Distributed Continuous-time Approximate Projection Protocols for Shortest Distance Optimization Problems

Youcheng Lou, Yiguang Hong, Shouyang Wang

Abstract

In this paper, we investigate the distributed shortest distance optimization problem for a multi-agent network to cooperatively minimize the sum of the quadratic distances from some convex sets, where each set is only associated with one agent. To deal with the optimization problem with projection uncertainties, we propose a distributed continuous-time dynamical protocol based on a new concept of approximate projection. Here each agent can only obtain an approximate projection point on the boundary of its convex set, and communicate with its neighbors over a time-varying communication graph. First, we show that no matter how large the approximate angle is, the system states are always bounded for any initial condition, and uniformly bounded with respect to all initial conditions if the inferior limit of the stepsize is greater than zero. Then, in the two cases, nonempty intersection and empty intersection of convex sets, we provide stepsize and approximate angle conditions to ensure the optimal convergence, respectively. Moreover, we give some characterizations about the optimal solutions for the empty intersection case and also present the convergence error between agents’ estimates and the optimal point in the case of constant stepsizes and approximate angles.

Keywords: distributed optimization; convex intersection; shortest distance optimization; approximate projection

*This work is partially supported by the NNSF of China under Grant 71401163, 61333001, and by Beijing Natural Science Foundation 4152057.

†Y. Lou, Y. Hong and S. Wang are with Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. Email: louyoucheng@amss.ac.cn, yghong@iss.ac.cn, sywang@amss.ac.cn
1 Introduction

In recent years, distributed optimization of a sum of convex functions has attracted much attention due to its wide applications in resource allocation, source localization, and robust estimation (referring to [5, 18, 19, 20, 22, 27, 28, 29]). The whole optimization behavior can be achieved cooperatively by a group of autonomous agents via simple local information exchange and distributed protocol design even when the communication graph is time-varying.

Although many existing distributed optimization works have been done by discrete-time algorithms, more and more attention has been paid to continuous-time algorithms in recent years [32, 34, 35, 30, 31, 17], partially because the continuous-time models can be studied by various well-developed continuous-time methods or make the algorithms easily implemented in physical systems. A distributed continuous-time computation model was proposed to solve an optimization problem for fixed undirected graph in [30], with the optimization achieved by controlling the sum of subgradients of convex functions to make the state enter the optimal solution set, and later this model was generalized to the weight balanced graph case in [32], for differentiable objective functions with globally Lipschitz continuous gradient. Another continuous-time distributed algorithm with constant stepsize was developed in [34] for optimization problems with positivity constraints in the fixed undirected graph case, where a lower bound of convergence rate and an upper bound on the agents’ estimate error were presented. Moreover, the relationship between the existing dual decomposition and consensus-based methods for distributed optimization was revealed in [35], where both approaches were based on the subgradient method, but one with a proportional control term and the other with an integral control term.

When the optimal solution sets of agents’ individual objective functions have a nonempty intersection, the distributed optimization problem is equivalent to the convex intersection problems (CIP) [12, 11, 19, 16, 17, 15, 14]. A projected consensus algorithm was proposed in [19] for a network to solve the CIP, and the authors showed that all agents converged to a common point in the intersection set for weight-balanced and jointly connected communication graphs. Later, a continuous-time dynamical system was designed and connectivity conditions were discussed for the optimal convergence in [17]. In addition, a random sleep algorithm was proposed with providing conditions to converge almost surely to a common point in the intersection set in [14], where agents randomly took the neighbor-based average or projection onto their individual sets based on a Bernoulli process. Almost all the existing optimization results were obtained
based on the assumption that the exact projection point onto the convex sets can be obtained on the assumption that the exact projection point onto the convex sets can be obtained [12, 11, 19, 16, 17, 21, 26].

On the other hand, the intersection of the considered convex sets may be empty in practice. In this case, how to seek a point with the shortest (quadratic) distance to these sets is also important. For instance, the supply center location problem is concerned with how to seek the location of raw materials supply center so that the average transportation cost from the supply center to the multiple factories is minimal ([9, 10]); the source localization in a sensor network is related to estimate the location of the source emitting a signal based on the received signals of multiple sensors in a noisy environment ([26, 4]). In fact, the problem for both the empty and nonempty intersection case is referred to as the shortest distance optimization problem (SDOP). Obviously, CIP is a special case of SDOP, and the average consensus problem is also a special case of SDOP since the optimal solution of the minimum of the sum of quadratic functions from some points is exactly the average of these points. Some distributed algorithms were proposed to discuss SDOP. For example, [26] formulated the source localization problem as the SDOP in a plane and proposed a discrete-time distributed algorithm, with the adjacency matrices of communication graphs required to be doubly stochastic. Moreover, [21] proposed two distributed continuous-time algorithms to solve SDOP in the empty intersection case for connected graphs: the first one was designed for optimal consensus based on sign functions, and the second one was modified to avoid chattering but only to achieve the optimal consensus approximately.

The objective of this paper is to design a continuous-time distributed protocol to solve SDOP based on approximate projection. Note that the exact projection is usually hard to obtain in practice. Therefore, approximate projection may have to be discussed in different situations, and, in fact, [15] proposed a discrete-time approximate projected consensus algorithm to solve CIP. The motivation of the current research aims at analysis and distributed design to cooperatively solve SDOP with projection uncertainties and continuous-time dynamics. For example, in a practical robotic network to solve the SDOP, a continuous-time robot may not always obtain the exact projection point of its own convex set, but only spot some point on the set surface near the exact projection point. The contribution of this paper can be summarized as follows.

- We propose a new concept of approximate projection, which is related to some points on the convex set’s boundary surface and close to the exact projection point when the exact projection is hard to obtain. In other words, we consider an approximate projection related to set boundary surfaces, different from that defined in a “triangle” in [13]. To overcome the
analysis difficulties resulting from this new approximate projection, we employ a geometric method to convert the original problem to a heterogeneous stepsize problem.

- Although many results on discrete-time algorithms are similar to those on continuous-time ones, there are still fundamental differences between the two algorithms in some situations. In fact, π/4 was shown to be the critical approximate angle for the boundedness of the discrete-time algorithm given in [15]. With the proposed continuous-time protocol in this paper, we show that, for any approximate angle, the agent states are always bounded for any initial condition, and uniformly bounded with respect to all initial conditions when the stepsize is not too small.

- We study SDOP in both the nonempty and empty intersection cases, and propose a unified protocol based on the approximate projection. The proof in the study of the general empty intersection case is totally different from that in the nonempty intersection case. In fact, the proposed convergence conditions in the two cases are different. Note that our result is different from that in [21] because we handle approximate projections without assuming that the communication graph is always connected, and ours tackles both the nonempty and empty intersection cases, while [15] only does the nonempty intersection case. Moreover, we also discuss the convergence error between agents’ estimates and the optimal point in the case of constant stepsizes and approximate angles. Our results are certainly consistent with those discrete-time algorithms in the literature such as [18, 19] based on the exact projection.

The paper is organized as follows. Section 2 shows some basic concepts and preliminary results. Section 3 defines an approximate projection concept and formulates our shortest distance optimization problem (SDOP), followed by Section 4 for the discussions on boundedness and stepsizes. Section 5 presents the main convergence results for the nonempty intersection case, while Section 6 for the empty intersection case. Section 7 discusses the constant stepsize and approximate angle case. Then Section 8 provides numerical simulations. Finally, Section 9 gives some concluding remarks.

Notations: ⊗ denotes the Kronecker product; 1 denotes the vector with all ones; (A)_{ij} denotes the i-th row and j-th column entry of matrix A; y^T denotes the transpose of a vector \( y \in \mathbb{R}^m \); \( |y| \) denotes the Euclidean norm of \( y \); \([v, z]\) denotes the line segment connecting the two points \( v, z \); \( \text{line}(v, z) \) denotes the line passing the two points \( v, z \); for a set \( K \subseteq \mathbb{R}^m \), \( \text{ri}(K) \), \( \text{int}(K) \)
and \( \text{bd}(K) = K/\text{int}(K) \) denote the sets of relative interior points, interior points and boundary points of \( K \), respectively; for a closed convex set \( K \subseteq \mathbb{R}^m \), \( P_K(\cdot) \) denotes the projection operator onto \( K \); \( |y|_K := |y - P_K(y)| \) denotes the distance between \( y \) and \( K \); \( \langle \cdot, \cdot \rangle \) denotes the Euclidean inner product in \( \mathbb{R}^m \); the angle between nonzero vectors \( y \) and \( z \) is denoted as \( \angle(y, z) \in [0, \pi] \), where \( \cos \angle(y, z) = \langle y, z \rangle /(|y||z|) \); \( \text{span}\{v_1, ..., v_p\} \) (aff\{\(v_1, ..., v_p\}\}) denotes the subspace (affine hull) generated by vectors \( v_1, ..., v_p \).

### 2 Preliminaries

In this section, we give preliminaries on graph theory [1], convex analysis [2] and the consensus model with disturbances [25].

#### 2.1 Graph Theory

A multi-agent network can be described by a directed graph \( G = (V, E) \), where \( V = \{1, 2, ..., n\} \) is the node (or agent) set and \( E \subseteq V \times V \) the arc set with the arc \((j, i) \in E\) describing that node \( i \) can receive the information of node \( j \). Here \((i, i) \notin E\) for all \( i \). Let \( N_i = \{j \in V | (j, i) \in E\} \) be the set of neighbors of node \( i \). A path from node \( i \) to node \( j \) in \( G \) is a sequence \((i, i_1), (i_1, i_2), ..., (i_p, j)\) of arcs in \( E \). Graph \( G \) is said to be strongly connected if there exists a path from \( i \) to \( j \) for each pair of nodes \( i, j \in V \). Graph \( G \) is undirected when \((j, i) \in E\) if and only if \((i, j) \in E\).

The communication over the network considered here is switching and characterized by a directed graph process \( G_{\sigma(t)} = (V, E_{\sigma(t)}), t \geq 0 \), with \( E_{\sigma(t)} \) the arc set of the graph at time \( t \). Here \( \sigma : [0, \infty) \rightarrow Q \) is a piecewise constant function to describe the time-varying graph process, where \( Q \) is the index set of all possible graphs on \( V \). As usual, we assume there is a dwell time \( \tau \) between two consecutive graph switching moments. The switching graph \( G_{\sigma} \) is uniformly jointly strongly connected (UJSC) if there exists \( T > 0 \) such that the union graph \((V, \bigcup_{t \leq s \leq t+T} E(s))\) is strongly connected for \( t \geq 0 \).

#### 2.2 Convex Analysis

A set \( K \subseteq \mathbb{R}^m \) is convex if \( \lambda z_1 + (1 - \lambda)z_2 \in K \) for any \( z_1, z_2 \in K \) and \( 0 < \lambda < 1 \). For a closed convex set \( K \) in \( \mathbb{R}^m \), we can associate with any \( z \in \mathbb{R}^m \) a unique element \( P_K(z) \in K \) satisfying \( |z - P_K(z)| = \inf_{y \in K} |z - y| := |z|_K \), where \( P_K \) is called the projection operator onto \( K \). Then we have the following properties for the projection operator \( P_K \).
Lemma 2.1 Let $K$ be a closed convex set in $\mathbb{R}^m$. Then

\begin{enumerate}[(i)]
    \item $\langle y - P_K(y), z - P_K(y) \rangle \leq 0$ for any $y$ and $z \in K$;
    \item $|P_K(y) - z| \leq |y - z|$ for any $y \in \mathbb{R}^m$ and any $z \in K$;
    \item $\langle y - P_K(y), z - y \rangle \leq |y|_K (|z|_K - |y|_K)$ for any $y$ and $z$;
    \item $|P_K(y) - P_K(z)| \leq |y - z|$ for any $y$ and $z$.
\end{enumerate}

Proof. (i) is an equivalent definition of convex projection; (ii) comes from Lemma 1 (b) in [19]. We now show (iii). First of all, $\langle y - P_K(y), P_K(z) - P_K(y) \rangle \leq 0$ by (i). It is also clear that $\langle y - P_K(y), z - P_K(z) \rangle \leq |y|_K |z|_K$. Then

$$
\langle y - P_K(y), z - y \rangle = \langle y - P_K(y), z - P_K(z) + P_K(z) - P_K(y) + P_K(y) - y \rangle
\leq |y|_K |z|_K - |y|^2_K.
$$

Thus, the inequality (iii) follows. (iv) is the standard non-expansive property. \hfill \Box

The following lemma characterizes the distance between convex sets and their nonempty intersection, which can be found from Proposition 5.6.1 on page 72 in [6].

Lemma 2.2 Let $K_1, \ldots, K_n$ be closed convex sets in $\mathbb{R}^m$. If $\bigcap_{i=1}^n \text{ri}(K_i) \neq \emptyset$, then for every bounded set $S$, there exists $\kappa_S > 0$ such that

$$
|x|^2_{\bigcap_{i=1}^n K_i} \leq \kappa_S \max_{1 \leq i \leq n} |x|^2_{K_i}, \forall x \in S.
$$

The following lemma can be found from Proposition 1 on page 24 in [3].

