Complex Constrained Consensus

Ji Liu  Angelia Nedić  Tamer Başar

Abstract—In a network of $m > 1$ agents, constrained consensus means that all $m$ agents reach an agreement on a specific value of some quantity via local interactions while their states are restricted to lie in different closed convex sets. This paper formulates and solves two generalized versions of the basic constrained consensus problem. The first version deals with the case when the constraint set of each agent is complex so that the projection operation on the whole constraint set is computationally expensive or even prohibitive. The second version models the constrained flocking problem in which each agent can only sense the current headings of its neighbors and independently updates its heading at times determined by its own clock. Two constrained consensus algorithms are proposed for the two versions. Both are guaranteed to reach a consensus under appropriate assumptions.

I. INTRODUCTION

Recently, there has been considerable interest in developing algorithms intended to cause a group of $m > 1$ agents to reach a consensus in a distributed manner [1]–[14]. In a typical consensus seeking process, the agents in a given group all aim to agree on a specific value of some quantity. Each agent initially has only limited information available. The agents then cooperatively reach a consensus via local interactions only with their neighboring agents. The interactions among the agents can be either communication-based [4] or sensing-based [3], depending on the problem of interest. The former allows active communications among agents and the latter case is sometimes called flocking in which each agent can only sense the variables (e.g., headings) of its neighboring agents. For a survey covering the works in this area, see [15].

One particular type of consensus process which has received much attention lately is called constrained consensus [16]–[20]. In [17], a continuous-time constrained consensus algorithm was proposed using logarithmic barrier functions. In [18] and [20], discrete-time constrained consensus algorithms were presented for a special case in which the variable of each agent is a scalar quantity. Among all the proposed strategies, there is one particular approach that has motivated the research in this paper. The work of [16] proposed a simple discrete-time model of constrained consensus problem in which the variable of each agent is restricted to lie in different closed convex sets. The algorithm in [16] relies on convergence properties of doubly stochastic matrices. As a result, the implementation of the algorithm requires each agent to know an upper bound on the number of neighbors of each of its current neighbors. An extension of the constrained consensus algorithm developed in [19] gets around this limitation and also takes into account bounded transmission delays.

In some cases, the constrained set $C_i$ for agent $i$ can be the intersection of a large number of nonempty closed convex sets which are called components of $C_i$. These components may not be known in advance, but are revealed through a pre-specified sequence, or these components make the projection operation on $C_i$ computationally expensive or even prohibitive. Thus, at each iteration, each agent can only perform the projection operation on some component of $C_i$. Such a situation has been considered recently in the literature of distributed optimization [21], [22] in which a probabilistic projection algorithm was proposed for solving convex optimization problems. Prompted by this, we introduce a deterministic projection algorithm for the constrained consensus problem which gets around expensive projection operations.

Most existing constrained consensus algorithms [16], [18]–[20] have implicitly imposed two assumptions: (1) Each agent can actively send out the values of its state to its nearby agents. (2) Each agent’s state is able to change discontinuously. Although these assumptions make sense in the context of distributed computation, they cannot be justified in flocking problems where each agent’s state represents the direction of its motion which must change continuously. An asynchronous version of the unconstrained flocking problem in which each agent’s heading changes (piece-wise) continuously, is addressed in [23]. Inspired by this, we consider an asynchronous constrained flocking problem in which each agent can only sense the current headings of its neighboring agents and independently updates its heading at times determined by its own clock, with the assumption that headings cannot change discontinuously.

The main contribution of this paper is to formulate and solve the two generalized versions of the constrained consensus problem just described, the first is communication-based and the second is sensing-based. Two constrained consensus algorithms are proposed for the two versions. Both are guaranteed to solve the problem under appropriate assumptions.

A. Preliminaries

For an integer $m \geq 1$, we write $[m]$ to denote the index set $\{1, 2, \ldots, m\}$. We view vectors as column vectors. We use $x'$ to denote the transpose of a vector $x$ and, similarly,
we use $A'$ for the transpose of a matrix $A$. We say that a vector is stochastic if its entries are nonnegative and sum to 1. A square matrix is said to be stochastic if both $A$ and its transpose $A'$ are stochastic. The $i$th entry of a matrix $A$ will be denoted by $a_{ij}$ and, also, by $[A]_{ij}$ when convenient. We use $I$ to denote the identity matrix and use $1$ to denote the vector with all entries equal to 1, while the size of the matrix and vector is to be understood from the context. We use $\| \cdot \|$ for the Euclidean norm.

