Reaching an Optimal Consensus: Dynamical Systems that Compute Intersections of Convex Sets

Guodong Shi, Karl Henrik Johansson, *Senior Member, IEEE,* and Yiguang Hong, *Senior Member, IEEE*

***Abstract*—In this paper, multi-agent systems minimizing a sum of objective functions, where each component is only known to a particular node, is considered for continuous-time dynamics with time-varying interconnection topologies. Assuming that each node can observe a convex solution set of its optimization component, and the intersection of all such sets is nonempty, the considered optimization problem is converted to an intersection computation problem. By a simple distributed control rule, the considered multi-agent system with continuous-time dynamics achieves not only a consensus, but also an optimal agreement within the optimal solution set of the overall optimization objective. Directed and bidirectional communications are studied, respectively, and connectivity conditions are given to ensure a global optimal consensus. In this way, the corresponding intersection computation problem is solved by the proposed decentralized continuous-time algorithm. We establish several important properties of the distance functions with respect to the global optimal solution set and a class of invariant sets with the help of convex and non-smooth analysis.**

***Index Terms*—Multi-agent systems, Optimal consensus, Connectivity Conditions, Distributed optimization, Intersection computation**

I. Introduction

In recent years, multi-agent dynamics has been intensively investigated in various areas including engineering, natural science, and social science. Cooperative control of multi-agent systems is an active research topic, and rapid developments of distributed control protocols via interconnected communication have been made to achieve the collective tasks, e.g., [25], [24], [21], [34], [19], [18], [29], [31], [26], [27]. However, fundamental challenges still lie in finding suitable tools to describe and design the dynamical behavior of these systems and thus providing insights in their functioning principles. Different from the classical control design, the multi-agent studies aim at fully exploiting, rather than avoiding, interconnection between agents in analysis and synthesis in order to deal with distributed design and large-scale information process.

Consensus is a basic problem in the study of multi-agent coordination. It requires that all the agents achieve the same state, such as a certain relative position or velocity. To achieve consensus, connectivity plays a key role, and consequently several connectivity conditions have been established to describe suitable switching topologies. *Jointly connected graph* and similar concepts are important in the analysis to guarantee convergence. *Uniformly jointly connected graph*, i.e., the joint graph is connected during all intervals which are longer than a constant, has been employed [25], [24], [33], [28], [15]. On the other hand, $[t, \infty)$-joint connectedness, i.e., the joint graph is connected in the time intervals $[t, \infty)$, is necessary [31], [34], and therefore the most general form to secure the global coordination.

Moreover, distributed optimization of a sum of convex objective functions, $\sum_{i=1}^{N} f_i(z)$, where each component f_i is known only to node i, has attracted much attention in recent years, due to its wide application in multi-agent systems and wireless networks [38], [39], [41], [40], [42]. A class of subgradient-based incremental methods, in which some estimate of the optimal solution can be passed over the network via deterministic or randomized iteration, were studied in [38], [39], [43]. Then a non-gradient-based algorithm was proposed in [42], where each node starts at its own optimal solution and updates using a pairwise equalizing protocol. In view of multi-agent systems, the local information transmitted over the neighborhood is usually limited to a convex combination of its neighbors [25], [24], [34]. Combining the ideas of consensus algorithms and subgradient methods, a number of significant results were obtained. A subgradient method in combination with consensus steps was given for solving coupled optimization problems with fixed undirected topology in [40]. Then, an important work on multi-agent optimization was [36], where a decentralized algorithm was proposed as a simple sum of an averaging (consensus) part and a subgradient part, and convergence bounds for a distributed multi-agent model under various connectivity conditions were shown. Constrained consensus and optimization were further studied in [37], where each agent was always restricted in its own convex set. A "projected consensus algorithm" was presented to solve the constrained consensus problem in which each agent takes averaging and projection steps alternatively, and it was generalized to "projected subgradient algorithm" with optimization goal also took into consideration [37].

Most of the literature on optimization and consensus algorithms is in discrete time, and it is usually hard for the considered agents to reach both consensus and optimum unless the weights rule of the links, the step size in the iteration and the connectedness of the communication graph are properly selected [36], [37], [39]. Few researchers have considered continuous-time agent dynamics that solves a distributed optimization problem. However, dynamical system solution to

G. Shi and K. H. Johansson are with ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Stockholm 10044, Sweden e-mail: guodong.shi@ee.kth.se, kallej@ee.kth.se

Y. Hong is with Key Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Sciences, Beijing 100190, China email: yghong@iss.ac.cn

Manuscript received ??; revised ??; accepted ??. Recommended by Associate Editor P. Shi. This work has been supported in part by the Knut and Alice Wallenberg Foundation, the Swedish Research Council, KTH SRA TNG, and the NNSF of China under Grants 61174071.

optimization problem is of great interest since a simple vector-field solution may provide important geometrical insights. The classical Arrow-Hurwicz-Uzawa flow was shown to converge to the set of saddle points for a constrained convex optimization problem [9]. Then in [10], a simple and elegant continuous-time protocol was presented which solves linear programming problems.

The goal of this paper is to establish a simple distributed continuous-time control law which can ensure consensus and minimize $\sum_{i=1}^{N} f_i(z)$ asymptotically. Each optimal solution set, X_i of optimization objective $f_i(z)$, is assumed to be a convex set observed only by node i. Assuming that the intersection set, $\bigcap_{i=1}^{N} X_i$, is nonempty, the optimal solution set of the group objective becomes this intersection set, and the considered optimization problem is then converted to a distributed intersection computation problem. In fact, computing several convex sets' intersection is a classical problem, and "alternating projection algorithm" was a standard solution, in which the algorithm is carried out by iteratively projecting onto each set [6], [7], [8]. The "projected consensus algorithm" presented in [37] can be viewed as its generalized version. The intersection computation problem is also of interest in the study of computational geometry, a branch of computer science [12], [13]. Hence, an important motivation for our work is to provide a system-theoretic insight into the convergence properties of certain distributed optimization problems. Similar to the continuous-time approximation of recursive algorithms [11] and constrained optimizations [9], [10], we establish a suitable dynamical model for such analysis. Also by itself, the considered continuous-time distributed optimization problem has many applications, e.g., wireless resource allocation [38], [39], formation control [18], [31], [23], and mobile sensing [19], [46].

In this paper, we present a simple dynamical system solution to this convex intersection computation problem, as the sum of a consensus part and a projection part. Since this projection part can be viewed as a special subgradient information, this protocol is actually a continuous-time version of the algorithm proposed in [36]. We show that an optimal consensus (i.e., consensus within the global optimal solution set), can be achieved under time-varying communications. Both directed and bidirectional cases are investigated, and sharp connectivity conditions are obtained in the sense that a general optimal consensus will no longer hold for a general model with weaker connectedness. Additionally, we use quite general weights rule which allow the weight of each arc in the communication graph to depend on time or system state.

The rest of the paper is organized as follows. In Section 2, some preliminary concepts are introduced. In Section 3, we formulate the considered optimal consensus problem, and the main results are shown. Then, in Sections 4 and 5, convergence to the optimal solution set and global consensus are analyzed, respectively, based on which the proofs of the main results are obtained. Finally, in Section 6 concluding remarks are given.

II. Preliminaries

In this section, we introduce some notations and theories on graph theory [4], convex analysis [1], [3] and nonsmooth analysis [5].

A directed graph (digraph) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a finite set $\mathcal{V}$ of nodes and an arc set $\mathcal{E}$, in which an arc is an ordered pair of distinct nodes of $\mathcal{V}$. An element $(i, j) \in \mathcal{E}$ describes an arc which leaves i and enters j. A *walk* in digraph $\mathcal{G}$ is an alternating sequence $\mathcal{W} : i_1 e_1 i_2 e_2 \ldots e_{m-1} i_m$ of nodes i_κ and arcs $e_\kappa = (i_\kappa, i_{\kappa+1}) \in \mathcal{E}$ for $\kappa = 1, 2, \ldots, m-1$. A walk is called a *path* if the nodes of this walk are distinct, and a path from i to j is denoted as $i \to j$. $\mathcal{G}$ is said to be *strongly connected* if it contains path $i \to j$ and $j \to i$ for every pair of nodes i and j. A digraph $\mathcal{G}$ is called to be *bidirectional* when for any two nodes i and j, $(i, j) \in \mathcal{E}$ if and only if $(j, i) \in \mathcal{E}$. Ignoring the direction of the arcs, the connectedness of a bidirectional digraph will be transformed to that of the corresponding undirected graph. A time-varying graph is defined as $\mathcal{G}_{\sigma(t)} = (\mathcal{V}, \mathcal{E}_{\sigma(t)})$ with $\sigma : [0, +\infty) \to \mathcal{Q}$ as a piecewise constant function, where $\mathcal{Q}$ is a finite set indicating all possible graphs. Moreover, the joint graph of $\mathcal{G}_{\sigma(t)}$ in time interval $[t_1, t_2)$ with $t_1 < t_2 \leq +\infty$ is denoted as $\mathcal{G}([t_1, t_2)) = \cup_{t\in[t_1,t_2)} \mathcal{G}(t) = (\mathcal{V}, \cup_{t\in[t_1,t_2)} \mathcal{E}_{\sigma(t)})$.

A set $K \subset \mathbb{R}^m$ is said to be *convex* if $(1-\lambda)x + \lambda y \in K$ whenever $x \in K, y \in K$ and $0 \leq \lambda \leq 1$. For any set $S \subset \mathbb{R}^m$, the intersection of all convex sets containing S is called the *convex hull* of S, denoted by $co(S)$. The next lemma can be found in [2].

Lemma 2.1: Let K be a subset of $\mathbb{R}^m$. The convex hull $co(K)$ of K is the set of elements of the form

$$x = \sum_{i=1}^{m+1} \lambda_i x_i,$$

where $\lambda_i \geq 0, i = 1, \ldots, m+1$ with $\sum_{i=1}^{m+1} \lambda_i = 1$ and $x_i \in K$.

Let K be a closed convex subset in $\mathbb{R}^m$ and denote $|x|_K \doteq \inf_{y\in K} |x-y|$ as the distance between $x \in \mathbb{R}^m$ and K, where $|\cdot|$ denotes the Euclidean norm. There is a unique element $P_K(x) \in K$ satisfying $|x - P_K(x)| = |x|_K$ associated to any $x \in \mathbb{R}^m$ [2]. The map P_K is called the *projector* onto K. We also have

$$\langle P_K(x) - x, P_K(x) - y\rangle \leq 0, \quad \forall y \in K. \tag{1}$$

Moreover, P_K has the following non-expansiveness property:

$$|P_K(x) - P_K(y)| \leq |x - y|, x, y \in \mathbb{R}^m. \tag{2}$$

Clearly, $|x|_K^2$ is continuously differentiable at point x, and (see [2])

$$\nabla |x|_K^2 = 2(x - P_K(x)). \tag{3}$$

The following lemma was obtained in [31], which is useful in what follows.