Lemma 2.3 Let $K$ be a closed convex set in $\mathbb{R}^m$. Then $|x|^2_K$ is continuously differentiable and

$$
\nabla|x|^2_K = 2(x - P_K(x)).
$$

A function $\varphi(\cdot): \mathbb{R}^m \to \mathbb{R}$ is said to be convex if $\varphi(\lambda z_1 + (1 - \lambda)z_2) \leq \lambda \varphi(z_1) + (1 - \lambda)\varphi(z_2)$ for any $z_1, z_2 \in \mathbb{R}^m$ and $0 < \lambda < 1$, and it is $\ell$-strongly convex if $\varphi(\lambda z_1 + (1 - \lambda)z_2) \leq \lambda \varphi(z_1) + (1 - \lambda)\varphi(z_2) - \frac{1}{2} \ell (1 - \lambda)|z_1 - z_2|^2$ for any $z_1, z_2 \in \mathbb{R}^m$ and $0 < \lambda < 1$. The following two inequalities hold for a continuously differentiable convex and $\ell$-strongly convex function $\varphi$, respectively:

\begin{align}
\varphi(y) &\geq \varphi(x) + \langle y - x, \nabla \varphi(x) \rangle, \forall x, y \in \mathbb{R}^m, \quad (1) \\
\varphi(y) &\geq \varphi(x) + \langle y - x, \nabla \varphi(x) \rangle + \frac{\ell}{2}|y - x|^2, \forall x, y \in \mathbb{R}^m. \quad (2)
\end{align}
The upper Dini derivative of function \( g : (a, b) \to \mathbb{R} \) at \( t \in (a, b) \) is defined as
\[
D^+ g(t) = \limsup_{s \to 0^+} \frac{g(t + s) - g(t)}{s}.
\]
g is non-increasing on \((a, b)\) if \( D^+ g(t) \leq 0, \forall t \in (a, b) \). The following result was shown in [8].

**Lemma 2.4** Let \( g_i(t, x) : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R} \), \( i = 1, ..., j \) be continuously differentiable and \( g(t, x) = \max_{1 \leq i \leq j} g_i(t, x) \). Then \( D^+ g(t, x(t)) = \max_{i \in \mathcal{I}(t)} \dot{g}_i(t, x(t)) \) with \( \mathcal{I}(t) = \{ i \mid g_i(t, x(t)) = g(t, x(t)), 1 \leq i \leq j \} \).

### 2.3 Consensus

Consider the following consensus model with disturbance \( w_i \),
\[
\dot{z}_i(t) = \sum_{j \in \mathcal{N}_i(t)} (z_j(t) - z_i(t)) + w_i(t), \; i = 1, ..., n, \tag{3}
\]
where the disturbance \( w_i(t) : [0, \infty) \to \mathbb{R} \) is Lebesgue integrable and system \([3]\) has a unique Caratheodory solution, that is, there is a unique absolutely continuous function to satisfy \([3]\) for almost all \( t \). Consensus is said to be achieved for system \([3]\) if for any initial condition, \( \lim_{t \to \infty} |z_i(t) - z_j(t)| = 0 \) for all \( 1 \leq i, j \leq n \).

The next two lemmas can be obtained from the proofs of Theorem 4.2 and Proposition 4.10 in [25], respectively.

**Lemma 2.5** If the switching graph \( G_\sigma \) is UJSC for system \([3]\), then there exist \( 0 < \beta < 1 \) and \( B_0, B_1 > 0 \) such that
\[
H((k + 1)B_0) \leq \beta H(kB_0) + B_1 \int_{kB_0}^{(k+1)B_0} \max_{1 \leq i \leq n} |w_i(t)| dt, \; \forall k \geq 0,
\]
\[
H(t) \leq H(kB_0) + B_1 \int_{kB_0}^{(k+1)B_0} \max_{1 \leq i \leq n} |w_i(t)| dt, \; \forall t : kB_0 \leq t < (k+1)B_0,
\]
where \( H(t) = \max_{1 \leq i,j \leq n} |z_i(t) - z_j(t)| \).

**Lemma 2.6** Suppose the switching graph \( G_\sigma \) of system \([3]\) is UJSC and \( \lim_{t \to \infty} w_i(t) = 0 \) for all \( i \). Then consensus is achieved for system \([3]\).

### 3 Approximate Projection and Problem Formulation

In this section, we introduce the distributed SDOP and the distributed continuous-time approximate projected algorithm.
Consider a network of \( n \) agents (or nodes) and bounded closed convex sets \( X_i \subseteq \mathbb{R}^m \) for \( i = 1, \ldots, n \), with \( X_i \) only associated with (or known by) agent \( i \). The goal of the network is to cooperatively find a point \( x^* \) with the shortest quadratic distance from the \( n \) closed convex sets:

\[
x^* \in \arg\min_f(x), \quad f(x) = \sum_{i=1}^n |x|^2_{X_i}.
\]  

(4)

Projection-based methods have been widely adopted in the literature to solve CIP and constrained optimization problems, and almost all methods require that the exact projection can be obtained \([6, 16, 17, 19, 21, 22, 26]\). Since the exact projection may be difficult to obtain in practice, each agent may only obtain an approximate projection point located on the convex set surface and near the exact projection point. To be strict, we give the following definition.

**Definition 3.1** Let \( 0 < \theta < \pi/2 \) and \( K \) be a closed convex set in \( \mathbb{R}^m \). Define sets

\[
C_K(v, \theta) = v + \{ z | \langle z, P_K(v) - v \rangle \geq |z||v|_K \cos \theta \},
\]

\[
b(v, K) = \{ z | z \in bd(K), [v, z] \cap bd(K) = z \}.
\]

The approximate projection \( P_K^a(v, \theta) \) of point \( v \) onto \( K \) is defined as the following set:

\[
P_K^a(v, \theta) = \begin{cases} 
C_K(v, \theta) \cap b(v, K), & \text{if } v \notin K; \\
\{v\}, & \text{otherwise.}
\end{cases}
\]

**Figure 1**: The approximate projection of point \( v \) onto closed convex set \( K \).

As shown in Fig. 1, the cone \( C_K(v, \theta) - v \) consists of all vectors having angle with the direction \( P_K(v) - v \) less than \( \theta \), and \( b(v, K) \) is the region on the boundary of \( K \) that the agent can “see” starting from point \( v \). Obviously, the exact projection \( P_K(v) \in P_K^a(v, \theta) \) for any \( v \in \mathbb{R}^m \) and \( 0 < \theta < \pi/2 \) and \( P_K^a(v, 0) = \{P_K(v)\} \).
Remark 3.1 The approximate projection is more “practical” than the exact projection. For example, a robot likes to get its exact projection point on its convex target set when it approaches the set. However, in reality, it may select another point on the set surface as the exact one by mistake or to avoid expensive measurement or tedious computation. Then the selected projection point becomes an approximate one. In other words, this concept captures the situation when agents can only obtain some point on the set surface, which may not be but close to the exact projection point. Note that this concept is different from that given in [15], where the approximate projection point was located in a “triangle” region specified by \( v \), the hyperplane of \( K \) on \( P_K(v) \) with \( v - P_K(v) \) as the normal direction and the approximate angle \( \theta \).

We next give some basic assumptions for our following analysis.

**A1 (Connectivity)** The switching graph \( G_\sigma \) is UJSC.

**A2 (Surface)** The boundary surfaces of sets \( X_i, i = 1, \ldots, n \) are regular (or smooth).

The definition of regularity or smoothness of a manifold can be found in the literature (see Definition 1 on page 52 in [7] for more details), which is beyond the scope of this paper. Note that the Gaussian curvature of regular (or smooth) surfaces of closed bounded sets are bounded, from Definition 6 on page 146 in [7]. In fact, **A2** is quite mild. The boundaries of many well-known sets, such as the surfaces of spheres, ellipsoids and toruses, are regular.

Let \( P^q_{X_i}(\cdot) : \mathbb{R}^m \to \mathbb{R}^m \) be a continuous map with \( P^q_{X_i}(v) \in P^q_{X_i}(v, \theta_i(v)) \) for any \( v \), where \( \theta_i(v) = \angle(P^q_{X_i}(v) - v, P_{X_i}(v) - v), 0 \leq \theta_i(v) < \pi/2 \). Let \( \theta_i(v) = 0 \) for simplicity when \( v \in X_i \).

In this paper, \( \theta_i(v) \) is referred to as the approximate angle of \( v \) onto \( X_i \).

**A3 (Approximate Angle)** (i) There exists \( 0 < \theta^* < \pi/2 \) such that \( 0 \leq \theta_i(v) \leq \theta^* \) for all \( i, v \);

(ii) \( \theta_i(v_1) \leq \theta_i(v_2) \) for all \( v_1, v_2 \) if \( P_{X_i}(v_1) = P_{X_i}(v_2) \) and \( |v_1|_{X_i} \leq |v_2|_{X_i} \).

**A3 (i)** was used in [15], and **A3 (ii)** is reasonable because it means that agents can obtain their approximate projection points with higher accuracies when agents are closer to their sets.

Here we propose a distributed continuous-time approximate projected algorithm:

\[
\dot{x}_i(t) = \sum_{j \in N_i(t)} (x_j(t) - x_i(t)) + \alpha_i(P^q_{X_i}(x_i(t)) - x_i(t)), \quad i = 1, \ldots, n,
\]

where \( x_i \in \mathbb{R}^m \) is the state estimate of agent \( i \) for the optimal solution, \( N_i(t) \) is the neighbor set of node \( i \) at time \( t \), \( \{\alpha_i\} \) is the stepsize \( (0 \leq \alpha_i \leq \alpha^*, \alpha^* > 0) \) and is uniformly continuous over \( t \). The continuity of maps \( P^q_{X_i}(\cdot) \) guarantees that the solution of the proposed continuous-time dynamics is well-defined.
Remark 3.2 The term \( P^a_{X_i}(x) - x \) can be viewed as a negative “approximate” gradient because it becomes the negative gradient of \( \frac{1}{2} |x|^2_{X_i} \) by noting that \( \nabla |x|^2_{X_i} = 2(x - P_{X_i}(x)) \) in the exact projection case (i.e., \( P^a_{X_i}(x_i(t)) = P_{X_i}(x_i(t)) \)). In fact, (5) with taking \( \alpha_t \equiv 1 \) and exact projection was proposed in [17] to solve the CIP.

The convergence analysis of (5) is not easy because the gradient term is corrupted with the state-dependent approximation and there is no explicit expression to describe the relationship between the approximate projection point and the exact one. To handle the problem, we make some transformation. Define by \( P^h_{X_i}(v) \) the intersection point of the hyperplane of \( X_i \) at \( P_{X_i}(v) \) with \( P_{X_i}(v) - v \) as the normal direction and the line segment \([v, P^a_{X_i}(v)]\), as shown in Fig. 2.

Clearly, \( P^h_{X_i}(v) = P_{X_i}(v) \) when \( P^a_{X_i}(v) = P_{X_i}(v) \). Then we write

\[
P^a_{X_i}(v) - v = \gamma_{X_i}(v)(P^h_{X_i}(v) - v),
\]

where \( \gamma_{X_i}(v) = \frac{|P^a_{X_i}(v) - v|}{|P^h_{X_i}(v) - v|} \geq 1 \) with \( \gamma_{X_i}(v) = 1 \) if \( P^a_{X_i}(v) = P_{X_i}(v) \) (= \( v \) or not).

![Figure 2: An illustrative figure for \( P^h_{X_i}(v) \).](image)

Rewrite \( \alpha_t(P^a_{X_i}(x_i(t)) - x_i(t)) = \alpha_{i,t}(P^h_{X_i}(x_i(t)) - x_i(t)) \), with the virtual stepsize of agent \( i \) defined as

\[
\alpha_{i,t} = \begin{cases} 
\gamma_{X_i}(x_i(t))\alpha_t = \frac{|P^a_{X_i}(x_i(t)) - x_i(t)|}{|P^h_{X_i}(x_i(t)) - x_i(t)|}\alpha_t, & \text{if } P^h_{X_i}(x_i(t)) \neq x_i(t); \\
\alpha_t, & \text{otherwise}.
\end{cases}
\]

Obviously, \( \alpha_{i,t} \geq \alpha_t \). Although the designed stepsize \( \alpha_t \) is the same for all agents, each agent \( i \) has its own virtual stepsize \( \alpha_{i,t} \) based on its own approximate projection. Because \( P^h_{X_i}(x_i(t)) = x_i(t) \) if and only if \( x_i(t) \in X_i \), we can express (6) in another form with heterogeneous virtual stepsizes:

\[
\dot{x}_i(t) = \sum_{j \in N_i(t)} (x_j(t) - x_i(t)) + \alpha_{i,t}(P^h_{X_i}(x_i(t)) - x_i(t)), \quad i = 1, ..., n.
\]

(6)

Then we give a definition for our problem as follows.
Definition 3.2 The shortest distance optimization problem (SDOP) is solved by (5) or (6) if, for any initial condition \( x_i(0) \in \mathbb{R}^m, i = 1, ..., n \), there exists \( x^* \in \arg\min \sum_{i=1}^n |x_i|^2 \) such that
\[
\lim_{t \to \infty} x_i(t) = x^*, \quad i = 1, ..., n.
\]

In the following three sections, we first establish some basic results, and then present the convergence results in the nonempty intersection and empty intersection cases.

4 Discussions on Boundedness and Stepsizes

In this section, we show the state boundedness and establish an “equivalent” relationship between the designed stepsize \( \alpha_t \) and the virtual stepsize \( \alpha_{i,t} \).

4.1 Boundedness of System States

Denote \( \theta_{i,t} = \theta_i(x_i(t)) \) for simplicity. Note that \( \theta_{i,t} \) is also equal to \( \angle(P_{X_i}(x_i(t)) - x_i(t), P_{X_i}(x_i(t)) - x_i(t)) \). Here we study the boundedness of \( x_i(t), i \in \mathcal{V}, t \geq 0 \) of (6) with the approximate angle \( \theta_{i,t} \). Particularly, an essential difference between our continuous-time algorithm and a similar discrete-time one given in [15] is found on the existence of critical approximate angle.