The Euclidean projection of a point $y$ on a convex closed set $\mathcal{Y}$ is denoted by $P_{\mathcal{Y}}[y]$, i.e., $P_{\mathcal{Y}}[y] = \arg\min_{z \in \mathcal{Y}} \|y - z\|$. We will use the following important non-expansiveness property.

**Lemma 1:** [Lemma 1 in [16]] Let $\mathcal{Y} \subset \mathbb{R}^n$ be a nonempty closed convex set. Then, for any $x \in \mathbb{R}^n$ and $z \in \mathcal{Y}$, there holds

$$\|P_{\mathcal{Y}}[x] - z\|^2 \leq \|x - z\|^2 - \|P_{\mathcal{Y}}[x] - x\|^2$$

We use $\mathcal{G}_o$ to denote the set of all directed graphs with $m$ vertices which have self-arcs at all vertices. The graph of an $m \times m$ nonnegative matrix $M$, denoted by $\gamma(M)$, is an $m$-vertex directed graph defined so that $(i, j)$ is an arc from $i$ to $j$ in the graph whenever the $j$th entry of $M$ is nonzero. Such a graph will be in $\mathcal{G}_o$ if and only if all diagonal entries of $M$ are positive.

Let $\mathcal{G}_p$ and $\mathcal{G}_q$ be two directed graphs with $m$ vertices. By the composition of $\mathcal{G}_p$ with $\mathcal{G}_q$, denoted by $\mathcal{G}_q \circ \mathcal{G}_p$, is meant that directed graph with $m$ vertices and arc set defined so that $(i, j)$ is an arc in the composition whenever there is a vertex $k$ such that $(i, k)$ is an arc in $\mathcal{G}_p$ and $(k, j)$ is an arc in $\mathcal{G}_q$. Note that composition is an associative binary operation; because of this, the definition extends unambiguously to any finite sequence of directed graphs with the same vertex set. Composition is defined so that for any pair of $m \times m$ nonnegative matrices $M_1$ and $M_2$, there holds $\gamma(M_2M_1) = \gamma(M_2) \circ \gamma(M_1)$. If we focus exclusively on graphs in $\mathcal{G}_o$, more can be said. In this case, the definition implies that the arcs of both $\mathcal{G}_p$ and $\mathcal{G}_q$ are arcs of $\mathcal{G}_q \circ \mathcal{G}_p$; the converse is false.

A directed graph $\mathcal{G}$ is strongly connected if there is a directed path between each pair of distinct vertices. We say that a finite sequence of directed graphs $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_p$ with the same vertex set is jointly strongly connected if the composition $\mathcal{G}_p \circ \mathcal{G}_{p-1} \circ \cdots \circ \mathcal{G}_1$ is strongly connected. We say that an infinite sequence of directed graphs $\mathcal{G}_1, \mathcal{G}_2, \ldots$ with the same vertex set is repeatedly jointly strongly connected if there exists a finite positive integer $l$ such that for any integer $k \geq 0$, the finite sequence $\mathcal{G}_{kl+1}, \mathcal{G}_{kl+2}, \ldots, \mathcal{G}_{(k+1)l}$ is jointly strongly connected. These notions of connectivity are more or less well known in the study of distributed averaging and consensus problems [24], although the form of the condition may vary slightly from publication to publication. See for example [25].

**B. Organization**

The remainder of this paper is structured as follows. We first review the basic constrained consensus problem and the model studied in [16] in Section II. Then, we consider two generalized versions of the model. A deterministic projection algorithm for complex constraints is presented in Section III and an asynchronous constrained flocking problem is studied in Section IV. Finally, the key technical results are given in Section V. Detailed proofs of the main theorems are omitted due to space limitations and will be given in the full length version of this paper.

