Network Synchronization with Convexity∗

Guodong Shi, Alexandre Proutiere, and Karl Henrik Johansson†

Abstract

In this paper, we establish some new synchronization conditions for complex networks with nonlinear and nonidentical self-dynamics with switching directed communication graphs. In light of the recent works on distributed sub-gradient methods, we impose some integral convexity for the nonlinear node self-dynamics in the sense that the self-dynamics of a given node is the gradient of some concave function corresponding to that node. The node couplings are assumed to be linear but with switching directed communication graphs. Several sufficient and/or necessary conditions are established regarding exact or approximate synchronization over the considered complex networks. These results show when and how nonlinear node self-dynamics may cooperate with the linear consensus coupling, which eventually leads to network synchronization conditions under much relaxed connectivity requirements.

Keywords: coupled oscillator, complex networks, synchronization, switching graphs

1 Introduction

The past few decades have witnessed tremendous research interest on the emergence of collective behaviors for dynamics over complex networks [9, 10, 11, 12]. The new understanding we have gained is that some global network-level tasks, such as synchronization or consensus, can be achieved by local interactions under cooperative couplings of individual node dynamics [14, 12].

∗This work has been supported in part by the Knut and Alice Wallenberg Foundation, the Swedish Research Council, and KTH SRA TNG. A preliminary version of this work has been presented at The 51st Annual Allerton Conference on Communication, Control, and Computing, Oct. 2013 [1].
†G. Shi is with the College of Engineering and Computer Science, The Australian National University, Canberra ACT 0200, Australia. A. Proutiere and K. H. Johansson are with ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Stockholm 10044, Sweden. Email: guodong.shi@anu.edu.au, alepro@kth.se, kallej@ee.kth.se
More advanced strategies have also been developed for problems like formation, swarming, optimization, or even signaling. Synchronization problems require the node states asymptotically reach a common trajectory or a common value all over the network. In [9], the master stability function method was proposed for the local synchronization of linearly coupled oscillators, where the dynamics of each node consists of a term of nonlinear self-dynamics and another term of local linear couplings. Then in [14], a thorough treatment was established for synchronization of linear diffusive couplings. When the self-dynamics of each node is nonlinear, it was shown that the coupling strength must dominate the influence of this self-dynamics in order for global synchronization [15, 12]. Further extensions for linearly coupled oscillators have been established under more restrictions on the individual self-dynamics, e.g., passivity or symmetry [16, 17], or linearity [18, 19]. All above works mainly focused on fixed interaction graph and identical self-dynamics, and efforts have also been made on synchronization under switching interactions, non-identical node self-dynamics, or nonlinear couplings, which turned out to be far more challenging [20, 21]. Some recent improvements to classical synchronization results include [22] and [23]. In [22], connectivity requirements are relaxed to jointly connected undirected graphs, where only the union of the switching communication graphs is assumed to be connected over certain intervals, for linear agent models. In [23], the authors provided a graph comparison perspective, based on which some new graphical conditions are obtained for synchronization conditions with nonlinear node self-dynamics but fixed communication graph.

The difficulty in analyzing synchronization conditions comes from the nontrivial coupling between node self-dynamics and the local interactions, as well as the coupling between different node states, especially for a switching communication graph. While without self-dynamics in each node, network synchronization falls into a pure distributed consensus problem. For consensus seeking, it has been shown that various convergence conditions can be derived based on much relaxed connectivity conditions with even directed node interactions [24, 25, 26, 27, 29, 30, 31, 32]. On the other hand, it has also been shown that if the node self-dynamics can be properly designed, this node self-dynamics can cooperate with the consensus couplings which lead to distributed solutions to certain network optimization problems [33, 34, 35, 36, 37, 38, 39], which generalized the classical incremental methods for distributed optimization [40, 41, 42].

In this paper, we try to borrow the insights from consensus-based distributed optimization
methods [33, 34, 35, 36, 39], in the aim of establishing some new synchronization conditions which can partially relax the in general strong assumptions on the nonlinear node self-dynamics [14, 23]. We assume that the network nodes have non-identical nonlinear self-dynamics as gradients of some concave functions. This allows functions which might not be passive nor globally Lipschitz. The node couplings are linear but with switching directed communication graphs. Then several sufficient and/or necessary conditions are established regarding exact or approximate synchronization of the overall node states. These results reveal when and how nonlinear node self-dynamics may cooperate with the linear consensus coupling, which leads to synchronization conditions under much relaxed connectivity requirements to the communication graphs.

The remainder of the paper is organized as follows. In Section 2, some preliminary mathematical concepts and lemmas are introduced. In Section 3, we formulate the considered network dynamics and define the problem of interest. Section 4 presents some results on fixed graphs, and then Section 5 discusses time-varying graphs. Finally some concluding remarks are given in Section 6.

2 Preliminaries

In this section, we introduce some notations and provide preliminary results that will be used in the rest of the paper.

2.1 Directed Graphs

A directed graph (digraph) \( G = (V, E) \) consists of a finite set \( V \) of nodes and an arc set \( E \), where an arc is an ordered pair of distinct nodes of \( V \) [4]. An element \((i, j) \in E\) describes an arc which leaves \( i \) and enters \( j \). A walk in \( G \) is an alternating sequence \( W : i_1e_1i_2e_2\ldots e_{m-1}i_m \) of nodes \( i_\kappa \) and arcs \( e_\kappa = (i_\kappa, i_{\kappa+1}) \in E \) for \( \kappa = 1, 2, \ldots, m - 1 \). A walk is called a path if the nodes of the walk are distinct, and a path from \( i \) to \( j \) is denoted as \( i \rightarrow j \). A digraph \( G \) is called bidirectional when for any two nodes \( i \) and \( j \), \((i, j) \in E\) if and only if \((j, i) \in E\); strongly connected if it contains path \( i \rightarrow j \) and \( j \rightarrow i \) for every pair of nodes \( i \) and \( j \). Ignoring the direction of the arcs, the connectivity of a bidirectional digraph is transformed to that of the corresponding undirected graph. A time-varying graph is defined as \( G_{\sigma(t)} = (V, E_{\sigma(t)}) \) where
σ : [0, +∞) → Q denotes a piecewise constant function, where Q is a finite set containing all possible graphs with node set V. Moreover, the joint graph of G_σ(t) in time interval [t_1, t_2) with t_1 < t_2 ≤ +∞ is denoted as G([t_1, t_2)) = ∪_{t ∈ [t_1, t_2)} G(t) = (V, ∪_{t ∈ [t_1, t_2)} E_σ(t)).

2.2 Dini Derivatives and Limit Sets

The upper Dini derivative of a continuous function h : (a, b) → R (−∞ ≤ a < b ≤ ∞) at t is defined as

\[ D^+ h(t) = \limsup_{s \to 0^+} \frac{h(t + s) - h(t)}{s}. \]

When h is continuous on (a, b), h is non-increasing on (a, b) if and only if \( D^+ h(t) \leq 0 \) for any \( t ∈ (a, b) \). The next result is convenient for the calculation of the Dini derivative \[6, 31\].

**Lemma 1** Let \( V_i(t, x) : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R} \) be \( C^1 \) and \( V(t, x) = \max_{i=1,...,n} V_i(t, x) \). If \( I(t) = \{i ∈ \{1, 2, ..., n\} : V(t, x(t)) = V_i(t, x(t))\} \) is the set of indices where the maximum is reached at \( t \), then \( D^+ V(t, x(t)) = \max_{i ∈ I(t)} \dot{V}_i(t, x(t)) \).

Next, consider the following autonomous system

\[ \dot{x} = f(x), \tag{1} \]

where \( f : \mathbb{R}^d \to \mathbb{R}^d \) is a continuous function. Let \( x(t) \) be a solution of (1) with initial condition \( x(t_0) = x^0 \). Then \( \Omega_0 \subset \mathbb{R}^d \) is called a positively invariant set of (1) if, for any \( t_0 ∈ \mathbb{R} \) and any \( x^0 ∈ Ω_0 \), we have \( x(t) ∈ Ω_0 \), \( t ≥ t_0 \), along every solution \( x(t) \) of (1).

We call \( y \) a ω-limit point of \( x(t) \) if there exists a sequence \( \{t_k\} \) with \( \lim_{k \to ∞} t_k = ∞ \) such that \( \lim_{k \to ∞} x(t_k) = y \). The set of all ω-limit points of \( x(t) \) is called the ω-limit set of \( x(t) \), and is denoted as \( Λ^+(x(t)) \). The following lemma is well-known \[5\].

**Lemma 2** Let \( x(t) \) be a solution of (1). Then \( Λ^+(x(t)) \) is positively invariant. Moreover, if \( x(t) \) is contained in a compact set, then \( Λ^+(x(t)) \neq ∅ \).

2.3 Convex Analysis

A set \( K ⊂ \mathbb{R}^d \) is said to be convex if \( (1 − λ)x + λy ∈ K \) whenever \( x ∈ K, y ∈ K \) and \( 0 ≤ λ ≤ 1 \). For any set \( S ⊂ \mathbb{R}^d \), the intersection of all convex sets containing \( S \) is called the convex hull of \( S \), denoted by \( \text{co}(S) \).
Let $K$ be a closed convex subset in $\mathbb{R}^d$ and denote $|x|_K = \inf_{y \in K} |x - y|$ as the distance between $x \in \mathbb{R}^d$ and $K$, where $| \cdot |$ is the Euclidean norm. There is a unique element $P_K(x) \in K$ satisfying $|x - P_K(x)| = |x|_K$ associated to any $x \in \mathbb{R}^d$ \cite{3}. The map $P_K$ is called the projector onto $K$. The following lemma holds \cite{3}.