Lemma 2.2: Suppose $K \subset \mathbb{R}^m$ is a convex set and $x_a, x_b \in \mathbb{R}^m$. Then

$$\langle x_a - P_K(x_a), x_b - x_a\rangle \leq |x_a|_K \cdot \left||x_a|_K - |x_b|_K\right|. \tag{4}$$

Particularly, if $|x_a|_K > |x_b|_K$, then

$$\langle x_a - P_K(x_a), x_b - x_a\rangle \leq -|x_a|_K \cdot (|x_a|_K - |x_b|_K). \tag{5}$$

Next, the upper *Dini derivative* of a continuous function $h : (a, b) \to \mathbb{R}$ $(-\infty \le a < b \le \infty)$ 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 $D^+h(t) \le 0$ for any $t \in (a, b)$. The next result is given for the calculation of Dini derivative (see [14], [33]).

Lemma 2.3: Let $V_i(t, x) : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ $(i = 1, \dots, n)$ be C^1 and $V(t, x) = \max_{i=1,\dots,n} V_i(t, x)$. If $\mathcal{I}(t) = \{i \in \{1, 2, \dots, n\} : V(t, x(t)) = V_i(t, x(t))\}$ is the set of indices where the maximum is reached at t, then $D^+V(t, x(t)) = \max_{i\in\mathcal{I}(t)} \dot{V}_i(t, x(t))$.

Finally, consider a system

$$\dot{x} = f(t, x), \tag{6}$$

where $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is piecewise continuous in t and continuous in x. Let $x(t) = x(t, t_0, x^0)$ be a solution of (6) with initial condition $x(t_0) = x^0$. Then $\Omega_0 \subset \mathbb{R}^d$ is called a *positively invariant set* of (6) if, for any $t_0 \in R$ and any $x^0 \in \Omega_0$, $x(t, t_0, x^0) \in \Omega_0$ when $t \ge t_0$.

III. Problem Formulation and Main Results

In this section, we first define the considered optimal consensus problem. We propose a multi-agent optimization model and a distributed control law to solve this optimization problem. Then the main results are presented on connectivity conditions which can ensure an optimal consensus globally.

A. Multi-agent Model

Consider a multi-agent system with agent set $\mathcal{V} = \{1, 2, \dots, N\}$, for which the dynamics of each agent is a first-order integrator:

$$\dot{x}_i = u_i, \quad i = 1, \dots, N \tag{7}$$

where $x_i \in \mathbb{R}^m$ represents the state of agent i, and u_i is the control input.

The communication in the multi-agent network is modeled as a time-varying graph $\mathcal{G}_{\sigma(t)} = (\mathcal{V}, \mathcal{E}_{\sigma(t)})$. Moreover, node j is said to be a *neighbor* of i at time t when there is an arc $(j, i) \in \mathcal{E}_{\sigma(t)}$, and $N_i(\sigma(t))$ represents the set of agent i's neighbors at time t. As usual in the literature [24], [33], [31], an assumption is given to the variation of $\mathcal{G}_{\sigma(t)}$.

A1 *(Dwell Time)* There is a lower bound constant $\tau_D > 0$ between two consecutive switching time instants of $\sigma(t)$.

We have the following definition.

Definition 3.1: (i) $\mathcal{G}_{\sigma(t)}$ is said to be *uniformly jointly strongly connected* (UJSC) if there exists a constant $T > 0$ such that $\mathcal{G}([t, t + T))$ is strongly connected for any $t \ge 0$.

(ii) Assume that $\mathcal{G}_{\sigma(t)}, t \ge 0$ is bidirectional. $\mathcal{G}_{\sigma(t)}$ is said to be *infinitely jointly connected* (IJC) if $\mathcal{G}([t, +\infty))$ is connected for all $t \ge 0$.

Remark 3.1: $[t, +\infty)$-joint connectedness for all $t \ge 0$ is equivalent to that there exists an unbounded time sequence $0 \le t_1 < \dots < t_k < t_{k+1} < \dots$ such that $\mathcal{G}([t_k, t_{k+1}))$ is connected for all $k = 1, 2, \dots$. Note that it does not require an upper bound for $|t_{k+1} - t_k|$ in the definition.

The objective for this group of autonomous agents is to reach a consensus, and meanwhile to cooperatively solve the following optimization problem

$$\min_{z\in\mathbb{R}^m} \quad \sum_{i=1}^{N} f_i(z) \tag{8}$$

where $f_i : \mathbb{R}^m \to \mathbb{R}$ represents the cost function of agent i, observed by agent i only, and z is a decision vector. We suppose the optimal solution set of each component f_i exists, denoted $X_i \doteq \{v \mid f_i(v) = \min_{z\in\mathbb{R}^m} f_i(z)\}$.

We impose the following assumptions.

A2 *(Convexity)* $X_1, \dots, X_N$, are closed convex sets.

A3 *(Nonempty Intersection)* $X_0 \doteq \bigcap_{i=1}^{N} X_i$ is nonempty and bounded.

Remark 3.2: The assumption that each X_i is a convex set is quite general, and it is not hard to see that this assumption will be satisfied as long as each f_i is a convex function. Moreover, since the intersection of convex sets is a convex set itself, X_0 is a convex set with the convexity of each X_i. Additionally, with A3, it is obvious to see that X_0 is compact, and it is the optimal solution set of (8).

B. Distributed Control

Denote $x = (x_1^T, \dots, x_N^T)^T \in \mathbb{R}^{mN}$ and let the continuous function $a_{ij}(x, t) > 0$ be the weight of arc (j, i), for $i, j \in \mathcal{V}$. Then we present the following distributed control law:

$$u_i = \sum_{j\in N_i(\sigma(t))} a_{ij}(x, t)(x_j - x_i) + P_{X_i}(x_i) - x_i, \ i = 1, \dots, N \tag{9}$$

Remark 3.3: We write the arc weight $a_{ij}(x, t)$ in a quite general form showing that this weight function can be time-varying and may depend nonlinearly on the state. Note that this doesn't mean global information is required for the control design.

Remark 3.4: When X_i can be observed by node i, $P_{X_i}(x_i(t)) - x_i(t)$ can be easily obtained. For instance, node i may first establish a local coordinate system, and then construct a function $h(z) = |z|^2_{X_i}$ to compute $\nabla h(z)$ within this coordinate system. Then by (3), we have $P_{X_i}(x_i(t)) - x_i(t) = -1/2\nabla h(z)|_{z=x_i(t)}$.

Another assumption is made on each $a_{ij}(x, t), i, j = 1, 2, ..., N$.

A4 *(Weights Rule)* There are $a^* > 0$ and $a_* > 0$ such that

$$a_* \le a_{ij}(x, t) \le a^*, \quad t \in \mathbb{R}^+, x \in \mathbb{R}^{mN}.$$

In this paper, we assume that Assumptions A1-A4 always hold. With (7) and (9), the closed loop system is expressed by

$$\dot{x}_i = \sum_{j\in N_i(\sigma(t))} a_{ij}(x, t)(x_j - x_i) + P_{X_i}(x_i) - x_i, \ i = 1, \dots, N. \tag{10}$$

Remark 3.5: By the non-expansiveness property (2), the convex projection $P_K(z)$ is continuous for all $z \in \mathbb{R}^m$ for

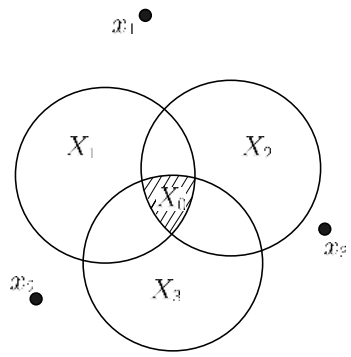

Fig. 1. The goal of the agents is to achieve a consensus in X_0.

any closed convex set $K \subseteq R^m$. Therefore, a Caratheodory solution of (10) exists at least over a finite interval for any initial condition $x(t_0)$ based on assumption A1 (see [16], [17]). Note that the solution is not necessarily unique. As will be shown in Remark 4.1, it also exists in $[t_0, +\infty)$.

Remark 3.6: Since the projection term can be viewed as a subgradient for the special case $f_i(z) = |z|^2_{X_i}/2$, (10) is actually a continuous-time version of the algorithm proposed in [36], which has the form of the sum of a consensus term and a subgradient term. On the other hand, in [37], a "projected consensus algorithm" was presented to solve the same intersection computation problem in which each agent takes consensus and projection steps alternatively. Note that there is some essential difference between (10) and the "projected consensus algorithm" in [37], because (10) takes advantage of the consensus and projection information at the same time instant. It is not hard to construct examples in which each node i would never enter its own set X_i along the trajectories of (10).

Let $x(t)$ be the trajectory of (10) with initial condition $x^0 = x(t_0) = (x_1^T(t_0), \ldots, x_N^T(t_0))^T \in \mathbb{R}^{mN}$. Then the considered optimal consensus is defined as following (see Fig. 1).

Definition 3.2: (i) A global *optimal set convergence* of (10) is achieved if for all $x^0 \in \mathbb{R}^{mN}$, we have

$$\lim_{t\to+\infty} |x_i(t)|_{X_0} = 0, \quad i = 1, \ldots, N. \tag{11}$$

(ii) A global *consensus* of (10) is achieved if for all $x^0 \in \mathbb{R}^{mN}$, we have

$$\lim_{t\to+\infty} |x_i(t) - x_j(t)| = 0, \quad i, j = 1, \ldots, N. \tag{12}$$

(iii) A global *optimal consensus* is achieved of (10) if both (i) and (ii) hold.

Remark 3.7: It is easy to find that, based on the analysis methods we provide, all the results obtained in this paper will still hold if the control law (10) is replaced by

$$\dot{x}_i = \sum_{j\in N_i(\sigma(t))} a_{ij}(x,t)(x_j - x_i) + b_i(x_i,t)(P_{X_i}(x_i) - x_i)$$

for $i = 1, \ldots, N$ and some scalar functions $0 < b_* \leq b_i(x_i, t), i = 1, \cdots, N$ with $b_* > 0$ being a constant. Here we just choose the form of (10) to make the statements and proofs simplified.

C. Main Results

In this subsection, we present the main results on optimal consensus.

First the following conclusion is our main result for directed graphs.

Theorem 3.1: System (10) achieves a global optimal consensus if $\mathcal{G}_{\sigma(t)}$ is UJSC.

We say the communications over the considered multi-agent network are bidirectional if $\mathcal{G}_{\sigma(t)}$ is a bidirectional graph for all $t \geq t_0$. Note that, this does not imply that the arc weights, $a_{ij}(x,t), i, j = 1, \ldots, N$, are symmetric. Then we have the following main result on optimal consensus for the bidirectional case.

Theorem 3.2: System (10) with bidirectional communications achieves a global optimal consensus if (and in general only if) $\mathcal{G}_{\sigma(t)}$ is IJC.

Theorem 3.2 shows that the connectedness conditions to reach an optimal consensus can be relaxed for bidirectional communications without requiring a uniform bound of the length of intervals in the definition of connectivities.