**II. Basic Constrained Consensus**

Consider a network of $m > 1$ autonomous agents with the constraint that each agent is able to receive information only from its “neighbors”. Neighbor relations are characterized by a time-dependent directed graph $\mathcal{N}(t)$ with $m$ vertices and a set of arcs defined so that there is an arc in the graph from vertex $j$ to vertex $i$ at time $t$ just in case agent $j$ is a neighbor of agent $i$ at time $t$. Thus, the directions of arcs represent the directions of information flow, and the neighbors of an agent $i$ at time $t$ have the same labels as the vertices in $\mathcal{N}(t)$ which are adjacent to vertex $i$. For simplicity we always take each agent to be a neighbor of itself. Thus, $\mathcal{N}(t)$ has self-arcs at all vertices for each time $t$. Each agent $i$ has a real, time-dependent state vector $x_i(t)$ taking values in $\mathbb{R}^n$, and we assume that the information agent $i$ can receive from a neighbor $j$ is only the state vector of neighbor $j$. We also assume that each agent $i$’s state vector $x_i(t)$ is constrained to lie in a nonempty closed convex set $\mathcal{C}_i \subset \mathbb{R}^n$ which is known only to agent $i$. The constrained consensus problem is to devise local algorithms, one for each agent, which will enable all $m$ agents to iteratively reach a consensus in the intersection $\mathcal{C} = \bigcap_{i=1}^m \mathcal{C}_i$. For this to be possible, it is clearly necessary that $\mathcal{C}$ be a nonempty set, which we assume to be the case.

A constrained consensus algorithm was proposed in [16] and its extension was developed in [19]. One implementation of the extended algorithm is as follows. Suppose that time is discrete in that $t$ takes values in $\{1, 2, \ldots\}$. Each agent $i$ iteratively updates its state by setting

$$x_i(t+1) = \Pi_{\mathcal{C}_i} \left( \frac{1}{n_i(t)} \sum_{j \in \mathcal{N}_i(t)} x_j(t) \right), \quad t \geq 1 \tag{1}$$

where $\mathcal{N}_i(t)$ is the set of labels of those agents which are neighbors of agent $i$ at time $t$ and $n_i(t)$ is the number of indices in $\mathcal{N}_i(t)$, or equivalently, the in-degree of vertex $i$ in $\mathcal{N}(t)$.

The following proposition states that the algorithm (1) can solve the basic constrained consensus problem under repeatedly jointly strongly connectedness.

**Proposition 1:** Suppose that the intersection $\mathcal{C} = \bigcap_{i=1}^m \mathcal{C}_i$ is nonempty. For any trajectory of the synchronous system defined by (1) whose associated sequence of neighbor graphs $\mathcal{N}(1), \mathcal{N}(2), \ldots$ is repeatedly jointly strongly connected, there

1465
is a constant vector \( x_{ss} \in \mathbb{C} \) for which
\[
\lim_{t \to \infty} x_i(t) = x_{ss}
\]
This result is a consequence of Proposition 2 in Section V.
In the sequel, we will consider two generalized but more complicated versions of the basic constrained consensus problem for which the proofs in [16], [19] become somewhat restrictive and possibly not applicable.

### III. Constrained Consensus with Complex Constraints

In this section, we consider the cases in which the constrained set \( C_i \) for agent \( i \) may not be known in advance or \( C_i \) may have a large number of components, which makes the projection operation on \( C_i \) computationally expensive and prohibitive.

Suppose that each constrained set \( C_i \) is the intersection of finitely many closed convex sets. To be more precise,
\[
C_i = \bigcap_{j=1}^{k_i} C_{ij}
\]
where \( k_i \) is a positive integer and each \( C_{ij} \) is a closed convex set. The component-based projection algorithm is as follows:
\[
x_i(t + 1) = P_{C_{hi(t)}} \left[ \frac{1}{n_i(t)} \sum_{j \in N_i(t)} x_j(t) \right], \quad t \geq 1
\]
where \( h_i(t) \) is an integer in \( [k_i] \). We assume that the specific sequence of \( h_i(t) \) which occurs during a given process is deterministic. Possibly the simplest implementation is to let \( h_i(t) \) take values from \( [k_i] \) periodically.