**Lemma 3**

(i). $\langle P_K(x) - x, P_K(y) - y \rangle \leq 0$, $\forall y \in K$.

(ii). $|P_K(x) - P_K(y)| \leq |x - y|, x, y \in \mathbb{R}^d$.

(iii) $|x|^2_K$ is continuously differentiable at $x$ with $\nabla |x|^2_K = 2(x - P_K(x))$.

Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a real-valued function. We call $f$ a convex function if for any $x, y \in \mathbb{R}^d$ and $0 \leq \lambda \leq 1$, it holds that $f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y)$. The following lemma states some well-known properties for convex functions.

**Lemma 4** Let $f : \mathbb{R}^d \rightarrow \mathbb{R} \in C^1$ be a convex function.

(i). $f(x) \geq f(y) + \langle x - y, \nabla f(y) \rangle$.

(ii). Any local minimum is a global minimum, i.e., $\arg \min f = \{z : \nabla f(z) = 0\}$.

### 3 Problem Definition

Consider a network with node set $V = \{1, 2, \ldots, N\}$. The interaction relations in the multi-agent network is modeled as a time-varying directed graph $G_{\sigma(t)} = (V, E_{\sigma(t)})$ with $\sigma : [0, +\infty) \rightarrow \mathcal{Q}$ being a piecewise constant function, where $\mathcal{Q}$ is the finite set indicating all possible digraphs over node set $V$. We assume that there is a lower bound $\tau_D > 0$ between two consecutive switching time instants of $\sigma(t)$.

A node $j$ is said to be a **neighbor** of $i$ at time $t$ when there is an arc $(j, i) \in E$, and we let $N_i(\sigma(t))$ represent the set of agent $i$’s neighbors at time $t$. Each node holds a state $x_i(t) \in \mathbb{R}^m$. Let $a_{ij}(t) > 0$ be a function marking the weight associated with arc $(j, i)$ at time $t$. The nodes’ dynamics are described by the following coupled oscillators \cite{14, 23}:

$$
\frac{dx_i(t)}{dt} = f_i(x_i(t)) + K \sum_{j \in N_i(\sigma(t))} a_{ij}(t)(x_j(t) - x_i(t)), \quad i = 1, \ldots, N.
$$

where $f_i(\cdot) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is a continuous function denoting the self-dynamics of node $i$ and $K \geq 0$ is a given constant. Let the weighted adjacency matrix be denoted as $A_{\sigma(t)}$ where
\[ [A_{\sigma(t)}]_{ij} = a_{ij}(t) \text{ if } j \in N_i(\sigma(t)) \text{ and } [A_{\sigma(t)}]_{ij} = 0 \text{ otherwise}. \]

The weighted degree matrix is then defined as
\[ D_{\sigma(t)} = \text{diag}(d_1(\sigma(t)), \ldots, d_N(\sigma(t))) \text{ with } d_i(\sigma(t)) = \sum_{j \in N_i(\sigma(t))} a_{ij}(t). \]

Then \[ P_{\sigma(t)} = D_{\sigma(t)} - A_{\sigma(t)} \]
is the time-varying Laplacian of the network representing the coupling of the node dynamics.

For the time-varying weight function \( a_{ij}(t) \), we assume that there are \( a^* > 0 \) and \( a_* > 0 \) such that \( a_* \leq a_{ij}(t) \leq a^*, t \in \mathbb{R}^+. \)

For the self-dynamics \( f_i \), we impose the following assumptions.

[1] There are \( F_i : \mathbb{R}^m \rightarrow \mathbb{R}, i = 1, \ldots, N \) such that \( f_i = -\nabla F_i \), where each \( F_i \) is a \( C^1 \) convex function with \( \text{arg min } F_i \neq \emptyset \).

Remark 1 Under Assumption A1, System (2) is indeed a continuous-time version of the algorithms for distributed sub-gradient optimization, e.g., [33, 34, 36]. It is indeed a rather strong assumption, but a considerably larger set of functions is covered by Assumption A1 compared to the existing synchronization literature [13, 16, 17, 18, 19, 20]: there requires no global Lipschitz condition, nor identical dynamics for the \( f_i \). For instance, we can take
\[ f_i(x) = -(x - m_i)^3 \]
with \( m_i \in \mathbb{R} \) being a constant under the assumption A1. Moreover, as will be shown in the coming results, this assumption allows us to deal with switching directed interaction graphs, which is in general a challenging problem [22].

The initial time is set to be 0. Let \( x(t) = (x_1^T(t), \ldots, x_N^T(t))^T \in \mathbb{R}^{mN} \) be the Caratheodory solution of system (2) for initial condition \( x^0 = x(0) \). We refer to [8] regarding the existence of the Caratheodory solution for (2). We introduce the following standard synchronization definition [23].

Definition 1 Global synchronization of System (2) is achieved if for all \( x^0 \in \mathbb{R}^{mN} \), we have
\[ \lim_{t \to +\infty} |x_i(t) - x_j(t)| = 0 \text{ for all } i, j = 1, \ldots, N. \]

Remark 2 By Assumption A1 itself there might be finite-time escape for the trajectory of System (2), i.e., \( x(t) \) approaches infinity in a finite time interval. With suitable assumptions finite-time escape can however be excluded. We refer to the coming Lemma [7], Eq. (14), and Lemma [7] respectively, which guarantee the existence of \( x(t) \) for the entire \([0, \infty) \) under the corresponding conditions.
4 Fixed Interaction Graphs

In this section, we consider the possibility of synchronization under fixed interaction graphs. We first discuss whether exact synchronization can be reached for directed graphs. Then we show the existence of an approximate synchronization over bidirectional graphs.

4.1 Exact Synchronization

We make an assumption on the $F_i$.

$\text{[A2]} \{z : f_i(z) = 0\} \neq \emptyset$ is a bounded set, and $\langle x_i - P_{\Theta_i}(x_i), f_i(x_i) \rangle \leq 0$ for all $x_i \in \mathbb{R}^m$ and $i = 1, \ldots, N$, where $\Theta_i = \text{co}(\bigcup_{z \in V} \{z : f_i(z) = 0\})$.

We present the following result.

**Theorem 1** Assume that A1 and A2 hold. Let $G_{\sigma(t)} \equiv G$ for some strongly connected digraph $G$, and let $a_{ij}(t) \equiv a_{ij}$ for some $a_{ij} > 0$, $i, j = 1, \ldots, N$. Then global synchronization for System (2) is achieved if and only if $\bigcap_{i=1}^N \{z : f_i(z) = 0\} \neq \emptyset$.

In the rest of this subsection, we first give the proof of the necessity claim of Theorem 1 and then we present a simple proof for the sufficiency part with bidirectional graphs. The sufficiency part of Theorem 1 in fact follows from the upcoming conclusion, Theorem 4 which does not rely on Assumption A2.

4.1.1 Necessity

We prove the necessity statement in Theorem 1 by a contradiction argument. Suppose global synchronization is reached under the condition that $\bigcap_{i=1}^N \{z : f_i(z) = 0\} = \emptyset$. Let $x(t)$ be a trajectory of system (2) $\Lambda^+(x(t))$ be its $\omega$-limit set.

First we show that $\Lambda^+(x(t))$ is a nonempty set. Introduce $\theta(x(t)) := \max_{i \in V} |x_i(t)|^2_{\Theta_i}$. The following lemma holds.

**Lemma 5** Let A1 and A2 hold. Then $\theta(x(t))$ is non-increasing along each solution of System (2).
Proof. From Lemmas 3 and 4, it is not hard to see that
\[
D^+ \theta(x(t)) = 2 \max_{i \in I(t)} \left\langle x_i(t) - P_{\Theta_s}(x_i(t)), f_i(x_i(t)) \right\rangle + K \sum_{j \in N_i(\sigma(t))} a_{ij}(t) (x_j(t) - x_i(t)) \leq 2 \max_{i \in I(t)} \left\langle x_i(t) - P_{\Theta_s}(x_i(t)), f_i(x_i(t)) \right\rangle \leq 0, \quad (3)
\]
where 
\[
I^*(t) := \{i \in V : |x_i(t)|^{\Theta_s} = \theta(x(t))\}. \quad \square
\]

From the above lemma we immediately know that each trajectory \(x(t)\) is contained in a compact set. Let
\[
\mathcal{M} = \{x = (x_1^T \ldots x_N^T)^T : x_1 = \cdots = x_N; x_i \in \mathbb{R}^m, i = 1, \ldots, N\} \quad (4)
\]
denote the consensus manifold. Based on Lemma 2, we conclude that \(\Lambda^+(x(t)) \subseteq \mathcal{M} \neq \emptyset\). Moreover, \(\Lambda^+(x(t))\) is positively invariant since (2) defines an autonomous system when the interaction graph is fixed. This is to say, any trajectory of system (2) must stay within \(\Lambda^+(x(t))\) for any initial value in \(\Lambda^+(x(t))\).

Now we take \(y \in \Lambda^+(x(t))\). Then we have \(y = (z_1^T \ldots z_N^T)^T\) for some \(z_s \in \mathbb{R}^m\). Suppose there exist two indices \(i_1, i_2 \in \{1, \ldots, N\}\) with \(i_1 \neq i_2\) such that \(f_{i_1}(z_s) \neq f_{i_2}(z_s)\). Consider the solution of (2) for initial time 0 and initial value \(y\). We have \(\dot{x}_{i_1}(0) \neq \dot{x}_{i_2}(0)\). As a result, there exists a constant \(\varepsilon > 0\) such that \(x_{i_1}(t) \neq x_{i_2}(t)\) for \(t \in (0, \varepsilon)\). In other word, the trajectory will leave the consensus manifold \(\mathcal{M}\) for \((0, \varepsilon)\), and therefore will also leave the set \(\Lambda^+(x(t))\). This contradicts the fact that \(\Lambda^+(x(t))\) is positively invariant. The necessity part of Theorem 1 has been proved.