Let \( X_{co} = \text{co}\{X_i, i = 1, ..., n\} \) be the convex hull of the sets \( X_i, i = 1, ..., n \), \( d_0 = \sup_{\omega_1, \omega_2 \in X_{co}} |\omega_1 - \omega_2| \), which is finite due to the boundedness of \( X_i \)s.

Theorem 4.1 Under \( A3 \) (i), we have
(i) for any initial condition \( x_i(0), i \in \mathcal{V}, \) the system states \( x_i(t), i \in \mathcal{V}, t \geq 0 \) are bounded;
(ii) if \( \lim \inf_{t \to \infty} \alpha_t > 0 \), then, for any initial condition \( x_i(0), i \in \mathcal{V}, \)
\[
\lim \sup_{t \to \infty} |x_i(t)|_{X_{co}} \leq (\tan \theta^* + \sqrt{(\tan \theta^*)^2 + 2 \tan \theta^*})d_0, \quad \forall i,
\]
where \( \theta^* \) is the angle upper bound given in \( A3 \) (i).

Remark 4.1 A discrete-time algorithm was proposed in [15] to solve CIP with approximate projection, where \( \pi/4 \) was found to be a critical approximate angle ensuring the boundedness of system states in the case \( \alpha_{i,k} = 1 \) and \( \theta_{i,k} = \theta, 0 \leq \theta < \pi/2 \). To be specific, the states are uniformly bounded with respect to all initial conditions when \( \theta < \pi/4 \) and unbounded for most all initial conditions when \( \theta > \pi/4 \). However, Theorem 4.1 shows that the continuous-time system
Moreover, recalling the definitions of $P_X$ according to Lemma 2.1 (i), the states are always bounded for any initial condition no matter how large $\theta$ is, and moreover, the states are uniformly bounded for all initial conditions with fixing $\alpha_{i,t} \equiv 1$.

**Remark 4.2** Notice that the boundedness results in Theorem 4.1 do not require any connectivity of communication graph. Moreover, when the exact projection is obtained ($\theta_{i,t} = 0$), Theorem 4.1 (ii) implies that all agents converge to the convex hull spanned by all the convex sets, which is related to the target aggregation and leader-following problems [24, 13, 23].

Here we present a sharp result for a simple case: there is only one node in the network. Its set is bounded and denoted as $X_\ast$. We denote the states of this node as $x_\ast(t), t \geq 0$ driven by the continuous-time approximate projected dynamical system:

$$\dot{x}_\ast(t) = \alpha_t(P_{X_\ast}^a(x_\ast(t)) - x_\ast(t)), $$

where $P_{X_\ast}^a(x_\ast(t)) \in P_{X_\ast}^a(x_\ast(t), \theta_\ast(x_\ast(t)))$. Noticing that $\langle x_\ast(t) - P_{X_\ast}(x_\ast(t)), P_{X_\ast}^a(x_\ast(t)) - P_{X_\ast}(x_\ast(t)) \rangle \leq 0$, we have

$$\frac{d|x_\ast(t)|_{X_\ast}^2}{dt} = 2\langle x_\ast(t) - P_{X_\ast}(x_\ast(t)), \dot{x}_\ast(t) \rangle \leq -2\alpha_t|x_\ast(t)|_{X_\ast}^2.$$

Then for any initial condition $x_\ast(0)$ and any stepsize $\{\alpha_t\}$, $|x_\ast(t)|_{X_\ast} \leq |x_\ast(0)|_{X_\ast}$ always holds.

Now we present the proof of Theorem 4.1.

**Proof.** Denote $h_i(t) = \frac{1}{2}|x_i(t)|_{X_{\ast,0}}^2$ and $h(t) = \max_{1 \leq i \leq n} h_i(t)$. By Lemmas 2.3, 2.4 we have

$$D^+ h(t) = \max_{i \in I(t)} \langle x_i(t) - P_{X_{\ast,0}}(x_i(t)), \dot{x}_i(t) \rangle$$

$$= \max_{i \in I(t)} \langle x_i(t) - P_{X_{\ast,0}}(x_i(t)), \sum_{j \in N_i(t)} (x_j(t) - x_i(t)) + \alpha_{i,t}(P_{X_\ast}^b(x_i(t)) - x_i(t)) \rangle$$

(7)

with $I(t) = \{i|i \in V, h_j(t) = h(t)\}$. Take $i \in I(t)$. Lemma 2.1 (iii) implies that, for any $j$,

$$\langle x_i - P_{X_{\ast,0}}(x_i), x_j - x_i \rangle \leq |x_i|_{X_0}(|x_j|_{X_{\ast,0}} - |x_i|_{X_{\ast,0}}) \leq 0.$$  \hspace{1cm} \text{ (8)}

According to Lemma 2.1 (i), $\langle x_i - P_{X_{\ast,0}}(x_i), P_{X_\ast}(x_i) - P_{X_{\ast,0}}(x_i) \rangle \leq 0$ due to $X_\ast \subseteq X_{\ast,0}$. Therefore,

$$\langle x_i - P_{X_{\ast,0}}(x_i), P_{X_\ast}(x_i) - x_i \rangle \leq -|x_i|_{X_{\ast,0}}.$$  \hspace{1cm} \text{ (9)}

Moreover, recalling the definitions of $P_{X_\ast}^b(x_i(t))$ and $\theta_{i,t}$, we have $\langle x_i(t) - P_{X_\ast}(x_i(t)), P_{X_\ast}^b(x_i(t)) -$
bounded by each other, it suffices to establish the boundedness of $\alpha$. To obtain the convergence conditions, we establish a relationship between the designed stepsize $\alpha$ and virtual stepsizes $\alpha_t$. Thus, the conclusion follows. □

where the last inequality follows from A3 (i) and $|x_i|_{X_{\infty}} \leq |x_i - P_{X_{\infty}}(x_i)| + |P_{X_{\infty}}(x_i) - P_{X_t}(x_i)| \leq |x_i|_{X_{\infty}} + d_0$. Thus, based on (9) and (10), we have

$$
\langle x_i(t) - P_{X_{\infty}}(x_i(t)), P^h_{X_t}(x_i(t)) - x_i(t) \rangle \leq -|x_i(t)|^2_{X_{\infty}} + d_0 \tan \theta^*(|x_i(t)|_{X_{\infty}} + d_0). \quad (11)
$$

With (7), (8), (11) and $i \in I(t)$, we obtain

$$
D^+ h(t) \leq \alpha_{i,t} \left(-2 \bar{h}(t) + d_0 \tan \theta^* \left(\sqrt{2 \bar{h}(t)} + d_0\right)\right), \quad t \geq 0, \; i \in I(t). \quad (12)
$$

We complete the proof by the following analysis:

(i) It is not hard to find that $D^+ h(t) \leq 0$ if

$$
\bar{h}(t) \geq \frac{d_0^2(\tan \theta^*\tan \theta^*)^2}{4} + \frac{d_0^2 \tan \theta^*}{2} \left(1 + \frac{\sqrt{(\tan \theta^*)^2 + 4 \tan \theta^*}}{2}\right).
$$

Hence, the system states are bounded for any initial conditions.

(ii) Let $\alpha_* = (\liminf_{t \to \infty} \alpha_t)/2 > 0$ and $t_0$ be the moment such that when $t \geq t_0$, $\alpha_{i,t} \geq \alpha_t \geq \alpha_*$. Then $D^+ h(t) \leq -\alpha_* h(t)$ once

$$
\bar{h}(t) \geq d_0^2(\tan \theta^*)^2 + d_0^2 \tan \theta^* \left(1 + \sqrt{\tan \theta^* + 4 \tan \theta^*}\right), \; t \geq t_0. \quad (13)
$$

Therefore, $h(t)$ is not greater than the number in (13) when $t \geq t_0$.

Thus, the conclusion follows.

4.2 Equivalence between Stepsizes

To obtain the convergence conditions, we establish a relationship between the designed stepsize $\alpha_t$ and virtual stepsizes $\alpha_{i,t}$. To show that they are equivalent in the sense that they can be bounded by each other, it suffices to establish the boundedness of $\gamma_{X_t}(\cdot)$.  

13
Let $S = \bigcup_{i=1}^{n} X_i + B(0, r_0)$ with sufficiently large $r_0 > 0$. Without loss of generality, in this subsection we assume that $X_i$ has nonempty interior points. Note that a convex set $X_i \subseteq \mathbb{R}^m$ has nonempty interior points is equivalent to $\dim(X_i) = m$. In fact, when $\dim(X_i) \leq m - 1$, from the regularity of the boundary of $X_i$ we can similarly show that $\sup_{v \in S} \gamma_{X_i}(v) < \infty$. Namely, $\sup_{v \in S} \frac{|P_{X_i}^a(v) - v|}{|P_{X_i}^a(v) - v|} < \infty$.

Denote $\mu_i(v) = \angle(P_{X_i}(v) - P_{X_i}^a(v), v - P_{X_i}(v))$ and $\text{cone}(v, M) = \{v + \lambda(z - v) | \lambda \geq 0, z \in M\}, M \subseteq \mathbb{R}^m$.

**Lemma 4.1** If the map $P_{X_i}(\cdot)$ satisfies

$$\inf_{v \in S \setminus X_i, P_{X_i}^a(v) \neq P_{X_i}(v)} \mu_i(v) > 0, \quad (14)$$

then $\sup_{v \in S} \gamma_{X_i}(v) < \infty$. Furthermore, (14) holds if $A3$ (ii) holds and there exists $\hat{\theta} > 0$ such that for any $v \in \text{bd}(S)$,

$$\text{cone}(v, C_{X_i}(v, \hat{\theta})) \cap \text{b}(v, X_i)) = C_{X_i}(v, \hat{\theta}). \quad (15)$$

Its proof is in the Appendix. The following result provides a condition to guarantee the condition (15).

**Lemma 4.2** Suppose $A2$ holds. Then there exists $\hat{\theta} > 0$ such that (15) holds for any $v \in \text{bd}(S)$.

The proof is also in the Appendix. Because the states of (6) are bounded by Theorem 4.1 we take sufficiently large $r_0$ such that $S = \bigcup_{i=1}^{n} X_i + B(0, r_0)$ contains all the system states. We next show that $A2$ and $A3$ (ii) imply the equivalence between $\{\alpha_t\}$ and $\{\alpha_{i,t}\}$.

Clearly, $\alpha_t = \alpha_{i,t}$ when $P_{X_i}(x_i(t)) = P_{X_i}^a(x_i(t))$. Without loss of generality, we assume $P_{X_i}(x_i(t)) \neq P_{X_i}^a(x_i(t))$ in the sequel of this subsection. By (50) in the Appendix, $\gamma_{X_i}(x_i(t)) \leq 1 + \frac{1}{\sin \mu_{i,t}} \sin \theta_{i,t}$, where $\mu_{i,t} := \mu_i(x_i(t))$. Then we have

**Theorem 4.2** Under $A2$ and $A3$ (ii),

$$\alpha_t \leq \alpha_{i,t} \leq C_{i,t} \alpha_t \leq C_i \alpha_t, \forall i, t, \quad (16)$$

where $C_{i,t} = 1 + \frac{1}{\sin \mu_{i,t}} \sin \theta_{i,t}$, $C_i = 1 + \frac{1}{\sin \mu_i}$,

$$\mu_i := \inf_{t \geq 0} \mu_{i,t} \geq \inf_{v \in S \setminus X_i, P_{X_i}^a(v) \neq P_{X_i}(v)} \mu_i(v) > 0. \quad (17)$$
Note that the first inequality of (17) follows from $x_i(t) \in S$ and the second one from Lemmas 4.1 and 4.2. In fact, (16) somehow characterizes the bounded bending property of smooth surfaces, which helps convert the convergence conditions on $\alpha_{i,t}$ to the conditions on $\alpha_t$.

**Remark 4.3** As Theorems 4.1 and 4.2 show, A3 (i) implies the boundedness of system states, while A2 and A3 (ii) guarantee the equivalence between the designed stepsize and the virtual stepsize. In fact, with (16), we found that under A1 and A3 (i), the optimal convergence established in the next two sections hold for general convex sets (not necessary to satisfy A2) and general continuous approximate projected maps (not necessary to satisfy A3 (ii)).

### 5 Nonempty Intersection Case

In this section, we show the convergence result in the nonempty intersection case, $\bigcap_{i=1}^{n} X_i \neq \emptyset$.

Clearly, $X_0 := \bigcap_{i=1}^{n} X_i$ is the optimal solution set of $\min \sum_{i=1}^{n} |x|^2_{X_i}$.

**Theorem 5.1** Suppose A1-A3 hold and $\bigcap_{i=1}^{n} X_i \neq \emptyset$. Then SDOP is solved by system (6) if $\int_0^\infty \alpha_t dt = \infty$, $\int_0^\infty \alpha_t \tan \theta^+_i dt < \infty$. Furthermore, in the special case when $\theta_{i,t} = 0$ $\forall i, t$, SDOP is solved by (6) if and only if $\int_0^\infty \alpha_t dt = \infty$.

**Remark 5.1** When the intersection set of all $X_i$s is nonempty, SDOP (4) is equivalent to CIP of finding a point in $X_0$ [11, 12, 14, 15, 16, 17, 19, 26]. The optimal consensus algorithm based on the exact projection presented in [17] is a special case of (5) with taking $\alpha_t \equiv 1$ and $\theta_{i,t} \equiv 0$, which is consistent with Theorem 5.1. Theorem 5.1 is also consistent with the convex intersection computation results of discrete-time algorithms in [13, 14, 20].