To state our result, let us say that a sequence of \( h_i(t) \) is complete if each integer \( j \in [k_i] \) appears in the sequence at least once. An infinite sequence of \( h_i(t) \) is repetitively complete with period \( T \) if each successive subsequence of \( h_i(t) \) of length \( T \) in the sequence is complete.

**Theorem 1:** Suppose that the intersection \( C = \bigcap_{i=1}^{m} C_i \) is nonempty and that the sequence of \( h_i(t) \) is repetitively complete for all \( i \in [m] \). For any trajectory of the synchronous system defined by (2) whose associated sequence of neighbor graphs \( \mathbb{N}(1), \mathbb{N}(2), \ldots \) is repeatedly jointly strongly connected, there is a constant vector \( x_{ss} \in \mathbb{C} \) for which
\[
\lim_{t \to \infty} x_i(t) = x_{ss}
\]

### IV. Asynchronous Constrained Flocking

In this section, we consider an asynchronous constrained flocking problem.

Suppose that all \( m \) agents are moving in the plane with different headings. Each agent \( i \)'s heading \( \theta_i \) is restricted to lie in a nonempty closed convex interval \( C_i \subset [0, 2\pi) \).

Following the work of [23], we associate with each agent \( i \) a strictly increasing, infinite sequence of event times \( t_{i1}, t_{i2}, \ldots \)

which is determined by its own clock. Each agent is able to compute its desired heading (or next “way-point” which will be defined later) at its own event times using a simple local rule based on its own current heading and the current headings of its neighbors. We assume that for \( i \in [m] \), agent \( i \)'s event times satisfy the constraints
\[
T_i \geq t_{i(k+1)} - t_{ik} \geq T_i, \quad k \geq 1
\]
where \( T_i \) and \( T_i \) are positive numbers such that \( T_i > T_i \).
Thus, the event times of agent \( i \) are distinct and the difference between any two successive event times cannot be too large. It is not assumed that the agents’ clocks are synchronized or that the event times are evenly spaced. Thus, two agents may have completely different event time sequences.

The update rules of each agent \( i \)'s heading are as follows.

At its \( k \)th event time \( t_{ik} \), agent \( i \) senses the current headings of its current neighbors \( \theta_j(t_{ik}), j \in N_i(t_{ik}) \), and computes its \( k \)th way-point
\[
w_i(t_{ik}) = P_{C_i} \left[ \frac{1}{n_i(t_{ik})} \sum_{j \in N_i(t_{ik})} \theta_j(t_{ik}) \right], \quad k \geq 1
\]

Agent \( i \) then changes its heading from \( \theta_i(t_{ik}) \) to \( w_i(t_{ik}) \) on the time interval \( (t_{ik}, t_{i(k+1)}) \). We assume that the change is monotonic, continuous, and can be completed within the interval. To be more precise, there is a continuous signal \( \mu_i : [0, \infty) \to [0, 1] \) such that
\[
\theta_i(t) = \theta_i(t_{ik}) + \mu_i(t)(w_i(t_{ik}) - \theta_i(t_{ik})), \quad t \in (t_{ik}, t_{i(k+1)})
\]
and \( \mu_i \) changes continuously from 0 to 1 on \( (t_{ik}, t_{i(k+1)}) \), \( k \geq 1 \).
Thus,
\[
\theta_i(t_{i(k+1)}) = w_i(t_{ik}), \quad k \geq 1
\]
and \( \theta_i \) is continuous on \( [t_{i1}, \infty) \).

To characterize the behavior of the above system, we need a common time scale on which the neighbor relations among the \( m \) agents can be redefined. For this, let \( t_1 = \max \{t_{i1}\} \) and write \( T_i \) for the event times of agent \( i \) which are greater than or equal to \( t_1 \). Let \( T \) denote the set of all event times of all \( m \) agents which are greater than or equal to \( t_1 \). Relabel the times in \( T \) so that \( t_k < t_{k+1} \) for \( k \geq 1 \). For each \( i \in [m] \) and \( \tau \in T \), we define the extended neighbor set
\[
\bar{N}_i(\tau) = N_i(\tau), \quad \tau \in T_i
\]
\[
\bar{N}_i(\tau) = \{i\}, \quad \tau \notin T_i
\]
Thus, \( \bar{N}_i(\tau) \) coincides with \( N_i(\tau) \) whenever \( \tau \) is an event time of agent \( i \) and the singleton \( \{i\} \) otherwise. These defined neighbor relations can be conveniently described by a time-dependent directed graph \( \bar{\mathbb{N}}(\tau), \tau \in \{1, 2, \ldots \} \), with vertex set \( \mathbb{V} = [m] \) and arc set \( \bar{A}(\tau) \subset \mathbb{V} \times \mathbb{V} \) which is defined so that \( (i, j) \) is an arc from \( i \) to \( j \) whenever label \( i \) is in \( \bar{N}_j(\tau) \).