4.1.2 Sufficiency: Bidirectional Case

We now provide an alternative proof of sufficiency for bidirectional graphs, which is based on some geometrical intuition of the vector field. Note that compared to the proof of Theorem 4 on directed graphs, this proof uses completely different arguments which indeed cannot be applied to directed graphs. Therefore, we believe the proof given in the following is interesting at its own right, because it reveals some fundamental difference between directed and bidirectional graphs.

Suppose \(G\) is bidirectional, i.e., \((i, j) \in E\) if and only if \((j, i) \in E\), and \(a_{ij} = a_{ji}\) for all \(i\) and \(j\). We use unordered pair \(\{i, j\}\) to denote the edge between node \(i\) and \(j\). Denote \(F(z) = \sum_{i=1}^N F_i(z)\)
and $\mathcal{F}_G(x; K) = \sum_{i=1}^{N} F_i(x_i) + \frac{K}{2} \sum_{(j,i) \in E} a_{ij} |x_j - x_i|^2$. Denote the $N$’th Cartesian product of a set $S$ as $S^N$. The following lemma holds.

**Lemma 6** Suppose $\bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \neq \emptyset$. Let the communication graph $G$ be fixed, bidirectional, and connected. Then $\arg \min \mathcal{F}_G(x; K) = \left( \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \right)^N \cap \mathcal{M}$.

**Proof.** When $\bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \neq \emptyset$, it is clear that $\arg \min \mathcal{F}(z) = \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \}$.

Now take $x_* = (p_1^T \ldots p_N^T)^T \in \left( \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \right)^N \cap \mathcal{M}$, where $p_* \in \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \}$. First we have $x_* \in \arg \min_x \sum_{i=1}^{N} F_i(x_i)$. Second we have $x_* \in \arg \min_x \frac{K}{2} \sum_{(j,i) \in E} a_{ij} |x_j - x_i|^2$. Therefore, we conclude that $x_* \in \arg \min \mathcal{F}_G(x; K)$. This gives

$$\arg \min \mathcal{F}_G(x; K) \supseteq \left( \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \right)^N \cap \mathcal{M}. \quad (5)$$

On the other hand, convexity gives

$$\arg \min \mathcal{F}_G(x; K) = \left\{ x : -K(P \otimes I_m)x = \left( (f_1(x_1))^T \ldots (f_N(x_N))^T \right)^T \right\}, \quad (6)$$

where $\otimes$ represents the Kronecker product, $I_m$ is the identity matrix in $\mathbb{R}^m$, and $P$ is the Laplacian of the graph $G$. Noticing that

$$(1_N^T \otimes I_m)(P \otimes I_m) = 1_N^T P \otimes I_m = 0,$$

where $1_N = (1 \ldots 1)^T \in \mathbb{R}^N$, we have

$$\left(1_N^T \otimes I_m\right)\left( (f_1(x_1))^T \ldots (f_N(x_N))^T \right)^T = \sum_{i=1}^{N} f_i(x_i) = 0 \quad (7)$$

for any $x \in \arg \min \mathcal{F}_G(x; K)$.

Now take $x^* = (q_1^T \ldots q_N^T)^T \in \arg \min \mathcal{F}_G(x; K)$. Suppose there exist two indices $i_*$ and $j_*$ such that

$$f_{i_*}(q_{i_*}) \neq f_{j_*}(q_{j_*}).$$

Then at least one of $f_{i_*}(q_{i_*})$ and $f_{j_*}(q_{j_*})$ must be nonzero. Taking $\hat{p} \in \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \}$, we have

$$\sum_{i=1}^{N} F_i(q_i) > \sum_{i=1}^{N} F_i(\hat{p})$$

because for $x = (x_1^T \ldots x_N^T)^T \in \arg \min \sum_{i=1}^{N} F_i(x_i)$, we have $f_i(x_i) = 0, i = 1, \ldots, N$. Consequently, for $w_* = (\hat{p}^T \ldots \hat{p}^T)^T$, we have

$$\mathcal{F}_G(x^*; K) > \mathcal{F}_G(w_*; K)$$
which is impossible according to the definition of $x^*$ so that such $i_*$ and $j_*$ cannot exist. In light of (7), this immediately implies $f_i(q_i) = 0$, $i = 1, \ldots, N$, or equivalently

$$q_i \in \{ z : f_i(z) = 0 \}, \quad i = 1, \ldots, N$$

for all $x^* = (q_1^T \ldots q_N^T)^T \in \arg \min \mathcal{F}_G(x)$. Therefore, we conclude from (8) that $\sum_{i=1}^{N} F_i(q_i) = \sum_{i=1}^{N} F_i(p_*)$, and this implies

$$\sum_{\{j,i\} \in E} a_{ij} |q_j - q_i|^2 = 0$$

as long as $x^* = (q_1^T \ldots q_N^T)^T \in \arg \min \mathcal{F}_G(x)$. The connectivity of the communication graph thus further guarantees that $q_1 = \cdots = q_N$, so we have proved that $x^* \in (\bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \}) \bigcap \mathcal{M}$. Consequently, we obtain

$$\arg \min \mathcal{F}_G(x; K) \subseteq \left( \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \right) \bigcap \mathcal{M}. \quad (9)$$

The desired lemma holds from (5) and (9). \qed

Note that

$$K \sum_{j \in N_i} a_{ij} (x_j - x_i) + f_i(x_i) = -\nabla x_i \mathcal{F}_G(x; K). \quad (10)$$

As a result,

$$\frac{d}{dt} \mathcal{F}_G(x(t); K) = -\left| \nabla \mathcal{F}_G(x; K) \right|^2 \quad (11)$$

along each trajectory of System (2). Then by LaSalle’s invariance principle we have

$$\lim_{t \to \infty} \text{dist}(x(t), \arg \min \mathcal{F}_G(x; K)) = 0.$$  

Lemma 6 further ensures

$$\lim_{t \to \infty} \text{dist}\left( x(t), \left( \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \right) \bigcap \mathcal{M} \right) = 0$$

if $G$ is bidirectional and connected. Equivalently, global synchronization is reached and we can even predict that each limit point of $x_i(t)$ lies in $\bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \}$ for all $i$.

**Remark 3** We see from the proof above that the construction of $\mathcal{F}_G(x)$ is critical because the convergence argument is based on the fact that the gradient of $\mathcal{F}_G(x)$ is consistent with the interaction graph. It can be easily verified that finding such a function is in general impossible for directed graphs.
4.2 Approximate Synchronization

Theorem 1 indicates that exact synchronization is impossible unless \( \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \neq \emptyset \) is fulfilled. In this subsection, we discuss the possibility of approximate synchronization in the absence of this nonempty interaction condition. We introduce the following definition.

**Definition 2** Global \( \epsilon \)-synchronization is achieved if for all \( x_0 \in \mathbb{R}^{mN} \), we have

\[
\lim_{t \to +\infty} \sup_{t} |x_i(t) - x_j(t)| \leq \epsilon, \quad i, j = 1, \ldots, N.
\]  

(12)

We use the following assumption.

[A3] (i) \( \arg \min F(z) \neq \emptyset \); (ii) \( \arg \min F_G(x; K) \neq \emptyset \) for all \( K \geq 0 \); (iii) \( \bigcup_{K \geq 0} \arg \min F_G(x; K) \) is bounded.

For \( \epsilon \)-synchronization, we present the following result.

**Theorem 2** Assume that A1 and A3 hold. Let the interaction graph \( G_{\sigma(t)} \equiv G \) for some fixed, bidirectional, and connected \( G \), and let \( a_{ij}(t) \equiv a_{ij} \) for some \( a_{ij} > 0, i, j = 1, \ldots, N \). Then for any \( \epsilon > 0 \), there exists \( K_{\epsilon} > 0 \) such that global \( \epsilon \)-synchronization is achieved for all \( K \geq K_{\epsilon} \).

**Proof.** Let’s fix \( \epsilon \). Again, since

\[
K \sum_{j \in N_i} a_{ij} (x_j - x_i) + f_i(x_i) = -\nabla x_i F_G(x; K),
\]  

(13)

the convexity of \( F_G(x; K) \) ensures that

\[
\lim_{t \to +\infty} \text{dist} \left( x(t), \arg \min F_G(x; K) \right) = 0.
\]  

(14)

Define \( \tilde{F}(x) = \sum_{i=1}^{N} F_i(x_i) \). Under Assumptions A1 and A3, we have that

\[
\tilde{L}_0 \doteq \sup \left\{ |\nabla \tilde{F}(x)| : x \in \bigcup_{K \geq 0} \arg \min F_G(x; K) \right\}
\]  

(15)

is a finite number. We also define

\[
D_0 \doteq \sup \left\{ |z_* - x_i| : i = 1, \ldots, N, \ x \in \bigcup_{K \geq 0} \arg \min F_G(x; K) \right\},
\]  

(16)

where \( z_* \in \arg \min \mathcal{F} \) is an arbitrarily chosen point.
Let $p = (p_1^T \ldots p_N^T)^T \in \arg \min \mathcal{F}_G(x; K)$ with $p_i \in \mathbb{R}^m, i = 1, \ldots, N$. Let $P$ be the Laplacian of the graph $G$. Since the graph is bidirectional and connected, we can sort the eigenvalues of the matrix $P \otimes I_m$ as

$$0 = \lambda_1 = \cdots = \lambda_m < \lambda_{m+1} \leq \cdots \leq \lambda_{mN}.$$ 

Let $l_1, \ldots, l_{mN}$ be the orthonormal basis of $\mathbb{R}^{mN}$ formed by the right eigenvectors of $P \otimes I_m$, where $l_1, \ldots, l_m$ are eigenvectors corresponding to the zero eigenvalue. Suppose $p = \sum_{k=1}^{mN} c_k l_k$ with $c_k \in \mathbb{R}, k = 1, \ldots, mN$.