Theorem 5.1 provides sufficient conditions to guarantee the optimal consensus and Theorem 4.1 shows the boundedness no matter how large the approximate angle is. However, the following example shows that the approximate angle somehow plays a key role for the optimal consensus, which may not be achieved for some approximate projection sequence if the approximate angle condition in Theorem 5.1 fails.

**Example 5.1** Consider a network consisting of two nodes 1, 2 in $\mathbb{R}^2$. The communication graph is fixed with two arcs (1, 2) and (2, 1). Here $X_1 = \{(z,0)^T | -1 \leq z \leq 0\}$, $X_2 = \{(0,z)^T | 0 \leq z \leq 1\}$ and $X_0 = \{(0,0)^T\}$. Let $\alpha_t \equiv 1$ and $\theta_{1,t} = \theta_{2,t} \equiv \frac{\pi}{4}$. Clearly, the first stepsize...
condition in Theorem 5.1 holds but the second one does not. The initial condition is \( x_1(0) = (-\frac{2}{3}, \frac{1}{3})^T \), \( x_2(0) = (-\frac{1}{3}, \frac{2}{3})^T \) and the approximate projection point sequence takes \( P_{X_1}^a(x_1(t)) = P_{X_1}^a(x_1(t)) = (1, 0)^T \), \( P_{X_2}^h(x_2(t)) = P_{X_2}^h(x_2(t)) = (0, 1)^T \). It is easy to verify that the two nodes’ states are time-invariant, i.e., \( x_i(t) \equiv x_i(0), i = 1, 2 \), and thus, the optimal consensus cannot be achieved.

![Figure 3: Approximate angle plays a key role for the optimal convergence.](image)

Denote \( \alpha_1^+ = \max_{1 \leq i \leq n} \alpha_{i,t}, \theta_1^+ = \max_{1 \leq i \leq n} \theta_{i,t} \), and the distance functions

\[
h(t) = \max_{1 \leq i \leq n} h_i(t), \quad h_i(t) = \frac{1}{2} |x_i(t)|^2_{X_0}, \quad i = 1, ..., n, \quad t \geq 0.
\]

**Lemma 5.1** Suppose \( \cap_{i=1}^n X_i \neq \emptyset \). Then \( D^+ h(t) \leq 2\alpha_1^+ \tan \theta_1^+ h(t) \).

**Proof.** Similar to (7), we have

\[
D^+ h(t) = \max_{i \in \mathcal{I}(t)} \left\langle x_i(t) - P_{X_0}^a(x_i(t)), \sum_{j \in \mathcal{V}_i(t)} (x_j(t) - x_i(t)) + \alpha_{i,t}(P_{X_i}^h(x_i(t)) - x_i(t)) \right\rangle, \tag{18}
\]

where \( \mathcal{I}(t) = \{ j | j \in \mathcal{V}, h_j(t) = h(t) \} \). Take \( i \in \mathcal{I}(t) \). Similar to (3), we also have

\[
\left\langle x_i(t) - P_{X_0}(x_i(t)), x_j(t) - x_i(t) \right\rangle \leq |x_i(t)|_{X_0} \left( |x_j(t)|_{X_0} - |x_i(t)|_{X_0} \right) \leq 0. \tag{19}
\]

From \( \left\langle x_i(t) - P_{X_i}(x_i(t)), P_{X_i}^h(x_i(t)) - P_{X_i}(x_i(t)) \right\rangle = 0 \), we have

\[
\left\langle x_i(t) - P_{X_0}(x_i(t)), P_{X_i}^h(x_i(t)) - P_{X_i}(x_i(t)) \right\rangle = \left\langle P_{X_i}(x_i(t)) - P_{X_0}(x_i(t)), P_{X_i}^h(x_i(t)) - P_{X_i}(x_i(t)) \right\rangle \leq |x_i(t)|_{X_0} \tan \theta_{i,t} |x_i(t)|_{X_i} \leq \tan \theta_{i,t} |x_i(t)|_{X_0}^2, \tag{20}
\]
where the inequalities follow from Lemma 2.1 (ii) by setting \( K = X_i, y = x_i(t), z = P_{X_0}(x_i(t)) \in X_i \), and \(|x_i(t)|_{X_i} \leq |x_i(t)|_{X_0}\) (due to \( X_0 \subseteq X_i \)). Moreover, it follows from Lemma 2.1 (i) that
\[
\langle P_{X_i}(x_i(t)) - P_{X_0}(x_i(t)), P_{X_i}(x_i(t)) - x_i(t) \rangle \leq 0
\]
and then
\[
\langle x_i(t) - P_{X_0}(x_i(t)), P_{X_i}(x_i(t)) - x_i(t) \rangle
\]
\[
= -|x_i(t)|_{X_i}^2 + \langle P_{X_i}(x_i(t)) - P_{X_0}(x_i(t)), P_{X_i}(x_i(t)) - x_i(t) \rangle
\]
\[
\leq -|x_i(t)|_{X_i}^2.
\]
Therefore, based on (20) and (21) we have
\[
\langle x_i(t) - P_{X_0}(x_i(t)), P_{X_i}(x_i(t)) - x_i(t) \rangle = \langle x_i(t) - P_{X_0}(x_i(t)), P_{X_i}(x_i(t)) - x_i(t) \rangle
\]
\[
+ \langle x_i(t) - P_{X_0}(x_i(t)), P_{X_i}(x_i(t)) - P_{X_i}(x_i(t)) \rangle
\]
\[
\leq -|x_i(t)|_{X_i}^2 + \tan \theta_i |x_i(t)|_{X_0}^2
\]
\[
\leq \tan \theta_i |x_i(t)|_{X_0}^2.
\]

From (18), (19) and (23), \( D^+ h(t) \leq 2\alpha_i \tan \theta_i h(t) \leq 2\alpha_i \tan \theta^+_i h(t) \). Thus, the conclusion follows.

\[\square\]

**Lemma 5.2** If \( \bigcap_{i=1}^n X_i \neq \emptyset \) and \( \int_0^\infty \alpha_i^+ \tan \theta_i^+ dt < \infty \), then \( \lim_{t \to \infty} h(t) \) is a finite number.

**Proof.** Lemma 5.1 implies \( h(t) \leq h(0) + e^{2 \int_0^\infty \alpha_i^+ \tan \theta_i^+ ds} h(0) \) and then \( \bar{h} = \sup_{t \geq 0} h(t) \) is a finite number. We then show the conclusion by contradiction. Hence suppose there are two limit points \( \bar{h}_1 \neq \bar{h}_2 \) of \( \{ h(t) \}_{t \geq 0} \) such that \( \lim_{k \to \infty} h(s_k^1) = \bar{h}_1 \) and \( \lim_{k \to \infty} h(s_k^2) = \bar{h}_2 \). Without loss of generality, we assume \( \bar{h}_1 < \bar{h}_2 \). Clearly, for any \( \varepsilon > 0 \) for which \( (1 + \varepsilon) (\bar{h}_1 + \varepsilon) \leq \frac{\bar{h}_1 + \bar{h}_2}{2} \), there is an integer \( T_0 > 0 \) such that \( e^{2 \int_{s_k^1}^\infty \alpha_i^+ \tan \theta_i^+ ds} \leq 1 + \varepsilon \) and \( |h(s_k^1) - \bar{h}_1| \leq \varepsilon \) for \( k \geq T_0 \). According to Lemma 5.1, for any \( t \geq s_k^1 \) with \( k \geq T_0 \),
\[
h(t) \leq e^{2 \int_{s_k^1}^t \alpha_i^+ \tan \theta_i^+ ds} h(s_k^1) \leq (1 + \varepsilon)(\bar{h}_1 + \varepsilon) \leq \frac{\bar{h}_1 + \bar{h}_2}{2} < \bar{h}_2,
\]
which contradicts that \( \bar{h}_2 \) is also a limit point of \( \{ h(t) \}_{t \geq 0} \). Thus, the conclusion follows. \[\square\]

By Lemma 5.2 if \( \int_0^\infty \alpha_i^+ \tan \theta_i^+ dt < \infty \), then the sequence \( \{ h(t) \}_{t \geq 0} \) converges to a finite number denoted as \( h^* \),
\[
\lim_{t \to \infty} h(t) = h^*.
\]
Denote \( h_i^+ = \limsup_{t \to \infty} h_i(t), h_i^- = \liminf_{t \to \infty} h_i(t), i \in \mathbb{V} \). Clearly, \( 0 \leq h_i^- \leq h_i^+ \leq h^* \) for all \( i \).
Lemma 5.3 Suppose A1 holds and $\bigcap_{i=1}^{n} X_i \neq \emptyset$. If $\int_0^{\infty} \alpha_i^+ \tan \theta_i^+ dt < \infty$ and there exists some node $i_0 \in V$ such that $h_{i_0}^- < h^*$, then $h^* = 0$.

The proof of Lemma 5.3 can be completed with similar arguments in Lemma 4.3 in [17], which is omitted here.

Now it is time to prove Theorem 5.1.

**Proof of Theorem 5.1.** Based on the similar arguments in Lemmas 5.1 and 5.2, we can show that, for any $z \in X_0$, the limit $\lim_{t \to \infty} \max_{1 \leq i \leq n} |x_i(t) - z|^2$ exists. Therefore, if consensus is achieved and $h^* = 0$, all agents will converge to a common point in $X_0$. Thus, it suffices to show that consensus is achieved and $h^* = 0$.

Because

$$|P_i^{\alpha}(x_i(t)) - x_i(t)| = |x_i(t)|_{X_i} / \cos \theta_i, t \leq |x_i(t)|_{X_i} / \cos \theta^* \leq \sqrt{2h(t) / \cos \theta^*},$$

(24)

it follows that, if $h^* = 0$, the second term on the right-hand side of (23) tends to zero as $t \to +\infty$ and then the consensus is achieved for system (6) by Lemma 2.6. Therefore, it suffices to show $h^* = 0$ in what follows.

In fact, if there is some node $i_0$ with $h_{i_0}^- < h^*$, then $h^* = 0$ by Lemma 5.3. Therefore, we need to prove $h^* = 0$ from $h_i^+ = h_i^- = h^*, \forall i$ by contradiction. Clearly, for any $\varepsilon > 0$, there is $T_1 > 0$ such that when $t \geq T_1$, $|x_i(t)|_{X_0} \leq \sqrt{2h^*} + \varepsilon := \phi$. We complete the proof by the following two steps.

Step (i). Suppose $h_i^+ = h_i^- = h^* > 0, \forall i$. We claim that consensus can be achieved for system (6).

We first show that $\limsup_{t \to \infty} \alpha_{i,t} |x_i(t)|_{X_i}^2 = 0$ by contradiction. Suppose there exist $i_0$ and an increasing time subsequence $\{s_k\}$ such that $\alpha_{i_0,s_k} |x_i(s_k)|_{X_{i_0}}^2 \geq c$ for some $c > 0$. Without loss of generality, we assume $s_0$ is sufficiently large such that $s_0 \geq T_1$ and $\int_{s_0}^{\infty} \alpha_i^+ \tan \theta_i^+ dt \leq \varepsilon / \sqrt{2h^*}$.

Because $\alpha_i$ is uniformly continuous, so is $\alpha_{i_0} |x_{i_0}(t)|_{X_{i_0}}^2$. Therefore, there is $\delta > 0$ such that $\alpha_{i_0,t} |x_{i_0}(t)|_{X_{i_0}}^2 \geq \alpha_{i_0} |x_{i_0}(t)|_{X_{i_0}}^2 \geq c/2$ when $s_k \leq t \leq s_k + \delta$. Recalling (19) and (22), we have

$$\frac{dh_{i_0}(t)}{dt} \leq \sum_{j \in N_{i_0}(t)} |x_{i_0}(t) - x_j(t)|_{X_{i_0}} + \alpha_{i_0,t} |x_{i_0}(t)|_{X_{i_0}}$$
and then for \( s_k \leq t \leq s_k + \delta \),

\[
D^+|x_{i_0}(t)|x_0 \leq \sum_{j \in N_{i_0}(t)} \left( |x_j(t)|x_0 - |x_{i_0}(t)|x_0 \right) - \frac{\alpha_{i_0,t}|x_{i_0}(t)|^2}{|x_{i_0}(t)|x_0} + \alpha_{i_0,t} \tan \theta_{i_0,t} |x_{i_0}(t)|x_0
\]

\[
\leq \sum_{j \in N_{i_0}(t)} \left( |x_j(t)|x_0 - |x_{i_0}(t)|x_0 \right) - \frac{\alpha_{i_0,t}|x_{i_0}(t)|^2}{\phi} + \alpha_{i_0,t} \tan \theta_{i}^+ \phi
\]

\[
\leq (n-1)(\phi - |x_{i_0}(t)|x_0) - \frac{c}{2\phi} + \alpha_{i_0,t}^+ \tan \theta_{i}^+ \phi,
\]

which leads to

\[
|x_{i_0}(t)|x_0 \leq e^{-(n-1)(t-s_k)} |x_{i_0}(s_k)|x_0 + (1 - e^{-(n-1)(t-s_k)}) \left( \phi - \frac{c}{2(n-1)\phi} \right) + \phi \int_{s_k}^t e^{-(n-1)(t-s)} \alpha_{i}^+ \tan \theta_{i}^+ ds
\]

and then

\[
|x_{i_0}(s_k + \delta)|x_0 \leq \xi(\sqrt{2h^*} + \varepsilon) + \frac{4(1-\xi)}{4(n-1)\sqrt{2h^*}} + \frac{c}{2(n-1)\phi} + \sqrt{\frac{\varepsilon}{2h^*}},
\]

where \( 0 < \xi = e^{-(n-1)\delta} < 1 \). We can find that right-hand side of (27) is less than \( \sqrt{2h^*} - \frac{4(1-\xi)}{4(n-1)\sqrt{2h^*}} + \frac{c}{2(n-1)\phi} + \sqrt{\frac{\varepsilon}{2h^*}} \), which contradicts \( \lim_{t \to \infty} h_{i_0}(t) = h^* \). Thus, \( \lim_{t \to \infty} \alpha_{i_0,t} x_{i_0}(t) |X_i| = 0, \forall i \). From Theorem 4.2 we have \( 0 \leq \alpha_{i,t} \leq C_i \alpha^* \), and hence \( \lim_{t \to \infty} \alpha_{i,t} |x_i(t)|x_i = 0, \forall i \).

