

Constrained Consensus and Alternating Projections

Asuman Ozdaglar
Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02142
Email: asuman@mit.edu

Abstract—We study the problem of reaching a consensus in the estimates of multiple agents forming a network with time-varying connectivity. Our main focus is on constrained consensus problems where the estimate of each agent is constrained to lie in a different closed convex constraint set. We consider a distributed “projected consensus algorithm” in which the local averaging operation is combined with projection onto the individual constraint sets. This algorithm can be viewed as an alternating projection method with weights that are varying over time and across agents. We study the convergence properties of the projected consensus algorithm. In particular, we show that under an interior point assumption, the estimates of each agent converge to the same vector, which lies in the intersection of the constraint sets.

I. INTRODUCTION

There has been much interest in distributed cooperative control problems, in which several autonomous agents collectively try to achieve a global objective. Most focus has been on the canonical consensus problem, where the goal is to develop distributed algorithms that can be used by a group of agents to reach a common decision or agreement. A widely studied algorithm in the consensus literature, proposed by Tsitsiklis [17] and Tsitsiklis *et al.* [18], involves at each time step every agent taking a weighted average of its own value with values received from some of the other agents. Despite much work on the consensus problem, the existing literature does not consider problems where the agent values are constrained to take values in a given set. Such constraints are significant in a number of applications including motion planning and alignment problems, where each agent’s position is limited to a certain region or range, and distributed constrained multi-agent optimization problems.

In this paper, we study consensus problems where the values of the agents are constrained. In particular, we assume that the values of each agent is constrained to lie in a nonempty closed convex constraint set. We propose an algorithm where the agents update their values subject to their constraints. More specifically, each agent linearly combines his value with those values received from the neighboring agents and projects the combination on his constraint set. We show that this update rule can be viewed as an alternating projection method where, at each iteration, the values are combined with weights that are varying in time and across agents, and projected on the respective constraint sets.

We provide a convergence analysis for this algorithm by decomposing the evolution of the estimates into two parts: one part involving a time-varying averaging operation (similar to the unconstrained consensus update) and can be represented in terms of *transition matrices* which is the product of the time-varying weight matrices; the second part involving the error due to projection. This decomposition allows us to represent the evolution of the estimates using linear dynamics and decouples the analysis of the effects of constraints from the convergence analysis of the transition matrices. We show that the transition matrices converge to a rank one matrix at a geometric rate. We also show that under some assumptions the projection error goes to 0. We combine these two facts to show that the agents reach consensus in the limit. We also show that under an additional interior point assumption on the intersection set X , the agent estimates converge to a common vector in X .

In addition to the works cited above, our paper is also related to recent work in providing mathematical models to understand the group behavior and flocking observed in dynamical and biological systems (see Vicsek *et al.* [19], Jadbabaie *et al.* [9], Boyd *et al.* [4], Olfati-Saber and Murray [14], Cao *et al.* [5], and Olshevsky and Tsitsiklis [15], [16]). These works assume that the agent values can be processed arbitrarily and are unconstrained. Our paper is also related to the recent game-theoretic distributed control paradigm in cooperative networks. In this approach, the agents are endowed with local utility functions that lead to a game form with a Nash equilibrium which is the same as or close to a global optimum. Various learning algorithms can then be used as distributed control schemes that will reach the equilibrium. In a recent paper, Marden *et al.* [10] used this approach for the consensus problem where agents have constraints on their values. Our paper provides an alternative approach for this problem. Finally, the constrained consensus policy developed in this paper is relevant to the recent distributed multi-agent optimization model presented in Nedić and Ozdaglar [12], [11] for unconstrained problems and can be used to extend this model to constrained optimization problems.

The paper is organized as follows. Section 2 describes the model. In Section 3, we study the convergence behavior of the transition matrices. Section 4 analyzes the projection error. Section 5 contains our main convergence result for the agent estimates. Section 6 contains concluding remarks and

future directions.

Regarding notation, a vector is viewed as a column, unless clearly stated otherwise. We denote by x_i or $[x]_i$ the i -th component of a vector x . When $x_i \geq 0$ for all components i of a vector x , we write $x \geq 0$. We denote the nonnegative orthant by \mathbb{R}_+^n , i.e., $\mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x \geq 0\}$. We write x' to denote the transpose of a vector x . The scalar product of two vectors $x, y \in \mathbb{R}^m$ is denoted by $x'y$. We use $\|x\|$ to denote the standard Euclidean norm, $\|x\| = \sqrt{x'x}$. A vector $a \in \mathbb{R}^m$ is said to be a *stochastic vector* when its components $a_j, j = 1, \dots, m$, are nonnegative and their sum is equal to 1, i.e., $\sum_{j=1}^m a_j = 1$. A set of vectors $\{a^i\}_{i \in \{1, \dots, m\}}$ is said to be *doubly stochastic* when each a^i is a stochastic vector and $\sum_{i=1}^m a_j^i = 1$ for all $j \in \{1, \dots, m\}$. A square $m \times m$ matrix A is said to be a *stochastic matrix* when each row of A is a stochastic vector. A square $m \times m$ matrix A is said to be *doubly stochastic* when both A and A' are stochastic matrices. We write $\text{dist}(\bar{x}, X)$ to denote the standard Euclidean distance of a vector \bar{x} from a set X , i.e.,

$$\text{dist}(\bar{x}, X) = \inf_{x \in X} \|\bar{x} - x\|.$$

We use the notation $P_X(\bar{x})$ to denote the projection of a vector $\bar{x} \in \mathbb{R}^n$ to a nonempty closed convex set X , i.e.,

$$P_X(\bar{x}) = \arg \min_{x \in X} \|\bar{x} - x\|.$$

II. MODEL

We consider a set of agents denoted by $V = \{1, \dots, m\}$. We assume that agent i has a nonempty closed convex constraint set X_i . We assume a slotted-time system, and we denote by $x^i(k)$ the vector estimate stored by agent i at time slot k .

