcal{F}(x) + D\mathcal{F}(x)^\top)$. Here, $D\mathcal{F}$ is the Jacobian of $\mathcal{F}$, or simply, the derivative of $\mathcal{F}$ when $\mathcal{F}$ is scalar. For example, for a Kuramoto oscillator $\bar{c}_{\mathcal{F}} = 0$ and for a FitzHugh-Nagumo model it is equal to 1.

In Condition (ii), the skew-symmetric condition is a generalization of diffusive coupling to nonlinear coupling. This condition guarantees the existence of an invariant synchronization manifold, i.e., the coupling dynamics vanish on the synchronization manifold, since $\mathcal{H}(x, x) = 0$. Also, for technical proofs, we assume that $\mathcal{H}$ is bounded below by a linear function. Here are three examples of these types of coupling: 1) A gap junction $\mathcal{H}(x, y) = x - y$ with $\underline{c}_{\mathcal{H}} = 1$; 2) A sinusoidal function $\mathcal{H}(x, y) = \sin(x - y)$ defined on $\Omega_1 = [-\pi/2\ \pi/2)^2$ with $\underline{c}_{\mathcal{H}} = 0$ as in coupled Kuramoto oscillators; and 3) A nonlinear coupling described by $\mathcal{H}(x, y) = H(x) - H(y)$, where H is a one-sided Lipschitz function (bounded below) and $\underline{c}_{\mathcal{H}}$ is the Lipschitz constant of H.

In Condition (iii), we assume a linear upper bound for the stochastic coupling $\mathcal{C}$ which ensures that the coupling vanishes on the synchronization manifold, $\mathcal{C}(x, x) = 0$. Later in Theorem 2, we will consider a class of coupling functions $\mathcal{C}$ which are lower bounded by linear functions and will show how it helps synchronization.

The Lipschitz condition given in (iv) is necessary for the existence and uniqueness of the solutions on the synchronization manifold. Note that when $\bar{c}_{\mathcal{K}} = 0$, $\mathcal{K}$ becomes constant, i.e., $\mathcal{K}$ becomes an additive noise. So the last term in c becomes zero and therefore it suggests that the additive noise has no detrimental or beneficial effects on a network synchronization. In Theorem 2 below, we will consider a class of multiplicative noise which are lower bounded by linear functions and show how these bounds aid synchronization.

The constant c consists of four terms related to deterministic and stochastic intrinsic and coupling dynamics, respectively, and the topology of the network graph. The first term in c depends on the intrinsic dynamics of isolated systems. The second term in c depends on the coupling term, coupling intensity, and the algebraic connectivity of the underlying graph, $\lambda = \lambda_{2,[c]}$ (in the case of positive $\underline{c}_{\mathcal{H}}$). The algebraic connectivity of a graph, which determines how well-connected the graph is, may increase or decrease when the size of the graph changes. For example, in a line graph, the algebraic connectivity decreases as N increases while in an all to all graph it increases. In an almost surely connected Erdös-Réyni graph, the algebraic connectivity increases as N increases [35]. Therefore, our condition guarantees that large random networks of systems which are connected through e.g., diffusive or sinusoidal coupling have a better chance to synchronize.

The third term in c reflects the stochastic coupling. So similar to the second term, it depends on the coupling intensity (δ), coupling dynamics ($\bar{c}_{\mathcal{C}}$) and the topology of the underlying graph (the largest eigenvalue and the number of the nodes). The fourth term in c reflects the intrinsic noise. Note that both the third and fourth terms are always negative and show that the noise could be detrimental for synchronization, as this might be intuitively correct. However, this is not always true. Indeed, noise can be beneficial for network synchronization. For example, if all the individuals in a network are driven by a common noise, this common noise can act as a driving force to all the systems and foster synchronization. In Theorem 2, we analytically state this intuitive idea.

In Theorem 1, we showed that if a network synchronizes in the absence of any noise (common noise or noise induced by the interactions among the nodes in the network), it could also synchronize in the presence of sufficiently small noise and we found an upper bound for the noise intensities which guarantee such behavior, i.e., we proved that if the noise intensities are such that $c > 0$, then the network preserves its synchronization behavior. However, $c > 0$ is a sufficient condition for synchronization, and so, a network may synchronize with a negative c. In the following theorem, we consider networks with negative c and find a new sufficient condition for synchronization. Indeed, the next result shows that multiplicative noise terms can be beneficial for networks synchronization, if they are lower bounded by some linear functions.

Theorem 2 (***Noise-induced synchronization***)**.** *Consider conditions (i-iv) of Theorem 1 and furthermore assume that*

i. $\mathcal{C} : \Omega_1 \times \Omega_1 \to \mathbb{R}$ *satisfies* $\mathcal{C}^2(x, y) = \mathcal{C}^2(y, x)$ *and there exists a non-negative constant* $\underline{c}_{\mathcal{C}}$ *such that for all* $x, y \in \Omega_1$, $\underline{c}_{\mathcal{C}}|x - y| \leq |\mathcal{C}(x, y)|$*; and*

ii. there exists a non-negative constant $\underline{c}_{\mathcal{K}}$ *such that for all* $x, y \in \Omega_1$ *and* $t \geq 0$*,*

$$\underline{c}_{\mathcal{K}}(x - y)^2 \leq (x - y)(\mathcal{K}(x, t) - \mathcal{K}(y, t)).$$

Let

$$\alpha_1 = -\frac{c}{2} = \bar{c}_{\mathcal{F}} - \epsilon \underline{c}_{\mathcal{H}} \lambda + \delta^2 \bar{c}_{\mathcal{C}}^2 \Big(1 - \frac{1}{N}\Big) \lambda_{N,[c^2]} + \frac{1}{2}\Big(1 - \frac{1}{N}\Big)\sigma^2 \bar{c}_{\mathcal{K}}^2, \quad \alpha_2^2 = (\sigma \underline{c}_{\mathcal{K}})^2 + \frac{\delta^2 \underline{c}_{\mathcal{C}}^2 \lambda_{2,[c]}^2}{4N},$$

and assume that $0 \leq \alpha_1 < \alpha_2^2$*. Then for* $0 < p < 2(1 - \frac{\alpha_1}{\alpha_2^2}) \leq 2$ *and* $\alpha := -p[(\frac{p}{2} - 1)\alpha_2^2 + \alpha_1]$ *(which is positive), (1) stochastically synchronizes, that is, for any solution* $(\phi_1, \ldots, \phi_N)^\top$*, there exists a solution on* $\mathcal{S} := \{x = (x_1, \ldots, x_N)^\top \mid x_1 = \cdots = x_N\}$*, namely* $(\psi(t), \ldots, \psi(t))^\top$*, where* $\psi(t) = \frac{1}{N}\sum_{i=1}^N \phi_i(t)$*, such that*

$$\mathbb{E}\|e(t)\|^p \leq \mathbb{E}\|e(0)\|^p \, e^{-\alpha t}, \quad p\text{-th moment exponential stability}$$

$$\limsup_{t\to\infty} \frac{1}{t} \log \|e(t)\| \leq \alpha_1 - \alpha_2^2, \quad \text{almost sure exponential stability}$$

where $e = (\phi_1 - \psi, \ldots, \phi_N - \psi)^\top$ *is the corresponding error.*

The sufficient condition for stochastic synchronization suggested by Theorem 2 is $0 \leq \alpha_1 < \alpha_2^2$. Since $c = -2\alpha_1 < 0$, the result of Theorem 2 is a completion to the result of Theorem 1 (where $c > 0$ implies Stochastic synchronization). Consider a deterministic network which does not synchronize, i.e., $\bar{c}_{\mathcal{F}} - \epsilon \underline{c}_{\mathcal{H}} \lambda \geq 0$. Theorem 2 guarantees that adding a common noise with $\bar{c}_{\mathcal{K}}^2 < \frac{2N}{N-1}\underline{c}_{\mathcal{K}}^2$ and sufficiently large intensity, aims the network to synchronize (no noisy coupling is considered here).

Under the conditions of Theorems 1 and 2, p-th moment exponential stability implies almost sure exponential stability. See [33, Section 4.4, Theorem 4.2].

In summary, Theorems 1 and 2 provide sufficient conditions for stochastic synchronization. Both theorems are stated for homogeneous noise and guarantee complete synchronization. Strong multiplicative noise can destroy synchronization in the first theorem, while in the second theorem, multiplicative noise with a linear lower bound can foster synchronization.

In the following section, we state two theorems similar to Theorems 1 and 2 in the sense of how noise can be detrimental or beneficial for network synchronization. We relax the homogeneity condition of intrinsic noise and allow an independent noise to drive the systems. This leads to *approximate* synchronization instead of *complete* synchronization.

3 Approximate synchronization in heterogeneous noisy networks

In this section, we consider (1), where we assume that each system is driven by an independent noise instead of a common noise, i.e., we consider a network of *heterogeneous* noisy systems:

$$d\phi_i = \underbrace{\mathcal{F}(\phi_i,t)dt + \sigma_i \mathcal{K}(\phi_i,t)dW_i}_{\text{heterogeneous intrinsic dynamics}} + \underbrace{\sum_{j=1}^{N} c_{ij}\left(\epsilon\mathcal{H}(\phi_j,\phi_i)dt + \delta\mathcal{C}(\phi_j,\phi_i)dW_{ij}(t)\right)}_{\text{coupling dynamics}}. \tag{6}$$

All the terms in (6) are as defined in (1), except that dW_i's are independent standard Wiener processes. The goal is to study the synchronization behavior of (6). However, the conditions of the previous section do not guarantee stochastic synchronization in such heterogeneous networks (see Example 3 in Section 4). Therefore, in what follows, we provide conditions that the heterogeneous noisy network given in (6) *approximately* synchronizes in the sense of the following definition:

Definition 4 (***Approximate synchronization***)**.** *Let $\mathcal{S}$ be the set of states defined by $\mathcal{S} := \{x = (x_1,\ldots,x_N)^\top \mid x_1 = \cdots = x_N\}$. A stochastic network approximately synchronizes if $\mathcal{S}$ is stochastically invariant and for any solution $x(t)$ there exist $s(t) \in \mathcal{S}$ and $\eta \geq 0$ such that*

$$\mathbb{E}\|x(t) - s(t)\|^p \leq \eta, \quad \text{as } t \to \infty, \tag{7}$$

that is, the $p-th$ moment of the error is upper bounded by a constant η.