Remark 3.8: Let us explain what "in general only if" means in Theorems 3.2. Clearly, the connectivity condition proposed in Theorem 3.2 is not a necessary condition to ensure a global optimal consensus for a particular optimization problem (8). However, in regard to a global optimal consensus for all possibilities of $X_1, \ldots, X_N$, simple examples could show that this IJC assumption is also necessary using the same idea studying state agreement problem in [34], [31]. In fact, as long as $\bigcap_{i=1}^N X_i$ is not a singleton, it can be easily shown that consensus cannot be guaranteed for all initial conditions. Therefore, from this perspective, Theorem 3.2 gives "sharp" connectivity conditions for a global optimal consensus of system (10).

Remark 3.9: If A3, the nonempty intersection assumption, is removed, control law (10) becomes a special case of the target aggregation controller studied in [31] with respect to $co(\bigcup_{i=1}^N X_i)$. In this case, under proper connectivity assumptions (even each node cannot always obtain the information of X_i), it can be shown that (10) will lead the network to converge into $co(\bigcup_{i=1}^N X_i)$ [31]. The dynamics within $co(\bigcup_{i=1}^N X_i)$ can be complicated, and the optimal consensus will fail since there is no longer a simple expression of X_*, the real optimal solution set of (8). However, we guess that in this case the control law (10) still implies a suboptimal convergence such that there will be a constant B, which does not depend on the initial condition, satisfying $\limsup_{t\to\infty} |x_i(t)|_{X_*} \leq B$ under UJSC connectivity conditions.

In order to prove Theorems 3.1 and 3.2, on one hand, we have to prove all the agents converge to the global optimal solution set, i.e., X_0; and, on the other hand, we have to verify that a consensus is also achieved. In fact, the convergence analysis is quite challenging, due to the nonlinearity nature of each weight function $a_{ij}(x,t)$ and the convex projection part in the control law. In the following two sections, we will focus on the optimal solution set convergence and the consensus analysis, respectively, by which complete the proofs for Theorems 3.1 and 3.2.

IV. Optimal Set Convergence

In this section, we prove the optimal solution set convergence for system (10). We first establish a method to analyze the distance between the agents and the global optimal set with the help of convex analysis, and then the convergence to X_0 for all the agents is proposed under directed and bidirectional communications, respectively.

A. Distance Function

Define $d_i(t) = |x_i(t)|^2_{X_0}$ and let

$$d(t) = \max_{i\in\mathcal{V}} d_i(t)$$

be the maximum among all the agents. Although $d(t)$ may not be continuously differentiable, it is still continuous. Thus, we can analyze the Dini derivative of $d(t)$ to study its convergence property. Moreover, it is easy to see that $d(t)$ is locally Lipschitz. Then the Dini derivative of $d(t)$ is finite for any t.

We prove several elementary lemmas for the following analysis. At first, the following lemma indicates that $d(t)$ is nonincreasing.

Lemma 4.1: $D^+d(t) \le 0$ for all $t \ge 0$.

Proof. According to (3), one has

$$\begin{aligned}\frac{d}{dt}d_i(t) &= 2\langle x_i - P_{X_0}(x_i), \dot{x}_i\rangle \\ &= 2\langle x_i - P_{X_0}(x_i), \\ &\qquad \sum_{j\in N_i(\sigma(t))} a_{ij}(x,t)(x_j - x_i) + P_{X_i}(x_i) - x_i\rangle.\end{aligned} \tag{13}$$

Then, based on Lemma 2.3 and denoting $\mathcal{I}(t)$ as the set containing all the agents that reach the maximum in the definition of $d(t)$ at time t, we obtain

$$\begin{aligned}D^+d(t) &= \max_{i\in\mathcal{I}(t)} \frac{d}{dt}d_i(t) \\ &= 2\max_{i\in\mathcal{I}(t)}[\langle x_i - P_{X_0}(x_i), \\ &\qquad \sum_{j\in N_i(\sigma(t))} a_{ij}(x_j - x_i) + P_{X_i}(x_i) - x_i\rangle].\end{aligned} \tag{14}$$

Furthermore, for any $i \in \mathcal{I}(t)$, according to (5) of Lemma 2.2, one has

$$\langle x_i - P_{X_0}(x_i), x_j - x_i\rangle \le 0 \tag{15}$$

for any $j \in L_i(\sigma(t))$ since it always holds that $|x_j|_{X_0} \le |x_i|_{X_0}$.

Moreover, in light of (1), we obtain

$$\langle P_{X_i}(x_i) - P_{X_0}(x_i), P_{X_i}(x_i) - x_i\rangle \le 0 \tag{16}$$

since we always have $P_{X_0}(x_i) \in X_i$ for all $i = 1,\dots,N$. Therefore, it is easy to see that for any $i \in \mathcal{V}$,

$$\begin{aligned}&\langle x_i - P_{X_0}(x_i), P_{X_i}(x_i) - x_i\rangle \\ &= \langle x_i - P_{X_i}(x_i) + P_{X_i}(x_i) - P_{X_0}(x_i), P_{X_i}(x_i) - x_i\rangle \\ &\le \langle x_i - P_{X_i}(x_i), P_{X_i}(x_i) - x_i\rangle \\ &= -|x_i|^2_{X_i}.\end{aligned} \tag{17}$$

Thus, with (14), (15) and (17), one has

$$D^+d(t) \le 2\max_{i\in\mathcal{I}(t)}[-|x_i|^2_{X_i}] \le 0. \tag{18}$$

Then the proof is completed. □

Remark 4.1: According to Lemma 4.1, $\{y : |y|^2_{X_0} \le d(t_0)\}$ is a positively invariant set for system (10). Since X_0 is compact, $\{y : |y|^2_{X_0} \le d(t_0)\}$ is also compact. This leads to that each solution of (10) exists in $[t_0, +\infty)$. Moreover, if the weight functions $a_{ij}, i, j = 1,\dots,N$, are only state-dependent, the continuity implies that there will be $a^* \ge a_* > 0$ such that

$$a_* \le a_{ij}(x(t)) \le a^*, \quad \forall t > 0,\ i,j = 1,2,...N \tag{19}$$

along trajectory $x(t)$ of system (10). In this case, A4 follows automatically, and then needs not to be assumed.

With Lemma 4.1, for any initial condition, there exists a constant $d^* \ge 0$ such that $\lim_{t\to\infty} d(t) = d^*$. Clearly, the optimal solution set convergence will be achieved for system (10) if and only if $d^* = 0$. Furthermore, since it always holds that $d_i(t) \le d(t)$, there exist constants $0 \le \theta_i \le \eta_i \le d^*, i = 1,\dots,N$ such that

$$\liminf_{t\to\infty} d_i(t) = \theta_i, \quad \limsup_{t\to\infty} d_i(t) = \eta_i.$$

To establish the optimal set convergence, we also need the following lemmas, whose proofs can be found in the appendices.

Lemma 4.2: Assume that $\theta_i = \eta_i = d^*, i = 1,\dots,N$. Then we have $\lim_{t\to+\infty} |x_i(t)|_{X_i} = 0$ for all $i = 1,\dots,N$.

Lemma 4.3: Assume that either $\mathcal{G}_{\sigma(t)}$ being UJSC or $\mathcal{G}_{\sigma(t)}$ being IJC with bidirectional communications. Then $\theta_i = \eta_i = d^*$ for all $i = 1,2,\dots,N$.

Remark 4.2: If the network communication graph is undirected, i.e., $i \in N_j(\sigma(t))$ if and only if $j \in N_i(\sigma(t))$ with $a_{ij}(x,t) \equiv a_{ji}(x,t), i,j = 1,\dots,N$, then according to (13) and (17), we have

$$\begin{aligned}\frac{d}{dt}\sum_{i=1}^N d_i(t) &\le 2\sum_{i=1}^N \sum_{j\in N_i(\sigma(t))} a_{ij}(x,t)\langle x_i - P_{X_0}(x_i), x_j - x_i\rangle \\ &\quad - 2\sum_{i=1}^N |x_i|^2_{X_i} \\ &= \sum_{i=1}^N \sum_{j\in N_i(\sigma(t))} a_{ij}(x,t)\langle x_i - P_{X_0}(x_i), x_j - x_i\rangle \\ &\quad + \sum_{j=1}^N \sum_{i\in N_j(\sigma(t))} a_{ji}(x,t)\langle x_j - P_{X_0}(x_j), x_i - x_j\rangle \\ &\quad - 2\sum_{i=1}^N |x_i|^2_{X_i} \\ &= \sum_{i=1}^N \sum_{j\in N_i(\sigma(t))} a_{ij}(x,t)\Big\langle x_i - x_j + \\ &\qquad P_{X_0}(x_j) - P_{X_0}(x_i), x_j - x_i\Big\rangle - 2\sum_{i=1}^N |x_i|^2_{X_i}.\end{aligned}$$

Furthermore, based on (1) and (2), we obtain

$$\begin{aligned}&\langle x_i - x_j + P_{X_0}(x_j) - P_{X_0}(x_i), x_j - x_i\rangle\\ &\le -|x_i - x_j|^2 + |x_i - x_j|\cdot|P_{X_0}(x_j) - P_{X_0}(x_i)| \le 0\end{aligned}$$

for all $i, j = 1, \dots, N$. Therefore, we have

$$\frac{d}{dt}\sum_{i=1}^{N} d_i(t) \le -2\sum_{i=1}^{N}|x_i(t)|^2_{X_i},$$

which implies

$$\sum_{i=1}^{N}\int_0^\infty |x_i(t)|^2_{X_i} < \frac{N}{2}d(t_0) \tag{20}$$

immediately based on Lemma 4.1.

As a result, with (20), we can apply Barbalat's lemma on $|x_i(t)|^2_{X_i}$, and then it follows immediately that $\lim_{t\to+\infty}|x_i(t)|_{X_i} = 0$, $i = 1, \dots, N$ without the assumptions of Lemma 4.2.

Remark 4.3: Note that, Lemmas 4.1 and 4.2 hold without requiring any connectivity of the system communication graph.

B. Directed Graphs

The following conclusion is for optimal set convergence with directed communications.

Proposition 4.1: System (10) achieves the global optimal solution set convergence if $\mathcal{G}_{\sigma(t)}$ is UJSC.

Proof. According to Lemmas 4.2 and 4.3, we have $\lim_{t\to\infty}|x_i(t)|_{X_i} = 0$, $i = 1, \dots, N$. As a result, for any $\varepsilon > 0$, there exists $T_1(\varepsilon) > 0$ such that when $t \ge T_1$,

$$|x_i(t)|_{X_i} \le \varepsilon, \quad i = 1, \dots, N. \tag{21}$$

Take $t_1 = T_1$ and $k_0 \in \mathcal{V}$. Defining

$$h_{k_0}(t) \doteq \max_{i\in\mathcal{V}}|x_i(t)|_{X_{k_0}},$$

similarly to the analysis of (14), we have that for all t,

$$\frac{d}{dt}h^2_{k_0}(t) \le 2h_{k_0}(t)\cdot\max_{i=1,\dots,N}|x_i(t)|_{X_i},$$

which implies $D^+h_{k_0}(t) \le \varepsilon, t \ge t_1$. Thus, $h_{k_0}(t) \le h_{k_0}(t_1) + (N-1)T_0\varepsilon, t \in [t_1, t_1 + (N-1)T_0]$.