The agents exchange and update their estimates as follows: To generate the estimate at time $k+1$, agent i forms a convex combination of its estimate $x^i(k)$ with the estimates received from other agents at time k , and takes the projection of this vector on the constraint set X_i . More specifically, the new estimate stored by agent i at time $k+1$ is generated according to the following relation:

$$x^i(k+1) = P_{X_i} \left[\sum_{j=1}^m a_j^i(k) x^j(k) \right], \quad (1)$$

where the vector $a^i = (a_1^i, \dots, a_m^i)' \in \mathbb{R}_+^m$ is a vector of weights and P_{X_i} denotes the projection on the set X_i .

This update rule has similarities with the classical alternating or cyclic projection methods studied in the optimization literature. The goal of the alternating projection method is, given a collection of closed and convex sets $\{X_i\}_{i \in \mathcal{I}}$, to find a vector in the intersection of these sets, $X = \bigcap_{i \in \mathcal{I}} X_i$. Alternating projection methods generate a sequence of vectors by projecting iteratively on the sets (either cyclically or with some given order), see Figure 1(a). The convergence behavior of these methods has been established by Von Neumann [13] and Aronszajn [1] for the case when the X_i are affine sets; and by Gubin, Polyak, and Raik [8] when

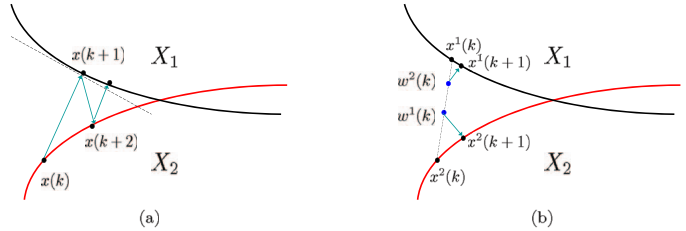


Fig. 1. Illustration of the connection between the alternating/cyclic projection method and the constrained consensus algorithm. Given two closed convex sets X_1 and X_2 , the alternating projection algorithm generates a sequence $\{x(k)\}$ by iteratively projecting onto sets X_1 and X_2 , i.e., $x(k+1) = P_{X_1}(x(k))$, $x(k+2) = P_{X_2}(x(k+1))$; see part (a). The constrained consensus algorithm generates sequences $\{x^i(k)\}$ for agent i by first combining the iterates with different weights and projecting onto respective sets X_i , i.e., $w^i(k) = \sum_{j=1}^m a_j^i(k) x^j(k)$ and $x^i(k+1) = P_{X_i}(w^i(k))$ for $i = 1, 2$.

the X_i are closed convex sets. Gubin, Polyak, and Raik [8] also provide convergence rate results for a particular form of alternating projection method. Similar rate results under different assumptions were also provided by Deutsch [6] and Deutsch and Hundal [7].

The constrained consensus algorithm [cf. Eq. (1)] generates a sequence of iterates for each agent as follows: at iteration k , agent i first forms a linear combination of the agent values using different weight vectors $a^i(k)$. Then, this combination is projected onto the constraint sets X_i of the agents to yield the iterates at iteration $k+1$. Therefore, the constrained consensus algorithm can be viewed as an alternating projection algorithm, where the iterates are combined with weights that are varying over time and across agents and then projected on individual constraint sets.

Following Blondel *et al.* [3], we adopt the following assumption on the weight vectors $a^i(k)$, $i \in \{1, \dots, m\}$.

Assumption 1: (Weights Rule) There exists a scalar η with $0 < \eta < 1$ such that for all $i \in \{1, \dots, m\}$,

- (i) $a_j^i(k) \geq \eta$ for all $k \geq 0$.
- (ii) If $a_j^i(k) > 0$, then $a_j^i(k) \geq \eta$.

This assumption states that agents at each iteration give significant weights to their own estimates and to the estimates of their neighbors. The next assumption imposes stochasticity assumptions on the weight vectors.

Assumption 2: (Doubly Stochastic Weights) The vectors $a^i(k) = (a_1^i, \dots, a_m^i)$ satisfy:

- (a) $\sum_{j=1}^m a_j^i(k) = 1$ for all i and k , i.e., the vectors $a^i(k)$ are stochastic.
- (b) $\sum_{i=1}^m a_j^i(k) = 1$ for all j and k .

We also impose some rules on the information exchange. At each update time t_k , the information exchange among the agents may be represented by a directed graph (V, E_k) with the set E_k of directed edges given by

$$E_k = \{(j, i) \mid a_j^i(k) > 0\}.$$

Note that, by Assumption 1(a), we have $(i, i) \in E_k$ for each agent i and all k . Also, we have $(j, i) \in E_k$ if and only if

agent i receives the information x^j from agent j in the time interval (t_k, t_{k+1}) .

We next formally state the connectivity assumption on the system. This assumption ensures that the information state of any agent i influences the information state of any other agent infinitely often in time.

Assumption 3: (Connectivity) The graph (V, E_∞) is connected, where E_∞ is the set of edges (j, i) representing agent pairs communicating directly infinitely many times, i.e.,

$$E_\infty = \{(j, i) \mid (j, i) \in E_k \text{ for infinitely many indices } k\}.$$

We also adopt an additional assumption that the intercommunication intervals are bounded for those agents that communicate directly. This is stated in the following.