Theorem 3 (***Approximate synchronization in mean square***)**.** *Assume that the conditions (i-iv) of Theorem 1 hold. Furthermore, assume that there exists $\gamma > 0$ such that for any solution $(\phi_1,\ldots,\phi_N)^\top$, $\psi(t) = \frac{1}{N}\sum_{i=1}^{N}\phi_i(t)$ satisfies*

$$\mathbb{E}\|\psi(t)\|^2 \leq \gamma^2.$$

Also, assume that $\mathcal{K}(0,t)$ is bounded, $\sup_t \|\mathcal{K}(0,t)\| = \mathcal{K}_0$. Then for any solution $(\phi_1,\ldots,\phi_N)^\top$,

$$\mathbb{E}\sum_{i=1}^{N}|\phi_i(t) - \psi(t)|^2 \leq \left(\mathbb{E}\sum_{i=1}^{N}|\phi_i(0) - \psi(0)|^2 - \eta\right)e^{-c_a t} + \eta, \quad \forall t > 0, \tag{8}$$

where

$$c_a = -2\bar{c}_{\mathcal{F}} + 2\epsilon \underline{c}_{\mathcal{H}}\lambda - 2\delta^2 \bar{c}_{\mathcal{C}}^2 \left(1 - \frac{1}{N}\right) \lambda_{N,[c^2]} - 2\Big(1 - \frac{1}{N}\Big) \bar{c}_{\mathcal{K}}^2 \max_i \sigma_i^2,$$

(the subscript "a" in c_a stands for approximate synchronization) and

$$\eta = \frac{1}{c_a}\Big(1 - \frac{1}{N}\Big)(\bar{c}_{\mathcal{K}}\gamma + \mathcal{K}_0)^2 \sum_i \sigma_i^2.$$

Therefore, when $c_a > 0$, the network approximately synchronizes (in the sense of (7) with $p = 2$).

Note that (8) can be written as

$$\mathbb{E}\sum_{i=1}^{N} |\phi_i(t) - \psi(t)|^2 \;\le\; \mathbb{E}\sum_{i=1}^{N} |\phi_i(0) - \psi(0)|^2 \, e^{-c_a t} + \frac{1 - e^{-c_a t}}{c_a}\, \bar{\eta}, \tag{9}$$

where $\bar{\eta} = c_a \eta$ is a positive constant and hence for any values of c_a, $\frac{1-e^{-c_a t}}{c_a}\,\bar{\eta}$ is always non-negative.

Theorem 3 is a generalization of Theorem 1 when either σdW is replaced by $\sigma_i dW$ or $\sigma_i dW_i$. In this case, roughly speaking, the solutions exponentially fast converge to a *tube* that surrounds the synchronization solution instead of converging to the synchronization solution. In Theorem 3, for equal σ_is, $c_a = c$, so the rate of convergence to the tube that surrounds the synchronization solution remains constant.

Next, we generalize Theorem 2 to heterogeneous noisy networks. The goal is to classify the networks which do not synchronize in the absence of noise, while they approximately synchronize in the presence of independent noise.

Theorem 4 (***Noise-induced expedited approximate synchronization***)**.** *Consider the conditions of Theorems 1, 2, and 3 where $\delta = 0$. Let*

$$\beta_1 = -\bar{c}_{\mathcal{F}} + \epsilon \underline{c}_{\mathcal{H}}\lambda - \frac{1}{2}\Big(1 - \frac{1}{N}\Big)\bar{c}_{\mathcal{K}}^2 \max_i \sigma_i^2, \qquad \beta_2^2 = \min_i \sigma_i^2 \frac{\underline{c}_{\mathcal{K}}^2}{2N}.$$

and assume that $0 \le \beta_1 \le \beta_2^2$. Then for $0 < p < 2(1 - \frac{\beta_1}{\beta_2^2})$, any solution $(\phi_1, \ldots, \phi_N)^\top$ satisfies,

$$\mathbb{E}\left(\sum_{i=1}^{N} |\phi_i(t) - \psi(t)|^2\right)^{\frac{p}{2}} \le \left(\mathbb{E}\left(\sum_{i=1}^{N} |\phi_i(0) - \psi(0)|^2\right)^{\frac{p}{2}} - \zeta\right) e^{-\beta t} + \zeta, \quad \forall t > 0, \tag{10}$$

where $\beta = -p\left(\beta_1 + \left(\frac{p}{2} - 1\right)\beta_2^2\right) > 0$, $\zeta = \frac{1}{\beta}(l_1 \mathbb{E}(e^\top e)^{\frac{p}{2}-1} + l_2 \mathbb{E}(e^\top e)^{\frac{p-1}{2}} + l_3 \mathbb{E}(e^\top e)^{\frac{p}{2}-2})$, $l_1 \ge 0, l_2 \ge 0$ and $l_3 \le 0$ are constants.

Consider a deterministic network such that $\bar{c}_{\mathcal{F}} - \epsilon \underline{c}_{\mathcal{H}}\lambda < 0$. Theorem 4 shows that the desynchronizing effect of adding independent intrinsic noise is mitigated to some extent under appropriate conditions.

The results in Theorems 3 and 4 carry the flavor of practical synchronization studied in networks of deterministic nonlinear oscillators [31]. Specifically, from the expression of η in Theorem 3, one can obtain the set of parameters (coupling strength, noise intensity, etc.) that can drive the system arbitrarily close to synchronization.

In summary, Theorems 3 and 4 provide sufficient condition for stochastic synchronization. Both theorems are stated for heterogeneous noise and guarantee approximate synchronization. Strong multiplicative noise can destroy synchronization in the first theorem, while in the second theorem, multiplicative noise with a linear lower bound can mitigate desynchronization to some extent.

4 Examples

In this section, we illustrate the results of Theorems 1-4 by a network of noisy Van der Pol oscillators described by

$$\begin{pmatrix} dx_1^{(i)} \\ dx_2^{(i)} \end{pmatrix} = \underbrace{\begin{pmatrix} x_1^{(i)} - \frac{1}{3}(x_1^{(i)})^3 - x_2^{(i)} \\ x_1^{(i)} \end{pmatrix}}_{\mathcal{F}} dt + \sigma\mathcal{K}(X^{(i)}) \underbrace{\begin{pmatrix} dW_1^{(i)} \\ dW_2^{(i)} \end{pmatrix}}_{dW^{(i)}}, \; i \in \{1, \dots, N\}, \tag{11}$$

that are coupled through a coupling function $H(\cdot)$ with coupling strength ϵ:

$$dX^{(i)} = \mathcal{F}(X^{(i)})dt + \sigma\mathcal{K}(X^{(i)})dW^{(i)} + \epsilon \left[\sum_{j=1}^{n} c_{ij} H(x_1^{(j)} - x_1^{(i)}) \quad 0\right]^\top dt, \tag{12}$$

where the state of oscillator i is denoted by $X^{(i)} = (x_1^{(i)}, x_2^{(i)})$, the intrinsic dynamics is given by $\mathcal{F}$, the state-dependent noise is given by $\mathcal{K}$ with constant noise intensity σ. c_{ij}'s are the edge weights in the interaction graph underlying the network. c_{ij} is set to 1 if an edge exists between nodes i and j; and is set to zero, otherwise. In Examples 1-2, we consider a common noise, i.e., $dW^{(i)} = dW$, for each i, as in Theorems 1-2. In contrast, in Examples 3-4, we consider independent noise as in Theorems 3-4. In the following examples, the noise terms $\sigma\mathcal{K}$ are chosen such that they satisfy the conditions of Theorems 1-4, respectively, and illustrate the corresponding results. We will illustrate synchronization among Van der Pol oscillators using the synchronization error defined by

$$e(t) = \frac{1}{N}\sum_{i=1}^{N}\left(\left\|X^{(i)}(t) - \frac{1}{N}\sum_{j=1}^{N} X^{(j)}(t)\right\|_1\right).$$

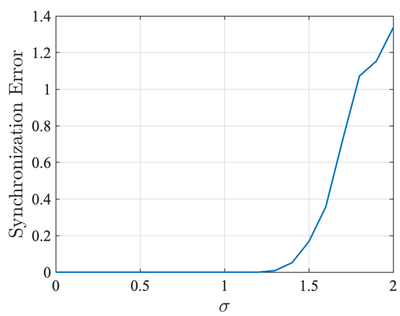
In the following, we consider the synchronization error $e(t)$ at $t = 200$ sec as the steady-state synchronization error. For the numerical simulations, we compute the steady-state synchronization error by averaging the steady-state synchronization error for 20 realizations of noise sequence.

Example 1. *In this example, we assume a common noise in* (11)-(12) *and let* $\epsilon = 2$. *We set* $H(x) = x$, *i.e., the coupling is diffusive. We select the interaction network as a (fixed) realization of Erdös-Réyni graph with* 10 *nodes in which an edge exists between any two distinct pair of nodes with probability* 0.5.