According to (24) and Lemma 2.6, consensus is achieved for system (6).

Step (ii). Suppose \( h_{i}^+ = h_{i}^- = h^* > 0, \forall i \). We show \( \lim_{t \to \infty} \sum_{i=1}^n |x_i(t)|^2_{X_i} = 0 \) by contradiction.

Hence suppose there is \( c > 0 \) such that \( \sum_{i=1}^n |x_i(t)|^2_{X_i} \geq c \) for all sufficiently large \( t \). Let

\[|x(t)|x_0 = (|x_1(t)|x_0, \ldots, |x_n(t)|x_0)^T, y(t) = (|x_1(t)|x_1^2, \ldots, |x_n(t)|x_n^2)^T, D(t) = \text{diag}\{\alpha_{1,t}, \ldots, \alpha_{n,t}\}.\]

Then by (25) we have

\[
D^+|x(t)|x_0 \leq -\mathcal{L}_{\sigma(t)}|x(t)|x_0 - \frac{1}{\phi} D(t)y(t) + \phi \alpha_t^+ \tan \theta_t^+ 1,
\]

where \( \mathcal{L}_{\sigma(t)} \) is the Laplacian of graph \( G_{\sigma(t)} \) with \( (\mathcal{L}_{\sigma(t)})_{ij} = -1 \) if \( j \in N_i(t) \) and \( (\mathcal{L}_{\sigma(t)})_{ij} = 0 \) otherwise. Let \( \{t_k\}_{k \geq 0} \) with \( t_0 = 0 \) be all the switching moments of switching graph \( G_{\sigma} \).

According to the dwell time assumption, \( t_{k+1} - t_k \geq \tau, \forall k \). It is easy to see that we can add some new “switching moments” in \( \{t_k\}_{k \geq 0} \), denoted as \( \{t'_k\}_{k \geq 0} \) such that \( 2\tau \geq t'_{k+1} - t'_k \geq \tau, \forall k \).

From (28) we have

\[
|x(t_{k+1}^{'})|x_0 \leq e^{-\mathcal{L}_{\sigma(t'_k)}(t'_{k+1}-t'_k)}|x(t'_k)|x_0 + \int_{t'_k}^{t'_{k+1}} e^{-\mathcal{L}_{\sigma(t'_k)}(t'_{k+1}-t)} \left( -\frac{1}{\phi} D(t)y(t) + \phi \alpha_t^+ \tan \theta_t^+ 1 \right) dt.
\]
Notice that for any \( s > 0 \), \( e^{-\mathcal{L}_{\sigma(t_k)}s} \) is a stochastic matrix and the graph \( \mathcal{G}_{\sigma(t_k)} \) is a subgraph of the graph associated with matrix \( e^{-\mathcal{L}_{\sigma(t_k)}s} \). Then applying the similar arguments given in the proof of Theorem 4.1 in [15] we can show that \( \lim_{t \to \infty} \sum_{i=1}^{n} |x_i(t)|^2_{X_i} = 0 \).

Then there is a time subsequence \( \{s_k\}_{k \geq 0} \) with \( s_k \to \infty \) such that \( \lim_{k \to \infty} |x_i(s_k)|_{X_i} = 0 \) for all \( i \). Because we have shown that consensus is achieved in Step (i), \( \lim_{k \to \infty} |x_i(s_k)|_{X_j} = 0 \) for all \( i, j \), which leads to \( \lim_{k \to \infty} h_i(s_k) = 0 \) for all \( i \). Thus, \( h^* = \lim_{t \to \infty} h_i(t) = 0 \), which contradicts the condition \( h^+_i = h^-_i = h^* > 0 \). It follows that \( h^+_i = h^-_i = h^* = 0 \) and then the first part is proved because, from (10), \( \int_0^\infty \alpha_t^+ dt = \infty \) and \( \int_0^\infty \alpha_t^+ \tan \theta_t^+ dt < \infty \) are equivalent to \( \int_0^\infty \alpha_t dt = \infty \) and \( \int_0^\infty \alpha_t \tan \theta_t^+ dt < \infty \), respectively.

Remark 6.1 In the case with the exact projection point (i.e., \( \theta_{i,t} = 0 \)), the stepsize conditions in Theorem 6.1 become \( \int_0^\infty \alpha_t dt = \infty \) and \( \int_0^\infty \alpha_t^2 dt < \infty \), which is a continuous-time version of the discrete-time stochastic approximation stepsize condition \( \sum_{k=0}^{\infty} \alpha_k = \infty \) and \( \sum_{k=0}^{\infty} \alpha_k^2 < \infty \).

6 Empty Intersection Case

In this section, we discuss the convergence in the empty intersection case (i.e., \( \bigcap_{i=1}^{n} X_i = \emptyset \)) and then show some properties of the optimal solution set in the following two subsections.

6.1 Convergence Analysis

The following is the convergence result for the case when \( \bigcap_{i=1}^{n} X_i = \emptyset \).

Theorem 6.1 Suppose A1-A3 hold, \( \mathcal{G}_{\sigma(t)}, t \geq 0 \) are undirected and \( \bigcap_{i=1}^{n} X_i = \emptyset \). Then SDOP is solved by system (6) if \( \int_0^\infty \alpha_t dt = \infty \), \( \int_0^\infty \alpha_t^2 dt < \infty \) and \( \int_0^\infty \alpha_t \tan \theta_t^+ dt < \infty \); Furthermore, if \( \theta_{i,t} = 0 \) for all \( i, t \), then it is necessary that \( \lim_{t \to \infty} \alpha_t = 0 \) for (6) to solve SDOP.

Remark 6.1 In the case with the exact projection point (i.e., \( \theta_{i,t} = 0 \)), the stepsize conditions in Theorem 6.1 become \( \int_0^\infty \alpha_t dt = \infty \) and \( \int_0^\infty \alpha_t^2 dt < \infty \), which is a continuous-time version of the discrete-time stochastic approximation stepsize condition \( \sum_{k=0}^{\infty} \alpha_k = \infty \) and \( \sum_{k=0}^{\infty} \alpha_k^2 < \infty \).
given in [19] to solve the optimization problem $\min \sum_{i=1}^{n} f_i$. Therefore, the result in Theorem 6.1 is consistent with those in [19]. Note that [21] proposed distributed continuous-time algorithms for the empty intersection case when the graphs kept connected, which is more restrictive than the UJSC given in this paper.

From Theorems 5.1 and 6.1 we find that the sufficient optimal consensus conditions are essentially different in the two cases. In the nonempty intersection case, besides the divergence condition, the square integrability condition is usually required in the empty intersection case. Moreover, the diminishing stepsize is also somewhat necessary for the optimal consensus.

Before presenting the proof of Theorem 6.1, we show two lemmas. The first one is taken from Lemma 7 in [19].

**Lemma 6.1** Let $0 < \lambda < 1$ and $\{b_k\}_{k \geq 1}$ be a positive sequence. If $\sum_{k=1}^{\infty} b_k < \infty$, then $\sum_{k=1}^{\infty} \sum_{r=1}^{k} \lambda^{k-r} b_r < \infty$.

Before introducing the second lemma, we rewrite (6) as

$$\dot{x}_i(t) = \sum_{j \in N_i(t)} (x_j(t) - x_i(t)) + \alpha_t (P_{X_i}^h(x_i(t)) - x_i(t)) + \Lambda_i(t),$$

where $\Lambda_i(t) = (\alpha_{i,t} - \alpha_t)(P_{X_i}^h(x_i(t)) - x_i(t))$. From the definition of $P_{X_i}^h$, we have

$$|P_{X_i}^h(x_i(t)) - x_i(t)| = |x_i(t)|_{\mathcal{X}_i} \leq \frac{d_1}{\cos \theta_{i,t}},$$

where $d_1 = \sup_{i,j,t} |x_i(t)|_{\mathcal{X}_j}$, which is finite by Theorem 4.1. Moreover, it follows from (16) that $|\alpha_{i,t} - \alpha_t| \leq (C_{i,t} - 1) \alpha_t \leq \frac{1}{\sin \mu_i} \alpha_t \sin \theta_{i,t}$. Then from the preceding two estimates,

$$|\Lambda_i(t)| \leq \frac{d_1}{\sin \mu_i} \alpha_t \tan \theta_{i,t}.$$  \hfill (29)

**Lemma 6.2** Under A1-A3, if $\int_{0}^{\infty} \alpha_t^2 dt < \infty$ and $\int_{0}^{\infty} \alpha_t \tan \theta_t^+ dt < \infty$, then

$$\int_{0}^{\infty} \alpha_t |x_i(t) - \bar{x}(t)| dt < \infty$$

for all $i$, where $\bar{x}(t) = \frac{1}{n} \sum_{i=1}^{n} x_i(t)$.

**Proof.** Let $H(t) = \max_{1 \leq i, j \leq n} |x_i(t) - x_j(t)|$. Since $|x_i(t) - \bar{x}(t)| \leq H(t)$, it suffices to show $\int_{0}^{\infty} \alpha_t H(t) dt < \infty$. 

21
It is not hard to see that \(|P_{X,i}^k(x_i) - x_i|^2 = |x_i|^2_{X,i} + |P_{X,i}^k(x_i) - P_{X,i}^k(x_i)|^2 \leq (1 + (\tan \theta^*)^2)|x_i|^2_{X,i} \leq (1 + (\tan \theta^*)^2)d_i^2\). By (29) and Lemma 2.5 for any \(k \geq 0\) and \(t, kB_0 \leq t < (k + 1)B_0\),

\[
H((k + 1)B_0) \leq \beta H(kB_0) + B_2\nu(kB_0),
\]

\[
H(t) \leq H(kB_0) + B_2\nu(kB_0),
\]

where \(B_2 = B_1d_1(\sqrt{1 + (\tan \theta^*)^2} + \tan \theta^*/\sin \mu)\), \(\mu = \min_{1 \leq i \leq n} \mu_i > 0\) and \(\nu(kB_0) = \int_{kB_0}^{(k+1)B_0} \alpha_t dt\). Thus, with (31) we have

\[
\int_0^\infty \alpha_t H(t) dt = \sum_{k=0}^\infty \int_{kB_0}^{(k+1)B_0} \alpha_t H(t) dt \leq \sum_{k=0}^\infty \int_{kB_0}^{(k+1)B_0} \alpha_t (H(kB_0) + B_2\nu(kB_0)) dt = \sum_{k=0}^\infty \nu(kB_0)(H(kB_0) + B_2\nu(kB_0)).
\]

Now we estimate the term in (32). First, by Cauchy-Schwarz inequality \(\int g_1g_2 \leq \sqrt{\int g_1^2 \int g_2^2}\), we have

\[
\sum_{k=0}^\infty \nu^2(kB_0) \leq B_0 \sum_{k=0}^\infty \int_{kB_0}^{(k+1)B_0} \alpha_t^2 dt = B_0 \int_0^\infty \alpha_t^2 dt < \infty.
\]

Second, by (31) we have \(H(kB_0) \leq \beta^k H(0) + B_2\sum_{r=1}^k \beta^{k-r} \nu((r-1)B_0), \forall k \geq 1\). Thus,

\[
\sum_{k=1}^\infty H(kB_0)\nu(kB_0) \leq H(0) \sum_{k=1}^\infty \beta^k \nu(kB_0) + B_2 \sum_{k=1}^\infty \nu(kB_0) \sum_{r=1}^k \beta^{k-r} \nu((r-1)B_0)
\]

\[
\leq \frac{H(0)\beta \alpha^* B_0}{1 - \beta} + \frac{B_2}{2(1 - \beta)} \sum_{k=1}^\infty \sum_{r=1}^k \beta^{k-r} (\nu^2(kB_0) + \nu^2((r-1)B_0))
\]

\[
\leq \frac{H(0)\beta \alpha^* B_0}{1 - \beta} + \frac{B_2}{2(1 - \beta)} \sum_{k=1}^\infty \nu^2(kB_0) + \frac{B_2}{2} \sum_{k=1}^\infty \sum_{r=1}^k \beta^{k-r} \nu^2((r-1)B_0)
\]

\[
< \infty,
\]

where the second inequality follows from \(\nu(kB_0) \leq \alpha^* B_0\), the third one from \(\sum_{r=1}^k \beta^{k-r} \leq \frac{1}{1 - \beta}, \forall k\), and the last one from (33) and Lemma 6.1. Thus, the conclusion follows from (32), (33) and (34). \(\square\)

It is time to give the proof of Theorem 6.1.