Assumption 4: (Bounded Intercommunication Interval) There exists an integer $B \geq 1$ such that for every $(j, i) \in E_\infty$, agent j sends his/her information to the neighboring agent i at least once every B consecutive time slots, i.e., at time t_k or at time t_{k+1} or ... or (at latest) at time t_{k+B-1} for any $k \geq 0$.

III. TRANSITION MATRICES

We introduce matrices $A(s)$, whose i -th column is the vector $a^i(s)$, and for all s and k with $k \geq s$, the matrices

$$\Phi(k, s) = A(s)A(s+1) \cdots A(k-1)A(k),$$

where

$$\Phi(k, k) = A(k) \quad \text{for all } k.$$

We will use these matrices to define the evolution of the agent estimates (see Section V). The convergence properties of these matrices as $k \rightarrow \infty$ have been studied in previous work (see [17], [9], [20]). Recent work [11], [12] has provided explicit convergence rate estimates for these matrices presented in the following.

Proposition 1: (Nedić and Ozdaglar [11], [12]) Let *Weights rule, Doubly Stochastic Weights, Connectivity, Bounded Intercommunication Interval assumptions hold [cf. Assumptions 1, 2, 3, 4].* We then have:

- (a) *The limit matrices $\Phi(s) = \lim_{k \rightarrow \infty} \Phi(k, s)$ are doubly stochastic and correspond to a uniform steady state distribution for all s , i.e.,*

$$\Phi(s) = \frac{1}{m} ee' \quad \text{for all } s.$$

- (b) *The entries $[\Phi(k, s)]_i^j$ converge to $\frac{1}{m}$ as $k \rightarrow \infty$ with a geometric rate uniformly with respect to i and j , i.e., for all $i, j \in \{1, \dots, m\}$, and s, k with $k \geq s$,*

$$\left| [\Phi(k, s)]_i^j - \frac{1}{m} \right| \leq 2 \frac{1 + \eta^{-B_0}}{1 - \eta^{B_0}} (1 - \eta^{B_0})^{\frac{k-s}{B_0}},$$

where η is the lower bound of Assumption 1, $B_0 = (m-1)B$, m is the number of agents, and B is the intercommunication interval bound of Assumption 4.

The geometric rate estimate will be used in studying the convergence of the estimates $\{x^i(k)\}$ in Section V.

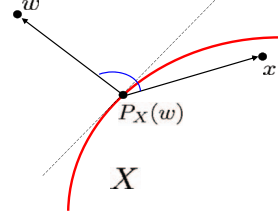


Fig. 2. Illustration of the relation between the projection error and feasible directions of a convex set at the projection vector.

IV. PROJECTION ERROR

We write the update rule in Eq. (1) as

$$x^i(k+1) = \sum_{j=1}^m a_j^i(k) x^j(k) - e^i(k), \quad (2)$$

where $e^i(k)$ represents the error due to projection given by

$$e^i(k) = \sum_{j=1}^m a_j^i(k) x^j(k) - \mathcal{P}_{X_i} \left[\sum_{j=1}^m a_j^i(k) x^j(k) \right]. \quad (3)$$

For notational convenience, let $w^i(k)$ denote

$$w^i(k) = \sum_{j=1}^m a_j^i(k) x^j(k), \quad (4)$$

i.e., $x^i(k+1) = \mathcal{P}_{X_i}(w^i(k))$. Hence the projection error $e^i(k)$ is given by

$$e^i(k) = w^i(k) - x^i(k+1).$$

In this section, we study the convergence behavior of the error sequences $\{e^i(k)\}$ under Assumption 2. The following lemma presents a standard relation about the projection error which will be key in subsequent analysis.

Lemma 1: Let X be a nonempty closed convex subset of \mathbb{R}^n and w be a vector in \mathbb{R}^n . We have for all $x \in X$,

$$\|w - P_X(w)\|^2 \leq \|w - x\|^2 - \|P_X(w) - x\|^2.$$

Proof: By the Projection Theorem (see [2] Proposition 2.2.1), we have the following relation:

$$(w - P_X(w))'(P_X(w) - x) \geq 0 \quad \text{for all } x \in X. \quad (5)$$

Let x be a vector that belongs to X , which is nonempty by assumption. We have

$$\begin{aligned} \|w - x\|^2 &= \|w - P_X(w) + P_X(w) - x\|^2 \\ &= \|w - P_X(w)\|^2 + \|P_X(w) - x\|^2 \\ &\quad + 2(w - P_X(w))'(P_X(w) - x) \\ &\geq \|w - P_X(w)\|^2 + \|P_X(w) - x\|^2, \end{aligned}$$

where the inequality follows from Eq. (5). ■

This relation is illustrated in Figure IV. The following lemma shows that under some assumptions, the error associated with the projection $e^i(k)$ goes to 0 as k goes to infinity.