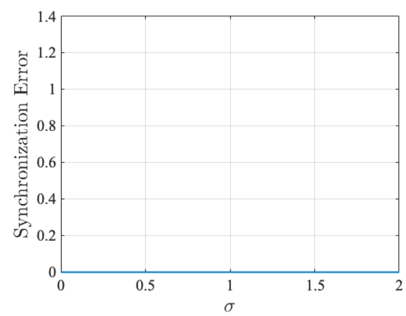
In Figure 1(a), we consider a multiplicative noise $\mathcal{K}(X^{(i)}) = \mathrm{diag}(\sin(10x_1^{(i)}), \sin(10x_2^{(i)}))$ *and show the steady-state synchronization error as a function of noise intensity* σ. *Observe that the network synchronizes in the absence of noise,* $\sigma = 0$; *it preserves its synchronization as noise increases slightly and it loses its synchronization as noise becomes large. In this example,* $\bar{c}_{\mathcal{F}} = 1$, $\underline{c}_{\mathcal{H}} = 1$,

$\lambda \approx 1$, and $\bar{c}_{\mathcal{K}} = 10$. Thus, Theorem 1 guarantees synchronization for $\sigma < 0.14$. However, as seen in Figure 1(a), the synchronization is preserved until $\sigma \approx 1.25$, which suggests that while the sufficient conditions in Theorem 1 capture the qualitative behavior of the system, they are conservative.

In Figure 1(b), we consider an additive noise $\mathcal{K}(X^{(i)}) = 1$. Observe that the network preserves its synchronization for all values of σ. Here, $\bar{c}_{\mathcal{K}} = 0$ and $c > 0$ for all values of σ. Thus, these numerical simulations are consistent with Theorem 1.



(a) Multiplicative noise.



(b) Additive noise.

Figure 1: Steady-state synchronization error for 10 Van der Pol oscillators coupled through a random graph with coupling strength $\epsilon = 2$ as described in Equation (12) with $H(x) = x$. (a) Synchronization error for *multiplicative* noise. The system retains synchronization for small noise, while it loses synchronization when the noise intensity increases (as expected by Theorem 1). (b) The network preserves its synchronization behavior even in the presence of large *additive* noise (as expected by Theorem 1, since $\bar{c}_{\mathcal{K}} = 0$). See Example 1.

Example 2. *In this example, we select the interaction graph as a line graph with 3 nodes. We assume a common noise in (11)-(12) and let $\epsilon = 0$ (no edge coupling) and*
$\mathcal{K}(X^{(i)}) = \begin{bmatrix} 1+x_1^{(i)}+\sin(x_1^{(i)}/10) & 0 \\ 0 & 1+x_2^{(i)}+\sin(x_2^{(i)}/10) \end{bmatrix}$*. It can be verified that each entry of $\mathcal{K}(X^{(i)})$ satisfies the assumptions of Theorem 2. Figure 2 shows the steady-state synchronization error as a function of the noise intensity σ. The system does not synchronize in the absence of edge coupling and small common noise, and synchronization is achieved for large common noise even in the absence of edge coupling. For this example $\underline{c}_{\mathcal{K}} \approx 0.9$ and $\bar{c}_{\mathcal{K}} \approx 1.1$. It can be verified that Theorem 2 requires σ to be at least $\sqrt{2}$ to guarantee synchronization; however, synchronization is achieved at $\sigma \approx 0.3$ suggesting that the conditions on Theorem 2 are conservative.*

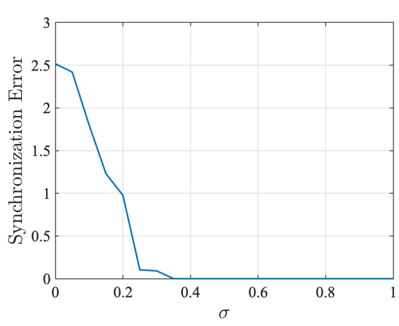


Figure 2: Noise induced synchronization of three Van der Pol oscillators as described in Equation (12) with no edge-coupling ($\epsilon = 0$) and common noise $\mathcal{K}(X^{(i)}) = \begin{bmatrix} 1+x_1^{(i)}+\sin(x_1^{(i)}/10) & 0 \\ 0 & 1+x_2^{(i)}+\sin(x_2^{(i)}/10) \end{bmatrix}$. No synchronization is achieved when σ is small, while synchronization is achieved when σ is large enough. See Example 2.

Example 3. *In this example, we assume heterogeneous noise (independent intrinsic noise) in (11)-(12). We select the same network and parameters as in Example 1. Figure 3 shows that the steady-state synchronization error increases with noise intensity σ, which is consistent with Theorem 3.*

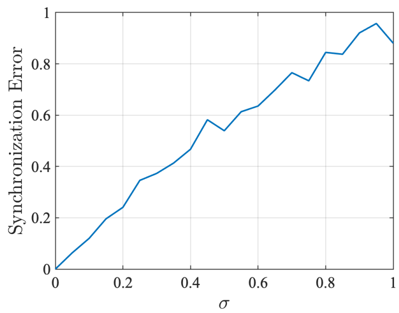


Figure 3: Synchronization error of van der Pol oscillators described in Equation (12) with $H(x) = x$ and heterogeneous noise of intensity σ. The steady-state synchronization error increases with the noise intensity σ. See Example 3.

Example 4. *In this example, we assume heterogeneous noise (independent intrinsic noise) in (11)-(12) and select the same network and parameters as in Example 2. Differently from Example 2, we let $\epsilon = 0.2$ and take $H(x) = \sin(x)$. Figure 4 shows asynchrony for small σ values. However, it should be noted that the evolution of the system for longer time (≈ 300 sec) does lead to synchronization. Approximate synchronization is achieved for moderate values of noise intensity $\sigma \approx 0.05$ and the steady-state synchronization error is reduced. Thus, adding moderate heterogeneous noise expedites convergence to achieve (approximate) synchronization. Further increase in σ results in a higher value of η and leads to a larger steady-state synchronization error.*

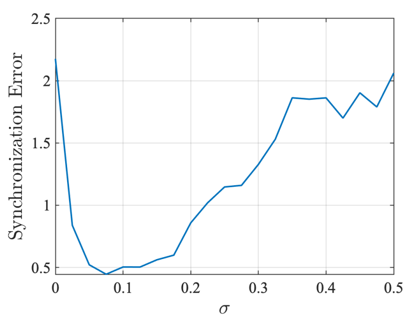


Figure 4: Noise induced expedited approximate synchronization of Van der Pol oscillators with small edge-coupling ($\epsilon = 0.2$) and heterogeneous noise with $\mathcal{K}(X^{(i)}) = \begin{bmatrix} 1+x_1^{(i)}+\sin(x_1^{(i)}/10) & 0 \\ 0 & 1+x_2^{(i)}+\sin(x_2^{(i)}/10) \end{bmatrix}$ as described in Equation (12). No synchronization at $\sigma = 0$ improves for approximate synchronization for moderate values of σ. For large σ, while approximate synchronization is achieved, the associated value of η may be quite large. See Example 4.

5 Discussion

Typically it is assumed that noise plays a destructive role and desynchronizes a network of synchronized oscillators (e.g. [36]). However, it has been observed both experimentally and theoretically that adding noise not only does not destroy the already synchronized networks but also can aid synchronization in non-synchronized networks; see [37] and [38] for a review.

In this paper, we studied the synchronization behavior of stochastic networks with nonlinear state-dependent noise terms described in Equations (1) and (6). These equations represent a broad range of network dynamics that can model many biological systems. For example, these frameworks cover the interconnected Kuramoto phase oscillators that model the brain's neural activity where the neural dynamics are subject to noise. The level of a functional connection between two regions is proportional to synchronization between the oscillators' phases associated with the two regions [39]. These frameworks also cover neuronal models such as Hodgkin-Huxley, Morris-Lecar, and FitzHugh-Nagumo which are connected through gap junctions. As another example, these frameworks cover coupled bursting models [40, 41] that approximate the dynamics of coupled central pattern generators (CPGs) [42, 43] which are complex networks of neurons that produce rhythmic behaviors, such as walking. Synchronization properties and cluster formation of coupled CPGs explain the generation of various gait patterns in animal locomotion [10, 11].

In Theorems 1 and 3 we studied destructive effects of noise on networks' synchronization properties: we identified a class of synchronized networks in which adding any additive noise or weak multiplicative noise does not ruin (approximate) synchronization while adding strong multiplicative noise desynchronizes the network. In Theorems 2 and 4, in contrast, we studied constructive effects of noise on networks' synchronization properties. In Theorem 2, we identified a class of multiplicative noise that can aid synchronization in desynchronized networks. Such behavior is reported for example in [18] for a large-scale model of the human brain network. In Theorem 4, we showed that heterogeneous multiplicative noise with a linear lower bound can mitigate desynchronization to some extent. The conditions of our theorems are easy to check. The only extra condition that we imposed to a noise term to foster synchronization was a linear lower bound (compare Theorems 1 and 2 or Theorems 3 and 4).

The ideas discussed in this paper can be further explored in several possible directions. First, we studied the cases of independent intrinsic noise. An interesting avenue is to explore the case with partially correlated noise. This is specifically interesting when the network is spatially embedded and the intrinsic noise is correlated due to the spatial proximity of systems. Another interesting direction of investigation is to understand the trade-off between the robustness of noise-induced synchronization and the precision of the oscillator. Specifically, adding common noise can make synchronization behavior more robust at the cost of the precision in the oscillator timing. Understanding the class of cost functions that may underlie handling this tradeoff can provide rich insights into biological systems.