Since $\mathcal{G}_{\sigma(t)}$ is UJSC, we can find a node k_1 such that $(k_0, k_1) \in \mathcal{E}_{\sigma(t)}$ for $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D) \subseteq [t_1, t_1 + T_0)$, where $T_0 = T + 2\tau_D$. In light of Lemma 2.2 and (21), we have

$$\begin{aligned}&\frac{d}{dt}|x_{k_1}(t)|^2_{X_{k_0}}\\ &= 2a_{k_1k_0}(x,t)\langle x_{k_1} - P_{X_{k_0}}(x_{k_1}), x_{k_0} - x_{k_1}\rangle\\ &\quad + 2\Big\langle x_{k_1} - P_{X_{k_0}}(x_{k_1}), \sum_{j\in N_{k_1}(\sigma(t))\setminus k_0} a_{k_1j}(x_j - x_{k_1})\\ &\qquad + P_{X_{k_1}}(x_{k_1}) - x_{k_1}\Big\rangle\\ &\le -2a_*|x_{k_1}(t)|_{X_{k_0}}\big(|x_{k_1}(t)|_{X_{k_0}} - \varepsilon\big)\\ &\quad + 2(N-2)a^*|x_{k_1}(t)|_{X_{k_0}}\cdot\Big(h_{k_0}(t_1) + (N-1)T_0\varepsilon\\ &\qquad - |x_{k_1}(t)|_{X_{k_0}}\Big) + 2|x_{k_1}(t)|_{X_{k_0}}\cdot\varepsilon, \quad t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D),\end{aligned} \tag{22}$$

from which we obtain that for any $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D)$,

$$\begin{aligned}&D^+|x_{k_1}(t)|_{X_{k_0}}\\ &\le -(a_* + (N-2)a^*)|x_{k_1}(t)|_{X_{k_0}} + (N-2)a^*\big[h_{k_0}(t_1)\\ &\quad + (N-1)T_0\varepsilon\big] + (1 + a_*)\varepsilon.\end{aligned}$$

Therefore, noticing that $|x_{k_1}(\tilde{t}_1)|_{X_{k_0}} \le h_{k_0}(t_1) + (N-1)T_0\varepsilon$ and denoting $\nu_0 = e^{-(a_*+(N-2)a^*)\tau_D}$, one has

$$\begin{aligned}&|x_{k_1}(\tilde{t}_1 + \tau_D)|_{X_{k_0}}\\ &\le \nu_0|x_{k_1}(\tilde{t}_1)|_{X_{k_0}} + (1 - \nu_0)\\ &\quad\times\frac{(N-2)a^*[h_{k_0}(t_1) + (N-1)T_0\varepsilon] + (1 + a_*)\varepsilon}{a_* + (N-2)a^*}\\ &\le w_0h_{k_0}(t_1) + M_0\varepsilon,\end{aligned} \tag{23}$$

where

$$w_0 = \frac{(N-2)a^* + a_*\nu_0}{(N-2)a^* + a_*} \tag{24}$$

and

$$M_0 = \frac{1 + [2(N-2)(N-1)T_0 + 1]a^*}{a_* + (N-2)a^*}. \tag{25}$$

Proceeding the estimation in time interval $[\tilde{t}_1 + \tau_D, t_1 + (N-1)T_0]$ will lead to

$$\begin{aligned}D^+|x_{k_1}(t)|_{X_{k_0}} &\le -(N-1)a^*|x_{k_1}(t)|_{X_{k_0}}\\ &\quad + (N-1)a^*[h_{k_0}(t_1) + (N-1)T_0\varepsilon] + \varepsilon,\end{aligned}$$

for all $t \in [\tilde{t}_1 + \tau_D, t_1 + (N-1)T_0]$. This implies

$$\begin{aligned}|x_{k_1}(t)|_{X_{k_0}} &\le \varsigma_0(w_0h_{k_0}(t_1) + M_0\varepsilon) + (1 - \varsigma_0)\Big(w_0h_{k_0}(t_1)\\ &\quad + (N-1)T_0\varepsilon + \frac{\varepsilon}{(N-1)a^*}\Big)\\ &\le w_0h_{k_0}(t_1) + \tilde{M}_0\varepsilon\end{aligned} \tag{26}$$

for $t \in [\tilde{t}_1 + \tau_D, t_1 + (N-1)T_0]$, where

$$\varsigma_0 = e^{-(N-1)^2a^*T_0}; \quad \tilde{M}_0 = \frac{2 + [3(N-1)^2T_0 + 1]a^*}{a_* + (N-2)a^*}. \tag{27}$$

Further, continuing the analysis on time interval $[t_1 + T_0, t_1 + 2T_0]$, k_2 can be found with a neighbor in $\{k_0, k_1\}$ during $[\tilde{t}_2, \tilde{t}_2 + \tau_D) \subseteq [t_1 + T_0, t_1 + 2T_0]$. An upper bound for $|x_{k_2}(t)|_{X_{k_0}}$ can be similarly obtained as

$$|x_{k_2}(t)|_{X_{k_0}} \le w_1h_{k_0}(t_1) + 2\tilde{M}_0\varepsilon, \quad t \in [\tilde{t}_2 + \tau_D, t_1 + (N-1)T_0] \tag{28}$$

where $w_1 = \frac{(N-2)a^* + a_*\nu_0^2}{(N-2)a^* + a_*}$.

Next, respectively, we repeat the analysis on time intervals $[t_1 + 2T_0, t_1 + 3T_0], \dots, [t_1 + (N-2)T_0, t_1 + (N-1)T_0]$ for $k_3, \dots, k_{N-1} \in \mathcal{V}$, and we finally reach

$$|x_i(t_1 + (N-1)T_0)|_{X_{k_0}} \le w_{N-1}h_{k_0}(t_1) + (N-1)\tilde{M}_0\varepsilon, \tag{29}$$

for $i = 1, \dots, N$, which implies

$$h_{k_0}(t_2) \le w_{N-1}h_{k_0}(t_1) + (N-1)\tilde{M}_0\varepsilon, \tag{30}$$

where $t_2 = t_1 + (N-1)T_0$ and $0 < w_{N-1} = \frac{(N-2)a^* + a_*\nu_0^N}{(N-2)a^* + a_*} < 1$.

Denoting $w_* = w_{N-1}$ and $t_{n+1} = t_n + (N-1)T_0$ for $n = 2, \dots,$ and by the same analysis on time intervals $[t_n, t_{n+1}], n = 2, \dots,$ one has

$$h_{k_0}(t_n) \le w_*^{n-1} h_{k_0}(t_1) + \sum_{j=1}^{n-1} w_*^{j-1}(N-1)\tilde{M}_0\varepsilon$$
$$\le w_*^{n-1} h_{k_0}(t_1) + \frac{(N-1)\tilde{M}_0}{1-w_*} \cdot \varepsilon \tag{31}$$

Since ε in (31) can be arbitrarily small, we see that $\lim_{t\to\infty} |x_i(t)|_{X_{k_0}} = 0$ for all $i, k_0 = 1, \dots, N$, which immediately implies $\lim_{t\to\infty} |x_i(t)|_{X_0} = 0$. The proof is completed. □

C. Bidirectional Graphs

The following conclusion is for optimal set convergence under bidirectional graphs.

Proposition 4.2: System (10) achieves the optimal solution set convergence with bidirectional communications if $\mathcal{G}_{\sigma(t)}$ is IJC.

Proof. Suppose $d^* > 0$. According to Lemmas 4.2 and 4.3, we have that for all $i = 1, \dots, N$,

$$\lim_{t\to\infty} |x_i(t)|_{X_0} = \sqrt{d^*}, \quad \lim_{t\to\infty} |x_i(t)|_{X_i} = 0. \tag{32}$$

This implies, for any $\varepsilon > 0$, we have that $x_i(t) \in \mathcal{B}_0(\varepsilon) \cap \mathcal{B}_i(\varepsilon)$ for sufficiently large t, where $\mathcal{B}_0(\varepsilon) \doteq \{y | \sqrt{d^* + \varepsilon} \le |y|_{X_0} \le \sqrt{d^* + \varepsilon}\}$ and $\mathcal{B}_i(\varepsilon) \doteq \{y | |y|_{X_i} \le \varepsilon\}, i = 1, \dots, N$. Then we see from (13) that the derivative of $|x_i(t)|^2_{X_0}$ is globally Lipschitz. Therefore, based on Barbalat's lemma, we know

$$\lim_{t\to\infty} \frac{d}{dt} |x_i(t)|^2_{X_0} = 0. \tag{33}$$

Define $\mathcal{E}_\infty \doteq \{(i,j) | (i,j) \in \mathcal{E}_{\sigma(t)}$ for infinitely long time$\}$. Then $\mathcal{G}_\infty = (\mathcal{V}, \mathcal{E}_\infty)$ is connected since $\mathcal{G}([t, +\infty))$ is connected for all $t \ge 0$.

Let N_i^∞ be the neighbor set of node i in graph $\mathcal{G}_\infty$. With Lemma 2.2, (32) and (33) yield that for any $i = 1, \dots, N$ and $j \in N_i^\infty$,

$$\lim_{t\to\infty} \langle x_i(t) - P_{X_0}(x_i(t)), x_j(t) - x_i(t) \rangle = 0. \tag{34}$$

Taking $i_0 \in \mathcal{V}$, we define two hyperplanes:

$$\mathcal{H}_1(t) \doteq \{v | \langle x_{i_0}(t) - P_{X_0}(x_{i_0}(t)), v - x_{i_0}(t) \rangle = 0\};$$
$$\mathcal{H}_2(t) \doteq \{v | \langle x_{i_0}(t) - P_{X_0}(x_{i_0}(t)), v - P_{X_0}(x_{i_0}(t)) \rangle = 0\}.$$

Then $\forall j \in N_{i_0}^\infty$, (34) implies that

$$\lim_{t\to\infty} |x_j(t)|_{\mathcal{H}_1(t)} = 0; \quad \lim_{t\to\infty} |x_j(t)|_{\mathcal{H}_2(t)} = \sqrt{g^*},$$

which leads to

$$\lim_{t\to\infty} |P_{X_0}(x_j(t)) - P_{\mathcal{H}_2(t)}(x_j(t))| = 0. \tag{35}$$

Because $\mathcal{G}_\infty$ is connected, we can repeat the analysis over the network, then arrive that (35) holds for all $j = 1, \dots, N$.