Lemma 2: Assume that the intersection $X = \cap_{i=1}^m X_i$ is nonempty. Let Doubly Stochastic Weights Assumption hold (cf. Assumption 2). We have

$$\lim_{k \rightarrow \infty} e^i(k) = 0 \quad \text{for all } i.$$

Proof: Since $x_i(k+1) = P_{X_i}(w_i(k))$ for all i and $k \geq 0$, using the nonexpansive property of the projection operation, we have

$$\|x^i(k+1) - x\| \leq \|w^i(k) - x\| \quad \text{for all } x \in X_i, \text{ all } i \text{ and } k.$$

Summing the preceding relation over all $i \in \{1, \dots, m\}$ yields for all k

$$\sum_{i=1}^m \|x^i(k+1) - x\| \leq \sum_{i=1}^m \|w^i(k) - x\| \quad \text{for all } x \in \cap_{i=1}^m X_i. \quad (6)$$

By Doubly Stochastic Weights Assumption (cf. Assumption 2), for all k , $w^i(k)$ is a convex combination of the $x^j(k)$. Therefore, we have for all $x \in X$,

$$\|w^i(k) - x\| = \left\| \sum_{j=1}^m a_j^i(k) x^j(k) - x \right\| = \left\| \sum_{j=1}^m a_j^i(k) [x^j(k) - x] \right\|,$$

where the second equality holds by the assumption that $\sum_{j=1}^m a_j^i(k) = 1$ for all i and $k \geq 0$. Summing the preceding relation over all i yields

$$\sum_{i=1}^m \|w^i(k) - x\| \leq \sum_{i=1}^m \sum_{j=1}^m a_j^i(k) \|x^j(k) - x\| = \sum_{j=1}^m \|x^j(k) - x\|, \quad (7)$$

where the equality holds by the assumption that $\sum_{i=1}^m a_j^i(k) = 1$ for all j and $k \geq 0$.

Let \bar{x} be a vector in the set X . We define the sequence $\rho(k)$ as

$$\rho(2k) = \sum_{i=1}^m \|x^i(k) - \bar{x}\|, \quad \rho(2k+1) = \sum_{i=1}^m \|w^i(k) - \bar{x}\|.$$

Using the relations in (6) and (7) for $x = \bar{x}$, it follows that the sequence $\rho(k)$ is nonincreasing. Since it is bounded from below, it has a limit, implying that

$$\lim_{k \rightarrow \infty} \sum_{i=1}^m \|w^i(k) - \bar{x}\| - \|x^i(k+1) - \bar{x}\| = 0. \quad (8)$$

Using Lemma 1 with the substitution $w = w^i(k)$, $x^i(k+1) = P_{X_i}(w^i(k))$ and $x = \bar{x} \in \cap_{i=1}^m X_i$, we have for all i and $k \geq 0$,

$$\|w^i(k) - x^i(k+1)\|^2 \leq \|w^i(k) - \bar{x}\|^2 - \|x^i(k+1) - \bar{x}\|^2.$$

Summing this relation over all i , we obtain

$$\sum_{i=1}^m \|w^i(k) - x^i(k+1)\|^2 \leq \sum_{i=1}^m \|w^i(k) - \bar{x}\|^2 - \sum_{i=1}^m \|x^i(k+1) - \bar{x}\|^2. \quad (9)$$

We rewrite the preceding relation as

$$\begin{aligned} & \sum_{i=1}^m \|w^i(k) - x^i(k+1)\|^2 \\ & \leq \sum_{i=1}^m \left(\|w^i(k) - \bar{x}\| - \|x^i(k+1) - \bar{x}\| \right) \\ & \quad \left(\|w^i(k) - \bar{x}\| + \|x^i(k+1) - \bar{x}\| \right) \\ & \leq \sum_{i=1}^m \left(\|w^i(k) - \bar{x}\| - \|x^i(k+1) - \bar{x}\| \right) \\ & \quad \sum_{i=1}^m \left(\|w^i(k) - \bar{x}\| + \|x^i(k+1) - \bar{x}\| \right), \end{aligned}$$

where the second inequality follows by the fact that $\|w^i(k) - \bar{x}\| - \|x^i(k+1) - \bar{x}\| \geq 0$ for all i [cf. Eq. (6)], and the relation $\sum_{i=1}^m c_i d_i \leq \sum_{i=1}^m c_i \sum_{i=1}^m d_i$ for any scalars $c_i \geq 0$, $d_i \geq 0$. Since the sequence $\{\rho(k)\}$ is nonincreasing, we can bound the second term on the right hand-side of the above relation by $2 \sum_{i=1}^m \|x^i(0) - \bar{x}\|$. Hence, we have

$$\begin{aligned} & \sum_{i=1}^m \|w^i(k) - x^i(k+1)\|^2 \leq \\ & \sum_{i=1}^m \left(\|w^i(k) - \bar{x}\| - \|x^i(k+1) - \bar{x}\| \right) 2 \sum_{i=1}^m \|x^i(0) - \bar{x}\|. \end{aligned}$$

Taking the limit as $k \rightarrow \infty$ in the preceding and using Eq. (8), this implies that

$$\lim_{k \rightarrow \infty} \sum_{i=1}^m \|w^i(k) - x^i(k+1)\|^2 = 0.$$

This establishes that $\|w^i(k) - x^i(k+1)\| = \|e^i(k)\|$ converges to 0, showing the desired result. \blacksquare

V. CONVERGENCE ANALYSIS

Exploiting the linearity of the update relation (1), we write the estimates $x^i(k+1)$ in terms of the transition matrices $\Phi(k, s)$ (see Section III). Recall that these matrices are defined as follows: Let $A(s)$ be a matrix whose i -th column is the vector $a^i(s)$, and the matrices $\Phi(k, s)$ be defined by

$$\Phi(k, s) = A(s)A(s+1) \cdots A(k-1)A(k),$$

where

$$\Phi(k, k) = A(k) \quad \text{for all } k.$$

We use these matrices to relate $x^i(k+1)$ to the estimates $x^j(s)$ at time $s \leq k$. In particular, this yields the following relation:

$$\begin{aligned} x^i(k+1) &= \sum_{j=1}^m [\Phi(k, s)]_j^i x^j(s) \\ &= \sum_{r=s+1}^k \left(\sum_{j=1}^m [\Phi(k, r)]_j^i e^j(r-1) \right) - e^i(k). \end{aligned} \quad (10)$$