Thus, each \( \bar{\mathbb{N}}(\tau) \) is a directed graph on \( m \) vertices with self-arcs at all vertices. We call \( \bar{\mathbb{N}}(\tau) \) the extended neighbor graph of the asynchronous system (3)-(4) at time \( \tau \).

The following theorem states that a consensus can be reached in the asynchronous constrained flocking problem.
Theorem 2: Suppose that the intersection \( C = \bigcap_{i=1}^{m} C_i \) is nonempty. For any trajectory of the asynchronous system defined by (3) and (4) whose associated sequence of extended neighbor graphs \( \mathcal{N}(1), \mathcal{N}(2), \ldots \) is repeatedly jointly strongly connected, there is a constant \( \theta_{ss} \in C \) for which
\[
\lim_{t \to \infty} \theta_i(t) = \theta_{ss}
\]

V. ANALYSIS

The aim of this section is to present some key technical results which will play important roles in the proofs of the main theorems. We first present some important features of stochastic matrices for unconstrained consensus problems.

A. Absolute Probability Sequence

Most of the discussion in this part is based on the work of Touri reported in his thesis [26]. We begin with the notion of an absolute probability sequence associated with a sequence of stochastic matrices. This notion was first introduced by Kolmogorov [27].

Definition 1: Let \( \{A(t)\} \) be a sequence of stochastic matrices. A sequence of stochastic vectors \( \{\pi(t)\} \) is an absolute probability sequence for \( \{A(t)\} \) if
\[
\pi'(t) = \pi'(t+1)A(t), \quad t \geq 1
\]

It has been shown by Blackwell [28] that every sequence of stochastic matrices has an absolute probability sequence. In general, a sequence of stochastic matrices may have more than one absolute probability sequence (see the example in [29]). If we focus exclusively on “ergodic” sequences of stochastic matrices, more can be said.

We say that a sequence of stochastic matrices \( \{A(t)\} \) is \textit{ergodic} if for any \( t \geq 1 \), there holds
\[
\lim_{\tau \to \infty} A(\tau)A(\tau-1) \cdots A(t+1)A(t) = 1\phi'(t)
\]
where \( \phi(t) \) is a stochastic vector. As an immediate consequence of Blackwell’s result [28], every ergodic sequence of stochastic matrices has an absolute probability sequence. In fact, if the sequence \( \{A(t)\} \) is ergodic, the vector sequence \( \{\phi(t)\} \) satisfying (6) is the unique absolute probability sequence for \( \{A(t)\} \) (see Lemma 1 in [29]).

The conditions under which a sequence of stochastic matrices is ergodic have been very extensively studied in connection with unconstrained consensus problems [3]. Specifically, as the iterates \( x(t) \) are related over time by the following linear dynamics:
\[
x(t) = A(t)A(t-1) \cdots A(s+1)A(s)x(s), \quad t \geq s \geq 1
\]
the convergence of the iterates generated by the algorithm is related to the convergence of the matrix products \( A(t)A(t-1) \cdots A(1) \), as \( t \to \infty \). In particular, when the matrices \( A(t)A(t-1) \cdots A(1) \) converge to a rank one matrix, the iterates \( x(t) \) reach a consensus. Concretely, some conditions on the matrices \( A(t) \) and the graphs of \( A(t) \) that yield an ergodic sequence \( \{\pi(t)\} \) are given in the following assumption.