According to (6), we have

$$|K(P \otimes I_m)p|^2 = K^2 \left| \sum_{k=m+1}^{mN} c_k \lambda_k l_k \right|^2 = K^2 \sum_{k=m+1}^{mN} c_k^2 \lambda_k^2 \leq L_0^2,$$

which yields

$$\sum_{k=m+1}^{mN} c_k^2 \leq \left( \frac{L_0}{K \lambda_2} \right)^2,$$

where $\lambda_2 > 0$ denotes the second smallest eigenvalue of $P$.

Now recall that

$$\mathcal{M} = \{ x = (x_1^T \ldots x_N^T)^T : x_1 = \cdots = x_N; x_i \in \mathbb{R}^m, i = 1, \ldots, N \}.$$ 

is the consensus manifold. Noticing that $\mathcal{M} = \text{span}\{l_1, \ldots, l_m\}$, we conclude from (18) that

$$\sum_{k=m+1}^{mN} c_k^2 = \left| \sum_{k=m+1}^{mN} c_k l_k \right|^2 = \left| p \right|^2_{\mathcal{M}} = \sum_{i=1}^{N} \left| p_i - \frac{\sum_{i=1}^{N} p_i}{N} \right|^2 \leq \left( \frac{L_0}{K \lambda_2} \right)^2.$$ 

The last equality in (20) is due to the fact that $1_N \otimes \left( \frac{\sum_{i=1}^{N} p_i}{N} \right)$ is the projection of $p$ on to $\mathcal{M}$. Thus, for any $\varsigma > 0$, there is $K_1(\varsigma) > 0$ such that when $K \geq K_1(\varsigma)$,

$$|p_i - p_{\text{ave}}| \leq \varsigma, \ i = 1, \ldots, N$$ 

and

$$|\mathcal{F}(p_i) - \mathcal{F}(p_{\text{ave}})| \leq \varsigma, \ i = 1, \ldots, N,$$

where $p_{\text{ave}} = \frac{\sum_{i=1}^{N} p_i}{N}$.

On the other hand, with (6), we have

$$\sum_{i=1}^{N} f_i(p_i) = \sum_{i=1}^{N} f_i(p_{\text{ave}} + \hat{p}_i) = 0,$$

where $p_{\text{ave}} = \frac{\sum_{i=1}^{N} p_i}{N}$.
where $\hat{p}_i = p_i - p_{\text{ave}}$. Now according to (21) and (23), since $F_i \in C^1$, for any $\varsigma > 0$, there is $K_2(\varsigma) > 0$ such that when $K \geq K_2(\varsigma)$,

$$\sum_{i=1}^N f_i(p_{\text{ave}}) \leq \frac{\varsigma}{D_0}. \tag{24}$$

This implies

$$F(p_{\text{ave}}) \leq F(z_*) + |z_* - p_{\text{ave}}| \times \sum_{i=1}^N f_i(p_{\text{ave}}) \leq F(z_*) + \varsigma. \tag{25}$$

Therefore, for any $\epsilon > 0$, we can take $K_0 = \max\{K_1(\epsilon/2), K_2(\epsilon/2)\}$. Then when $K \geq K_0$,

$$|p_i - p_j| \leq \epsilon; \quad F(p_i) \leq \min_z F(z) + \epsilon \tag{26}$$

for all $i$ and $j$. Now with (14), every limit point of system (2) is contained in the set $\arg \min F_G(x; K)$. Noting that $p$ is arbitrarily chosen from $\arg \min F_G(x; K)$, $\epsilon$-synchronization is achieved as long as we choose $K_\epsilon \geq K_0$. This completes the proof.

From Theorems 1 and 2, we conclude that even though without the nonempty intersection condition, it is impossible to reach exact synchronization for the considered coupled dynamics, it is still possible to find a control law that guarantees approximate synchronization with arbitrary accuracy.

### 4.3 Assumption Feasibility

This subsection discusses the feasibility of Assumptions A2 and A3.

**Proposition 1** If $\tilde{F}(x) = \sum_{i=1}^N F_i(x_i)$ is coercive, i.e., $\tilde{F}(x) \to \infty$ as long as $|x| \to \infty$, then $\{z : f_i(z) = 0\} \neq \emptyset$ is a bounded set for all $i = 1, \ldots, N$, and A3 holds.

**Proof.** First of all, since $\tilde{F}(x) = \sum_{i=1}^N F_i(x_i)$ is coercive, it follows straightforwardly that $F(z) = \sum_{i=1}^N F_i(z)$ and each $F_i(z)$ are also coercive. This implies immediately that $\{z : f_i(z) = 0\} \neq \emptyset$ is a bounded set for all $i = 1, \ldots, N$ and A3(i) hold.

Next, Observing that $\frac{K}{2} \sum_{\{j,i\} \in E} a_{ij} |x_j - x_i|^2 \geq 0$ for all $x = (x_1^T \ldots x_N^T)^T \in \mathbb{R}^{mN}$ and that $\tilde{F}(x) = \sum_{i=1}^N F_i(x_i)$ is coercive, we obtain that $\arg \min F_G(x; K) \neq \emptyset$ for all $K \geq 0$. Thus, A3.(ii) holds.
Finally, we denote $F^* = \min_z F(z) = F(z_*)$. Since $\sum_{i=1}^N F_i(x_i)$ is coercive, there exists a constant $M(F_*) > 0$ such that $\sum_{i=1}^N F_i(x_i) > F_*$ for all $|x| > M(F_*)$. This implies

$$F_G(x; K) > F_G(1_N \otimes z_*; K) = F_*$$

(27)

for all $|x| > M$. That is to say, the global minimum of $F_G(x; K)$ is reached within the set $\{|x| \leq M\}$ for all $K > 0$. Therefore, we have

$$\bigcup_{K \geq 0} \arg \min_{K \geq 0} F_G(x; K) \subseteq \{|x| \leq M(F_*)\}.$$  

(28)

This proves A3.(iii).

**Proposition 2** Suppose $\{z : f_i(z) = 0\} \neq \emptyset$ is a bounded set for all $i = 1, \ldots, N$ and the node state space is $\mathbb{R}$, i.e., $m = 1$. Then A2 and A3 hold.

**Proof.** Since each $\{z : f_i(z) = 0\}$ is a finite interval when the node state is one dimensional, it is straightforward to verifying that $\langle x_i - P_{\Theta_*}(x_i), f_i(x_i) \rangle \leq 0$ for all $x_i \in \mathbb{R}$. Thus A2 holds. We now prove A3 also holds.

(i). Let $x_i^* \in \arg \min F_i$. Denote $y_* = \min\{x_1^*, \ldots, x_N^*\}$. Then for any $i = 1, \ldots, N$, we have

$$0 \geq F_i(x_i^*) - F_i(y_*) \geq -(x_i^* - y_*) f_i(y_*)$$

(29)

according to inequality (i) of Lemma 4. This immediately yields $f_i(y_*) \geq 0$ for all $i = 1, \ldots, N$.

Thus, for any $y < y_*$, we have

$$F(y) - F(y_*) \geq (y - y_*) \nabla F(y_*) = -\sum_{i=1}^N (y - y_*) f_i(y_*) \geq 0,$$

(30)

which implies $F(y) \geq F(y_*)$ for all $y < y_*$. This immediately yields $f_i(y_*) \geq 0$ for all $i = 1, \ldots, N$.

A symmetric analysis leads to that $F(y) \geq F(y^*)$ for all $y > y^*$ with $y^* = \max\{x_1^*, \ldots, x_N^*\}$. Therefore, we obtain $F(y) \geq \min\{F(y_*), F(y^*)\}$ for all $y \neq [y_*, y^*]$. This implies that a global minimum is reached within the interval $[y_*, y^*] = \text{co}\{x_1^*, \ldots, x_N^*\}$ and A3.(i) thus follows.

(ii). Introduce the following cube in $\mathbb{R}^N$:

$$C_\eta^y = \left\{ x = (x_1^T, \ldots, x_N^T)^T : x_i \in [y_* - \eta, y^* + \eta], i = 1, \ldots, N \right\}.$$  

where $\eta > 0$ is a given constant.
Claim. For any $K \geq 0$, $C^0_\eta$ is an invariant set of System (2).

Define $\Psi(x(t)) = \max_{i \in V} x_i(t)$. Then based on Lemma 1, we have

$$D^+ \Psi(x(t)) = \max_{i \in I_0(t)} \frac{d}{dt} x_i(t)$$

$$= \max_{i \in I_0(t)} \left[ \sum_{j \in N_i} a_{ij}(x_j - x_i) + f_i(x_i) \right]$$

$$\leq \max_{i \in I_0(t)} \left[ f_i(x_i) \right], \quad (31)$$

where $I_0(t)$ denotes the index set which contains all the nodes reaching the maximum for $\Psi(x(t))$.

Since

$$0 \geq F_i(x^*_i) - F_i(y^* + \eta) \geq -(x^*_i - y^* - \eta)f_i(y^* + \eta), \quad i = 1, \ldots, N \quad (32)$$

we have $f_i(y^* + \eta) \leq 0$ for all $i = 1, \ldots, N$. As a result, we obtain

$$D^+ \Psi(x(t)) \bigg|_{\Psi(x(t)) = y^* + \eta} \leq 0, \quad (33)$$

which implies $\Psi(x(t)) \leq y^* + \eta$ for all $t \geq t_0$ under initial condition $\Psi(x(t_0)) \leq y^* + \eta$. Similar analysis ensures that $\min_{i \in V} x_i(t) \geq y^* - \eta$ for all $t \geq t_0$ as long as $\min_{i \in V} x_i(t_0) \geq y^* - \eta$. This proves the claim.