Let $\mathcal{C}_{i_0}(t) = co\{P_{X_{i_0}}(x_{i_0}(t)), P_{X_0}(x_1(t)), \dots, P_{X_0}(x_N(t))\}$. Then $\mathcal{C}_{i_0}(t) \subseteq X_{i_0}, \forall t \ge 0$.

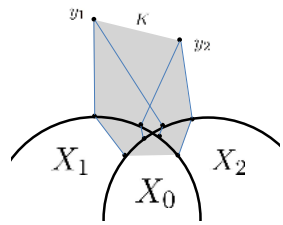

Fig. 2. Constructing an invariant set from $K = co\{y_1, y_2\}$.

Therefore, with (32) and (35) and according to the structure of $\mathcal{H}_1(t)$ and $\mathcal{H}_2(t)$, there will be a point $z_* \in \bigcap_{i_0=1}^N \mathcal{C}_{i_0}(t) \subseteq X_0$ for sufficiently large t such that

$$\langle x_{i_0}(t) - P_{X_0}(x_{i_0}(t)), z_* - P_{X_0}(x_{i_0}(t)) \rangle > 0,$$

which contradicts (1). Therefore, $d^* > 0$ does not hold, and then the optimal set convergence follows. □

V. Global Consensus

In this section, we present the consensus analysis. In order to show the consensus, we have to present a clear estimation of the influence on state agreement by terms $x_i - P_{X_i}(x_i), i = 1, \dots, N$.

We first introduce a class of positively invariant set for system (10) which characterizes the agreement property in Subsection 5.1. Then the consensus analysis is investigated for directed and bidirectional communication cases, respectively in Subsection 5.2.

A. Invariant Set

We define a multi-projection function: $P_{i_k i_{k-1} \dots i_1} : \mathbb{R}^m \to \bigcup_{i=1}^N X_i$ with $i_1, \dots, i_k \in \{1, \dots, N\}, k = 1, 2, \dots,$ by

$$P_{i_k i_{k-1} \dots i_1}(x) = P_{X_{i_k}} P_{X_{i_{k-1}}} \dots P_{X_{i_1}}(x).$$

Particularly, $P_\emptyset$ is denoted by $P_\emptyset(x) = x$ as the case for $k = 0$. Let

$$\Gamma \doteq \{P_{i_k i_{k-1} \dots i_1} | i_1, \dots, i_k \in \{1, \dots, N\}, k = 0, 1, 2, \dots\}$$

be the set which contains all the multi-projection functions we define.

Furthermore, let K be a convex set in $\mathbb{R}^m$, and define Δ_K as $\Delta_K \doteq co\{P(y) | y \in K, P \in \Gamma\}$. Denoting $\hat{g}(t) = \max_{i=1,\dots,N} |x_i(t)|^2_{\Delta_K}$, based on a similar analysis as the proof of Lemma 4.1, it is not hard to find that

$$D^+\hat{g}(t) \le 0, \quad t \ge 0.$$

This implies, $\hat{g}(t) \equiv 0$ for all $t \ge t_0$ once we have $\hat{g}(t_0) = 0$, which leads to the following conclusion immediately (see Fig. 2).

Lemma 5.1: Let K be a convex set in $\mathbb{R}^m$. Then $\Delta_K^N \doteq \Delta_K \times \dots \times \Delta_K$ is positively invariant for system (10).

We next establish an important property of the constructed invariant set Δ_K^N.

Lemma 5.2: $|y|_K \le 2\max_{z\in K} |z|_{X_0}, \ \forall y \in \Delta_K$.

Proof. With Lemma 2.1, any $y \in \Delta_K$ has the following form

$$y = \sum_{i=1}^{m+1} \lambda_i P^{\langle i \rangle}(z_i),$$

where $\sum_{i=1}^{m+1} \lambda_i = 1$ with $\lambda_i \geq 0$, $P^{\langle i \rangle} \in \Gamma$ and $z_i \in K$, $i = 1, \ldots, m+1$. Then, by the non-expansiveness property (2), we have that for any $z \in \mathbb{R}^m$ and $P_* \in \Gamma$,

$$\begin{aligned} |P_{X_0}(z) - P_*(z)| &= |P_*(P_{X_0}(z)) - P_*(z)| \\ &\leq |P_{X_0}(z) - z| \\ &= |z|_{X_0}. \end{aligned}$$

This leads to

$$\begin{aligned} &\Big| \sum_{i=1}^{m+1} \lambda_i P^{\langle i \rangle}(z_i) - \sum_{i=1}^{m+1} \lambda_i z_i \Big| \\ &\leq \sum_{i=1}^{m+1} \lambda_i \big| z_i - P^{\langle i \rangle}(z_i) \big| \\ &\leq \sum_{i=1}^{m+1} \lambda_i \big| z_i - P_{X_0}(z_i) \big| + \sum_{i=1}^{m+1} \lambda_i \big| P_{X_0}(z_i) - P^{\langle i \rangle}(z_i) \big| \\ &\leq 2 \max_{z \in K} |z|_{X_0}, \end{aligned}$$

which implies the conclusion because $\sum_{i=1}^{m+1} \lambda_i z_i \in K$. □

Now we are ready to reach the global consensus for system (10). Let us focus on each coordinate, and denote $x_i^\ell(t)$ as the ℓ-th coordinate of $x_i(t)$. Moreover, let

$$\phi(t) = \min_{i \in \mathcal{V}} \{x_i^\ell(t)\}, \quad \varphi(t) = \max_{i \in \mathcal{V}} \{x_i^\ell(t)\}$$

be the minimum and the maximum within all the agents. Denote $H(t) \triangleq \varphi(t) - \phi(t)$. Then a consensus is achieved for system (10) if and only if $\lim_{t \to \infty} H(t) = 0$.

In the next subsection, we will prove the global consensus for system (10) with directed and bidirectional communications, respectively by showing that $\lim_{t \to \infty} H(t) = 0$.

B. Consensus Analysis

In this subsection, we propose the consensus analysis. First we study the directed case.

Proposition 5.1: System (10) achieves a global consensus if $\mathcal{G}_{\sigma(t)}$ is UJSC.

Proof. Based on Proposition 4.1, we have that $\lim_{t \to \infty} |x_i(t)|_{X_0} = 0$ for all $i = 1, \ldots, N$. Therefore, for any $\varepsilon > 0$, there exists $T_1(\varepsilon) > 0$ such that, when $t \geq T_1$,

$$|x_i(t)|_{X_0} \leq \frac{1}{2}\varepsilon, \ i = 1, \ldots, N \tag{36}$$

As a result, according to Lemma 5.2, for any $y \in \Delta_{co\{x_1(t), \ldots, x_N(t)\}}$ with $t > T_1(\varepsilon)$, we have

$$\operatorname{dist}\big(y, co\{x_1(t), \ldots, x_N(t)\}\big) \leq \varepsilon.$$

Moreover, by Lemma 5.1, we see that $x_i(\hat{t}) \in \Delta_{co\{x_1(t), \ldots, x_N(t)\}}, i = 1, \ldots, N$ for all $t \leq \hat{t} \leq \infty$, which implies that for all $\hat{t} \geq t \geq T_1$, we have

$$\operatorname{dist}\big(x_i(\hat{t}), \Delta_{co\{x_1(t), \ldots, x_N(t)\}}\big) \leq \varepsilon, \quad i = 1, \ldots, N. \tag{37}$$

We divide the following proof into three steps.

Step 1: Take $t_1 = T_1$ with $x_{i_0}^\ell(t_1) = \phi(t_1)$ and denote $T_0 = T + 2\tau_D$. In this step, we give bound to $x_{i_0}^\ell(t)$ during $t \in [t_1, t_1 + (N-1)T_0]$.

Based on (37), we see that for all $T_1 \leq t < \hat{t} \leq \infty$

$$\phi(\hat{t}) \geq \phi(t) - \varepsilon; \ \varphi(\hat{t}) \leq \varphi(t) + \varepsilon. \tag{38}$$

Noting the fact that

$$\frac{d}{dt} x_{i_0}^\ell(t) \leq -(N-1)a^* x_{i_0}^\ell(t) + (N-1)a^*(\varphi(t_1) + \varepsilon) + \varepsilon, \tag{39}$$

for all $t \geq t_1$, we obtain

$$x_{i_0}^\ell(t) \leq \mu_1 \triangleq \varsigma_0 \phi(t_1) + (1 - \varsigma_0)\varphi(t_1) + \frac{(N-1)a^* + 1}{(N-1)a^*} \cdot \varepsilon, \tag{40}$$

where $t \in [t_1, t_1 + (N-1)T_0]$.

Step 2: Since $\mathcal{G}_{\sigma(t)}$ is UJSC, we can find $i_1 \in \mathcal{V}$ and $\tilde{t}_1 \geq t_1$ such that $(i_0, i_1) \in \mathcal{G}_{\sigma(t)}$ for $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D) \subseteq [t_1, t_1 + T_0)$. In this step, we give bound to $x_{i_1}^\ell(t_1)$ during $t \in [\tilde{t}_1 + \tau_D, t_1 + (N-1)T_0]$.

Similarly to the analysis of (22), when $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D)$, one has

$$\begin{aligned} &\frac{d}{dt} x_{i_1}^\ell(t) \\ &\leq a_*(\mu_1 - x_{i_1}^\ell(t)) + (N-2)a^*(\varphi(t_1) + \varepsilon - x_{i_1}^\ell(t)) + \varepsilon, \end{aligned}$$

which yields

$$\begin{aligned} &x_{i_1}^\ell(\tilde{t}_1 + \tau_D) \\ &\leq \nu_0(\varphi(t_1) + \varepsilon) + (1 - \nu_0) \\ &\quad \times \frac{a_*\mu_1 + (N-2)a^*(\varphi(t_1) + \varepsilon) + \varepsilon}{a_* + (N-2)a^*} \\ &= (1 - w_0)\varsigma_0 \phi(t_1) + [1 - (1 - w_0)\varsigma_0]\varphi(t_1) + \hat{L}_0 \varepsilon \\ &\triangleq \theta_1 \end{aligned} \tag{41}$$

after some simple manipulations by combining (40) and (41), where $\hat{L}_0 = 1 + \frac{N}{[a_* + (N-2)a^*](N-1)}$.

Then, applying (39) on node i_1 during $t \in [\tilde{t}_1 + \tau_D, t_1 + (N-1)T_0]$ will lead to

$$\begin{aligned} x_{i_1}^\ell(t) &\leq \varsigma_0 \theta_1 + (1 - \varsigma_0)\varphi(t_1) + \frac{(N-1)a^* + 1}{(N-1)a^*} \cdot \varepsilon \\ &= m_1 \phi(t_1) + [1 - m_1]\varphi(t_1) + L_0 \varepsilon, \end{aligned} \tag{42}$$

for all $t \in [\tilde{t}_1 + \tau_D, t_1 + NT_0]$, where $m_1 = (1 - w_0)\varsigma_0^2$ and $L_0 = \varsigma_0 \hat{L}_0 + 1 + \frac{1}{(N-1)a^*}$.