Assumption 1: Let \( \{A(t)\} \) be a sequence of \( m \times m \) stochastic matrices that satisfy the following conditions:

(a) (Aperiodicity) The diagonal entries of each \( A(t) \) are positive, i.e., \( a_{ii}(t) > 0 \) for all \( t \) and \( i \in [m] \).
(b) (Uniform Positivity) There is a scalar \( \beta > 0 \) such that \( a_{ij}(t) \geq \beta \) whenever \( a_{ij}(t) > 0 \).
(c) (Irreducibility) The sequence of graphs \( \{\gamma(A(t))\} \) is repeatedly jointly strongly connected.

Under Assumption 1, the following result is well known.

Lemma 2: [Lemma 5.2.1 in [26], Lemma 5 in [25]] Suppose that Assumption 1 holds. Then,
\[
\lim_{t \to \infty} A(t) \cdots A(k+1)A(k) = 1\phi'(k)
\]
for all \( k \geq 1 \), where each \( \phi(k) \) is a stochastic vector. Furthermore, the convergence rate is geometric, i.e., for all \( t \geq k \geq 1 \),
\[
\|A(t) \cdots A(k+1)A(k) - 1\phi'(k)\| \leq Cq^{t-k}
\]
where the constants \( C > 0 \) and \( q \in [0, 1) \) depend only on \( m \) and \( \beta \).

It is worth emphasizing that the results in Lemma 2 still hold if Assumption 1(c) is replaced by the assumption that the sequence of graphs \( \{\gamma(A(t))\} \) is repeatedly jointly rooted (see Theorem 3 in [8]).

We use \( \Phi(t, \tau) \) to denote the discrete-time state transition matrix of \( A(t) \), i.e.,
\[
\Phi(t, \tau) = \begin{cases}
A(t-1) \cdots A(\tau+1)A(\tau) & \text{if } k > j \\
I & \text{if } k = j
\end{cases}
\]
Note that the sequence \( \{\phi(t)\} \) in Lemma 2 is the unique absolute probability sequence for the sequence \( \{A(t)\} \). The following result is a direct consequence of Lemma 2.

Corollary 1: Suppose that Assumption 1 holds. Then, there exist finite constants \( c > 0 \) and \( \lambda \in [0, 1) \) such that
\[
|\Phi(t, \tau)|_{ij} - \pi_j(\tau) - c\lambda^{t-\tau}
\]
for all \( t \geq \tau \) and \( i, j \in [m] \), where \( \{\pi(t)\} \) is the unique absolute probability sequence for \( \{A(t)\} \) and \( \pi_j(t) \) denotes the \( j \)th entry of \( \pi(t) \).

More can be said. It has been shown by Touri [26] that the (unique) absolute probability sequence of an ergodic sequence of stochastic matrices has the property that the entries of the absolute probability vectors are uniformly bounded away from zero.

Lemma 3: [Theorem 4.8 in [26]] Let \( \{\pi(t)\} \) be an absolute probability sequence for \( \{A(t)\} \). Suppose that Assumption 1 holds. Then, there is a positive scalar \( \delta > 0 \) such that \( \pi_i(t) \geq \delta \) for all \( i \in [m] \) and \( t \).

A class of stochastic matrices which have this property was introduced in [26] (as class \( P^* \)).
B. Technical Results

In this part, we consider an extension of the algorithm (1) in which each agent \( i \) iteratively updates its state by setting

\[
x_i(t + 1) = P_{C_i} \left[ \sum_{j=1}^{m} a_{ij}(t)x_j(t) \right], \quad t \geq 1 \tag{7}
\]

where \( a_{ij}(t) \) are nonnegative weights compliant with the neighbor graph \( \mathbb{N}(t) \), i.e., \( a_{ij}(t) > 0 \) whenever \( (j, i) \) is an arc in \( \mathbb{N}(t) \). It can be seen that the algorithm (1) is a special case of (7). Let \( A(t) = [a_{ij}(t)] \) be the \( m \times m \) matrix whose \( ij \) entry is \( a_{ij}(t) \). A key technical result is as follows.