22
Proof of Theorem 6.1. Take \( x^* \in \text{arg}\ min \sum_{i=1}^{n} |x_i|^2 \). Clearly,

\[
\frac{d|x_i(t) - x^*|^2}{dt} = 2\langle x_i(t) - x^*, \dot{x}_i(t) \rangle = 2\langle x_i(t) - x^*, \sum_{j \in N_i(t)} (x_j(t) - x_i(t)) + \alpha_t(P_{X_i}^h(x_i(t)) - x_i(t)) + \Lambda_i(t) \rangle.
\]

(35)

Because \( G_{\sigma(t)} \) is undirected,

\[
\frac{d \sum_{i=1}^{n} |x_i(t) - x^*|^2}{dt}
= -2 \sum_{j \in N_i(t)} |x_j(t) - x_i(t)|^2 + 2 \sum_{i=1}^{n} \langle x_i(t) - x^*, \alpha_t(P_{X_i}^h(x_i(t)) - x_i(t)) + \Lambda_i(t) \rangle
\leq 2\alpha_t \sum_{i=1}^{n} \langle x_i(t) - x^*, P_{X_i}^h(x_i(t)) - x_i(t) \rangle + 2 \sum_{i=1}^{n} \langle x_i(t) - x^*, \Lambda_i(t) \rangle.
\]

(36)

Then we estimate the first term in (36). Note that

\[
\langle x_i(t) - x^*, P_{X_i}^h(x_i(t)) - P_{X_i}(x_i(t)) \rangle \leq |x_i(t) - x^*| \tan \theta_i \| x_i(t) \|_{X_i} \leq d_2 \tan \theta^+_i,
\]

(37)

where \( d_2 = \sup_{i \in \mathcal{I}} \| x_i(t) - x^* \|_{X_i} \| \bar{x}(t) \|_{X_i} \| x_i(t) \|_{X_i} \) is finite by Theorem 4.1. We also have

\[
\sum_{i=1}^{n} \langle x_i(t) - x^*, P_{X_i}(x_i(t)) - x_i(t) \rangle = -\langle \bar{x}(t) - x^*, \sum_{i=1}^{n} (\bar{x}(t) - P_{X_i}(\bar{x}(t))) \rangle + \Delta(t),
\]

(38)

where \( \bar{x}(t) = \frac{1}{n} \sum_{i=1}^{n} x_i(t) \),

\[
\Delta(t) = \sum_{i=1}^{n} \langle x_i(t) - \bar{x}(t), P_{X_i}(\bar{x}(t)) - \bar{x}(t) \rangle + \sum_{i=1}^{n} \langle x_i(t) - x^*, P_{X_i}(x_i(t)) - P_{X_i}(\bar{x}(t)) + \bar{x}(t) - x_i(t) \rangle.
\]

Clearly, the first term in \( \Delta(t) \) is not greater than \( d_2 \sum_{i=1}^{n} |x_i(t) - \bar{x}(t)| \) and the second term in \( \Delta(t) \) is not greater than \( 2d_2 \sum_{i=1}^{n} |x_i(t) - \bar{x}(t)| \) by Lemma 2.1 (iv). Moreover, by (29), the second term in (36) is not greater than \( 2nd_1d_2\alpha_t \tan \theta^+_i / \sin \mu_- \). In light of (36), (37) and (38), we have

\[
\frac{d \sum_{i=1}^{n} |x_i(t) - x^*|^2}{dt} \leq -2\alpha_t \langle \bar{x}(t) - x^*, \sum_{i=1}^{n} (\bar{x}(t) - P_{X_i}(\bar{x}(t))) \rangle + 6d_2 \left( \sum_{i=1}^{n} \alpha_t |x_i(t) - \bar{x}(t)| \right)
\]

(39)

\[
+ 2nd_2(d_2 + d_1 / \sin \mu_-) \alpha_t \tan \theta^+_i
\]

\[
\leq 6d_2 \left( \sum_{i=1}^{n} \alpha_t |x_i(t) - \bar{x}(t)| \right) + 2nd_2(d_2 + d_1 / \sin \mu_-) \alpha_t \tan \theta^+_i
\]

(40)

because the first term of the right-hand side in (39) is not positive, following from (11) and the convexity of objective function \( f \), that is,

\[
\langle \bar{x}(t) - x^*, \sum_{i=1}^{n} (\bar{x}(t) - P_{X_i}(\bar{x}(t))) \rangle = \langle \bar{x}(t) - x^*, \frac{1}{2} \nabla f(\bar{x}(t)) \rangle \geq \frac{1}{2} (f(\bar{x}(t)) - f(x^*)) \geq 0.
\]

(41)
According to \( \Box \). Barbalat’s Lemma (see Lemma 4.2 in [36]). Therefore, because the network achieves a consensus, \( \lim_{t \to \infty} |x_i(t) - x^*|^2 \) is a finite number by contradiction. Let us suppose there exist \( \{t_k\}_{k \geq 0} \) with \( t_k \to \infty \) and \( \varepsilon > 0 \) such that \( \sum_{i=1}^{n} |x_i(t_{2k+1}) - x^*|^2 \geq \varepsilon \) for all \( k \). According to Lemma 6.2, \( \int_0^\infty \sum_{i=1}^{n} \alpha_t |x_i(t) - \bar{x}(t)| dt < \infty \). Therefore, there is \( K_0 > 0 \) such that \( \int_{K_0}^\infty \sum_{i=1}^{n} \alpha_t |x_i(t) - \bar{x}(t)| dt \leq \frac{2\varepsilon}{kd_2} \) and \( \int_{K_0}^\infty \alpha_t \tan \theta_t^+ dt \leq \frac{2\varepsilon}{3d_2(2d_1/\sin \mu_\infty)} \).

By (40), we have that, for sufficiently large \( k \),

\[
\varepsilon \leq \sum_{i=1}^{n} |x_i(t_{2k+1}) - x^*|^2 - \sum_{i=1}^{n} |x_i(t_{2k}) - x^*|^2
\]

\[
\leq 6d_2 \int_{K_0}^\infty \alpha_t (\sum_{i=1}^{n} |x_i(t) - \bar{x}(t)|) dt + 2nd_2 \int_{K_0}^\infty \alpha_t \tan \theta_t^+ dt
\]

\[
\leq \frac{\varepsilon}{2},
\]

which yields a contradiction. Hence, \( \lim_{t \to \infty} \sum_{i=1}^{n} |x_i(t) - x^*|^2 \) is a finite number.

Thus, it follows from (39) that

\[
2 \int_0^\infty \alpha_t (\bar{x}(t) - x^*, \sum_{i=1}^{n} (\bar{x}(t) - P_{X_i}(\bar{x}(t)))) = \int_0^\infty \alpha_t (\bar{x}(t) - x^*, \nabla f(\bar{x}(t))) < \infty.
\]

Due to \( \int_0^\infty \alpha_t dt = \infty \), there is a subsequence \( \{t_r\}_{r \geq 0} \) such that \( \lim_{r \to \infty} (\bar{x}(t_r) - x^*, \nabla f(\bar{x}(t_r))) = 0 \). Since the system states are bounded, without loss of generality we assume \( \lim_{r \to \infty} \bar{x}(t_r) = \hat{x} \) for some \( \hat{x} \) (otherwise we can further find a subsequence of \( \{t_r\}_{r \geq 0} \). Since \( \nabla f \) is continuous, \( \bar{x} - x^*, \nabla f(\hat{x}) \) = 0, which leads to \( f(x^*) \geq f(\hat{x}) + (x^* - \hat{x}, \nabla f(\hat{x})) = f(\hat{x}) \). Thus, \( \hat{x} \in \text{arg min} f \).

Replacing \( x^* \) with \( \hat{x} \), we can similarly show that \( \lim_{t \to \infty} \sum_{i=1}^{n} |x_i(t) - \hat{x}|^2 \) is also a finite number, denoted as \( \rho^* \). Moreover, the uniform continuity of \( \alpha_t \) and \( \int_0^\infty \alpha_t^2 dt < \infty \) imply \( \lim_{t \to \infty} \alpha_t = 0 \). Therefore, consensus is achieved by Lemma 2.6 and then \( \lim_{r \to \infty} x_i(t_r) = \hat{x} \).

Hence \( \rho^* = 0 \), which in return implies \( \lim_{t \to \infty} x_i(t) = \hat{x} \) for all \( i \). Then the first part is completed.

Now we show the second part. Notice that \( \theta_{i,t} \equiv 0 \) implies that \( \alpha_{i,t} = \alpha_{j,t} = \alpha_t \forall i,j,t \) and \( P^h_{X_i}(x_i(t)) = P^a_{X_i}(x_i(t)) = P_{X_i}(x_i(t)) \). Let \( x^* \in \text{arg min} \sum_{i=1}^{n} |x|_{X_i}^2 \) be the point with \( \lim_{t \to \infty} x_i(t) = x^* \) for all \( i \). According to Theorem 4.1 and the uniform continuity of \( \alpha_t \), \( \hat{x}_i(t) \) is uniformly continuous, which, along with \( \lim_{t \to \infty} x_i(t) = x^* \), leads to \( \lim_{t \to \infty} \hat{x}_i(t) = 0 \) by Barbalat’s Lemma (see Lemma 4.2 in [36]). Therefore, because the network achieves a consensus,

\[
\lim_{t \to \infty} \alpha_t (P_{X_i}(x_i(t)) - x_i(t)) = \lim_{t \to \infty} \alpha_t (P_{X_i}(x^*) - x^*) = 0, \; i = 1, \ldots, n.
\]

According to \( \bigcap_{i=1}^{n} X_i = \emptyset \), \( x^* \notin X_i \) for at least one \( i \). Thus, \( \lim_{t \to \infty} \alpha_t = 0 \). \qed
6.2 Optimal Solutions

Theorem 6.1 shows that under certain conditions, all agents will converge to a consensus point, which is the optimal solution of \( \min \sum_{i=1}^{n} |x|_{X_i}^2 \). Here we show some properties of the optimal solution set. According to Lemma 2.3, the optimal solution \( x \) of \( \min \sum_{i=1}^{n} |x|_{X_i}^2 \) must satisfy

\[
\nabla \sum_{i=1}^{n} |x|_{X_i}^2 = 2 \sum_{i=1}^{n} (x^* - P_{X_i}(x^*)) = 0, \quad \text{or equivalently,}
\]

\[
x^* = \frac{\sum_{i=1}^{n} P_{X_i}(x^*)}{n}.
\]

Before showing some properties of the optimal solutions, we give a lemma first.

**Lemma 6.3** Let \( K \) be a closed convex set in \( \mathbb{R}^m \). Then

(i) \( \langle y - z, P_K(y) - P_K(z) \rangle \geq |P_K(y) - P_K(z)|^2 \) for any \( y \) and \( z \);

(ii) \( |P_K(y) - P_K(z)| = |y - z| \) if and only if \( y - P_K(y) = z - P_K(z) \).

**Proof.** (i) follows from

\[
\langle y - z, P_K(y) - P_K(z) \rangle = (y - P_K(y), P_K(y) - P_K(z)) + |P_K(y) - P_K(z)|^2
\]

\[
+ \langle P_K(z) - z, P_K(y) - P_K(z) \rangle
\]

\[
\geq |P_K(y) - P_K(z)|^2
\]

because \( \langle y - P_K(y), P_K(y) - P_K(z) \rangle \geq 0 \) and \( \langle P_K(z) - z, P_K(y) - P_K(z) \rangle \geq 0 \) by Lemma 2.1 (i).

For (ii), the sufficiency is obvious. The necessity can be obtained from

\[
|y - P_K(y) - (z - P_K(z))|^2 = |y - z|^2 + |P_K(z) - P_K(y)|^2 + 2\langle y - z, P_K(z) - P_K(y) \rangle
\]

\[
= 2|y - z|^2 + 2\langle y - z, P_K(z) - P_K(y) \rangle
\]

\[
\leq 2|y - z|^2 - 2|P_K(y) - P_K(z)|^2
\]

\[
= 0,
\]

where the inequality follows from (i) of this lemma. \( \square \)

Let \( X^* \) be the optimal solution set of \( \min \sum_{i=1}^{n} |x|_{X_i}^2 \). Then we have the following result.

**Theorem 6.2** (i) For any \( x^*, y^* \in X^* \), we have \( x^* - P_{X_i}(x^*) = y^* - P_{X_i}(y^*) \), \( i = 1, \ldots, n \);

(ii) For any \( i \), either \( X^* \subseteq X_i \) or \( X^* \cap X_i = \emptyset \);

(iii) Let \( x^* \in X^* \), \( x^* \notin X_i \) for some \( i \). Then \( X^* \cap \text{line}(x^*, P_{X_i}(x^*)) = \{x^*\} \).
Proof. (i) Since \( x^* = \frac{\sum_{i=1}^{n} P_X(x^*)}{n} \) and \( y^* = \frac{\sum_{i=1}^{n} P_X(y^*)}{n} \),
\[
|x^* - y^*| = \left| \frac{\sum_{i=1}^{n} (P_X(x^*) - P_X(y^*))}{n} \right| 
\leq \frac{\sum_{i=1}^{n} |P_X(x^*) - P_X(y^*)|}{n} 
\leq |x^* - y^*|
\]
from Lemma 2.1 (iv). Therefore, \( |P_X(x^*) - P_X(y^*)| = |x^* - y^*| \) for all \( i \), which implies the conclusion by Lemma 6.3 (ii).

(ii) is straightforward from (i).

(iii) Let \( z^* \in X^* \cap \text{line}(x^*, P_X(x^*)) \), \( z^* \neq x^* \). If \( z^* \) locates the half-line with \( P_X(x^*) \) as the starting point and \( x^* - P_X(x^*) \) as the direction, then \( P_X(z^*) = P_X(x^*) \) by Lemma 2.1 (i).

Therefore, \( x^* - P_X(x^*) \neq z^* - P_X(z^*) \), which contradicts what we have proven in (i) since both \( x^* \) and \( z^* \) are optimal solutions. If \( z^* \) locates the half-line with \( P_X(x^*) \) as the starting point and \( P_X(x^*) - x^* \) as the direction, then \( P_X(x^*) \) is also an optimal solution since the optimal solution set \( X^* \) is a convex set and \( P_X(x^*) \) can be written as a convex combination of \( x^*, z^* \).