Step 3: We proceed the analysis for $i_2, \ldots, i_{N-1}$ with $m_k = ((1 - w_0)\varsigma_0^2)^k, k = 2, \ldots, N-1$. Denoting $t_2 \triangleq t_1 + (N-1)T_0$, we obtain

$$x_{i_\varrho}^\ell(t_2) \leq m_{N-1}\phi(t_1) + (1 - m_{N-1})\varphi(t_1) + (N-1)L_0 \varepsilon, \tag{43}$$

for $\varrho = 0, \ldots, N-1$, which implies

$$\varphi(t_2) \leq m_{N-1}\phi(t_1) + (1 - m_{N-1})\varphi(t_1) + (N-1)L_0 \varepsilon. \tag{44}$$

(38) and (44) lead to

$$\begin{aligned} H(t_2) &= \varphi(t_2) - \phi(t_2) \\ &\le m_{N-1}\phi(t_1) + (1-m_{N-1})\varphi(t_1) \\ &\quad + (N-1)L_0\varepsilon - (\phi(t_1) - \varepsilon) \\ &= (1-m_{N-1})H(t_1) + [(N-1)L_0+1]\varepsilon \end{aligned} \tag{45}$$

Define a time sequence $T_1 = t_1 < t_2 < \dots$ with $t_k = t_{k-1} + (N-1)T_0$. Applying the same analysis on each interval $[t_{k-1}, t_k)$ will lead to

$$H(t_k) \le (1-m_{N-1})H(t_{k-1}) + [(N-1)L_0+1]\varepsilon \tag{46}$$

for all $k = 1, 2, \dots$. As a result, we obtain

$$\begin{aligned} H(t_{k+1}) &\le (1-m_{N-1})^k H(t_1) \\ &\quad + \sum_{j=0}^{k-1}(1-m_{N-1})^j[(N-1)L_0+1]\varepsilon \\ &\le (1-m_{N-1})^k H(t_1) + \frac{(N-1)L_0+1}{m_{N-1}}\cdot\varepsilon, \end{aligned} \tag{47}$$

for all $k = 1, 2, \dots$. Therefore, noting the fact that $0 < m_{N-1} < 1$, (38) and (47) yield

$$\limsup_{t\to\infty} H(t) \le (2 + \frac{(N-1)L_0+1}{m_{N-1}})\cdot\varepsilon.$$

Then $\lim_{t\to\infty} H(t) = 0$ since ε can be arbitrarily small. This completes the proof. □

Then the global consensus for bidirectional case is proved by the following conclusion.

Proposition 5.2: System (10) achieves a global consensus with bidirectional communications if $\mathcal{G}_{\sigma(t)}$ is IJC.

Proof. Take $t_1 = T_1$ with $x^\ell_{i_0}(t_1) = \phi(t_1)$ as the proof of Proposition 5.1. Then (37) and (38) still hold.

Denote the first time when i_0 has at least one neighbor during $t \ge t_1$ as $\tilde{t}_1$, and denote the neighbor set of i_0 for $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D)$ as $\mathcal{V}_1$. Next, we show the bound for i_0 and $j \in \mathcal{V}_1$ during $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D)$.

Note that when i_0 has no neighbor during $t \in (t_1, s)$ for $t_1 \le s \le \infty$, one has that for any $t \in [t_1, s)$,

$$|x^\ell_{i_0}(t) - x^\ell_{i_0}(s)| \le \varepsilon. \tag{48}$$

Then, we see that

$$x^\ell_{i_0}(t) \le \hat{\mu}_1 \triangleq \hat{\varsigma}_0\phi(t_1) + (1-\hat{\varsigma}_0)\varphi(t_1) + \frac{(N-1)a^*+1}{(N-1)a^*}\cdot\varepsilon$$

for all $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D]$, where $\hat{\varsigma}_0 = e^{-(N-1)a^*\tau_D}$

By similar analysis with (41), we have that for any $j \in \mathcal{V}_1$,

$$x^\ell_j(\tilde{t}_1 + \tau_D) \le \hat{\theta}_1 \triangleq \hat{m}_1\phi(t_1) + (1-\hat{m}_1)\varphi(t_1) + \hat{L}_0\varepsilon \tag{49}$$

with $\hat{m}_1 = (1-w_0)\hat{\varsigma}_0$.

When there is no link between $\mathcal{V}\setminus(\{i_0\}\cup\mathcal{V}_1)$ and $\{i_0\}\cup\mathcal{V}_1$ for $t \in [\tilde{t}_1 + \tau_D, \breve{t})$, applying Lemma 5.1 on the subsystem formed by nodes in $\{i_0\}\cup\mathcal{V}_1$, (37) leads to

$$x^\ell_j(t) \le \hat{\theta}_1 + \varepsilon,\ \ t \in [\tilde{t}_1 + \tau_D, \breve{t}),\ j \in \{i_0\}\cup\mathcal{V}_1. \tag{50}$$

Therefore, defining $\tilde{t}_2$ as the first moment during $t \in [\tilde{t}_1 + \tau_D, \infty)$ when there is an edge between $j \in \{i_0\}\cup\mathcal{V}_1$ and $\mathcal{V}\setminus(\{i_0\}\cup\mathcal{V}_1)$, we have

$$x^\ell_j(t) \le \hat{\varsigma}_0(\hat{\theta}_1 + \varepsilon) + (1-\hat{\varsigma}_0)\varphi(t_1) + \frac{(N-1)a^*+1}{(N-1)a^*}\cdot\varepsilon \tag{51}$$

for $t \in [\tilde{t}_2, \tilde{t}_2 + \tau_D]$.

Denoting $\mathcal{V}_2 = \{k \in \mathcal{V} | \text{there is a link between } k \text{ and } \{i_0\}\cup\mathcal{V}_1 \text{ at } \tilde{t}_2\}$, bounds for $x^\ell_k(\tilde{t}_2 + \tau_D), k \in \mathcal{V}_2$ can be similarly given by

$$x^\ell_k(\tilde{t}_2 + \tau_D) \le \hat{m}_2\phi(t_1) + (1-\hat{m}_2)\varphi(t_1) + L_0\varepsilon, \tag{52}$$

where $\hat{m}_2 = ((1-w_0)\hat{\varsigma}_0^2)^2$.

Next, $\mathcal{V}_3, \dots, \mathcal{V}_{j_0}$ can be defined until $\mathcal{V} = \{i_0\}\cup\mathcal{V}_1\cup\dots\cup\mathcal{V}_{j_0}$ since $\mathcal{G}_{\sigma(t)}$ is IJC. Moreover, with $\hat{m}_{j_0} = ((1-w_0)\hat{\varsigma}_0^2)^{j_0}$, we have

$$x^\ell_{i_\varrho}(\tilde{t}_{j_0} + \tau_D) \le \hat{m}_{j_0}\phi(t_1) + (1-\hat{m}_{j_0})\varphi(t_1) + L_0(N-1)\varepsilon$$

for all $\varrho = 1, \dots, N$.

Therefore, denoting $t_2 \triangleq \tilde{t}_{j_0} + \tau_D$, we obtain

$$\varphi(t_2) \le \hat{m}_{j_0}\phi(t_1) + (1-\hat{m}_{j_0})\varphi(t_1) + L_0(N-1)\varepsilon, \tag{53}$$

which implies

$$H(t_2) \le (1-\hat{m}_{j_0})H(t_1) + (N-1)L_0\varepsilon. \tag{54}$$

Then similar to the analysis in the proof of Proposition 5.1, we have $\lim_{t\to\infty} H(t) = 0$. This completes the proof. □

With Propositions 4.1, 4.2, 5.1 and 5.2, it is straightforward to see that the main results of the paper, Theorems 3.1 and 3.2 hold.

VI. Conclusions

This paper addressed an optimal consensus problem for multi-agent systems. With jointly connected graphs, the considered multi-agent system achieved not only consensus, but also optimum by agreeing within the global solution set of a sum of objective functions. Assuming that each agent can observe the projection information onto the solution set of its own optimization component and the intersection of all solution sets is nonempty, the original unconstrained optimization problem was converted to an intersection computation problem. Control laws applied to the agents were simple and distributed. The results showed that a global optimization problem can be solved over a multi-agent network under time-varying communications and limited interactions. Future work includes randomization in the nodes' decision-making and event-based methods in the optimization algorithm design.

Appendix A
Proof of Lemma 4.2

Based on the definitions of θ_i and η_i, when $\theta_i = \eta_i = d^*$ holds for all $i = 1, \dots, N$, one has

$$\lim_{t\to+\infty} d_i(t) = d^*, \quad i = 1, \dots, N$$

Thus, for any $\varepsilon > 0$, there exists $T_1(\varepsilon) > 0$ such that, when $t \ge T_1(\varepsilon)$,

$$d_i(t) \in [d^* - \varepsilon, d^* + \varepsilon], \quad i = 1, \dots, N. \tag{55}$$

When $d^* = 0$, then it is easy to see that the conclusion holds because $|x_i(t)|_{X_i} \le |x_i(t)|_{X_0}$ for all $t \ge 0$. Therefore, we just assume $d^* > 0$ in the following.

According to (13) and (17), it is not hard to find that

$$\frac{d}{dt}d_i(t) \le -2|x_i|^2_{X_i} + 2\langle x_i - P_{X_0}(x_i), \sum_{j\in N_i(\sigma(t))} a_{ij}(x,t)(x_j - x_i)\rangle. \tag{56}$$

Furthermore, based on (55) and Lemmas 2.2 and 4.1, one has that when $t > T_1(\varepsilon)$,

$$\begin{aligned}\langle x_i - P_{X_0}(x_i), x_j - x_i\rangle &\le |x_i|_{X_0}\cdot\big|\,|x_i|_{X_0} - |x_j|_{X_0}\big| \\ &\le 2\sqrt{d(t_0)}\varepsilon\end{aligned} \tag{57}$$

for all $i = 1,\dots,N$ and $j \in N_i(\sigma(t))$.