**Proposition 2:** Suppose that Assumption 1 holds and that the intersection \( C = \bigcap_{i=1}^{m} C_i \) is nonempty. Then, for any trajectory of the synchronous system defined by (7), there is a constant vector \( x_{ss} \in C \) for which

\[
\lim_{t \to \infty} x_i(t) = x_{ss}
\]

Note that Proposition 1 is an immediate consequence of this proposition. The proof for a more restrictive case, in which \( C \) has nonempty interior, was given in [19]. Here, we relax the assumption of the set having nonempty interior thus allowing it to be a singleton.

To prove Proposition 2, we first rewrite (7) as follows. For each \( i \in [m] \), set

\[
v_i(t + 1) = \sum_{j=1}^{m} a_{ij}(t)x_j(t) \tag{8}
\]

\[
e_i(t) = P_{C_i}[v_i(t)] - v_i(t) \tag{9}
\]

Then,

\[
x_i(t) = P_{C_i}[v_i(t)] = v_i(t) + e_i(t) \tag{10}
\]

**Lemma 4:** Suppose that Assumption 1 holds and that the intersection \( C = \bigcap_{i=1}^{m} C_i \) is nonempty. Then, for each \( i \in [m] \), there holds

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

**Proof:** Let \( z \) be a vector in \( C \). Such a vector exists since \( C \) is nonempty. Since \( C = \bigcap_{i=1}^{m} C_i \), it follows that \( z \in C_i \) for all \( i \in [m] \). Then, by Lemma 1, for all \( i \in [m] \), there holds

\[
\|P_{C_i}[v_i(t)] - z\| \leq \|v_i(t) - z\| - \|P_{C_i}[v_i(t)] - v_i(t)\|
\]

which, from (10), simplifies to

\[
\|x_i(t) - z\| \leq \|v_i(t) - z\|^2 - \|e_i(t)\|^2 \tag{11}
\]

Since Assumption 1 holds, by Lemma 3, the sequence of stochastic matrices \( \{A(t)\} \) has an absolute probability sequence \( \{\pi(t)\} \) for which there exists a positive scalar \( \delta > 0 \) such that \( \pi_i(t) \geq \delta \) for all \( i \in [m] \) and \( t \geq 1 \). Note that since \( A(t) \) is stochastic, from (8) it is straightforward to verify that

\[
\|v_i(t + 1) - z\|^2 \leq \sum_{j=1}^{m} a_{ij}(t)\|x_j(t) - z\|^2
\]

With this inequality, we have

\[
\sum_{i=1}^{m} \pi_i(t + 1)\|v_i(t + 1) - z\|^2
\]

\[
\leq \sum_{i=1}^{m} \pi_i(t + 1)\sum_{j=1}^{m} a_{ij}(t)\|x_j(t) - z\|^2
\]

\[
= \sum_{i=1}^{m} \pi_i(t)\|x_i(t) - z\|^2 \tag{12}
\]

where, in the last equality, we use

\[
\pi_j(t) = \sum_{i=1}^{m} \pi_i(t + 1) a_{ij}(t)
\]

which is an immediate consequence of (5). From (11) and (12), it follows that

\[
\sum_{i=1}^{m} \pi_i(t + 1)\|x_i(t + 1) - z\|^2
\]

\[
\leq \sum_{i=1}^{m} \pi_i(t)\|x_i(t) - z\|^2 - \sum_{i=1}^{m} \pi_i(t + 1)\|e_i(t + 1)\|^2
\]

which implies that

\[
\sum_{i=1}^{m} \sum_{t=1}^{\infty} \pi_i(t)\|e_i(t)\|^2 \leq \sum_{i=1}^{m} \pi_i(1)\|x_i(1) - z\|^2 < \infty
\]

Since \( \pi_i(t) \geq \delta > 0 \) for all \( i \in [m] \) and \( t \geq 1 \), it follows that \( \lim_{t \to \infty} \|e_i(t)\| = 0 \) for all \( i \in [m] \). ■

**Remark 1:** The proof of Lemma 4 introduces a Lyapunov comparison function which makes use of the property of the (unique) absolute probability sequence of an ergodic sequence of stochastic matrices. This Lyapunov comparison function is more general than those in [16] and [19], and is the key idea with which we will be able to prove the main result of this paper. It is worth emphasizing that the proof of Lemma 4 does not require \( C_i \) to be fixed. It is straightforward to verify that in the more general case when the constrained set of each agent \( i \) is a time-varying set \( C_i(t) \), the lemma still holds with the same arguments as long as \( C \subset C_i(t) \) for all \( i \in [m] \) and \( t \).