To analyze the preceding model, we consider a related model where after some time, the agents stop projecting on

the local constraint sets, i.e., after some time, the errors $e^i(k)$ are equal to 0. More specifically, assume that agents stop projecting after some time $\hat{k} \geq s + 1$, so that

$$e^i(k) = 0 \quad \text{for all } i \text{ and all } k \text{ with } k \geq \hat{k}.$$

Let $\{\hat{x}^i(k)\}$, $i = 1, \dots, m$, be the sequence of estimates generated by this model. Then by using the preceding relation, we have for all i ,

$$\hat{x}^i(k) = x^i(k) \quad \text{for all } k \leq \hat{k},$$

and

$$\hat{x}^i(k) = \sum_{j=1}^m [\Phi(k-1, s)]_j^i x^j(s) - \sum_{r=s+1}^{\hat{k}} \left(\sum_{j=1}^m [\Phi(k-1, r)]_j^i e^j(r-1) \right) \quad (11)$$

for all $k > \hat{k}$.

By letting $k \rightarrow \infty$ and by using Lemma 1(b), we see that the limit vector $\lim_{k \rightarrow \infty} \hat{x}^i(k)$ exists. Furthermore, the limit vector does not depend on i , but does depend on \hat{k} . We denote this limit by $y(\hat{k})$, i.e.,

$$\lim_{k \rightarrow \infty} \hat{x}^i(k) = y(\hat{k}).$$

By Eq. (11), we have for all $k \geq s + 1$

$$y(\hat{k}) = \frac{1}{m} \sum_{j=1}^m x^j(s) - \frac{1}{m} \sum_{r=s+1}^{\hat{k}} \left(\sum_{j=1}^m e^j(r-1) \right).$$

Note that this relation holds for any \hat{k} , so we may re-index these relations by using k , and thus obtain

$$y(k) = \frac{1}{m} \sum_{j=1}^m x^j(s) - \frac{1}{m} \sum_{r=s+1}^k \left(\sum_{j=1}^m e^j(r-1) \right). \quad (12)$$

We next show that agent values “reach consensus” in the limit under some assumptions. This is formally stated in the following.

Proposition 2: (Consensus) Let Weights Rule, Doubly Stochastic Weights, Connectivity, and Bounded Intercommunication Interval assumptions hold [cf. Assumptions 1, 2, 3, and 4]. We then have:

$$\lim_{k \rightarrow \infty} (x^i(k) - y(k)) = 0 \quad \text{for all } i.$$

Proof: By Lemma 2, it follows that $e^i(k) \rightarrow 0$ as $k \rightarrow \infty$ for all i . Therefore, for any $\epsilon > 0$, we can choose some integer s such that $\|e^i(k)\| \leq \epsilon$ for all $k \geq s$ and i . Using the relations in Eqs. (10) and (12), we obtain for all i and $k \geq s + 1$

$$\|x^i(k) - y(k)\|$$

$$\begin{aligned} &= \left\| \sum_{j=1}^m \left([\Phi(k-1, s)]_j^i - \frac{1}{m} \right) x^j(s) \right. \\ &\quad \left. - \sum_{r=s+1}^{k-1} \sum_{j=1}^m \left([\Phi(k-1, r)]_j^i - \frac{1}{m} \right) e^j(r-1) \right. \\ &\quad \left. - \left(e^i(k-1) - \frac{1}{m} \sum_{j=1}^m e^j(k-1) \right) \right\| \\ &\leq \sum_{j=1}^m \left| [\Phi(k-1, s)]_j^i - \frac{1}{m} \right| \|x^j(s)\| \\ &\quad + \sum_{r=s+1}^{k-1} \sum_{j=1}^m \left| [\Phi(k-1, r)]_j^i - \frac{1}{m} \right| \|e^j(r-1)\| \\ &\quad + \frac{1}{m} \sum_{j=1}^m \|e^i(k-1) - e^j(k-1)\|. \end{aligned}$$

Using the estimates for $\left| [\Phi(k-1, s)]_j^i - \frac{1}{m} \right|$ of Proposition 1(b), it follows that

$$\begin{aligned} &\|x^i(k) - y(k)\| \\ &\leq 2 \frac{1 + \eta^{-B_0}}{1 - \eta^{B_0}} (1 - \eta^{B_0})^{\frac{k-1-s}{B_0}} \sum_{j=1}^m \|x^j(s)\| \\ &\quad + \sum_{r=s+1}^{k-1} 2 \frac{1 + \eta^{-B_0}}{1 - \eta^{B_0}} (1 - \eta^{B_0})^{\frac{k-1-r}{B_0}} \left(\sum_{j=1}^m \|e^j(r-1)\| \right) \\ &\quad + \frac{1}{m} \sum_{j=1}^m \|e^i(k-1) - e^j(k-1)\|. \end{aligned}$$

Using the relation $\|e^i(k)\| \leq \epsilon$ for all $k \geq s$ and i in the preceding, we obtain

$$\begin{aligned} \|x^i(k) - y(k)\| &\leq 2 \frac{1 + \eta^{-B_0}}{1 - \eta^{B_0}} (1 - \eta^{B_0})^{\frac{k-1-s}{B_0}} \sum_{j=1}^m \|x^j(s)\| \\ &\quad + 2m\epsilon \frac{1 + \eta^{-B_0}}{1 - \eta^{B_0}} \frac{1}{1 - (\frac{1}{\eta})^{\frac{1}{B_0}}} + 2\epsilon. \end{aligned}$$

Since the choice of ϵ was arbitrary, this yields

$$\lim_{k \rightarrow \infty} (x^i(k) - y(k)) = 0, \quad \text{for all } i,$$

thus showing that agents reach a consensus. \blacksquare

We next show that under an interior point assumption, the agent estimates $x^i(k)$ converge to a limit as k goes to infinity. The interior point assumption is stated in the following. This assumption allows us to obtain a nontrivial improvement in terms of the distance of the iterates to the projection set (see Figure V).