6 Proofs of Theorems

Proof of Theorem 1. The proof has three main steps:

Step 1. Introducing a synchronization manifold. Let $(\phi_1, \ldots, \phi_N)^\top$ be a solution of (1), $\psi(t) := \frac{1}{N}\sum_{i=1}^{N}\phi_i(t)$ be the average of ϕ_i's, and $e_i := \phi_i - \psi$ be the corresponding error. The dynamics of $(e_1, \dots, e_N, \psi)$ can be written as:

$$\begin{pmatrix} de_1 \\ \vdots \\ de_N \end{pmatrix} = \begin{pmatrix} 1-\frac{1}{N} & -\frac{1}{N} & \cdots & -\frac{1}{N} \\ & \ddots & & \\ -\frac{1}{N} & -\frac{1}{N} & \cdots & 1-\frac{1}{N} \end{pmatrix}_{N\times N} \left(\begin{pmatrix} \mathcal{F}(e_1+\psi, t) \\ \vdots \\ \mathcal{F}(e_N+\psi, t) \end{pmatrix} + \epsilon \begin{pmatrix} H_1(e,\psi) \\ \vdots \\ H_N(e,\psi) \end{pmatrix} \right) dt \tag{13a}$$

$$+ \begin{pmatrix} 1-\frac{1}{N} & -\frac{1}{N} & \cdots & -\frac{1}{N} \\ & \ddots & & \\ -\frac{1}{N} & -\frac{1}{N} & \cdots & 1-\frac{1}{N} \end{pmatrix}_{N\times N} \begin{pmatrix} \sigma\mathcal{K}(e_1+\psi, t)dW \\ \vdots \\ \sigma\mathcal{K}(e_N+\psi, t)dW \end{pmatrix} \tag{13b}$$

$$+ \begin{pmatrix} 1-\frac{1}{N} & -\frac{1}{N} & \cdots & -\frac{1}{N} \\ & \ddots & & \\ -\frac{1}{N} & -\frac{1}{N} & \cdots & 1-\frac{1}{N} \end{pmatrix}_{N\times N} \begin{pmatrix} C_1(e,\psi) \mid \cdots \mid C_N(e,\psi) \end{pmatrix} \begin{pmatrix} d\mathbf{W}_1 \\ \vdots \\ d\mathbf{W}_N \end{pmatrix}, \tag{13c}$$

$$d\psi = \frac{1}{N}\sum_{i=1}^{N} \left(\mathcal{F}(e_i+\psi,t) + \epsilon H_i(e,\psi)\right) dt + \frac{\delta}{N}\sum_{i,j=1}^{N} c_{ji}\mathcal{C}(e_j, e_i) dW_{ij} + \frac{\sigma}{N}\sum_{i=1}^{N}\mathcal{K}(e_i+\psi,t)dW, \tag{13d}$$

where in (13a), for $i = 1, \dots, N$,

$$H_i(e,\psi) = \sum_{j=1}^{N} c_{ij}\mathcal{H}(e_j+\psi, e_i+\psi),$$

and in (13c), $C_i(e,\psi)$ is an $N \times N$ matrix with its $i-$th row $\delta(c_{i1}\mathcal{C}(e_1+\psi, e_i+\psi), \dots, c_{iN}\mathcal{C}(e_N+\psi, e_i+\psi)$ and its other rows are zero row vectors, and $d\mathbf{W}_i = (dW_{i1}, \dots, dW_{iN})^\top$ is an $N-$dimensional Wiener increment. We denote the $N \times N$ matrix in (13a)-(13c) by A.

Let $e = (e_1, \dots, e_N)^\top$ and $\boldsymbol{y} = (e_1, \dots, e_N, \psi)^\top$, and define $V(\boldsymbol{y}, t) = \frac{1}{2}e^\top e$. Note that the set of zeros of V is

$$\mathcal{S} := \left\{ (e_1, \dots, e_N, \psi, t)^\top \in \Omega_2^N \times \Omega_1 \times [0,\infty) \;\middle|\; e_1 = \cdots = e_N = 0, \text{ almost surely} \right\}.$$

This set is a candidate for the desired synchronization manifold. In the following two steps we show that if $c > 0$, then $\mathcal{S}$ is an exponentially stable invariant set for (13a)-(13d) and therefore it is the synchronization manifold.

Step 2. Invariant property of the synchronization manifold. Note that the Itô derivative of V is equal to

$$dV(\boldsymbol{y}, t) = \mathcal{L}V(\boldsymbol{y}, t)dt + V_{\boldsymbol{y}}(\boldsymbol{y}, t)^\top g(\boldsymbol{y}, t) d\mathcal{W},$$

where $d\mathcal{W}$ is a one dimensional Wiener increment and

$$\mathcal{L}V(\boldsymbol{y}, t) := V_t(\boldsymbol{y}, t) + V_{\boldsymbol{y}}(\boldsymbol{y}, t)^\top f(\boldsymbol{y}, t) + \frac{1}{2}\text{tr}\left[g^\top(\boldsymbol{y}, t) V_{\boldsymbol{yy}}(\boldsymbol{y}, t) g(\boldsymbol{y}, t)\right]. \tag{14}$$

The $(N+1)-$dimensional vectors $f(\boldsymbol{y}, t)$ and $g(\boldsymbol{y}, t)$ are respectively the drift and diffusion terms of (13a)-(13d), $V_t = \frac{\partial V}{\partial t} = 0$, $V_{\boldsymbol{y}} = \frac{\partial V}{\partial \boldsymbol{y}} = (e^\top, 0)^\top$, and $V_{\boldsymbol{yy}}(\boldsymbol{y}, t)$ is the $(N+1) \times (N+1)$ Hessian matrix of V which is a diagonal matrix with all entries equal to 1 except the last diagonal entry which is equal to 0. The trace operator is denoted by $\text{tr}[\cdot]$. We show that there exists $c_{\mathcal{L}} > 0$ such that $\mathcal{L}V \leq -c_{\mathcal{L}}V$. Then by [44, Theorem 1] we conclude that $\mathcal{S}$ is an invariant set for (13a)-(13d).

- Because $e_1 + \cdots + e_N = 0$, $e^\top A = e^\top$, and $e^\top \begin{pmatrix} \mathcal{F}(\psi,t) \\ \vdots \\ \mathcal{F}(\psi,t) \end{pmatrix} = 0$. Therefore, the second term of the right hand side of (14) becomes:

$$\begin{aligned} V_{\boldsymbol{y}}(\boldsymbol{y},t)^\top f(\boldsymbol{y},t) &= (e^\top, 0)^\top f(e,\psi,t) \\ &= (e_1,\ldots,e_N)\left\{ \begin{pmatrix} \mathcal{F}(e_1+\psi,t) \\ \vdots \\ \mathcal{F}(e_N+\psi,t) \end{pmatrix} - \begin{pmatrix} \mathcal{F}(\psi,t) \\ \vdots \\ \mathcal{F}(\psi,t) \end{pmatrix} + \epsilon \begin{pmatrix} H_1(e,\psi) \\ \vdots \\ H_N(e,\psi) \end{pmatrix} \right\} \\ &= \sum_{i=1}^N e_i(\mathcal{F}(e_i+\psi,t) - \mathcal{F}(\psi,t)) + \epsilon \sum_{i=1}^N e_i H_i(e,\psi) \end{aligned}$$

The first sum satisfies $\sum_{i=1}^N e_i(\mathcal{F}(e_i+\psi,t) - \mathcal{F}(\psi,t)) \le \bar{c}_{\mathcal{F}} \sum_{i=1}^N e_i^2 = 2\bar{c}_{\mathcal{F}} V(\boldsymbol{y},t)$, following condition (i) and the definition of V. By condition (ii) and using $c_{ij} = c_{ji}$, the second sum satisfies

$$\begin{aligned} \epsilon \sum_{i=1}^N e_i H_i(e,\psi) &= \epsilon \sum_{i=1}^N e_i \sum_{j=1}^N c_{ji} \mathcal{H}(e_j+\psi, e_i+\psi) \\ &= \frac{\epsilon}{2} \sum_{i=1}^N \sum_{j=1}^N c_{ji}(e_i \mathcal{H}(e_j+\psi, e_i+\psi) + e_j \mathcal{H}(e_i+\psi, e_j+\psi)) \\ &= -\frac{\epsilon}{2} \sum_{i=1}^N \sum_{j=1}^N c_{ji}(e_i - e_j)\mathcal{H}(e_i+\psi, e_j+\psi) && \text{condition (ii)} \\ &< -\frac{\epsilon}{2} \sum_{i=1}^N \sum_{j=1}^N c_{ji} \underline{c}_{\mathcal{H}} (e_i - e_j)^2 && \text{condition (ii)} \\ &= -\epsilon \underline{c}_{\mathcal{H}} e^\top L_{[c]} e. \end{aligned}$$

Since $e^\top v_1 = 0$, where $v_1 = (1,\ldots,1)^\top$ is the eigenvector of $L_{[c]}$ corresponding to $\lambda_{1,[c]} = 0$, by min-max theorem, $\lambda_{2,[c]} e^\top e \le e^\top L_{[c]} e \le \lambda_{N,[c]} e^\top e$. Therefore, depending on the sign of $\underline{c}_{\mathcal{H}}$, we have:

$$\begin{aligned} \epsilon \sum_{i=1}^N e_i H_i(e,\psi) &< -\epsilon \underline{c}_{\mathcal{H}} e^\top L_{[c]} e \le -\epsilon \underline{c}_{\mathcal{H}} \lambda_{2,[c]} e^\top e = -2\epsilon \underline{c}_{\mathcal{H}} \lambda_{2,[c]} V(\boldsymbol{y},t) && \text{for } \underline{c}_{\mathcal{H}} > 0, \text{ or} \\ \epsilon \sum_{i=1}^N e_i H_i(e,\psi) &< -\epsilon \underline{c}_{\mathcal{H}} e^\top L_{[c]} e \le -\epsilon \underline{c}_{\mathcal{H}} \lambda_{N,[c]} e^\top e = -2\epsilon \underline{c}_{\mathcal{H}} \lambda_{N,[c]} V(\boldsymbol{y},t) && \text{for } \underline{c}_{\mathcal{H}} < 0. \end{aligned}$$

Therefore, $V_{\boldsymbol{y}}(\boldsymbol{y},t)^\top f(\boldsymbol{y},t) \le (2\bar{c}_{\mathcal{F}} - 2\epsilon \underline{c}_{\mathcal{H}} \lambda) V(\boldsymbol{y},t)$.