**Lemma 5:** Suppose that Assumption 1 holds. Let \( \{\pi(t)\} \) be an absolute probability sequence for \( \{A(t)\} \) and set

\[
y(t) = \sum_{i=1}^{m} \pi_i(t)x_i(t).
\]

If \( \lim_{t \to \infty} e_i(t) = 0 \) for all \( i \in [m] \), then for each \( i \in [m] \), there holds

\[
\lim_{t \to \infty} \|x_i(t) - y(t)\| = 0
\]

The proof of this lemma is omitted due to space limitations and will be given in the full length version of this paper.

**Remark 2:** Lemma 5 does not require that the entries of the absolute probability vectors are uniformly bounded away from zero, though this is a consequence of Assumption 1.

We are now in a position to prove Proposition 2.

**Proof of Proposition 2:** Since Assumption 1 holds, by Lemma 3, the sequence of stochastic matrices \( \{A(t)\} \) has an absolute probability sequence \( \{\pi(t)\} \) for which there exists
a positive scalar $\delta > 0$ such that $\pi_i(t) \geq \delta$ for all $i \in [m]$ and $t \geq 1$.

Let $z$ be a vector in $C$. Such a vector exists since $C$ is nonempty. From the proof of Lemma 4, it can be seen that the sequence $\{\sum_{i=1}^{m} \pi_i(t)\|x_i(t) - z\|^2\}$ is a non-increasing sequence for any $z \in C$. Then, the sequence $\{x_i(t)\}$ is bounded for all $i \in [m]$. Thus, each sequence $\{x_i(t)\}$, $i \in [m]$, has accumulation points. By Lemma 5, the sequences $\{x_i(t)\}$, $i \in [m]$, have the same accumulation points. Since the sequence $\{x_i(t)\}$ lies in the closed set $C$, $i \in [m]$, it follows that the accumulation points of the $x_i(t)$ lie in $C$. Since the accumulation points of the $m$ sequences coincide, the accumulation points must lie in the intersection $C$.

Next we show that the sequence $\{x_i(t)\}$ has only one accumulation point. To prove this, we suppose that, to the contrary, there are two accumulation points for the sequence. Let $\{s_\tau\}$ and $\{\tau_s\}$ be the time subsequences along which $\{x_i(t)\}$ and $\{x_i(t)\}$ converge to $z_1 \in C$ and $z_2 \in C$ respectively. Without loss of generality, we assume that $t_s > \tau_s$ for all $s \geq 1$. From the preceding discussion, it follows that for any $z \in C$, there holds

$$\sum_{i=1}^{m} \pi_i(t_s)\|x_i(t_s) - z\|^2 \leq \sum_{i=1}^{m} \pi_i(\tau_s)\|x_i(\tau_s) - z\|^2$$

Set $z = z_2$. Since $\pi_i(t) \geq \delta$ for all $i \in [m]$ and $t \geq 1$, it follows that

$$\delta \sum_{i=1}^{m} \|x_i(t_s) - z_2\|^2 \leq \sum_{j=1}^{m} \pi_j(\tau_s)\|x_j(\tau_s) - z_2\|^2 \leq \sum_{j=1}^{m} \|x_j(\tau_s) - z_2\|^2$$

Let $s \to \infty$. Then, there holds

$$\delta \sum_{i=1}^{m} \|z_1 - z_2\|^2 \leq 0$$

which implies that $z_1 = z_2$. But this is a contradiction. Therefore, each sequence $\{x_i(t)\}$, $i \in [m]$, has only one accumulation point and thus must be convergent.

VI. CONCLUDING REMARKS

This paper has formulated two generalized versions of the constrained consensus problem. Two deterministic constrained consensus algorithms have been proposed for the two problems. A new proof method has been introduced which is used to generalize existing results in the literature and prove convergence of the two proposed algorithms. To quantify convergence rates of the proposed algorithms is a subject of future work.

REFERENCES