If the conclusion does not hold, there exist a node i_0 and a constant $M_0 > 0$ such that

$$|x_{i_0}(t_k)|_{X_{i_0}} = M_0 \tag{58}$$

for a time serial $0 < t_1 < \cdots < t_k < t_{k+1} < \dots$ with $\lim_{k\to\infty} t_{k+1} = \infty$. Noting the fact that there is a constant $L > 0$ such that $|a - b| \le L$ for all $a, b \in \{y|\ |y|^2_{X_0} \le d(t_0)\}$ since X_0 is compact, we have that for all for all $i = 1,\dots,N$,

$$\begin{aligned}&\Big|\frac{d}{dt}|x_i(t)|^2_{X_i}\Big| \\ &= \Big|2\sum_{j\in N_i(\sigma(t))} a_{ij}\langle x_i - P_{X_i}(x_i), x_j - x_i\rangle - 2|x_i(t)|^2_{X_i}\Big| \\ &\le 2|x_i(t)|^2_{X_0} + 2(N-1)a^*|x_i(t)|_{X_0}\cdot\big|x_j(t) - x_i(t)\big| \\ &\le 2d(t_0) + 2(N-1)a^*\sqrt{d(t_0)}L.\end{aligned} \tag{59}$$

Denoting $\tau_0 \triangleq \frac{M_0}{2\sqrt{d(t_0)+(N-1)a^*\sqrt{d(t_0)}L}}$ and according to (58) and (59), we obtain

$$|x_{i_0}(t_k)|^2_{X_{i_0}} \ge \frac{1}{2}M_0^2, \quad t \in [t_k, t_k + \tau_0], \tag{60}$$

which leads to

$$\frac{d}{dt}d_{i_0}(t) \le -\frac{1}{2}M_0^2 + 2\sqrt{d(t_0)}\varepsilon \le -\frac{1}{4}M_0^2, \quad t \in [t_k, t_k + \tau_0]. \tag{61}$$

for all $t_k > T_1$ and $\varepsilon \le \frac{M_0^2}{8\sqrt{d(t_0)}}$. As a result, we have

$$d_{i_0}(t_k + \tau_0) \le d(t_0) - \frac{M_0^2\tau_0}{4} + \varepsilon \tag{62}$$

Therefore, (62) contradicts (55) when $\varepsilon < \frac{M_0^2\tau_0}{8}$, which completes the proof. □

Appendix B
Proof of Lemma 4.3

We prove the conclusion by contradiction. Suppose there exists a node $i_0 \in \mathcal{V}$ such that $0 \le \theta_{i_0} < \eta_{i_0} \le d^*$. Then for any $\varepsilon > 0$, there exists $T_1(\varepsilon) > 0$ such that, when $t \ge T_1(\varepsilon)$,

$$d_i(t) \in [0, d^* + \varepsilon], \quad i = 1,\dots,N. \tag{63}$$

Take $\zeta_0 = \sqrt{\frac{1}{2}(\theta_{i_0} + \eta_{i_0})}$. Then there exists a time serial

$$0 < \hat{t}_1 < \cdots < \hat{t}_k < \dots$$

with $\lim_{t\to\infty}\hat{t}_k = \infty$ such that $|x_{i_0}(\hat{t}_k)|_{X_0} = \zeta_0$ for all $k = 1, 2, \dots$.

According to (63) and Lemma 2.2, we have that for all $t > \hat{t}_{k_0}$,

$$\begin{aligned}\frac{d}{dt}d_{i_0}(t) &\le 2\sum_{j\in N_{i_0}(\sigma(t))} a_{i_0 j}(x,t)\langle x_{i_0} - P_{X_0}(x_{i_0}), x_j - x_{i_0}\rangle \\ &\le 2(N-1)a^*|x_{i_0}(t)|_{X_0}(\sqrt{d^*+\varepsilon} - |x_{i_0}(t)|_{X_0}),\end{aligned}$$

which will lead to

$$D^+|x_{i_0}(t)|_{X_0} \le -(N-1)a^*|x_{i_0}(t)|_{X_0} + (N-1)a^*\sqrt{d^*+\varepsilon}. \tag{64}$$

As a result, for $t \in [s, \infty)$ with $s \ge \hat{t}_{k_0}$, we have

$$\begin{aligned}|x_{i_0}(t)|_{X_0} &\le e^{-(N-1)a^*(t-s)}|x_{i_0}(s)|_{X_0} \\ &\quad + (1 - e^{(N-1)a^*(t-s)})\sqrt{d^*+\varepsilon}.\end{aligned} \tag{65}$$

We divide the following proof into two cases: directed communications and bidirectional communications.

Directed Case: Denote $T_0 = T + 2\tau_D$. Since $\mathcal{G}_{\sigma(t)}$ is UJSC, it is not hard to find that there exist $i_1 \in \mathcal{V}$ and $\tilde{t}_1$ such that $(i_0, i_1) \in \mathcal{G}_{\sigma(t)}$ for $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D) \subseteq [\hat{t}_{k_0}, \hat{t}_{k_0} + T_0)$. Then based on (65), we obtain

$$|x_{i_0}(t)|_{X_0} \le \xi_1 \triangleq \varsigma_0\zeta_0 + (1 - \varsigma_0)\sqrt{d^*+\varepsilon} \tag{66}$$

for $t \in [\hat{t}_{k_0}, \hat{t}_{k_0} + (N-1)T_0]$, where $\varsigma_0 = e^{-(N-1)^2a^*T_0}$. Thus, for $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D)$, one has

$$\begin{aligned}&\frac{d}{dt}d_{i_1}(t) \\ &\le 2\Big(\sum_{j\in N_{i_1}(\sigma(t))\setminus i_0} a_{i_1 j}\big\langle x_{i_1} - P_{X_0}(x_{i_1}), x_j - x_{i_1}\big\rangle \\ &\quad + a_{i_1 i_0}\big\langle x_{i_1} - P_{X_0}(x_{i_1}), x_{i_0} - x_{i_1}\big\rangle\Big) \\ &\le 2(N-2)a^*|x_{i_1}(t)|_{X_0}\Big(\sqrt{d^*+\varepsilon} - |x_{i_1}(t)|_{X_0}\Big) \\ &\quad - a_*|x_{i_1}(t)|_{X_0}\Big(|x_{i_1}(t)|_{X_0} - \xi_1\Big),\end{aligned} \tag{67}$$

which leads to

$$\begin{aligned}D^+|x_{i_1}(t)|_{X_0} &\le -\big((N-2)a^* + a_*\big)|x_{i_1}(t)|_{X_0} \\ &\quad + (N-2)a^*\sqrt{d^*+\varepsilon} + a_*\xi_1.\end{aligned} \tag{68}$$

Therefore, we obtain

$$\begin{aligned}|x_{i_1}(t)|_{X_0} &\le e^{-((N-2)a^*+a_*)(t-\tilde{t}_1)}|x_{i_1}(\tilde{t}_1)|_{X_0} \\ &\quad + (1 - e^{-((N-2)a^*+a_*)(t-\tilde{t}_1)})\cdot\frac{(N-2)a^*\sqrt{d^*+\varepsilon} + a_*\xi_1}{(N-2)a^* + a_*}\end{aligned}$$

for $t \in [\tilde{t}_1, \tilde{t}_1 + \tau_D)$, which implies

$$|x_{i_1}(\tilde{t}_1 + \tau_D)|_{X_0} \le \zeta_1 \triangleq w_0\sqrt{d^*+\varepsilon} + (1 - w_0)\xi_1, \tag{69}$$

where w_0 is defined in (24). Furthermore, applying the same analysis of (65) on node i_1, one has that when $t \in [\tilde{t}_1 + \tau_D, \infty)$,

$$\begin{aligned}|x_{i_1}(t)|_{X_0} &\le e^{-(N-1)a^*(t-(\tilde{t}_1+\tau_D))}\zeta_1 \\ &\quad + (1 - e^{-(N-1)a^*(t-(\tilde{t}_1+\tau_D))})\sqrt{d^*+\varepsilon},\end{aligned} \tag{70}$$

Combing (66), (69) and (70), we obtain

$$|x_{i_1}(t)|_{X_0} \leq m_1\zeta_0 + (1-m_1)\sqrt{d^*+\varepsilon}, \tag{71}$$

for all $t \in [\tilde{t}_1+\tau_D, \hat{t}_{k_0}+(N-1)T_0]$, where $m_1 = (1-w_0)\varsigma_0^2$. (71) also holds for i_0 since $0 < \varsigma_0 < m_1 < 1$.

We can proceed to find a node $i_2 \in \mathcal{V}$ such that there is an arc leaving from $\{i_0, i_1\}$ entering i_2 in $\mathcal{G}([\hat{t}_{k_0}+T_0, \hat{t}_{k_0}+2T_0))$ because $\mathcal{G}_{\sigma(t)}$ is uniformly jointly strongly connected. Meanwhile, similar analysis will result in estimations for agent i_2 with the form (71) by $m_2 = ((1-w_0)\varsigma_0^2)^2$.

Repeating similar analysis on time intervals $[\hat{t}_{k_0}+2T_0, \hat{t}_{k_0}+3T_0], \ldots, [\hat{t}_{k_0}+(N-2)T_0, \hat{t}_{k_0}+(N-1)T_0]$ respectively, and finally, by $m_{N-1} = ((1-w_0)\varsigma_0^2)^{N-1}$, we obtain

$$|x_i(\hat{t}_{k_0}+NT_0)|_{X_0} \leq m_{N-1}\zeta_0 + (1-m_{N-1})\sqrt{d^*+\varepsilon} \tag{72}$$

for all $i = 1, \ldots, N$, which yields

$$d(\hat{t}_{k_0}+NT_0) \leq m_{N-1}\zeta_0 + (1-m_{N-1})\sqrt{d^*+\varepsilon}. \tag{73}$$

Note that, (73) contradicts the definition of d^* since $m_{N-1}\zeta_0 + (1-m_{N-1})\sqrt{d^*+\varepsilon} < \sqrt{d^*}$ for sufficiently small ε. The conclusion holds.

Bidirectional Case: When i_0 has no neighbor for $t \in [\hat{t}_{k_0}, s]$, by (17) we see that

$$|x_{i_0}(t)|_{X_0} \leq |x_{i_0}(\hat{t}_{k_0})|_{X_0} = \zeta_0,\ t \in [\hat{t}_{k_0}, s]. \tag{74}$$

Denote the first moment when i_0 has at least one neighbor during $t \in [\hat{t}_{k_0}, \infty)$ as $\tilde{t}_1$, and denote the neighbor set of i_0 for $t \in [\tilde{t}_1, \tilde{t}_1+\tau_D)$ as $\mathcal{V}_1$. Then, by a similar analysis as (66), one has

$$|x_{i_0}(t)|_{X_0} \leq \hat{\xi}_1 \triangleq \hat{\varsigma}_0\zeta_0 + (1-\hat{\varsigma}_0)\sqrt{d^*+\varepsilon},\ \ t \in [\tilde{t}_1, \tilde{t}_1+\tau_D] \tag{75}$$

with $\hat{\varsigma}_0 = e^{-(N-1)a^*\tau_D}$. Thus, according to the same process by which we obtain (69), one also obtains

$$\begin{aligned} |x_{i_1}(\tilde{t}_1+\tau_D)|_{X_0} &\leq w_0\sqrt{d^*+\varepsilon} + (1-w_0)\hat{\xi}_1 \\ &= \hat{m}_1\zeta_0 + (1-\hat{m}_1)\sqrt{d^*+\varepsilon}, \end{aligned} \tag{76}$$

where $\hat{m}_1 = \hat{\varsigma}_0(1-w_0)$.