Note that every trajectory of system (2) asymptotically reaches $\arg \min F_G(x; K)$. This immediately leads to that $F_G(x; K)$ reaches its minimum within $C^0_\eta$ for any $K \geq 0$ since $C^0_\eta$ is an invariant set. Then A3.(ii) holds.

(iii). Since $\arg \min F_i$ is bounded for $i = 1, \ldots, N$, there exist $b_i \leq d_i, i = 1, \ldots, N$ such that $\arg \min F_i = [b_i, d_i]$. Define $b_* = \min\{b_1, \ldots, b_N\}$ and $d^* = \max\{d_1, \ldots, d_N\}$. We will prove the conclusion by showing $\arg \min F_G(x; K) \subseteq C_*$ for all $K \geq 0$, where

$$C_* = \left\{ x = (x^T_1 \ldots x^T_N)^T : x_i \in [b_*, d^*], i = 1, \ldots, N \right\}.$$

Let $z = (z_1, \ldots, z_N)^T \in \arg \min F_G(x; K)$. First we show $\max\{z_1, \ldots, z_N\} \leq d^*$ by a contradiction argument. Suppose $\max\{z_1, \ldots, z_N\} > d^*$.

Now let $i_1, \ldots, i_k$ be the nodes reaching the maximum state, i.e., $z_{i_1} = \cdots = z_{i_k} = \max\{z_1, \ldots, z_N\}$. There will be two cases.
Let \( k = N \). We have \( z_1 = \cdots = z_N = y \) in this case. Then for all \( i \) and \( x_i^* \in \arg \min F_i \), we have

\[
0 > F_i(x_i^*) - F_i(y) \geq -(x_i^* - y)f_i(y)
\]

which yields \( f_i(y) > 0, i = 1, \ldots, N \) since \( y > d^* \). This immediately leads to

\[
\mathcal{F}_G(z; K) = \mathcal{F}(y) > \min \mathcal{F} \geq \min \mathcal{F}_G(z; K),
\]

which contradicts the fact that \( z \in \arg \min \mathcal{F}_G(x; K) \).

Let \( k < N \). Then we denote \( s_* = \max \{z_i : i \notin \{i_1, \ldots, i_k\}, i = 1, \ldots, N\} \), which is actually the second largest value in \( \{z_1, \ldots, z_N\} \). We define a new point \( \hat{z} = (\hat{z}_1, \ldots, \hat{z}_N)^T \) by \( \hat{z}_i = z_i, i \notin \{i_1, \ldots, i_k\} \) and

\[
\hat{z}_i = \begin{cases} 
  d^*, & \text{if } s_* < d^* \\
  s_*, & \text{otherwise}
\end{cases}
\]

for \( i \in \{i_1, \ldots, i_k\} \). Then it is easy to obtain that \( \mathcal{F}_G(z; K) > \mathcal{F}_G(\hat{z}; K) \), which again contradicts the choice of \( z \).

Therefore, we have proved that \( \max\{z_1, \ldots, z_N\} \leq d^* \). Based on a symmetric analysis we also have \( \min\{z_1, \ldots, z_N\} \geq b_* \). Therefore, we obtain \( \arg \min \mathcal{F}_G(x; K) \subseteq C_* \) for all \( K \geq 0 \) and A3.(iii) follows.

\section{Time-varying Interaction Graphs}

In this section, we consider time-varying graphs. We introduce the following definition \cite{25, 31}.

**Definition 3** \( G_{\sigma(t)} \) is said to be uniformly jointly strongly connected if there exists a constant \( T > 0 \) such that \( G([t, t+T]) \) is strongly connected for any \( t \geq 0 \).

We present the following result.

**Theorem 3** Let A1 hold. Suppose \( G_{\sigma(t)} \) is uniformly jointly strongly connected and \( \bigcap_{i=1}^{N} \{z : f_i(z) = 0\} \neq \emptyset \) contains at least one interior point. Then global synchronization is achieved for System (2). In fact, for any initial value \( x^0 \), there exists \( x_* \in \bigcap_{i=1}^{N} \{z : f_i(z) = 0\} \) such that \( \lim_{t \to \infty} x_i(t) = x_* \) for all \( i \in V \).
Note that the condition \( \lim_{t \to \infty} x_i(t) = x^*_i \) is indeed a stronger conclusion than our definition of synchronization as Theorem 3 guarantees that all the node states converge to a common point. We will see from the proof of Theorem 3 that this state convergence highly relies on the existence of an interior point of \( \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \). In the absence of such an interior point condition, it turns out that global synchronization still stands. We present another theorem stating the fact.

**Theorem 4** Let A1 hold. Suppose \( G_{\sigma(t)} \) is uniformly jointly strongly connected and \( \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \neq \emptyset \). Then global synchronization is achieved for System (2).

For \( \epsilon \)-synchronization under switching interactions, we present the following result.

**Theorem 5** Let A1 and A2 hold. Suppose \( G_{\sigma(t)} \) is uniformly jointly strongly connected. Then for any \( \epsilon > 0 \) and any initial value \( x^0 \in \mathbb{R}^{mN} \), there exist a sufficiently small \( T^\dagger_\epsilon(x^0) > 0 \) and a sufficiently large \( K^\dagger_\epsilon(x^0) \) such that \( \epsilon \)-synchronization is achieved under \( x^0 \) for all \( T \leq T^\dagger_\epsilon(x^0) \) and \( K \geq K^\dagger_\epsilon(x^0) \).

Note that compared to the results under discrete-time dynamics [33, 34], Theorems 3 and 4 stand on quite general assumptions, which applies to the case when the \( \{ z : f_i(z) = 0 \} \) are unbounded. In fact, even Theorem 5 does not rely on the boundedness of the \( f_i \).

**Remark 4** Compared to Theorem 3, Theorem 5 is semi-global in the sense that the control gain \( K^\dagger_\epsilon(x^0) \) depends on the initial value. With switching interaction graphs, it becomes fundamentally difficult to characterize the limit set of the trajectories.

The remaining of this section presents the proofs of the above results. We establish some useful lemmas first, and then the proofs of Theorems 3, 4, and 5 will be given, respectively.

### 5.1 Preliminary Lemmas

We establish three useful lemmas in this subsection. Suppose \( \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \neq \emptyset \) and take \( z^*_i \in \bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \). We define

\[
V_i(t) = \left| x_i(t) - z^*_i \right|^2, \quad i = 1, \ldots, N,
\]

and

\[
V(t) = \max_{i=1,\ldots,N} V_i(t).
\]

17
The following lemma holds.

**Lemma 7** Let A1 hold. Suppose $\bigcap_{i=1}^{N} \{z : f_i(z) = 0\} \neq \emptyset$. Then along any trajectory of System (3), we have $D^+ V(t) \leq 0$ for all $t \in \mathbb{R}^+$.  

**Proof.** Based on Lemma 1, we have 

$$D^+ V(t) = \max_{i \in \mathcal{I}(t)} \frac{d}{dt} V_i(t) = \max_{i \in \mathcal{I}(t)} 2\left\langle x_i(t) - z_s, \sum_{j \in \mathcal{N}_i(\sigma(t))} a_{ij}(t)(x_j - x_i) + f_i(x_i) \right\rangle,$$  

(39) 

where $\mathcal{I}(t)$ denotes the index set which contains all the nodes reaching the maximum for $V(t)$.  

Let $m \in \mathcal{I}(t)$. Denote 

$$Z_t = \{z : |z - z_s| \leq \sqrt{V(t)}\}$$ 

as the disk centered at $z_s$ with radius $\sqrt{V(t)}$. Take $y = x_m(t) + (x_m(t) - z_s)$. Then from some simple Euclidean geometry it is obvious to see that $P_{Z_t}(y) = x_m(t)$, where $P_{Z_t}$ is the projection operator onto $Z_t$. Thus, for all $j \in \mathcal{N}_m(\sigma(t))$, we obtain 

$$\left\langle x_m(t) - z_s, x_j(t) - x_m(t) \right\rangle = \left\langle y - x_m(t), x_j(t) - x_m(t) \right\rangle 
= \left\langle y - P_{Z_t}(y), x_j(t) - P_{Z_t}(y) \right\rangle 
\leq 0$$ 

(40) 

according to inequality (i) in Lemma 3, since $x_j(t) \in Z_t$. On the other hand, based on inequality (i) in Lemma 4, we also have 

$$\left\langle x_m(t) - z_s, f_m(x_m(t)) \right\rangle \leq F_m(z_s) - F_m(x_m(t)) \leq 0$$ 

(41) 

in light of the definition of $z_s$.  

With (39), (40) and (41), we conclude that 

$$D^+ V(t) = \max_{i \in \mathcal{I}(t)} 2\left\langle x_i(t) - z_s, \sum_{j \in \mathcal{N}_i(\sigma(t))} a_{ij}(t)(x_j - x_i) + f_i(x_i) \right\rangle \leq 0,$$ 

(42) 

which completes the proof. □  

A direct consequence of Lemma 7 is that when $\bigcap_{i=1}^{N} \{z : f_i(z) = 0\} \neq \emptyset$, we have 

$$\lim_{t \to \infty} V(t) = d_s^2$$ 

(43)
for some $d_s \geq 0$ along any trajectory of system (2) with control law $J_i(n_i, g_i)$. However, it is still unclear whether $V_i(t)$ converges or not. We establish another lemma indicating that with proper connectivity condition for the communication graph, all $V_i(t)$’s have the same limit $d_s^2$.