Assumption 5: (Interior Point) There exists a vector \bar{x} such that

$$\bar{x} \in \text{int}(X) = \text{int}\left(\bigcap_{i=1}^m X_i\right),$$

i.e., there exists some scalar $\delta > 0$ such that $\{z \mid \|z - \bar{x}\| \leq \delta\} \subset X$.

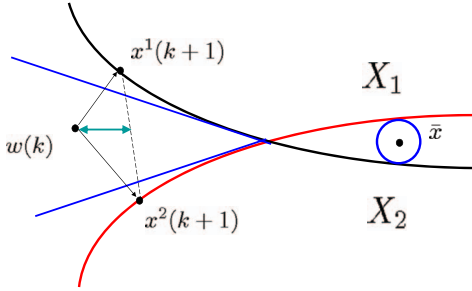


Fig. 3. Illustration of the Interior Point Assumption.

The next lemma provides an *error bound* relating the distance of the iterates to set X with the sum of distances of iterates to each set under the Interior Point Assumption.

Lemma 3: Let Doubly Stochastic Weights and Interior Point assumptions hold (cf. Assumptions 2 and 5). For all i , let $\{x^i(k)\}$ be generated by the update rule (1) and $\{w^i(k)\}$ be given by Eq. (4). For all $k \geq 0$, we have

$$\begin{aligned} \sum_{j=1}^m \text{dist}(w^j(k), X) &\leq \\ \sum_{j=1}^m \|w^j(k) - x^j(k+1)\| &+ R \sum_{j=1}^m \|w^j(k+1) - x^j(k+2)\| \\ + \sum_{j=1}^m \left\| w^j(k) - \frac{1}{m} \sum_{i=1}^m w^i(k) \right\| & \\ + R \sum_{j=1}^m \left\| w^j(k+1) - \frac{1}{m} \sum_{i=1}^m w^i(k+1) \right\|, & \end{aligned}$$

where the constant R is given by

$$R = \frac{1}{\delta} \sum_{i=1}^m \|x^i(0) - \bar{x}\|,$$

with \bar{x} and δ given in Assumption 5.

Proof: For a fixed k , we let y denote $y = \frac{1}{m} \sum_{i=1}^m x^i(k+1)$, and let ϵ be given by

$$\epsilon = \sum_{j=1}^m \text{dist}(y, X_j). \quad (13)$$

We define the vector

$$z = \frac{\epsilon}{\epsilon + \delta} \bar{x} + \frac{\delta}{\epsilon + \delta} y. \quad (14)$$

We first show that the vector z belongs to the intersection $X = \cap_{i=1}^m X_i$. To see this, note that for all i , we can write z as

$$z = \frac{\epsilon}{\epsilon + \delta} \left(\bar{x} + \frac{\delta}{\epsilon} (y - P_{X_i}(y)) \right) + \frac{\delta}{\epsilon + \delta} P_{X_i}(y).$$

By the definition of ϵ [cf. Eq. (13)], it follows that $\|y - P_{X_i}(y)\| \leq \epsilon$, implying by the Interior Point Assumption (cf. Assumption 5) that the vector $\bar{x} + \frac{\delta}{\epsilon} (y - P_{X_i}(y))$ belongs to set X , and therefore to set X_i . Since z is the convex combination of two vectors in the set X_i , it follows by the convexity of X_i that $z \in X_i$. The preceding argument can be repeated for all i , showing that $z \in X$.

Using the fact that $z \in X$, we can write for all i

$$\text{dist}(w^i(k), X) \leq \|w^i(k) - z\| \leq \|w(k) - y\| + \|y - z\|,$$

where $y = \frac{1}{m} \sum_{i=1}^m x^i(k+1)$. Using the definition of the vector z , this implies

$$\begin{aligned} \text{dist}(w^i(k), X) &\leq \left\| w^i(k) - \frac{1}{m} \sum_{i=1}^m x^i(k+1) \right\| + \\ &\quad \frac{\epsilon}{\epsilon + \delta} \left\| \frac{1}{m} \sum_{i=1}^m x^i(k+1) - \bar{x} \right\| \\ &\leq \left\| w^i(k) - \frac{1}{m} \sum_{i=1}^m x^i(k+1) \right\| + \frac{\epsilon}{\delta} \frac{1}{m} \sum_{i=1}^m \|x^i(k+1) - \bar{x}\|. \end{aligned}$$

Since $\bar{x} \in X$, the sequence $\{\sum_{i=1}^m \|x^i(k+1) - \bar{x}\|\}$ is a nonincreasing sequence (see the proof of Lemma 2). Therefore, we have $\sum_{i=1}^m \|x^i(k+1) - \bar{x}\| \leq \sum_{i=1}^m \|x^i(0) - \bar{x}\|$ for all k . Defining the constant $R = \frac{1}{\delta} \sum_{i=1}^m \|x^i(0) - \bar{x}\|$ and substituting in the preceding relation, we obtain

$$\begin{aligned} \text{dist}(w^h(k), X) &\leq \left\| w^h(k) - \frac{1}{m} \sum_{i=1}^m x^i(k+1) \right\| + \frac{R}{m} \epsilon \\ &\leq \left\| w^h(k) - \frac{1}{m} \sum_{i=1}^m x^i(k+1) \right\| + \frac{R}{m} \sum_{j=1}^m \text{dist}(y, X_j), \end{aligned}$$

where the second inequality follows by using the definition of ϵ in (13).