Then \( 0 = P_X(x^*) - P_X(P_X(x^*)) \neq x^* - P_X(x^*) \), which also yields a contradiction since both \( x^* \) and \( P_X(x^*) \) are optimal solutions. Thus, the conclusion follows. \( \square \)

7 Fixed Stepsize and Approximate Angle

In this section, we consider the constant stepsize and approximate angle case. The following result is about the convergence error between the agents’ estimates and the optimal point in terms of the stepsize and approximate angle.

Theorem 7.1 Consider system 2 with \( \alpha_t \equiv \alpha > 0 \) and \( 0 \leq \theta_{i,t} \equiv \theta_i < \theta^* < \pi/2 \), under A1, A2, A3 (ii) and that \( G_{\sigma(t)} \) is undirected for \( t \geq 0 \).

(i) Suppose \( \bigcap_{i=1}^{n} ri(X_i) \neq \emptyset \) (implies \( \bigcap_{i=1}^{n} X_i \neq \emptyset \)). Then
\[
\limsup_{t \to \infty} |x_i(t)|_{X_0} \leq \sqrt{\kappa_{Sn}(d_3)} \left( \frac{4 - 2\beta}{1 - \beta} B_0 B_1 C \sqrt{1 + (\tan \theta^*)^2 d_1 \alpha + C d_1 \tan \theta^*} \right)
+ \frac{2 - \beta}{1 - \beta} B_0 B_1 C \sqrt{1 + (\tan \theta^*)^2 d_1 \alpha}.
\]

(ii) Suppose \( \bigcap_{i=1}^{n} X_i = \emptyset \) and \( f(x) = \sum_{i=1}^{n} |x|^2_{X_i} \) is \( \ell \)-strongly convex. Let \( x^* \) be the unique
optimal solution of \( \min f \).

\[
\limsup_{t \to \infty} |x_i(t) - x^*| \leq \sqrt{\frac{4nd_2}{\ell} \left( 4 - \frac{2\beta}{1 - \beta} \right) B_0 B_1 C \sqrt{1 + (\tan \theta^+)^2 d_1 \alpha + C d_1 \tan \theta^+}} + 2 \frac{\beta}{1 - \beta} B_0 B_1 C \sqrt{1 + (\tan \theta^+)^2 d_1 \alpha}
\]

with \( S = \bigcup_{i=1}^n X_i + B(0, r_0) \), \( C = \max_{1 \leq i \leq n} C_i \), \( \theta^+ = \max_{1 \leq i \leq n} \theta_i \), \( C_i \) defined in [13], \( \kappa \) defined in Lemma 2.2, and \( \beta, B_0 \) and \( B_1 \) defined in Lemma 2.5, where \( d_1 = \sup_{t \in [0, \infty)} |x_i(t)|_{X_j} \), \( d_2 = \sup_{t \in [0, \infty)} \{ |x_i(t) - x^*|, |\bar{x}(t)|_{X_i}, |x_i(t)|_{X_i} \} \), \( d_3 = \sup_{t \in [0, \infty)} |x_i(t)|_{X_0} \), which are finite by Theorem 4.1.

**Proof.** Similar to (30),

\[
H((k + 1)B_0) \leq \beta^k H(B_0) + B_0 B_1 C \sqrt{1 + (\tan \theta^+)^2 d_1 \alpha},
\]

along with \( H(t) \leq H(kB_0) + B_0 B_1 C \sqrt{1 + (\tan \theta^+)^2 d_1 \alpha} \), \( \forall kB_0 \leq t < (k + 1)B_0 \), implies

\[
H(t) \leq \beta^k H(0) + (2 + \beta + \cdots + \beta^{k-1}) B_0 B_1 C \sqrt{1 + (\tan \theta^+)^2 d_1 \alpha}
\]

\[
\leq \frac{H(0)}{\beta^{k_0}} \left( \beta^{k_0} \right)^t + 2 \frac{\beta}{1 - \beta} B_0 B_1 C \sqrt{1 + (\tan \theta^+)^2 d_1 \alpha}
\]

(42)

when \( kB_0 \leq t < (k + 1)B_0 \). Then

\[
\limsup_{t \to \infty} H(t) \leq \frac{2 - \beta}{1 - \beta} B_0 B_1 C \sqrt{1 + (\tan \theta^+)^2 d_1 \alpha}.
\]

(43)

Clearly, (3) can be written as

\[
\dot{x}_i(t) = \sum_{j \in X_i(t)} (x_j(t) - x_i(t)) + \alpha(P_{X_i}(\bar{x}(t)) - \bar{x}(t)) + \omega_i(t), \quad i = 1, \ldots, n,
\]

(44)

where \( \bar{x}(t) = \frac{1}{n} \sum_{i=1}^n x_i(t) \),

\[
\omega_i(t) = \alpha(P_{X_i}(x_i(t)) - P_{X_i}(\bar{x}(t)) + \bar{x}(t) - x_i(t) + \alpha(P_{X_i}(x_i(t)) - P_{X_i}(x_i(t))).
\]

The first term in \( \omega_i(t) \) is not greater than \( 2\alpha H(t) \) by Lemma 2.1 (iv) and the inequality \( |\bar{x}(t) - x_i(t)| \leq H(t) \), and the second term is not greater than \( C d_1 \alpha \tan \theta^+ \) due to the relation

\[
\frac{|P_{X_i}(x_i(t)) - P_{X_i}(x_i(t))|}{|P_{X_i}(x_i(t)) - P_{X_i}(x_i(t))|} \leq \frac{\sin(\frac{\theta_i}{2})}{\sin \mu_i} \leq \frac{1}{\sin \mu_i} \leq C_i,
\]

where the equality follows from the well-known law of sines: \( a_1/\sin A_1 = a_2/\sin A_2 = a_3/\sin A_3 \) with \( a_1, a_2, a_3 \) the lengths of the sides of a triangle, and \( A_1, A_2, A_3 \) the opposite angles. Hence,

\[
|\omega_i(t)| \leq 2\alpha H(t) + C d_1 \alpha \tan \theta^+.
\]

(45)
Therefore, from (44) and the undirectedness of $G_{\sigma(t)}$, we have

$$\dot{x}(t) = \frac{1}{n} \sum_{i=1}^{n} \dot{x}_i(t) = \frac{\alpha}{n} \sum_{i=1}^{n} (P_{X_i} (\bar{x}(t)) - \bar{x}(t)) + \frac{1}{n} \sum_{i=1}^{n} \omega_i(t).$$

(46)

We complete the proof for both the nonempty intersection and empty intersection case.

(i) If $\bigcap_{i=1}^{n} x_i(t) \neq \emptyset$, then, from Lemma 2.2,

$$\lim_{t \to +\infty} ||x(t)||_{X_0} = 0$$

and (43); the third one from Lemma 2.2. As a result, we obtain that for any $t \geq t_0 \geq 0$,

$$||x(t)||_{X_0} \leq e^{-\frac{2\alpha}{\kappa_{\max} \theta^+}} ||x(t_0)||_{X_0} + e^{-\frac{2\alpha}{\kappa_{\max} \theta^+}} \int_{t_0}^{t} e^{-\frac{2\alpha}{\kappa_{\max} \theta^+}} (4\alpha H(s) + 2Cd_1 \alpha \tan \theta^+) ds,$$

(47)

which combines with (46) imply $\lim_{t \to +\infty} ||x(t)||_{X_0} \leq \frac{\kappa_{\max} \theta^+}{2\alpha} (4\alpha \frac{2-\beta}{1-\beta} B_0 B_1 C \sqrt{1 + (\tan \theta^+)^2} d_1 + 2Cd_1 \alpha \tan \theta^+).$ Thus, this implies the conclusion by noticing the relation $|x(t)|_{X_0} \leq ||x(t)||_{X_0} \leq ||x(t)||_{X_0} \leq H(t)$ and (43).

(ii) If $\bigcap_{i=1}^{n} x_i(t) = \emptyset$ and $f(x) = \sum_{i=1}^{n} |x|_{X_i}$ is $\ell$-strongly convex, then

$$\frac{d||x(t) - x^*||_{X_0}^2}{dt} = \frac{2\alpha}{n} \langle x(t) - x^*, \sum_{i=1}^{n} (P_{X_i} (\bar{x}(t)) - \bar{x}(t)) \rangle + \frac{2\alpha}{n} \langle x(t) - x^*, \sum_{i=1}^{n} \omega_i(t) \rangle$$

$$\leq -\frac{\alpha}{n} \langle f(\bar{x}(t)) - f(x^*) \rangle + \langle 4\alpha H(t) + 2Cd_1 \alpha \tan \theta^+ \rangle d_2$$

$$\leq -\frac{\alpha \ell}{2n} ||x(t) - x^*||_{X_0}^2 + \langle 4\alpha H(t) + 2Cd_1 \alpha \tan \theta^+ \rangle d_2,$$

(49)

where the first inequality follows from (46) and the second one from (2). Thus, the conclusion can be obtained with a proof similar to that for (i).

Remark 7.1 From the proof (ii) of Theorem 4.7, $|x_i(t)|_{X_{co}} \leq \left( \tan \theta^* + \sqrt{(\tan \theta^*)^2 + 2 \tan \theta^*} \right) d_0$ for all $i, t$. Then we take sufficiently large $r_0$ with

$$r_0 \geq \left( \tan \theta^* + \sqrt{(\tan \theta^*)^2 + 2 \tan \theta^*} \right) d_0 + \max_{1 \leq i \leq n} \sup_{z_1 \in x_i, z_2 \in X_{co}} |z_1 - z_2|$$

28
such that $S$ contains all states $x_i(t), i, t \geq 0$. Since set $S$ depends only on $\theta^*$ and $X_i, i = 1, ..., n$, not on the parameters $\alpha$ and $\theta_i$, $d_i \leq \sup_{z \in S} \{|z - x^*|, |z|_{X_0}, |z|_{X_i}, i = 1, ..., n\}$ for $i = 1, 2, 3$. From the proof of Proposition 4.10 in [25], we find that $\beta, B_0$ and $B_1$ depend only on the graphs $G_{\sigma(t)}, t \geq 0$. Moreover, from (17) we know that $C$ depends only on the sets $X_i$, the maps $P^a_{X_i}$ and $r_0$. Thus, the upper bound of convergence errors given in Theorem 7.1 can be made arbitrarily small when $\alpha$ and $\theta_i, i = 1, ..., n$ are sufficiently small.

8 Numerical Example

In this section, we provide an example to illustrate the above convergence results.

Consider a network of three agents with node set $\mathcal{V} = \{1, 2, 3\}$. The convex set $X_i$ of each agent $i$ is the ball in $\mathbb{R}^2$ with center $c_i$ and radius $r_i$. Let $\alpha_t = \frac{20}{t+20}, \theta_{i,t} = \frac{1}{t+50}$, which satisfy the conditions in Theorems 5.1 and 6.1. We next present the state trajectories of the three agents for the nonempty and empty intersection cases from time $t = 0$ to $t = 2000$, respectively.

(i) Nonempty intersection case with $c_1 = (-1,0)^T, c_2 = (1,0)^T, c_3 = (0,-2)^T, r_1 = 2, r_2 = 1, r_3 = 2$.

The graphs are periodically switching over the two directed graphs $G_1 = (\mathcal{V}, \mathcal{E}_1), G_2 = (\mathcal{V}, \mathcal{E}_2)$ with period 1, where $\mathcal{E}_1 = \{(2,1), (3,2)\}, \mathcal{E}_2 = \{(1,3)\}$. The initial conditions are $x_1(0) = (-4,3), x_2(0) = (3,5), x_3(0) = (-6,-3)$, which are marked as $\circ$ in Fig. 4.

(ii) Empty intersection case with $c_1 = (-\sqrt{3},0)^T, c_2 = (\sqrt{3},0)^T, c_3 = (0,-3)^T, r_1 = r_2 = r_3 = 1$.

In this case, the (unique) optimal solution is $(0,-1)$. The graphs are periodically switching over the two undirected graphs $G_1 = (\mathcal{V}, \mathcal{E}_1), G_2 = (\mathcal{V}, \mathcal{E}_2)$ with period 1, where $\mathcal{E}_1 = \{(3,2)\}, \mathcal{E}_2 = \{(1,2)\}$. The initial conditions are $x_1(0) = (-3,3), x_2(0) = (4,2), x_3(0) = (-5,-3)$, which are marked as $\circ$ in Fig. 4.

9 Conclusions

In this paper, a continuous-time method was proposed to cooperatively solve the SDOP by a group of agents with the help of graph theory, convex analysis and geometric technique. Here
Figure 4: All agents converge to a common point in the intersection set when it is not empty.

Figure 5: In the empty intersection case, all agents converge to the unique optimal solution $(0, -1)$. 
agents could only obtain their approximate projections and the communication graph among agents was UJSC. It was shown that the system states were always bounded for any constant approximate angle, and uniformly bounded for any stepsize with inferior limit greater than zero. Both nonempty intersection and empty intersection cases of convex sets were investigated with respective sufficient conditions. Moreover, the convergence error between agents’ estimates and the optimal point was also obtained for the constant stepsize and approximate angle case.

Acknowledgment

The authors would like to thank Dr. Guilin Yang for discussions about geometric analysis and Mr. Peng Yi for his generous help on numerical simulations.