Similarly, we can define $\tilde{t}_2$ as the first moment when there is another node connected to $\{i_0\} \cup \mathcal{V}_1$ during $t \geq \tilde{t}_1+\tau_D$. Let $\mathcal{V}_2$ be the node set which connect to $\{i_0\} \cup \mathcal{V}_1$ at $\tilde{t}_2$. Since we have the dwell time for $\sigma(t)$, without loss of generality, we can always assume that all the links between $\{i_0\} \cup \mathcal{V}_1$ and $\mathcal{V}_2$ last for at least τ_D time starting from $\tilde{t}_2$. Moreover, similar estimations will lead to

$$|x_{i_2}(\tilde{t}_2+\tau_D)|_{X_0} \leq \hat{m}_2\zeta_0 + (1-\hat{m}_2)\sqrt{d^*+\varepsilon}$$

for all $i_2 \in \{i_0\} \cup \mathcal{V}_1 \cup \mathcal{V}_2$, where $\hat{m}_2 = (\hat{\varsigma}_0(1-w_0))^2$.

Furthermore, since $\mathcal{G}_{\sigma(t)}$ is IJC, we can always proceed the upper process until $\mathcal{V} = \{i_0\} \cup \mathcal{V}_1 \cup \cdots \cup \mathcal{V}_{j_0}$, and then we obtain

$$|x_i(\tilde{t}_{j_0}+\tau_D)|_{X_0} \leq \hat{m}_{j_0}\zeta_0 + (1-\hat{m}_{j_0})\sqrt{d^*+\varepsilon},$$

with $\hat{m}_{j_0} = (\hat{\varsigma}_0(1-w_0))^{j_0}$, which contradicts the definition of d^*. Then the conclusion holds for bidirectional case.

The proof is completed. □

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the anonymous Reviewers for their valuable suggestions.

REFERENCES

[1] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY: Cambridge University Press, 2004.
[2] J. Aubin and A. Cellina, *Differential Inclusions*. Berlin: Speringer-Verlag, 1984.
[3] R. T. Rockafellar, *Convex Analysis.* New Jersey: Princeton University Press, 1972.
[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] N. Aronszajn, "Theory of reproducing kernels," *Trans. Amer. Math. Soc.*, vol. 68, no. 3, pp. 337-404, 1950.
[7] L. G. Gubin, B. T. Polyak, and E. V. Raik, "The method of projections for finding the common point of convex sets," *U.S.S.R Comput. Math. Math. Phys.*, vol. 7, no. 6, pp. 1211-1228, 1967.
[8] F. Deutsch, "Rate of convergence of the method of alternating projections," in *Parametric Optimization and Approximation*, B. Brosowski and F. Deutsch, Eds. Basel, Switzerland: Birkhäuser, vol. 76, pp. 96-107, 1983.
[9] K. Arrow, L. Hurwicz and H. Uzawa. *Studies in linear and non-linear programming*. Stanford University Press, 1958.
[10] R. W. Brockett, "Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems," vol. 146, pp. 79-91, 1991.
[11] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Trans. Autom. Control*, vol. 22, no. 4, pp. 551-575, 1977.
[12] M. Shamos and D. Hoey, "Geometric intersection problems," in *17th IEEE Conf. Foundations of Computer Science (FOCS 1976)*, pp. 208-215, 1976.
[13] J. L. Bentley and T. Ottmann, "Algorithms for reporting and counting geometric intersections", *IEEE Trans. Computers*, vol. C-28, no. 9, pp. 643647, 1979.
[14] J. Danskin, "The theory of max-min, with applications," *SIAM J. Appl. Math.*, vol. 14, 641-664, 1966.
[15] D. Cheng, J. Wang, and X. Hu, "An extension of LaSalle's invariance principle and its applciation to multi-agents consensus," *IEEE Trans. Autom. Control*, vol. 53, no. 7, pp. 1765-1770, 2008.
[16] F. Clarke, Yu.S. Ledyaev, R. Stern, and P. Wolenski, *Nonsmooth Analysis and Control Theory*. Berlin: Speringer-Verlag, 1998.
[17] J. Cortés, "Discontinuous dynamical systems–a tutorial on solutions, nonsmooth analysis, and stability," *IEEE Control Systems Magazine*, vol. 28, no. 3, 36-73, 2008.
[18] S. Martinez, J. Cortés, and F. Bullo, "Motion coordination with distributed information," *IEEE Control Syst. Mag.*, vol. 27, no. 4, pp. 75-88, 2007.
[19] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401-420, 2006.
[20] M. Cao, D. A. Spielman and A. S. Morse, "A lower bound on convergence of a distributed network consensus algorithm," in *Proc. IEEE Conf. on Decision and Control*, pp. 2356-2361, 2005.
[21] R. Olfati-Saber and R. Murray, "Consensus problems in the networks of agents with switching topology and time delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520-1533, 2004.
[22] J. Fax and R. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Automatic Control*, vol. 49, no. 9, 1465-1476, 2004.
[23] H. G. Tanner, A. Jadbabaie, G. J. Pappas, "Flocking in fixed and switching networks," *IEEE Trans. Autom. Control*, vol. 52, no.5, pp. 863-868, 2007.
[24] 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.
[25] 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.
[26] F. Xiao and L. Wang, "State consensus for multi-agent systems with swtiching topologies and time-varying delays," *Int. J. Control*, vol. 79, no. 10, pp. 1277-1284, 2006.

[27] F. Xiao and L. Wang, "Asynchronous consensus in continuous-time multi-agent systems with switching topology and time-varying delays," *IEEE Trans. Autom. Control*, vol. 53, no. 8, pp. 1804-1816, 2008.
[28] Y. Hong, L. Gao, D. Cheng, and J. Hu, "Lyapuov-based approach to multiagent systems with switching jointly connected interconnection," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 943-948, 2007.
[29] W. Ren and R. Beard. *Distributed Consensus in Multi-vehicle Cooperative Control*, Springer-Verlag, London, 2008.
[30] 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.
[31] G. Shi and Y. Hong, "Global target aggregation and state agreement of nonlinear multi-agent systems with switching topologies," *Automatica*, vol. 45, no. 5, pp. 1165-1175, 2009.
[32] G. Shi, Y. Hong and K. H. Johansson, "Connectivity and set tracking of multi-agent systems guided by multiple moving leaders," *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 663-676, 2012.
[33] 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.
[34] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169-182, 2005.
[35] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "Distributed subgradient methods and quantization effects," in *Proc. IEEE Conf. on Decision and Control, Cancun, Mexico*, pp. 4177-4184, 2008.
[36] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48-61, 2009.
[37] 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.
[38] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in *IPSN'04*, pp. 20-27, 2004.
[39] B. Johansson, M. Rabi, and M. Johansson, "A simple peer-to-peer algorithm for distributed optimization in sensor networks, in *Proc. IEEE Conf. on Decision and Control, New Orleans, LA*, pp. 4705-4710, 2007.
[40] B. Johansson, T. Keviczky, M. Johansson, and K. H. Johansson, "Subgradient methods and consensus algorithms for solving convex optimization problems," in *Proc. IEEE Conf. on Decision and Control*, Cancun, Mexico, pp. 4185-4190, 2008.
[41] S. S. Ram, A. Nedić, and V. V. Veeravalli, "Stochastic incremental gradient descent for estimation in sensor networks," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, pp. 582-586, 2007.
[42] J. Lu, C. Y. Tang, P. R. Regier, and T. D. Bow, "Gossip algorithms for convex consensus optimization over networks," *IEEE Trans. Autom. Control*, vol. 56, pp. 2917-2923, 2011.
[43] S. S. Ram, A. Nedić, and V. V. Veeravalli, "Incremental stochastic subgradient algorithms for convex optimization," *SIAM J. on Optimization*, vol. 20, no. 2, 691-717, 2009.
[44] B. Johansson, A. Speranzon, M. Johansson, K. H. Johansson, "On decentralized negotiation of optimal consensus," *Automatica*, vol. 44, pp. 1175-1179, 2008.
[45] B. Johansson, M. Rabi and M. Johansson. "A randomized incremental subgradient method for distributed optimization in networked systems," *SIAM J. on Optimization*, vol. 20, no. 3, pp. 1157-1170, 2009.
[46] D. Jakovetić, J. Xavier and J. M. F. Moura, "Cooperative convex optimization in networked systems: augmented lagrangian algorithms with directed gossip communication," http://arxiv.org/abs/1007.3706, 2011.

Guodong Shi received his B.Sc. degree in Pure and Applied Mathematics from Shandong University, Jinan, China, in 2005, and Ph.D. degree in Systems Theory from Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in 2010, respectively. Currently he is a postdoctoral researcher at the ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Sweden. His research interests include distributed control, computation, and optimization methods.

Karl Henrik Johansson (SM'08) is Director of the KTH ACCESS Linnaeus Centre and Professor at the School of Electrical Engineering, Royal Institute of Technology, Sweden. He is a Wallenberg Scholar and has held a Senior Researcher Position with the Swedish Research Council. He received MSc and PhD degrees in Electrical Engineering from Lund University. He has held visiting positions at UC Berkeley (1998-2000) and California Institute of Technology (2006-2007). His research interests are in networked control systems, hybrid and embedded control, and control applications in automotive, automation and communication systems. He was a member of the IEEE Control Systems Society Board of Governors 2009 and the Chair of the IFAC Technical Committee on Networked Systems 2008-2011. He has been on the Editorial Boards of Automatica (2003-2006) and IEEE Transactions on Automatic Control (2008-2010), and is currently on the Editorial Boards of IET Control Theory and Applications and the International Journal of Robust and Nonlinear Control. He was the General Chair of the ACM/IEEE Cyber-Physical Systems Week (CPSWeek) 2010 in Stockholm. He has served on the Executive Committees of several European research projects in the area of networked embedded systems. He was awarded an Individual Grant for the Advancement of Research Leaders from the Swedish Foundation for Strategic Research in 2005. He received the triennial Young Author Prize from IFAC in 1996 and the Peccei Award from the International Institute of System Analysis, Austria, in 1993. He received Young Researcher Awards from Scania in 1996 and from Ericsson in 1998 and 1999.

Yiguang Hong (M'99-SM'02) received his B.Sc. and M.Sc. degrees from Peking University, China, and his Ph.D. degree from Chinese Academy of Sciences (CAS). He is currently a professor in Institute of Systems Science, Academy of Mathematics and Systems Science, CAS, and serves as the director of Key Lab of Systems and Control, CAS and the director of Information Technology Division of National Center for Mathematics and Interdisciplinary Sciences, CAS.

Prof. Hong is also serving as Deputy Editor-in-Chief of Acta Automatica Sinca, Editor-at-Large of Journal of Control Theory and Applications, and Associate Editors for several journals including IEEE Transactions on Automatic Control, Journal of Systems Science and Complexity, and Kybernetika. He is a recipient of Guangzhaozhi Award of Chinese Control Conference (in 1997), Young Author Prize of IFAC World Congress (in 1999), Young Scientist Award of CAS (in 2001), and Youth Award for Science and Technology of China (in 2006), and National Natural Science Prize of China (in 2008, jointly with D. Cheng, et al).

His research interests include nonlinear dynamics and control, multi-agent systems, robotics, and reliability of software and communication systems.