- Because $A \begin{pmatrix} \sigma\mathcal{K}(\psi,t)dW \\ \vdots \\ \sigma\mathcal{K}(\psi,t)dW \end{pmatrix} = 0$, (13b) can be replaced by $AK(e,\psi,t)dW$ where

$$K(e,\psi,t) = \Big(\sigma(\mathcal{K}(e_1+\psi,t) - \mathcal{K}(\psi,t)), \ldots, \sigma(\mathcal{K}(e_N+\psi,t) - \mathcal{K}(\psi,t))\Big)^\top.$$

A straightforward matrix multiplication implies that the third term of $\mathcal{L}V$ satisfies:

$$\frac{1}{2}\mathrm{tr}\left[g^\top(\boldsymbol{y},t) V_{\boldsymbol{yy}}(\boldsymbol{y},t) g(\boldsymbol{y},t)\right] = \frac{\delta^2}{2}\left(1 - \frac{1}{N}\right) \sum_{i=1}^N \sum_{j=1}^N c_{ij}^2 \mathcal{C}^2(e_j+\psi, e_i+\psi) + \frac{1}{2}\|AK(e,\psi,t)\|^2,$$

where by condition (iii)

$$\frac{\delta^2}{2}\left(1-\frac{1}{N}\right)\sum_{i=1}^{N}\sum_{j=1}^{N} c_{ij}^2\mathcal{C}^2(e_j+\psi, e_i+\psi) \le \frac{\delta^2}{2}\left(1-\frac{1}{N}\right)\bar{c}_{\mathcal{C}}^2\sum_{i=1}^{N}\sum_{j=1}^{N} c_{ij}^2(e_j-e_i)^2$$
$$= \delta^2\left(1-\frac{1}{N}\right)\bar{c}_{\mathcal{C}}^2 e^\top L_{[c^2]} e$$
$$\le \delta^2\left(1-\frac{1}{N}\right)\bar{c}_{\mathcal{C}}^2 \lambda_{N,[c^2]} e^\top e$$
$$= 2\delta^2\left(1-\frac{1}{N}\right)\bar{c}_{\mathcal{C}}^2 \lambda_{N,[c^2]} V(\boldsymbol{y},t).$$

Using $A^\top A = A$ and simple multiplication, we can show that $K^\top A^\top A K = \left(1-\frac{1}{N}\right)K^\top K$. Then by condition (iv)

$$\frac{1}{2}\|AK(e,\psi,t)\|^2 = \frac{1}{2}\left(1-\frac{1}{N}\right)\|K(e,\psi,t)\|^2$$
$$= \frac{\sigma^2}{2}\left(1-\frac{1}{N}\right)\sum_{i=1}^{N}(\mathcal{K}(e_i+\psi,t)-\mathcal{K}(\psi,t))^2$$
$$\le \frac{\sigma^2}{2}\left(1-\frac{1}{N}\right)\sum_{i=1}^{N}\bar{c}_{\mathcal{K}}^2 e_i^2$$
$$= \sigma^2\left(1-\frac{1}{N}\right)\bar{c}_{\mathcal{K}}^2 V(\boldsymbol{y},t).$$

Therefore, $\mathcal{L}V(\boldsymbol{y},t) \le -c_{\mathcal{L}}V(\boldsymbol{y},t)$ where $c_{\mathcal{L}} = c = -2\bar{c}_{\mathcal{F}} + 2\epsilon \underline{c}_{\mathcal{H}}\lambda - 2\delta^2\bar{c}_{\mathcal{C}}^2(1-\frac{1}{N})\lambda_{N,[c^2]} - \left(1-\frac{1}{N}\right)\sigma^2\bar{c}_{\mathcal{K}}^2$. If $c > 0$ then $\mathcal{L}V \le -cV < 0$, and by [44, Theorem 1], $\mathcal{S}$ becomes an invariant set for (13a)-(13d). In Step 3 below, we use $\mathbb{E}\mathcal{L}V(\boldsymbol{y},t) \le -c_{\mathcal{L}}\mathbb{E}V(\boldsymbol{y},t)$.

Step 3. Stability of the synchronization manifold. As we discussed in Step 2, the Itô derivative of V is $dV(\boldsymbol{y},t) = \mathcal{L}(V(\boldsymbol{y},t))dt + V_{\boldsymbol{y}}^\top g(\boldsymbol{y},t)d\mathcal{W}$. By Dynkin's formula [45, Theorem 7.4.1]:

$$\mathbb{E}V(\boldsymbol{y}(t),t) - \mathbb{E}V(\boldsymbol{y}(0),0) = \mathbb{E}\int_0^t \mathcal{L}V(\boldsymbol{y}(\tau),\tau)\,d\tau \qquad \text{Dynkin's formula,}$$
$$= \int_0^t \mathbb{E}\mathcal{L}V(\boldsymbol{y}(\tau),\tau)\,d\tau \qquad \text{Fubini's Theorem,}$$
$$\le -c\int_0^t \mathbb{E}V(\boldsymbol{y}(\tau),\tau)\,d\tau \qquad \text{Step 2.}$$

The second equality holds because $\mathbb{E}\mathcal{L}V(\boldsymbol{y}(\tau),\tau)$ is a continuous function of τ and hence its integral on $[0,t]$ is finite. Let $h(t) = \mathbb{E}V(\boldsymbol{y}(t),t)$, then for $\delta t > 0$

$$h(t+\delta t) - h(t) \le -c\int_t^{t+\delta t} h(\tau)\,d\tau.$$

Dividing both sides by δt and letting $\delta t \to 0^+$, we obtain

$$D^+h(t) \le -ch(t), \tag{15}$$

where D^+ is the upper Dini derivative of h. Applying comparison lemma, [46, Lemma 3.4]:

$$h(t) \le h(0)e^{-ct}.$$

Hence,

$$\mathbb{E}V(\boldsymbol{y}(t),t) \leq \mathbb{E}V(\boldsymbol{y}(0),0)e^{-ct} \quad \Rightarrow \quad \mathbb{E}\|e(t)\|^2 \leq \mathbb{E}\|e(0)\|^2 \, e^{-ct},$$

or equivalently,

$$\mathbb{E}\sum_{i=1}^{N} |\phi_i(t) - \psi(t)|^2 \leq \mathbb{E}\sum_{i=1}^{N} |\phi_i(0) - \psi(0)|^2 \, e^{-ct}.$$

If $c > 0$, then $\mathcal{S}$ becomes exponentially stable. Since $e_i = \phi_i - \psi = 0$ almost surely, by the definition of stochastic synchronization and Step 2, $\mathcal{S}$ becomes a synchronization manifold.

Note that Step 3 can also be followed from [33, Chapter 4, Theorem 4.4]. We provided Step 3, which contains a different approach than [33, Chapter 4, Theorem 4.4], for a self-contained proof.

□

Proof of Theorem 2. To prove Theorem 2, we use the following lemma which is a modified version of [33, Chapter 4, Corollary 4.6].

Lemma 1. *Consider $dx = f(x,t)dt + g(x,t)dW$ and assume that there exist constants α_1 and α_2 such that for any $t \geq 0$,*

$$x^\top f(x,t) + \frac{1}{2}\mathrm{tr}[g^\top(x,t)g(x,t)] \;\leq\; \alpha_1 x^\top x, \quad \text{and} \tag{16}$$

$$\alpha_2 x^\top x \leq \|x^\top g(x,t)\|. \tag{17}$$

If $0 \leq \alpha_1 < \alpha_2^2$, then the trivial solution of $dx = f(x,t)dt + g(x,t)dW$ is $p-$th moment exponentially stable for $0 < p < 2(1 - \frac{\alpha_1}{\alpha_2^2}) \leq 2$ and $\alpha := -p[(\frac{p}{2} - 1)\alpha_2^2 + \alpha_1] > 0$, i.e., $\forall t > 0$

$$\mathbb{E}\|x(t)\|^p < \mathbb{E}\|x(0)\|^p e^{-\alpha t}.$$

Under the conditions of Theorems 1 and 2, we apply Lemma 1 to (13a). The left hand side of (16) is equivalent to $\mathcal{L}(V(\boldsymbol{y},t))$ which we showed $\mathcal{L}(V(\boldsymbol{y},t)) \leq -cV(\boldsymbol{y},t) = -\frac{c}{2}e^\top e$. Therefore,

$\alpha_1 = -c/2$. Straightforward matrix multiplications yield:

$$
\begin{aligned}
&\|e^\top g\|^2 \\
&= |e^\top AK(e,\psi,t)|^2 + \|e^\top A\left(\ C_1(e,\psi) \mid \cdots \mid C_N(e,\psi)\ \right)\|^2 \\
&= |e^\top K(e,\psi,t)|^2 + \|e^\top \left(\ C_1(e,\psi) \mid \cdots \mid C_N(e,\psi)\ \right)\|^2 && e^\top A = e^\top \\
&= \sigma^2 \left(\sum_{i=1}^N e_i(\mathcal{K}(e_i+\psi,t) - \mathcal{K}(\psi,t)) \right)^2 \\
&\quad + \delta^2 \sum_{i=1}^N e_i^2 \sum_{j=1}^N c_{ij}^2 \mathcal{C}^2(e_j+\psi, e_i+\psi) \\
&= \sigma^2 \left(\sum_{i=1}^N (e_i+\psi-\psi)(\mathcal{K}(e_i+\psi,t) - \mathcal{K}(\psi,t)) \right)^2 \\
&\quad + \frac{\delta^2}{2} \sum_{i,j=1}^N c_{ij}^2 (e_i^2+e_j^2)\mathcal{C}^2(e_i+\psi, e_j+\psi) && \mathcal{C}^2(x,y) = \mathcal{C}^2(y,x) \\
&\geq (\sigma \underline{c}_{\mathcal{K}})^2 \left(\sum_{i=1}^N e_i^2 \right)^2 && \text{conditions (ii) \& (i) of Theorem 2} \\
&\quad + \frac{\delta^2}{2} \frac{\underline{c}_{\mathcal{C}}^2}{2} \sum_{i,j=1}^N c_{ij}^2 (e_i-e_j)^4 && \text{and } e_i^2 + e_j^2 \geq \frac{(e_i-e_j)^2}{2} \\
&\geq \left((\sigma \underline{c}_{\mathcal{K}})^2 + \frac{\delta^2}{2} \frac{\underline{c}_{\mathcal{C}}^2}{2} \frac{\lambda_{2,[c]}^2}{N} \right) (e^\top e)^2.
\end{aligned}
$$