By the Doubly Stochastic Weights assumption, i.e., $\sum_{i=1}^m a_j^i(k+1) = 1$ for all j , we have

$$\begin{aligned} \frac{1}{m} \sum_{i=1}^m w^i(k+1) &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^m a_j^i(k+1) x^j(k+1) = \\ &= \frac{1}{m} \sum_{j=1}^m x^j(k+1). \end{aligned}$$

This implies that $y = \frac{1}{m} \sum_{i=1}^m x^i(k+1) = \frac{1}{m} \sum_{i=1}^m w^i(k+1)$. Combining these relations we obtain

$$\begin{aligned}
& \text{dist}(w^h(k), X) \\
& \leq \left\| w^h(k) - \frac{1}{m} \sum_{i=1}^m x^i(k+1) \right\| \\
& \quad + \frac{R}{m} \sum_{j=1}^m \left\| \frac{1}{m} \sum_{i=1}^m w^i(k+1) - x^j(k+2) \right\| \\
& \leq \left\| \frac{1}{m} \sum_{i=1}^m w^i(k) - \frac{1}{m} \sum_{i=1}^m x^i(k+1) \right\| \\
& \quad + \left\| w^h(k) - \frac{1}{m} \sum_{i=1}^m w^i(k) \right\| \\
& \quad + \frac{R}{m} \sum_{j=1}^m \left\| w^j(k+1) - x^j(k+2) \right\| \\
& \quad + \frac{R}{m} \sum_{j=1}^m \left\| \frac{1}{m} \sum_{i=1}^m w^i(k+1) - w^j(k+1) \right\|.
\end{aligned}$$

Summing the preceding relation over all $h \in \{1, \dots, m\}$ yields

$$\begin{aligned}
& \sum_{j=1}^m \text{dist}(w^j(k), X) \\
& \leq \sum_{j=1}^m \|w^j(k) - x^j(k+1)\| \\
& \quad + R \sum_{j=1}^m \|w^j(k+1) - x^j(k+2)\| \\
& \quad + \sum_{j=1}^m \left\| w^j(k) - \frac{1}{m} \sum_{i=1}^m w^i(k) \right\| \\
& \quad + R \sum_{j=1}^m \left\| w^j(k+1) - \frac{1}{m} \sum_{i=1}^m w^i(k+1) \right\|,
\end{aligned}$$

completing the proof. \blacksquare

The next proposition contains our main convergence result.

Proposition 3: Let Weights Rule, Doubly Stochastic Weights, Connectivity, and Bounded Intercommunication Interval assumptions hold [cf. Assumptions 1, 2, 3, and 4]. Let also Interior Point Assumption hold (cf. Assumption 5). For all i , let $\{x^i(k)\}$ be generated by the update rule (1). We then have:

$$\lim_{k \rightarrow \infty} x^i(k) = \tilde{x} \quad \text{for all } i,$$

where \tilde{x} is a vector in X .

Proof: Let $\{w^i(k)\}$ be defined as in Eq. (4), i.e., $w^i(k) = \sum_{j=1}^m a_j^i(k) x^j(k)$. By Lemma 3, we have $\sum_{j=1}^m \text{dist}(w^j(k), X)$

$$\leq \sum_{j=1}^m \|e^j(k)\| + \sum_{j=1}^m \|e^j(k+1)\| \quad (15)$$

$$\begin{aligned}
& + \sum_{j=1}^m \left\| w^j(k) - \frac{1}{m} \sum_{i=1}^m w^i(k) \right\| \\
& + R \sum_{j=1}^m \left\| w^j(k+1) - \frac{1}{m} \sum_{i=1}^m w^i(k+1) \right\|,
\end{aligned} \quad (16)$$

where $e^j(k)$ is the projection error, i.e., $e^j(k) = w^j(k) - x^j(k+1)$. We can rewrite the third term on the right hand-side of the preceding relation as

$$\begin{aligned}
& \sum_{j=1}^m \left\| w^j(k) - \frac{1}{m} \sum_{i=1}^m w^i(k) \right\| \\
& \leq \sum_{j=1}^m \|w^j(k) - y(k)\| + \left\| y(k) - \frac{1}{m} \sum_{i=1}^m w^i(k) \right\|,
\end{aligned} \quad (17)$$

where $y(k)$ is given by Eq. (12). Using the relations

$$\sum_{j=1}^m \|w^j(k) - z\| \leq \sum_{j=1}^m \|x^j(k) - z\| \quad \text{for all } z \in \mathbb{R}^n,$$

and $\frac{1}{m} \sum_{i=1}^m w^i(k) = \frac{1}{m} \sum_{i=1}^m x^i(k)$ (which follow by the Doubly Stochastic Weights Assumption), Eq. (17) implies that

$$\begin{aligned}
& \sum_{j=1}^m \left\| w^j(k) - \frac{1}{m} \sum_{i=1}^m w^i(k) \right\| \\
& \leq \sum_{j=1}^m \|x^j(k) - y(k)\| + \left\| y(k) - \frac{1}{m} \sum_{i=1}^m x^i(k) \right\|.
\end{aligned}$$

Combined with Proposition 2, it follows from this relation that

$$\lim_{k \rightarrow \infty} \sum_{j=1}^m \left\| w^j(k) - \frac{1}{m} \sum_{i=1}^m w^i(k) \right\| = 0.$$