Appendix

Denote \( \vartheta_i(v) = \angle (v - P_{X_i}(v), P_{X_i}^a(v) - P_{X_i}(v)). \) By Lemma 2.1 (i), \( \vartheta_i(v) \geq \pi/2 \) when \( P_{X_i}^a(v) \neq P_{X_i}(v). \) In the following proofs, we omit all subscript \( i \) and simplify \( \vartheta, \vartheta_i, \theta_i, \mu_i, X_i \) as \( \vartheta, \theta, \mu, X. \)

Proof of Lemma 4.1. Let \( v \in S \setminus X \) and \( P_{X_i}^a(v) \neq P_X(v) \). We obtain

\[
\gamma_X(v) = \frac{|P_{X_i}^b(v) - v| + |P_{X_i}^a(v) - P_{X_i}^b(v)|}{|P_{X_i}^b(v) - v|} = 1 + \frac{|P_{X_i}^a(v) - P_{X_i}^b(v)|}{|P_{X_i}^b(v) - P_X(v)|} \sin \vartheta(v) \\
= 1 + \frac{\sin(\theta(v) - \frac{\pi}{2})}{\sin \mu(v)} \sin \vartheta(v) \\
\leq 1 + \frac{1}{\sin \mu(v)} \sin \vartheta(v) \\
\leq 1 + \frac{1}{\sin \mu(v)},
\]

where the third equality follows from the law of sines. Then the first part follows.

We now show the second part. Let \( y \in S \setminus X \) and \( P_X^a(y) \neq P_X(y) \). Let the intersection point of \( \text{bd}(S) \) and the half-line starting from \( P_X(y) \) with \( y - P_X(y) \) as the direction be \( z(y)(:= z). \) Clearly, \( P_X(z) = P_X(y). \) Therefore, from (15) there exists point \( \hat{z}(y)(:= \hat{z}) \in \text{bd}(C_X(z, \hat{\theta})) \cap b(z, X) \cap \text{aff}\{y, P_X(y), P_X^a(y)\} \) with \( \angle(\hat{z} - z, P_X(z) - z) = \hat{\theta}. \)

Since (15) also holds for any \( \theta \leq \hat{\theta} \) and \( X_i \)s are bounded, without loss of generality we assume \( \theta_i(v) \leq \hat{\theta} \forall v \in \text{bd}(S) \) otherwise we can take a larger \( r_0 \) to guarantee this assumption. This along
with A3 (ii) implies \( \theta_i(v) \leq \hat{\theta} \) for any \( v \in S \setminus X \). Then \( \theta(y) \leq \hat{\theta} \). Clearly, \( P^a_X(y) \) must locate the triangle with vertex points \( z, \hat{z} \) and \( P_X(z) \) since the line segment \([\hat{z}, P_X(z)]\) is in \( X \), which implies \( \theta(y) \leq \angle(z - P_X(y), \hat{z} - P_X(z)) \). Therefore, \( \mu(y) \geq \angle(P_X(z) - \hat{z}, z - \hat{z}) \).

For any \( \bar{z} \in \Gamma := \bigcup_{v \in \text{bd}(S)} (\text{bd}(C_X(v, \hat{\theta})) \cap \text{b}(v, X)) \), let \( z \in \text{bd}(S) \) be the point with \( \bar{z} \in \text{bd}(C_X(z, \hat{\theta})) \cap \text{b}(z, X) \). Denote \( \bar{\theta}(\bar{z}) = \angle(P_X(z) - \bar{z}, z - \bar{z}) \). Clearly, we can show \( \inf_{\bar{z} \in \Gamma} \bar{\theta}(\bar{z}) > 0 \) by contradiction. Thus, the conclusion follows from \( \bigcup_{y \in S \setminus X, P_X^a(y) \neq P_X(y)} \hat{z}(y) \subseteq \Gamma \). \( \square \)

---

**Proof of Lemma 4.2.** The conclusion follows from Lemma 4.1 and the following three claims:

(i) If for any \( v \in \text{bd}(S) \), there exists \( \theta(v) > 0 \) such that (15) holds for \( v, \theta(v) \), then there exists \( \hat{\theta} > 0 \) such that (15) holds for any \( v \in \text{bd}(S) \).

(ii) If \( \text{int}(X) \cap \text{line}(v, P_X(v)) \neq \emptyset \) for any \( v \in \text{bd}(S) \), the sufficient condition in (i) holds.
(iii) If the boundary surface of $X$ is regular, then the sufficient condition in (ii) holds.

We complete the proof of (i) by contradiction. Suppose $\{v_k\}_{k \geq 0} \subseteq \text{bd}(S)$ is a sequence with $\text{cone}(v_k, C_X(v_k, 1/k) \cap b(v_k, X)) \neq C_X(v_k, 1/k)$. Without loss of generality, we assume $\lim_{k \to \infty} v_k = v^* \in \text{bd}(S)$. Let $\theta_0 > 0$ be the angle to satisfy (15) for $v^*$, and $z_j \in C_X(v^*, \theta_0) \cap b(v^*, X)$ be all the points with $\angle(z_j - v^*, P_X(v^*) - v^*) = \theta_0$ (i.e., $z_j \in \text{bd}(C_X(v^*, \theta_0)) \cap b(v^*, X)$). Then

$$\lim_{k \to \infty} \angle(z_j - v_k, P_X(v_k) - v_k) = \theta_0, \forall j,$$

and the convergence is uniform with respect to all $j$. This implies that for sufficiently large $k$, (15) holds for $v_k, \theta_0/2$, which contradicts $\text{cone}(v_k, C_X(v_k, 1/k) \cap b(v_k, X)) \neq C_X(v_k, 1/k)$ based on the fact that (15) holds for $v, \theta_1$ implies that it also holds for $v, \theta_2, \theta_2 \leq \theta_1$.

For (ii), let $z \in \text{int}(X) \cap \text{line}(v, P_X(v))$. Then there exists $\epsilon > 0$ such that $B(z, \epsilon) \subseteq X$. Let $y \in \text{bd}(B(z, \epsilon))$ be the point for which $\angle(y - z, v - z) = \pi/2$. Obviously, (15) holds for $v, \theta(v)$, where $\theta(v) = \angle(y - v, z - v) > 0$.

We now prove (iii). According to Proposition 1 on page 83 in [7], the tangent plane $H_v$ at $P_X(v)$ of $\text{bd}(X)$ at $P_X(v)$ consists of the tangent vectors at $P_X(v)$ of all curves passing $P_X(v)$. Namely, for any unit vector $u$ for which $P_X(v) + u \in H_v$, there is $\{u_k\}_{k \geq 0} \subseteq \text{bd}(X)$ with $\lim_{k \to \infty} \frac{u_k - P_X(v)}{||u_k - P_X(v)||} = u$. We complete the proof by contradiction. Hence suppose there is $v_0 \in \text{bd}(S)$ with $\text{int}(X) \cap \text{line}(v_0, P_X(v_0)) = \emptyset$. Then by the convex set separation Theorem 11.3 on page 97 in [2], there is a hyperplane $H$ separating $X$ and line($v_0, P_X(v_0)$) properly. As a result, $H$ must contain line($v_0, P_X(v_0)$). Let $n$ be the unit normal vector of $H$ with $\angle(n, z - P_X(v_0)) \geq \pi/2$ for $z \in X$. Then $n \in H_{v_0}$ since $v_0 - P_X(v_0)$ is a normal vector of $H_{v_0}$. Thus, it is not possible to find a sequence $\{u_k\}_{k \geq 0} \subseteq \text{bd}(X)$ with $\lim_{k \to \infty} \frac{u_k - P_X(v_0)}{||u_k - P_X(v_0)||} = n$, which yields a contradiction. \hfill \Box

References

[1] Godsil, C., & Royle, G. (2001). Algebraic Graph Theory. Springer-Verlag, New York.

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

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

[4] Zhang, Y., Lou, Y., Hong, Y., & Xie, L. (2015). Distributed projection-based algorithms for source localization in wireless sensor networks. IEEE Trans. Wireless Commun., to appear.
[5] Bertsekas, D. P., & Tsitsiklis. J. (1989). *Parallel and Distributed Computation: Numerical Methods*. Princeton Hall, Englewood Cliffs, NJ.

[6] Bauschke, H. (1990). Projection algorithms and monotone operators. The Thesis, Available at: [http://docserver.carma.newcastle.edu.au/173/](http://docserver.carma.newcastle.edu.au/173/)

[7] do Carmo, M. (1976). *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Englewood Cliffs, NJ.

[8] Danskin, J. (1966). The theory of max-min, with applications. *SIAM J. Applied Mathematics*, 14(4), 641-664.

[9] Pardalos, P. M., & Romeijn. H. E. (2002). *Handbook of Global Optimization*. Volume 2. Kluwer Academic Publishers, London.

[10] Francis, R. L., McGinnis, L. F., & White. J. A. (1992). *Facility Layout and Location: An Analytical Approach*. 2 edition. Prentice Hall, Englewood Cliffs, NJ.

[11] Deutsch, F. (1983). Rate of convergence of the method of alternating projections. in *Parametric Optim. Approx.*, B. Brosowski and F. Deutsch, Eds. Birkhäuser, Basel, Switzerland, 76, 96-107.

[12] Gubin, L., Polyak, B., & Raik. E. (1967). The method of projections for finding the common point of convex sets. *U.S.S.R Comput. Math. Math. Phys.* 7(6), 1211-1228.

[13] Lou, Y., & Hong, Y. (2012). Target containment control of multi-agent systems with random switching interconnection topologies. *Automatica*, 48(5), 879-885.

[14] Lou, Y., Shi, G., Johansson, K. H., & Hong, Y. (2013). Convergence of random sleep algorithms for optimal consensus. *Systems & Control Letters*, 62(12), 1196-1202.

[15] Lou, Y., Shi, G., Johansson, K. H., & Hong, Y. (2014). Approximate projected consensus for convex intersection computation: Convergence analysis and critical error angle. *IEEE Trans. Autom. Control*, 59(7), 1722-1736.

[16] Shi, G., & Johansson, K. H. (2012). Randomized optimal consensus of multi-agent systems. *Automatica*, 48(12), 3018-3030.
[17] Shi, G., Johansson, K. H., & Hong, Y. (2013). Reaching an optimal consensus: Dynamical systems that compute intersections of convex sets. *IEEE Trans. Autom. Control*, 58(3), 610-622.

[18] Nedić, A., & Ozdaglar, A. (2009). Distributed subgradient methods for multi-agent optimization. *IEEE Trans. Autom. Control*, 54(1), 48-61.

[19] Nedić, A., Ozdaglar, A., & Parrilo, P. A. (2010). Constrained consensus and optimization in multi-agent networks. *IEEE Trans. Autom. Control*, 55(4), 922-938.

[20] Nedić, A., & Ozdaglar, A. (2008). Approximate primal solutions and rate analysis for dual subgradient methods. *SIAM J. Optim.*, 19(4), 1757-1780.

[21] Lin, P., & Ren, W. (2012). Distributed shortest distance consensus problem in multi-agent systems. In: *51st IEEE Conference on Decision and Control*, 2012, pp. 4696-4701.

[22] Johansson, B., Rabi, M., & Johansson, M. (2009). A randomized incremental subgradient method for distributed optimization in networked systems. *SIAM J. Optim.*, 20(3), 1157-1170.

[23] Hu, J., & Feng, G. (2010). Distributed tracking control of leader-follower multi-agent systems under noisy measurement. *Automatica*, 46(8), 1382-1387.

[24] Shi, G., & Hong, Y. (2009). Global target aggregation and state agreement of nonlinear multi-agent systems with switching topologies. *Automatica*, 45(5), 1165-1175.

[25] Shi, G., & Johansson, K. H. (2013). Multi-agent robust consensus: Convergence analysis and application. *SIAM J. Control Optim.*, 51(5), 3673-3691.

[26] Meng, W., Xiao, W., & Xie, L. (2013). A projection based fully distributed approach for source localization in wireless sensor networks. *Adhoc & Sensor Wireless Networks*, 18(1/2), 131-158.

[27] Lu, J., Tang, C. Y., Regier, P. R., & Bow, T. D. (2011). Gossip algorithms for convex consensus optimization over networks. *IEEE Trans. Autom. Control*, 56(12), 2917-2923.

[28] Lu, J., & Tang, C. Y. (2012). Zero-gradient-sum algorithms for distributed convex optimization: The continuous-time case. *IEEE Trans. Autom. Control*, 57(9), 2348-2354.
[29] Jakovetic, D., Xavier, J., & Moura, J. M. F. (2011). Cooperative convex optimization in networked systems: Augmented lagrangian algorithms with directed gossip communication. *IEEE Trans. Signal Processing*, 59(8), 3889-3902.

[30] Wang, J., & Elia, N. (2010). Control approach to distributed optimization. in: *Forty-Eighth Annual Allerton Conference*, 557-561.

[31] Wang, J., & Elia, N. (2011). A control perspective for centralized and distributed convex optimization. in *50th IEEE Conference on Decision and Control and European Control Conference*, 3800-3805.

[32] Gharesifard, B., & Cortés, J. (2014). Distributed continuous-time convex optimization on weight-balanced digraphs. *IEEE Trans. Autom. Control*, 59(3), 781-786.

[33] Gharesifard, B., & Cortés, J. (2013). Distributed convergence to Nash equilibria in two-network zero-sum games. *Automatica*, 49(6), 1683-1692.

[34] Kvaternik, K., & Pavel, L. (2012). A continuous-time decentralized optimization scheme with positivity constraints. in *51st IEEE Conference on Decision and Control*, 6801-6807.

[35] Droge, G., Kawashima, H., & Egerstedt, M. (2014). Continuous-time proportional-integral distributed optimization for networked systems. *Journal of Decision and Control*, 1(3), 191-213.

[36] Slotine, J.-J. E., & Li, W. (1991). *Applied Nonlinear Control*. Prentice Hall, Englewood Cliffs, NJ.