Lemma 8 Let A1 hold. Suppose $\bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \neq \emptyset$ and $G_{\sigma(t)}$ is uniformly jointly strongly connected. Then along any trajectory of System (2), we have $\lim_{t \to \infty} V_i(t) = d_s^2$ for all $i$.

Proof. In order to prove the desired conclusion, we just need to show $\lim \inf_{t \to \infty} V_i(t) = d_s^2$ for all $i$. With Lemma 7, we conclude that $\forall \varepsilon > 0, \exists M(\varepsilon) > 0$, s.t.,

$$\sqrt{V_i(t)} \leq d_s + \varepsilon$$

for all $i$ and $t \geq M$.

Claim. For all $t \geq M$ and all $i, j \in V$, we have

$$\langle x_i(t) - z_s, x_j(t) - x_i(t) \rangle \leq -V_i(t) + (d_s + \varepsilon)\sqrt{V_i(t)}.$$  (45)

If $x_i(t) = z_s$ (15) follows trivially from (14). Otherwise we take $y_s = z_s + (d_s + \varepsilon)\frac{x_i(t) - z_s}{|x_i(t) - z_s|}$ and $B_t = \{ z : |z - z_s| \leq d_s + \varepsilon \}$. Here $B_t$ is the disk centered at $z_s$ with radius $d_s + \varepsilon$, and $y_s$ is a point within the boundary of $B_t$ and falls the same line with $z_s$ and $x_{i_0}(t)$. Take also $q_s = y_s + x_i(t) - z_s$. Then we have

$$\langle x_i(t) - z_s, x_j(t) - y_s \rangle = \langle q_s - y_s, x_j(t) - y_s \rangle = \langle q_s - P_{B_t}(q_s), x_j(t) - P_{B_t}(q_s) \rangle \leq 0$$  (46)

according to inequality (i) in Lemma 3 which leads to

$$\langle x_i(t) - z_s, x_j(t) - x_i(t) \rangle = \langle x_i(t) - z_s, x_j(t) - y_s \rangle + \langle x_i(t) - z_s, y_s - x_i(t) \rangle$$

$$\leq \langle x_i(t) - z_s, y_s - x_i(t) \rangle$$

$$= -V_i(t) + (d_s + \varepsilon)\sqrt{V_i(t)}.$$  (47)

This proves the claim.

Now suppose there exists $i_0 \in V$ with $\lim \inf_{t \to \infty} V_i(t) = \theta_{i_0}^2 < d_s^2$. Then we can find a time sequence $\{t_k\}_1^\infty$ with $\lim_{k \to \infty} t_k = \infty$ such that

$$\sqrt{V_{i_0}(t_k)} \leq \frac{\theta_{i_0} + d_s}{2}.$$  (48)
We divide the rest of the proof into three steps.

**Step 1.** Take \( t_{k_0} > M \). We bound \( V_{i_0}(t) \) in this step.

With the weights rule A8, [43] and inequality (i) in Lemma 4, we see that

\[
\frac{d}{dt} V_{i_0}(t) = 2 \left\langle x_{i_0}(t) - z_*, \sum_{j \in N_{i_0}(\sigma(t))} a_{i_0j}(t) (x_j - x_{i_0}) + f_{i_0}(x_{i_0}(t)) \right\rangle \\
\leq 2 \sum_{j \in N_{i_0}(\sigma(t))} a_{i_0j}(t) \left\langle x_{i_0}(t) - z_*, x_j(t) - x_{i_0}(t) \right\rangle + F_{i_0}(z_*) - F_{i_0}(x_{i_0}(t)) \\
\leq 2(N-1) \alpha^* \left( - V_{i_0}(t) + (d_* + \varepsilon) \sqrt{V_{i_0}(t)} \right),
\]

for all \( t \geq t_{k_0} \), which implies

\[
\frac{d}{dt} \sqrt{V_{i_0}(t)} \leq -(N-1) \alpha^* \left( \sqrt{V_{i_0}(t)} - (d_* + \varepsilon) \right), \quad t \geq t_{k_0}.
\]

In light of Grönwall’s inequality, [48] and [50] yield

\[
\sqrt{V_{i_0}(t)} \leq e^{-(N-1)2 \alpha^* T_D} \sqrt{V_{i_0}(t_{k_0})} + \left( 1 - e^{-(N-1)2 \alpha^* T_D} \right) (d_* + \varepsilon) \\
\leq \frac{2}{2} \left( 1 - e^{-(N-1)2 \alpha^* T_D} \right) (d_* + \varepsilon) \\
\leq \Lambda_* 
\]

for all \( t \in [t_{k_0}, t_{k_0} + (N-1)T_D] \) with \( T_D = T + \tau_D \), where \( T \) comes from the definition of uniformly jointly strongly connected graphs and \( \tau_D \) represents the dwell time.

**Step 2.** Since the graph is uniformly jointly strongly connected, we can find an instant \( \hat{t} \in [t_{k_0}, t_{k_0} + T] \) and another node \( i_1 \in \mathcal{V} \) such that \((i_0, i_1) \in \mathcal{G}_{\sigma(t)} \) for \( t \in [\hat{t}, \hat{t} + \tau_D] \). In this step, we continue to bound \( V_{i_1}(t) \).

Similar to [43], for all \( t \geq M \) and all \( i, j \in \mathcal{V} \), we also have

\[
\left\langle x_i(t) - z_*, x_j(t) - x_i(t) \right\rangle \leq - \sqrt{V_i(t)} \left( \sqrt{V_i(t)} - \sqrt{V_j(t)} \right) 
\]

when \( V_j(t) \leq V_i(t) \). Then based on [43], [51], and [52], we obtain

\[
\frac{d}{dt} V_{i_1}(t) \leq 2 \left\langle x_{i_1}(t) - z_*, x_j(t) - x_{i_1}(t) \right\rangle \\
= 2 \sum_{j \in N_{i_1}(\sigma(t)) \setminus \{i_0\}} a_{i_1j}(t) \left\langle x_{i_1}(t) - z_*, x_j(t) - x_{i_1}(t) \right\rangle + 2a_{i_1i_0}(t) \left\langle x_{i_1}(t) - z_*, x_{i_0}(t) - x_{i_1}(t) \right\rangle \\
\leq 2(N-2) \alpha^* \left( - V_{i_1}(t) + (d_* + \varepsilon) \sqrt{V_{i_1}(t)} \right) - 2 \alpha_* \sqrt{V_{i_1}(t)} \left( \sqrt{V_{i_1}(t)} - \sqrt{V_{i_0}(t)} \right) \\
\leq -2 \left( (N-2) \alpha^* + a_* \right) V_{i_1}(t) + 2 \sqrt{V_{i_1}(t)} \left( (N-2) \alpha^*(d_* + \varepsilon) + \Lambda_* \alpha_* \right)
\]

(53)
for $t \in [\hat{t}, \hat{t} + \tau_D]$, where without loss of generality we assume $V_{i_1}(t) \geq V_{i_0}(t)$ during all $t \in [\hat{t}, \hat{t} + \tau_D]$. Then (53) gives

$$\frac{d}{dt} \sqrt{V_{i_1}(t)} \leq -\left( (N-2)a^* + a_* \right) \sqrt{V_{i_1}(t)} + \left( (N-2)a^*(d_\varepsilon + \varepsilon) + \Lambda_* a_* \right), \quad t \in [\hat{t}, \hat{t} + \tau_D]$$

which yields

$$\sqrt{V_{i_1}(\hat{t} + \tau_D)} \leq e^{-\left( (N-2)a^* + a_* \right) \tau_D} \left( d_\varepsilon + \varepsilon \right) + \left( 1 - e^{-\left( (N-2)a^* + a_* \right) \tau_D} \right) \frac{(N-2)a^*(d_\varepsilon + \varepsilon) + \Lambda_* a_*}{(N-2)a^* + a_*}$$

again by Grönwall’s inequality and some simple algebra.

Next, applying the estimate of node $i_0$ in Step 1 on $i_1$ during time interval $[\hat{t} + \tau_D, t_{k_0} + (N-1)T_D]$, we arrive at

$$\sqrt{V_{i_1}(t)} \leq \left( \frac{\left( a_* \left( 1 - e^{-\left( (N-2)a^* + a_* \right) \tau_D} \right) \right.}{(N-2)a^* + a_*} \right) \times \frac{e^{-2(N-1)^2a^*T_D}}{2} \theta_{i_{k_0}}$$

for all $t \in [t_{k_0} + T_D, t_{k_0} + (N-1)T_D]$.

**Step 3.** Noticing that the graph is uniformly jointly strongly connected, the analysis of steps 1 and 2 can be repeatedly applied to nodes $i_3, \ldots, i_{N-1}$, and eventually we have that for all $i_0, \ldots, i_{N-1}$,

$$\sqrt{V_{i_m}(t_{k_0} + (N-1)T_D)} \leq \left( \frac{\left( a_* \left( 1 - e^{-\left( (N-2)a^* + a_* \right) \tau_D} \right) \right.}{(N-2)a^* + a_*} \right) \times \frac{e^{-2(N-1)^2a^*T_D}}{2} \theta_{i_{k_0}}$$

for sufficiently small $\varepsilon$ because $\theta_{i_{k_0}} < d_\varepsilon$ and

$$\left( \frac{\left( a_* \left( 1 - e^{-\left( (N-2)a^* + a_* \right) \tau_D} \right) \right.}{(N-2)a^* + a_*} \right) \times \frac{e^{-2(N-1)^2a^*T_D}}{2} < 1$$

is a constant. This immediately leads to that

$$V(t_{k_0} + (N-1)T_D) < d_*^2,$$
which contradicts the definition of $d_*$.