The last inequality holds because of Hölder inequality $\sum c_{ij}^2(e_i-e_j)^4 = \sum(c_{ij}(e_i-e_j)^2)^2 \geq \frac{1}{N}\left(\sum c_{ij}(e_i-e_j)^2\right)^2$. Therefore, $\alpha_2^2 = (\sigma \underline{c}_{\mathcal{K}})^2 + \dfrac{\delta^2 \underline{c}_{\mathcal{C}}^2 \lambda_{2,[c]}^2}{4N}$. By Lemma 1, for $p < 2 - 2\alpha_1/\alpha_2^2$,

$$
\mathbb{E}\left(\sum_{i=1}^N |\phi_i(t) - \psi(t)|^2 \right)^{p/2} \leq \mathbb{E}\left(\sum_{i=1}^N |\phi_i(0) - \psi(0)|^2 \right)^{p/2} e^{-\alpha t},
$$

where $\alpha := -p[(\frac{p}{2}-1)\alpha_2^2 + \alpha_1] > 0$.

The proof of almost sure exponential stability is straightforward by [33, Chapter 4, Theorem 3.3]. □

Proof of Theorem 3. The proof is very similar to the proof of Theorem 1, except that (13b) becomes

$$
A \begin{pmatrix} \sigma_1 \mathcal{K}(e_1+\psi,t) dW_1 \\ \vdots \\ \sigma_N \mathcal{K}(e_N+\psi,t) dW_N \end{pmatrix},
$$

or equivalently, $A\hat{K}(e,\psi,t)\begin{pmatrix} dW_1 \\ \vdots \\ dW_N \end{pmatrix}$, where

$$\hat{K}(e,\psi,t) := \operatorname{diag}\big(\sigma_i \mathcal{K}(e_i+\psi,t)\big). \tag{18}$$

Next, we compute $\frac{1}{2}\text{tr}\left[(A\hat{K})^\top A\hat{K}\right]$. Since $A^\top A = A$, elementary calculations show that

$$\begin{aligned}
\frac{1}{2}\text{tr}\left[(A\hat{K})^\top A\hat{K}\right] &= \frac{1}{2}\Big(1-\frac{1}{N}\Big)\sum_{i=1}^{N}\sigma_i^2(\mathcal{K}(e_i+\psi,t))^2 && \text{(19a)}\\
&= \frac{1}{2}\Big(1-\frac{1}{N}\Big)\sum_{i=1}^{N}\sigma_i^2(\mathcal{K}(e_i+\psi,t)-\mathcal{K}(\psi,t)+\mathcal{K}(\psi,t))^2 && \text{(19b)}\\
&\le \Big(1-\frac{1}{N}\Big)\Big(\sum_{i=1}^{N}\sigma_i^2(\mathcal{K}(e_i+\psi,t)-\mathcal{K}(\psi,t))^2+\sum_{i=1}^{N}\sigma_i^2\mathcal{K}(\psi,t)^2\Big). && \text{(19c)}
\end{aligned}$$

(19c) holds by elementary inequality $\frac{1}{2}(a+b)^2 \le a^2+b^2$. The first term of (19c) satisfies:

$$\begin{aligned}
\Big(1-\frac{1}{N}\Big)\mathbb{E}\sum_{i=1}^{N}\sigma_i^2(\mathcal{K}(e_i+\psi,t)-\mathcal{K}(\psi,t))^2 &\le \Big(1-\frac{1}{N}\Big)\mathbb{E}\sum_{i=1}^{N}\sigma_i^2\bar{c}_{\mathcal{K}}^2 e_i^2\\
&\le 2\Big(1-\frac{1}{N}\Big)\bar{c}_{\mathcal{K}}^2\max_i \sigma_i^2\ \mathbb{E}V(\boldsymbol{y},t),
\end{aligned}$$

where we use the Lipschitz property of $\mathcal{K}$. The second term of (19c) satisfies:

$$\begin{aligned}
\Big(1-\frac{1}{N}\Big)\mathbb{E}\sum_{i=1}^{N}\sigma_i^2\mathcal{K}(\psi,t)^2 &\le \Big(1-\frac{1}{N}\Big)\mathbb{E}\sum_{i=1}^{N}\sigma_i^2(\mathcal{K}(\psi,t)-\mathcal{K}(0,t)+\mathcal{K}(0,t))^2 && \text{(21a)}\\
&\le \Big(1-\frac{1}{N}\Big)\mathbb{E}\sum_{i=1}^{N}\sigma_i^2(\bar{c}_{\mathcal{K}}|\psi|+\mathcal{K}_0)^2 && \text{(21b)}\\
&\le \Big(1-\frac{1}{N}\Big)\mathbb{E}\sum_{i=1}^{N}\sigma_i^2(\bar{c}_{\mathcal{K}}^2|\psi|^2+\mathcal{K}_0^2+2\bar{c}_{\mathcal{K}}\mathcal{K}_0|\psi|) && \text{(21c)}\\
&\le \Big(1-\frac{1}{N}\Big)\sum_{i=1}^{N}\sigma_i^2(\bar{c}_{\mathcal{K}}\gamma+\mathcal{K}_0)^2, && \text{(21d)}
\end{aligned}$$

where we use Jensen's inequality $(\mathbb{E}|\psi|)^2 \le \mathbb{E}|\psi|^2 \le \gamma^2$. Therefore,

$$\frac{1}{2}\mathbb{E}\ \text{tr}\left[(A\hat{K})^\top A\hat{K}\right] \le 2\Big(1-\frac{1}{N}\Big)\bar{c}_{\mathcal{K}}^2\max_i\sigma_i^2\ \mathbb{E}V(\boldsymbol{y},t)+\Big(1-\frac{1}{N}\Big)\sum_{i=1}^{N}\sigma_i^2(\bar{c}_{\mathcal{K}}\gamma+\mathcal{K}_0)^2.$$

Following the proof of Theorem 1, $\mathbb{E}\mathcal{L}V(\boldsymbol{y},t) \le -c_a\mathbb{E}V(\boldsymbol{y},t)+\bar{\eta}$ where

$$c_a = -2\bar{c}_{\mathcal{F}}+2\epsilon\underline{c}_{\mathcal{H}}\lambda-2\delta^2\bar{c}_{\mathcal{C}}^2\Big(1-\frac{1}{N}\Big)\lambda_{N,[c^2]}-2\Big(1-\frac{1}{N}\Big)\bar{c}_{\mathcal{K}}^2\max_i\sigma_i^2,$$

and $\bar{\eta} = \left(1 - \frac{1}{N}\right) \sum_{i=1}^{N} \sigma_i^2 (\bar{c}_{\mathcal{K}} \gamma + \mathcal{K}_0)^2$.

In what follows, we prove that if $\mathbb{E}\mathcal{L}V(\boldsymbol{y}, t) \leq -c_a \mathbb{E}V(\boldsymbol{y}, t) + \bar{\eta}$, for some constants c_a and $\bar{\eta}$, then

$$\mathbb{E} \sum_{i=1}^{N} |\phi_i(t) - \psi(t)|^2 \leq \left(\mathbb{E} \sum_{i=1}^{N} |\phi_i(0) - \psi(0)|^2 - \frac{\bar{\eta}}{c_a} \right) e^{-c_a t} + \frac{\bar{\eta}}{c_a}, \quad \forall t > 0,$$

which is a generalization of Step 3 in the proof of Theorem 1.

Equation (15) from Step 3 of Theorem 1 becomes $D^+ h(t) \leq -c_a h(t) + \bar{\eta}$. Multiplying both sides by $e^{c_a t}$, we obtain $D^+(e^{c_a t} h(t)) \leq \bar{\eta}\, e^{c_a t}$. Note that since D^+ is subadditive [47, Appendix I], i.e., $D^+(f + g) \leq D^+ f + D^+ g$,

$$\begin{aligned} D^+ \left(e^{c_a t} h(t) - e^{c_a t} \frac{\bar{\eta}}{c_a} \right) &\leq D^+(e^{c_a t} h(t)) + D^+ \left(-e^{c_a t} \frac{\bar{\eta}}{c_a} \right) \\ &\leq \bar{\eta} e^{c_a t} - c_a e^{c_a t} \frac{\bar{\eta}}{c_a} = 0, \end{aligned}$$

which implies that $e^{c_a t} h(t) - e^{c_a t} \frac{\bar{\eta}}{c_a}$ is non-increasing:

$$e^{c_a t} h(t) - e^{c_a t} \frac{\bar{\eta}}{c_a} \leq h(0) - \frac{\bar{\eta}}{c_a},$$

and therefore (8) holds, as desired. □

Proof of Theorem 4. The proof follows from a generalization of Lemma 1. Let $W(e) = (e^\top e)^{\frac{p}{2}}$. It is easy to verify that

$$\mathcal{L}W = p(e^\top e)^{\frac{p}{2}-1} \left(e^\top f + \frac{1}{2} \mathrm{tr}[g^\top g] \right) + p \left(\frac{p}{2} - 1 \right) (e^\top e)^{\frac{p}{2}-2} \|e^\top g\|^2. \tag{23}$$

In what follows, we show that $\mathbb{E}\mathcal{L}W \leq -\beta\, \mathbb{E}W + \bar{\zeta}$. Then, similar to the argument that we made at the end of the proof of Theorem 3, we conclude our desired result.