This completes the proof. □

Finally, the next lemma shows that each $x_i(t)$ asymptotically reaches $\arg\min F_i$ along the trajectories of system (2).

**Lemma 9** Let A1 hold. Suppose $\bigcap_{i=1}^{N} \{ z : f_i(z) = 0 \} \neq \emptyset$ and $G_{\sigma(t)}$ is uniformly jointly strongly connected. Then along any trajectory of system (2), we have

$$\limsup_{t \to \infty} |x_i(t)|_{\arg\min F_i} = 0$$

for all $i$.

**Proof.** With Lemma 8 we have that $\lim_{t \to \infty} V_i(t) = d_2^*$. Thus, $\forall \varepsilon > 0, \exists M(\varepsilon) > 0$, s.t.,

$$d_* \leq \sqrt{V_i(t)} \leq d_* + \varepsilon$$

(59)

for all $i$ and $t \geq M$. If $d_* = 0$, the desired conclusion follows straightforwardly. Now we suppose $d_* > 0$.

Assume that there exists a node $i_0$ satisfying $\limsup_{t \to \infty} |x_{i_0}(t)|_{\arg\min F_{i_0}} > 0$. Then we can find a time sequence $\{t_k\}^\infty_1$ with $\lim_{k \to \infty} t_k = \infty$ and a constant $\delta$ such that

$$|x_{i_0}(t_k)|_{\arg\min F_{i_0}} \geq \delta, \ k = 1, \ldots ,$$

(60)

Denote also $B_1 \equiv \{ z : |z - z_*| \leq d_* + 1 \}$ and $G_1 = \max \{|f_{i_0}(y)| : y \in B_1\}$. Assumption A1 ensures that $G_1$ is a finite number since $B_1$ is compact. By taking $\varepsilon = 1$ in (59), we see that $x_i(t) \in B_1$ for all $i$ and $t \geq M(1)$. As a result, we have

$$\left| \frac{d}{dt} x_{i_0}(t) \right| = \left| \sum_{j \in N_{i_0}(\sigma(t))} a_{i_0,j}(t)(x_j - x_{i_0}) + f_{i_0}(x_{i_0}) \right| \leq 2(n-1)a^*(d_* + 1) + G_1.$$  

(61)

Combining (60) and (61), we conclude that

$$|x_{i_0}(t)|_{\arg\min F_{i_0}} \geq \frac{\delta}{2}, \ t \in [t_k, t_k + \tau] ,$$

(62)

for all $k = 1, \ldots$, where by definition $\tau = \frac{\delta}{2(2(n-1)a^*(d_* + 1) + G_1)}$.

Now we introduce

$$D_\delta \equiv \min \left\{ F_{i_0}(y) - f_{i_0}(z_*) : |y|_{\arg\min F_{i_0}} \geq \frac{\delta}{2} \text{ and } y \in B_1 \right\}.$$
Then we know $D_\delta > 0$ again by the continuity of $F_{i_0}$. According to (49), (59), and (62), we obtain

$$\frac{d}{dt} V_{i_0}(t) \leq 2(N-1)a^*\left( -V_{i_0}(t) + (d_\ast + \varepsilon)\sqrt{V_{i_0}(t)} \right) + F_{i_0}(z_\ast) - F_{i_0}(x_{i_0}(t))$$

$$\leq 2(N-1)a^*(2d_\ast + \varepsilon)\varepsilon - D_\delta,$$

(63)

for $t \in [t_k, t_k + \tau]$, $k = 1, \ldots$. This leads to

$$V_{i_0}(t_k + \tau) \leq V_{i_0}(t_k) + \left( 2(N-1)a^*(2d_\ast + \varepsilon)\varepsilon - D_\delta \right)\tau$$

$$\leq (d_\ast + \varepsilon)^2 + \left( 2(N-1)a^*(2d_\ast + \varepsilon)\varepsilon - D_\delta \right)\tau$$

$$< d_\ast^2$$

(64)

as long as $\varepsilon$ is chosen sufficiently small. We see that (64) contradicts (59). The desired conclusion thus follows.

5.2 Proofs of Statements

5.2.1 Proof of Theorem 3

The proof of Theorem 3 relies on the following lemma.

Lemma 10 Let $z_1, \ldots, z_{m+1} \in \mathbb{R}^m$ and $d_1, \ldots, d_{m+1} \in \mathbb{R}^+$. Suppose there exist solutions to equations (with variable $y$)

$$\begin{cases}
|y - z_1|^2 = d_1; \\
\vdots \\
|y - z_{m+1}|^2 = d_{m+1}.
\end{cases}$$

(65)

Then the solution is unique if $\text{rank}(z_2 - z_1, \ldots, z_{m+1} - z_1) = m$.

Proof. Take $j > 1$ and let $y$ be a solution to the equations. Noticing that

$$\langle y - z_1, y - z_1 \rangle = d_1; \quad \langle y - z_j, y - z_j \rangle = d_j$$

we obtain

$$\langle y, z_j - z_1 \rangle = \frac{1}{2}\left( d_1 - d_j + |z_j|^2 - |z_1|^2 \right), \quad j = 2, \ldots, m+1.$$  

(66)

The desired conclusion follows immediately.
Let $r_\ast = (r_1 T \ldots r_N T)^T$ be a limit point of a trajectory of System (2). Based on Lemma 8, we have $\lim_{t \to \infty} V_i(t) = d_\ast$ for all $z_\ast \in \bigcap_{i=1}^N \{ z : f_i(z) = 0 \}$. This is to say, $|r_i - z_\ast| = d_\ast$ for all $i$ and $z_\ast \in \bigcap_{i=1}^N \{ z : f_i(z) = 0 \}$. Since $\bigcap_{i=1}^N \{ z : f_i(z) = 0 \} \neq \emptyset$ contains at least one interior point, it is obvious to see that we can find $z_1, \ldots, z_{m+1} \in \bigcap_{i=1}^N \{ z : f_i(z) = 0 \}$ with rank$(z_2 - z_1, \ldots, z_{m+1} - z_1) = m$ and $d_1, \ldots, d_{m+1} \in \mathbb{R}^+$, such that each $r_i, i = 1, \ldots, N$ is a solution of equations (65). Then based on Lemma 10, we conclude that $r_1 = \cdots = r_N$. Next, with Lemma 9, we have $|r_i|_{\arg \min F_i} = 0$. This implies that $r_1 = \cdots = r_N \in \bigcap_{i=1}^N \{ z : f_i(z) = 0 \}$, i.e., global synchronization is achieved.

We turn to state convergence. We only need to show that $r_\ast$ is unique along any trajectory of System (2). Now suppose $r_1^\ast = 1_N \otimes r_1$ and $r_2^\ast = 1_N \otimes r_2$ are two different limit points with $r_1^\ast \neq r_2^\ast \in \bigcap_{i=1}^N \{ z : f_i(z) = 0 \}$. According to the definition of a limit point, we have that for any $\varepsilon > 0$, there exists a time instant $t_\varepsilon$ such that $|x_i(t_\varepsilon) - r^1_\ast| \leq \varepsilon$ for all $i$. Note that Lemma 7 indicates that the disc $B(r^1_\ast, \varepsilon) = \{ y : |y - r^1_\ast| \leq \varepsilon \}$ is an invariant set for initial time $t_\varepsilon$. While taking $\varepsilon = |r^1_\ast - r^2_\ast|/4$, we see that $r^2_\ast \notin B(r^1_\ast, |r^1_\ast - r^2_\ast|/4)$. Thus, $r^2_\ast$ cannot be a limit point.

Now that the limit point is unique along any trajectory of System (2), we denote it as $1_N \otimes x_\ast$ with $x_\ast \in \bigcap_{i=1}^N \{ z : f_i(z) = 0 \}$. Then we have $\lim_{t \to \infty} x_i(t) = x_\ast$ for all $i = 1, \ldots, N$. This completes the proof.

5.2.2 Proof of Theorem 4

In this subsection, we prove Theorem 4. We need the following lemma on robust consensus, which is a special case of the results in [32] (cf., Theorem 4.1 and Proposition 4.10).

**Lemma 11** Consider the following dynamics for the considered network model:

$$
\dot{x}_i = K \sum_{j \in \mathcal{N}_i(\sigma(t))} a_{ij}(t)(x_j - x_i) + w_i(t), \ i \in V \tag{67}
$$

where $K > 0$ is a given constant, $a_{ij}(t)$ are weight functions satisfying our network model, and $w_i(t)$ is a piecewise continuous function. Let $G_{\sigma(t)}$ be uniformly jointly strongly connected with respect to $T > 0$.

(i). There holds $\lim_{t \to \infty} |x_i(t) - x_j(t)| = 0$ for all $i, j \in V$ if $\lim_{t \to \infty} w_i(t) = 0, i \in V$.

(ii). For any $\varepsilon > 0$, there exist a sufficiently small $T_\varepsilon > 0$ and sufficiently large $K_\varepsilon$ such that

$$
\limsup_{t \to \infty} |x_i(t) - x_j(t)| \leq \varepsilon \|w(t)\|_{\infty}
$$
for all initial value \( x^0 \) when \( K \geq K_c \) and \( T \leq T_c \), where \( \|w(t)\|_\infty := \max_{i \in V} \sup_{t \in [0, \infty)} |w_i(t)| \).

Lemma 9 indicates that \( \limsup_{t \to \infty} \left| x_i(t) \right|_{\arg \min F_i} = 0 \) for all \( i \), which yields

\[
\lim_{t \to \infty} f_i(x_i(t)) = 0
\]

for all \( i \) according to Assumption A1. Then global synchronization follows immediately from Lemma 11(i). Again by Lemma 9 we further conclude that \( \limsup_{t \to \infty} \dist(x_i(t), \bigcap_{i=1}^N \{ z : f_i(z) = 0 \}) = 0 \). The desired conclusion thus follows.

5.2.3 Proof of Theorem 5

From Lemma 5 we know that \( \theta(x(t)) = \max_{i \in V} \left| x_i(t) \right|^2_{\Theta_i} \) is non-increasing under A2. As a result, we conclude that

\[
x(t) \in \Gamma(x^0) := \left\{ z \in \mathbb{R}^{mN} : \theta(z) \leq \theta(x^0) \right\}
\]

for all \( t \geq 0 \). Again by Assumption A2, \( \Gamma(x^0) \) is a compact set. We can thus define

\[
h(x^0) := \max_{i \in V} \sup \left\{ \left| f_i(z_i) \right| : z = (z_1 \ldots z_N)^T \in \Gamma(x^0) \right\}.
\]

Now along the trajectory \( x(t) \) of (2) with initial value \( x^0 \), we have

\[
\left| f_i(x_i(t)) \right| \leq h(x^0)
\]

for all \( t \geq 0 \). Then the desired \( \epsilon \)-synchronization result follows immediately from Lemma 11(ii).

6 Conclusions

In insights of recent works on consensus-based distributed optimization methods, we have established some conditions on the synchronization problems of coupled oscillators. We assumed that the network nodes have non-identical nonlinear self-dynamics which are gradients of some concave functions. This allowed functions which might not be passive nor globally Lipschitz. The node interactions were under switching directed communication graphs. Some sufficient and/or necessary conditions are established regarding exact or approximate synchronization of the overall node states. These results revealed when and how nonlinear node self-dynamics can cooperate with the linear consensus coupling and reach synchronization with much relaxed connectivity conditions. They also naturally served a generalization of the recent studies on
continuous-time distributed optimization. Some interesting future generalizations include the exact convergence rate to a synchronization under strict convexity, and synchronization conditions with constrained node states.

References

[1] G. Shi, A. Proutiere, and K. H. Johansson, “Continuous-time distributed optimization of homogenous dynamics,” The 51st Annual Allerton Conference on Communication, Control, and Computing, Oct. 2013.

[2] R. T. Rockafellar. Convex Analysis. New Jersey: Princeton University Press, 1972.

[3] J. Aubin and A. Cellina. Differential Inclusions. Berlin: Springer-Verlag, 1984.

[4] C. Godsil and G. Royle, Algebraic Graph Theory. New York: Springer-Verlag, 2001.

[5] N. Rouche, P. Habets, and M. Laloy, Stability Theory by Liapunov’s Direct Method. New York: Springer-Verlag, 1977.

[6] J. Danskin, “The theory of max-min, with applications,” SIAM J. Appl. Math., vol. 14, 641-664, 1966.

[7] F. Clarke, Yu. S. Ledyaev, R. Stern, and P. Wolenski. Nonsmooth Analysis and Control Theory. Berlin: Springer-Verlag, 1998.

[8] J. Cortés, “Discontinuous dynamical systems–A tutorial on solutions, nonsmooth analysis, and stability,” IEEE Control Systems Magazine, vol. 28, no. 3, 36-73, 2008.

[9] L. M. Pecora and T. L. Carroll, “Master stability functions for synchronized coupled systems,” Physical Review Letters, vol. 80, pp. 2019-2112, 1998.

[10] T. Vicsek, A. Czirok, E. B. Jacob, I. Cohen, and O. Schochet, “Novel type of phase transitions in a system of self-driven particles,” Physical Review Letters, vol. 75, pp. 1226–1229, 1995.

[11] S. Martinez, J. Cortes, and F. Bullo, “Motion coordination with distributed information,” IEEE Control Syst. Mag., vol. 27, no. 4, pp. 75-88, 2007.
[12] C. W. Wu. *Synchronization in complex networks of nonlinear dynamical systems*. World Scientific, 2007.

[13] S. Kar and J. M. F. Moura, “Consensus + innovations distributed inference over networks: cooperation and sensing in networked systems,” *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 99-109, May 2013.

[14] C. W. Wu and L. O. Chua, “Synchronization in an array of linearly coupled dynamical systems,” *IEEE Transactions on Circuits and Systems I*, vol. 42, pp. 430-447, 1995.

[15] V. N. Belykh, I. V. Belykh, and M. Hasler, “Connection graph stability method for synchronized coupled chaotic systems,” *Physica D.*, vol. 195, pp. 159-187, 2004.

[16] A. Y. Pogromsky and H. Nijmeijer, “Cooperative oscillatory behavior of mutually coupled dynamical systems,” *IEEE Trans. Circ. and Syst.-I Fundamental theory and applications*, vol. 48, no. 2, pp. 152-162, 2001.

[17] A. Y. Pogromsky, G. Santoboni, and H. Nijmeijer, “Partial synchronization: from symmetry towards stability,” *Physica D*, vol. 172, pp. 65-87, 2002.

[18] Z. Li, Z. Duan, G. Chen and L. Huang, “Consensus of multi-agent systems and synchronization of complex networks: A unified viewpoint,” *IEEE Trans. Circ. Syst.-I Regular papers*, vol. 57, no. 1, pp. 213-224, 2010.

[19] K. You and L. Xie, “Network topology and communication data rate for consensusability of discrete-time multi-agent systems,” *IEEE Trans. Aut. Contr.*, vol. 56, no. 10, pp. 2262-2275, 2011.

[20] L. Scardovi and R. Sepulchre, “Synchronization in networks of identical linear systems,” *Automatica*, vol. 45, no. 11, pp. 2557-2562, 2009.

[21] F. Dörfler and F. Bullo, “Synchronization and transient stability in power networks and non-uniform kuramoto oscillators,” *SIAM J. Contr. Opt.*, vol. 50, no. 3, pp. 1616-1642, 2012.

[22] Y. Su and J. Huang, “Stability of a class of linear switching systems with applications to two consensus problem,” *IEEE Trans. Aut. Contr.*, vol. 57, no. 6, pp. 1420-1430, 2012.
[23] H. Liu, M. Cao, and C. W. Wu, “Coupling strength allocation for synchronization in complex networks using spectral graph theory,” *IEEE Transactions on Circuits and Systems-I*, in press, available online.

[24] J. Tsitsiklis, D. Bertsekas, and M. Athans, “Distributed asynchronous deterministic and stochastic gradient optimization algorithms,” *IEEE Trans. Autom. Control*, vol. 31, no. 9, pp. 803-812, 1986.

[25] A. Jadbabaie, J. Lin, and A. S. Morse, “Coordination of groups of mobile autonomous agents using nearest neighbor rules,” *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988-1001, 2003.

[26] J. Fax and R. Murray, “Information flow and cooperative control of vehicle formations,” *IEEE Trans. Automatic Control*, vol. 49, no. 9, 1465-1476, 2004.

[27] W. Ren and R. Beard, “Consensus seeking in multi-agent systems under dynamically changing interaction topologies,” *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655-661, 2005.

[28] R. Olfati-Saber, “Flocking for multi-agent dynamic systems: algorithms and theory,” *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401-420, 2006.

[29] M. Cao, A. S. Morse and B. D. O. Anderson, “Reaching a consensus in a dynamically changing environment: a graphical approach,” *SIAM J. Control Optim.*, vol. 47, no. 2, 575-600, 2008.

[30] L. Moreau, “Stability of multiagent systems with time-dependent communication links,” *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169-182, 2005.

[31] Z. Lin, B. Francis, and M. Maggiore, “State agreement for continuous-time coupled nonlinear systems,” *SIAM J. Control Optim.*, vol. 46, no. 1, pp. 288-307, 2007.

[32] G. Shi and K. H. Johansson, “Robust consensus for continuous-time multi-agent dynamics,” *SIAM J. Control Optim.*, vol. 51, no. 5, pp. 3673-3691, 2013.

[33] A. Nedić and A. Ozdaglar, “Distributed subgradient methods for multi-agent optimization,” *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48-61, 2009.

[34] A. Nedić, A. Ozdaglar and P. A. Parrilo, “Constrained consensus and optimization in multi-agent networks,” *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922-938, 2010.
[35] J. Wang and N. Elia, “Control approach to distributed optimization,” *The 48th Annual Allerton Conference*, pp. 557-561, 2010.

[36] G. Shi, K. H. Johansson and Y. Hong, “Reaching an optimal consensus: dynamical systems that compute intersections of convex sets,” *IEEE Transactions on Automatic Control*, vol. 58, no.3, pp. 610-622, 2013.

[37] D. Jakovetić, J. Xavier and J. M. F. Moura, “Cooperative convex optimization in networked systems: augmented Lagrangian algorithms with directed gossip communication,” *IEEE Trans. Signal Processing*, vol. 59, no. 8, pp. 3889-3902, 2011.

[38] H.-B. Dürr and C. Ebenbauer, “A smooth vector field for saddle point problems,” *The 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 4654-4660, 2011.

[39] B. Gharesifard, J. Cortés, “Distributed continuous-time convex optimization on weight-balanced digraphs,” *IEEE Trans. Automatic Control*, in press, available online.

[40] M. V. Solodov, “Incremental gradient algorithms with stepsizes bounded away from zero,” *Computational Optimization and Applications*, vol. 11, pp. 23-35, 1998.

[41] A. Nedić and D.P. Bertsekas, “Incremental subgradient methods for nondifferentiable optimization,” *SIAM J. on Optimization*, vol. 12, pp. 109-138, 2001.

[42] S. S. Ram, A. Nedić, and V. V. Veeravalli, “Incremental stochastic subgradient algorithms for convex optimization,” *SIAM Journal on Optimization*, vol. 20, no. 2, 691-717, 2009.