Step 1. In Theorem 1, we showed that $e^\top f \leq (2\bar{c}_{\mathcal{F}} - 2\epsilon \underline{c}_{\mathcal{H}} \lambda) e^\top e$. Therefore,

$$\mathbb{E} \left[p(e^\top e)^{\frac{p}{2}-1} e^\top f \right] \leq p(\bar{c}_{\mathcal{F}} - \epsilon \underline{c}_{\mathcal{H}} \lambda) \mathbb{E}(e^\top e)^{\frac{p}{2}} = p(\bar{c}_{\mathcal{F}} - \epsilon \underline{c}_{\mathcal{H}} \lambda) \mathbb{E}W. \tag{24}$$

Step 2. In this step, we show that for some l_1 and l_2:

$$\frac{p}{2} \mathbb{E}(e^\top e)^{\frac{p}{2}-1} \mathrm{tr}[g^\top g] \leq \frac{p}{2} \Big(1 - \frac{1}{N}\Big) \bar{c}_{\mathcal{K}}^2 \max_i \sigma_i^2 \mathbb{E}W + l_1 \mathbb{E}(e^\top e)^{\frac{p}{2}-1} + l_2 \mathbb{E}(e^\top e)^{\frac{p-1}{2}},$$

where $g = A\hat{K}$ and $\hat{K}$ is as defined in (18).

$$\mathrm{tr}\left[(A\hat{K})^\top A\hat{K} \right] = \Big(1 - \frac{1}{N}\Big) \sum_{i=1}^{N} \sigma_i^2 (\mathcal{K}(e_i + \psi, t) - \mathcal{K}(\psi, t) + \mathcal{K}(\psi, t))^2 \tag{25a}$$

$$\leq \Big(1 - \frac{1}{N}\Big) \sum_{i=1}^{N} \sigma_i^2 (\mathcal{K}(e_i + \psi, t) - \mathcal{K}(\psi, t))^2 + \Big(1 - \frac{1}{N}\Big) \sum_{i=1}^{N} \sigma_i^2 \mathcal{K}(\psi, t)^2 \tag{25b}$$

$$+ 2\Big(1 - \frac{1}{N}\Big)^2 \Big(\sum_{i=1}^{N} \sigma_i^2 (\mathcal{K}(e_i + \psi, t) - \mathcal{K}(\psi, t))^2 \Big)^{\frac{1}{2}} \Big(\sum_{i=1}^{N} \sigma_i^2 \mathcal{K}(\psi, t)^2 \Big)^{\frac{1}{2}}, \tag{25c}$$

where we use (19a), expand $(a_i+b_i)^2$, and apply Hölder inequality $\sum_i a_i b_i \le (\sum_i a_i^2)^{\frac{1}{2}} (\sum_i b_i)^{\frac{1}{2}}$. The first term of (25b) satisfies:

$$\mathbb{E}(e^\top e)^{\frac{p}{2}-1} \sum_{i=1}^N \sigma_i^2 (\mathcal{K}(e_i+\psi,t) - \mathcal{K}(\psi,t))^2 \le \mathbb{E}(e^\top e)^{\frac{p}{2}-1} \sum_{i=1}^N \sigma_i^2 \bar{c}_{\mathcal{K}}^2 e_i^2 \le \bar{c}_{\mathcal{K}}^2 \max_i \sigma_i^2 \ \mathbb{E}(e^\top e)^{\frac{p}{2}},$$

where we use the Lipschitz property of $\mathcal{K}$. The second term of (25b) satisfies:

$$\begin{aligned}
\mathbb{E}(e^\top e)^{\frac{p}{2}-1} \sum_{i=1}^N \sigma_i^2 \mathcal{K}(\psi,t)^2 &\le \mathbb{E}(e^\top e)^{\frac{p}{2}-1} \mathbb{E} \sum_{i=1}^N \sigma_i^2 (\mathcal{K}(\psi,t) - \mathcal{K}(0,t) + \mathcal{K}(0,t))^2 \\
&\le \mathbb{E}(e^\top e)^{\frac{p}{2}-1} \mathbb{E} \sum_{i=1}^N \sigma_i^2 (\bar{c}_{\mathcal{K}}|\psi| + \mathcal{K}_0)^2 \\
&\le \mathbb{E}(e^\top e)^{\frac{p}{2}-1} \sum_{i=1}^N \sigma_i^2 (\bar{c}_{\mathcal{K}}\gamma + \mathcal{K}_0)^2.
\end{aligned}$$

We use the fact that e and ψ are independent and Jensen's inequality $(\mathbb{E}|\psi|)^2 \le \mathbb{E}|\psi|^2 \le \gamma^2$. Therefore,

$$l_1 = \frac{p}{2}\Big(1 - \frac{1}{N}\Big) \sum_{i=1}^N \sigma_i^2 (\bar{c}_{\mathcal{K}}\gamma + \mathcal{K}_0)^2.$$

Finally, (25c) satisfies:

$$\begin{aligned}
&\Big(\mathbb{E}(e^\top e)^{\frac{p}{2}-1} \sum_{i=1}^N \sigma_i^2 (\mathcal{K}(e_i+\psi,t) - \mathcal{K}(\psi,t))^2\Big)^{\frac{1}{2}} \Big(\mathbb{E}(e^\top e)^{\frac{p}{2}-1} \sum_{i=1}^N \sigma_i^2 \mathcal{K}(\psi,t)^2\Big)^{\frac{1}{2}} \\
&\le \bar{c}_{\mathcal{K}} \max_i \sigma_i \ (\bar{c}_{\mathcal{K}}\gamma + \mathcal{K}_0) \Big(\sum_{i=1}^N \sigma_i^2\Big)^{\frac{1}{2}} \mathbb{E}(e^\top e)^{\frac{p-1}{2}}.
\end{aligned}$$

Therefore,

$$l_2 = p\Big(1 - \frac{1}{N}\Big)^2 \bar{c}_{\mathcal{K}} \max_i \sigma_i \ (\bar{c}_{\mathcal{K}}\gamma + \mathcal{K}_0) \Big(\sum_{i=1}^N \sigma_i^2\Big)^{\frac{1}{2}}.$$

Step 3. In this step, we show that

$$p\left(\frac{p}{2} - 1\right) \mathbb{E}(e^\top e)^{\frac{p}{2}-2} \|e^\top g\|^2 \le p\left(\frac{p}{2} - 1\right) (\min_i \sigma_i^2 \underline{c}_{\mathcal{K}}^2) \mathbb{E}W + l_3 \ \mathbb{E}(e^\top e)^{\frac{p-1}{2}},$$

for some constant l_3.

$$\|e^\top A\hat{K}(e,\psi,t)\|^2 = \|e^\top \hat{K}(e,\psi,t)\|^2 \tag{29a}$$

$$= \sum_{i=1}^{N} \sigma_i^2 |e_i(\mathcal{K}(e_i+\psi,t) - \mathcal{K}(\psi,t)) + e_i\mathcal{K}(\psi,t)|^2 \tag{29b}$$

$$\geq \sum_{i=1}^{N} \frac{\sigma_i^2}{2} \left|e_i(\mathcal{K}(e_i+\psi,t) - \mathcal{K}(\psi,t))\right|^2 - \sigma_i^2\left|e_i\mathcal{K}(\psi,t)\right|^2 \tag{29c}$$

$$\geq \sum_{i=1}^{N} \frac{\sigma_i^2}{2} \underline{c}_{\mathcal{K}}^2 e_i^4 - \sum_{i=1}^{N} \sigma_i^2 (\bar{c}_{\mathcal{K}}|\psi| + \mathcal{K}_0)^2 \tag{29d}$$

$$\geq \min_i \sigma_i^2 \frac{\underline{c}_{\mathcal{K}}^2}{2N} \left(\sum_{i=1}^{N} e_i^2\right)^2 - \sum_{i=1}^{N} \sigma_i^2 (\bar{c}_{\mathcal{K}}|\psi| + \mathcal{K}_0)^2, \tag{29e}$$

where we use $e^\top A = e^\top$ in (29a) and the fact that $(a+b)^2 \geq \frac{a^2}{2} - b^2$ in (29c). (29d) follows from Condition (ii) of Theorem 2 and (21b). (29e) follows using Hölder inequality. Multiplying above inequality by $p\left(\frac{p}{2}-1\right)(e^\top e)^{\frac{p}{2}-2}$ and taking expectation

$$p\left(\frac{p}{2}-1\right)\mathbb{E}(e^\top e)^{\frac{p}{2}-2}\|e^\top A\hat{K}(e,\psi,t)\|^2 \geq p\left(\frac{p}{2}-1\right)\beta_2^2\mathbb{E}(e^\top e)^{\frac{p}{2}} + l_3\mathbb{E}(e^\top e)^{\frac{p}{2}-2},$$

where $\beta_2^2 = \min_i \sigma_i^2 \frac{\underline{c}_{\mathcal{K}}^2}{2N}$ and $l_3 = -p\left(\frac{p}{2}-1\right)\sum_{i=1}^{N}\sigma_i^2(\bar{c}_{\mathcal{K}}\gamma + \mathcal{K}_0)^2$ follows from (21d).

Combining Step1-Step3, we obtain:

$$\mathbb{E}\mathcal{L}W \leq -\beta\mathbb{E}W + \bar{\zeta},$$

where

$$\beta = -p\left(\bar{c}_{\mathcal{F}} - \epsilon\underline{c}_{\mathcal{H}}\lambda + \frac{1}{2}\Big(1-\frac{1}{N}\Big)\bar{c}_{\mathcal{K}}^2\max_i \sigma_i^2 + \left(\frac{p}{2}-1\right)\frac{\underline{c}_{\mathcal{K}}^2}{2N}\min_i \sigma_i^2\right) > 0,$$

and $\bar{\zeta} = l_1\mathbb{E}(e^\top e)^{\frac{p}{2}-1} + l_2\mathbb{E}(e^\top e)^{\frac{p-1}{2}} + l_3\mathbb{E}(e^\top e)^{\frac{p}{2}-2}$.

Since $\bar{\zeta}$ comprises exponents of $e^\top e$ that are smaller than $p/2$, for sufficiently large $\mathbb{E}W$, the $-\beta\mathbb{E}W$ term will dominate $\bar{\zeta}$, which makes $\mathbb{E}\mathcal{L}W$ non-positive. Thus, $\mathbb{E}W$ and $\bar{\zeta}$ will remain bounded.

The remainder of the proof follows similar to Theorem 3 with $h(t) = \mathbb{E}W$. □

Acknowledgement

This work was supported in part by Simon Foundations grant 712522, NSF grant IOS-2037828, and ARO grant W911NF-18-1-0325. The authors thank the editor and the anonymous reviewers for their valuable comments and constructive suggestions.

References

[1] A. T. Winfree, *The Geometry of Biological Time*, vol. 12. Springer Science & Business Media, 2001.

[2] F. C. Hoppensteadt and E. M. Izhikevich, *Weakly Connected Neural Networks*, vol. 126. Springer Science & Business Media, 2012.

[3] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Transactions on Circuits and Systems—I: Fundamental Theory and Applications*, vol. 47, no. 5, p. 655, 2000.

[4] A. Hajimiri and T. H. Lee, "A general theory of phase noise in electrical oscillators," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 2, pp. 179–194, 1998.

[5] A. Goldbeter, "Computational approaches to cellular rhythms," *Nature*, vol. 420, no. 6912, p. 238, 2002.

[6] G. B. Ermentrout and D. H. Terman, *Mathematical Foundations of Neuroscience*, vol. 35. Springer Science & Business Media, 2010.

[7] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*. Courier Corporation, 2003.

[8] A. T. Winfree, "Biological rhythms and the behavior of populations of coupled oscillators," *Journal of Theoretical Biology*, vol. 16, no. 1, pp. 15–42, 1967.

[9] C. Liu, D. R. Weaver, S. H. Strogatz, and S. M. Reppert, "Cellular construction of a circadian clock: Period determination in the suprachiasmatic nuclei," *Cell*, vol. 91, no. 6, pp. 855–860, 1997.

[10] Z. Aminzare, V. Srivastava, and P. Holmes, "Gait transitions in a phase oscillator model of an insect central pattern generator," *SIAM Journal on Applied Dynamical Systems*, vol. 17, no. 1, pp. 626–671, 2018.

[11] Z. Aminzare and P. Holmes, "Heterogeneous inputs to central pattern generators can shape insect gaits," *SIAM Journal on Applied Dynamical Systems*, vol. 18, no. 2, pp. 1037–1059, 2019.

[12] N. E. Leonard, T. Shen, B. Nabet, L. Scardovi, I. D. Couzin, and S. A. Levin, "Decision versus compromise for animal groups in motion," *Proceedings of the National Academy of Sciences*, vol. 109, no. 1, pp. 227–232, 2012.

[13] D. A. Paley, N. E. Leonard, R. Sepulchre, D. Grunbaum, and J. K. Parrish, "Oscillator models and collective motion," *IEEE Control Systems Magazine*, vol. 27, no. 4, pp. 89–105, 2007.

[14] D. J. Klein, *Coordinated Control and Estimation for Multi-agent Systems: Theory and Practice*. PhD thesis, University of Washington, 2008.

[15] F. Dorfler, M. Chertkov, and F. Bullo, "Synchronization in complex oscillator networks and smart grids," *Proceedings of the National Academy of Sciences*, vol. 110, no. 6, pp. 2005–2010, 2013.

[16] K. Wiesenfeld, P. Colet, and S. H. Strogatz, "Frequency locking in Josephson arrays: connection with the Kuramoto model," *Physical Review E*, vol. 57, no. 2, p. 1563, 1998.

[17] D. Michaels, E. P. Matyas, and J. Jalife, "Mechanisms of sinoatrial pacemaker synchronization: A new hypothesis.," *Circulation Research*, vol. 61, no. 5, pp. 704–714, 1987.

[18] J. C. Pang, L. L. Gollo, and J. A. Roberts, "Stochastic synchronization of dynamics on the human connectome," *NeuroImage*, vol. 229, p. 117738, 2021.

[19] J. D. Touboul, C. Piette, L. Venance, and G. B. Ermentrout, "Noise-induced synchronization and antiresonance in interacting excitable systems: Applications to deep brain stimulation in Parkinson's disease," *Physical Review X*, vol. 10, no. 1, p. 011073, 2020.

[20] M. Jafarian, M. H. Mamduhi, and K. H. Johansson, "Stochastic stability of discrete-time phase-coupled oscillators over uncertain and random networks," *arXiv preprint arXiv:2104.05477*, 2021.

[21] J. H. Meng and H. Riecke, "Synchronization by uncorrelated noise: interacting rhythms in interconnected oscillator networks," *Scientific reports*, vol. 8, no. 1, pp. 1–14, 2018.

[22] I. Aradi and I. Soltesz, "Modulation of network behaviour by changes in variance in interneuronal properties," *The Journal of Physiology*, vol. 538, no. 1, pp. 227–251, 2002.

[23] H. Sompolinsky, A. Crisanti, and H. J. Sommers, "Chaos in random neural networks," *Phys. Rev. Lett.*, vol. 61, pp. 259–262, Jul 1988.

[24] G. Hermann and J. Touboul, "Heterogeneous connections induce oscillations in large-scale networks," *Phys. Rev. Lett.*, vol. 109, p. 018702, Jul 2012.

[25] G. Russo and R. Shorten, "On common noise-induced synchronization in complex networks with state-dependent noise diffusion processes," *Physica D: Nonlinear Phenomena*, vol. 369, pp. 47–54, 2018.

[26] G. Russo, F. Wirth, and R. Shorten, "On synchronization in continuous-time networks of nonlinear nodes with state-dependent and degenerate noise diffusion," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 389–395, 2019.

[27] Q.-C. Pham, N. Tabareau, and J.-J. Slotine, "A contraction theory approach to stochastic incremental stability," *IEEE Transactions on Automatic Control*, vol. 54, no. 4, pp. 816–820, 2009.

[28] J.-N. Teramae and D. Tanaka, "Robustness of the noise-induced phase synchronization in a general class of limit cycle oscillators," *Physical Review Letters*, vol. 93, no. 20, p. 204103, 2004.

[29] C. Ly and G. B. Ermentrout, "Coupling regularizes individual units in noisy populations," *Phys. Rev. E*, vol. 81, p. 011911, Jan 2010.

[30] Z. Aminzare and V. Srivastava, "Phase reduction and synchronization of coupled noisy oscillators," *arXiv preprint arXiv:2103.04492*, 2021.

[31] J. M. Montenbruck, M. Bürger, and F. Allgöwer, "Practical synchronization with diffusive couplings," *Automatica*, vol. 53, pp. 235–243, 2015.

[32] W. He, W. Du, F. Qian, and J. Cao, "Synchronization analysis of heterogeneous dynamical networks," *Neurocomputing*, vol. 104, pp. 146–154, 2013.

[33] X. Mao, *Stochastic Differential Equations and Applications*. Woodhead Publishing, second ed., 2011.

[34] P. DeLellis, M. di Bernardo, and G. Russo, "On quad, Lipschitz, and contracting vector fields for consensus and synchronization of networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 3, pp. 576–583, 2011.

[35] T. Von, "Algebraic connectivity of Erdös-Rényi graphs near the connectivity threshold." Available at: https://www.mathstat.dal.ca/ tkolokol/papers/critscaling4.pdf, 2014.

[36] J. Teramae and Y. Kuramoto, "Strong desynchronizing effects of weak noise in globally coupled systems.," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 63 3 Pt 2, p. 036210, 2001.

[37] G. Ermentrout, R. Galán, and N. Urban, "Reliability, synchrony and noise," *Trends Neurosci.*, vol. 31, pp. 428–434, 2008.

[38] M. D. McDonnell and L. M. Ward, "The benefits of noise in neural systems: bridging theory and experiment," *Nature Reviews Neuroscience*, vol. 12, pp. 415–425, 2011.

[39] T. Menara, G. Baggio, D. S. Bassett, and F. Pasqualetti, "A framework to control functional connectivity in the human brain," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 4697–4704, 2019.

[40] R. M. Ghigliazza and P. Holmes, "Minimal models of bursting neurons: How multiple currents, conductances, and timescales affect bifurcation diagrams," *SIAM Journal on Applied Dynamical Systems*, vol. 3, no. 4, pp. 636–670, 2004.

[41] R. M. Ghigliazza and P. Holmes, "A minimal model of a central pattern generator and motoneurons for insect locomotion," *SIAM Journal on Applied Dynamical Systems*, vol. 3, no. 4, pp. 671–700, 2004.

[42] E. Marder and D. Bucher, "Central pattern generators and the control of rhythmic movements," *Current Biol.*, vol. 11, no. 23, pp. R986–R996, 2001.

[43] A. Ijspeert, "Central pattern generators for locomotion control in animals and robots: A review," *Neural Networks*, vol. 21, no. 4, pp. 642–653, 2008.

[44] A. M. Stanzhitskii, "Investigation of invariant sets of Itô stochastic systems with the use of Lyapunov functions," *Ukrainian Mathematical Journal*, vol. 53, pp. 323–327, 2001.

[45] E. B. Dynkin, *Markov Processes*, vol. 1. Springer-Verlag, 1965.

[46] H. K. Khalil, *Nonlinear systems*. Prentice Hall, third ed., 2002.

[47] N. Rouche, P. Habets, and M. Laloy, *Stability Theory by Liapunov's Direct Method*. Springer, 1977.