A similar argument shows that the fourth term on the right hand-side of Eq. (15) also goes to 0 as $k \rightarrow \infty$. Combined with the fact that $e^i(k)$ goes to 0 as $k \rightarrow \infty$ (cf. Lemma 2), Eq. (15) implies that

$$\lim_{k \rightarrow \infty} \sum_{j=1}^m \text{dist}(w^j(k), X) = 0. \quad (18)$$

Since the sequence $\left\{ \sum_{i=1}^m \|w^i(k) - x\| \right\}$ is nondecreasing for all $x \in X$ (see the proof of Lemma 2), it follows that the sequence $\{w^i(k)\}$ is bounded for all i . Let \tilde{w} be a limit point of the sequence $\{w^i(k)\}$. Since $\|w^i(k) - w^j(k)\| \rightarrow 0$ as $k \rightarrow \infty$ for all $j \neq i$ (cf. Proposition 2), we have that \tilde{w} is also a limit point of the sequence $\{w^j(k)\}$ for all $j \neq i$. By Eq. (18), we have $\sum_{j=1}^m \text{dist}(w^j(k), X) \rightarrow 0$, which implies that $\tilde{w} \in X$. Because $\tilde{w} \in X$, it follows that the sequence $\sum_{i=1}^m \|w^i(k) - \tilde{w}\|$ is nonincreasing. Since along some subsequence $w^i(k) \rightarrow \tilde{w}$ for all i , this implies that

$$\lim_{k \rightarrow \infty} w^i(k) = \tilde{w} \quad \text{for all } i.$$

Since $x^i(k+1) = P_{X_i}(w^i(k))$, this implies that $x^i(k) \rightarrow \tilde{w} \in X$ for all i , completing the proof. \blacksquare

VI. ACKNOWLEDGEMENTS

The author would like to thank Professor Angelia Nedić and Professor Pablo Parrilo for comments and useful discussions.

This work was supported by DARPA's ITMANET program under grant 18870740-37362-C.

VII. CONCLUSIONS

We studied the constrained consensus problem where agent i 's estimate is constrained to be in a closed convex set X_i . We presented a distributed constrained consensus algorithm and studied its convergence properties. We showed that the agent values reach consensus asymptotically as the number of iterations goes to infinity. Under an interior point assumption on the intersection set $X = \cap_{i=1}^m X_i$, we proved that each of the agent estimates converge to the same limit, which belongs to set X . Future work includes analyzing the convergence rate of this algorithm and incorporating optimization of different objective functions into the constrained consensus model.

REFERENCES

- [1] N. Aronszajn, *Theory of reproducing kernels*, Transactions of the American Mathematical Society **68** (1950), no. 3, 337–404.
- [2] D.P. Bertsekas, A. Nedić, and A.E. Ozdaglar, *Convex analysis and optimization*, Athena Scientific, Cambridge, Massachusetts, 2003.
- [3] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis, *Convergence in multiagent coordination, consensus, and flocking*, Proceedings of IEEE CDC, 2005.
- [4] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, *Gossip algorithms: Design, analysis, and applications*, Proceedings of IEEE INFOCOM, 2005.
- [5] M. Cao, D.A. Spielman, and A.S. Morse, *A lower bound on convergence of a distributed network consensus algorithm*, Proceedings of IEEE CDC, 2005.
- [6] F. Deutsch, *Rate of convergence of the method of alternating projections*, Parametric Optimization and Approximation (B. Brosowski and F. Deutsch, eds.), vol. 76, Birkhuser, Basel, 1983, pp. 96–107.
- [7] F. Deutsch and H. Hundal, *The rate of convergence for the cyclic projections algorithm i: Angles between convex sets*, Journal of Approximation Theory **142** (2006), 36–55.
- [8] L.G. Gubin, B.T. Polyak, and E.V. Raik, *The method of projections for finding the common point of convex sets*, U.S.S.R Computational Mathematics and Mathematical Physics **7** (1967), no. 6, 1211–1228.
- [9] A. Jadbabaie, J. Lin, and S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Transactions on Automatic Control **48** (2003), no. 6, 988–1001.
- [10] J.R. Marden, G. Arslan, and J.S. Shamma, *Connections between cooperative control and potential games illustrated on the consensus problem*, Proceedings of the 2007 European Control Conference, 2007.
- [11] A. Nedić and A. Ozdaglar, *Distributed subgradient methods for multi-agent optimization*, Preprint, 2007.
- [12] ———, *On the rate of convergence of distributed subgradient methods for multi-agent optimization*, Proceedings of IEEE CDC, 2007.
- [13] J. Von Neumann, *Functional operators*, Princeton University Press, Princeton, 1950.
- [14] R. Olfati-Saber and R.M. Murray, *Consensus problems in networks of agents with switching topology and time-delays*, IEEE Transactions on Automatic Control **49** (2004), no. 9, 1520–1533.
- [15] A. Olshevsky and J.N. Tsitsiklis, *Convergence rates in distributed consensus averaging*, Proceedings of IEEE CDC, 2006.
- [16] ———, *Convergence speed in distributed consensus and averaging*, Preprint, 2006.
- [17] J.N. Tsitsiklis, *Problems in decentralized decision making and computation*, Ph.D. thesis, 1984.
- [18] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans, *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Transactions on Automatic Control **31** (1986), no. 9, 803–812.
- [19] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and Schochet O., *Novel type of phase transitions in a system of self-driven particles*, Physical Review Letters **75** (1995), no. 6, 1226–1229.
- [20] J. Wolfowitz, *Products of indecomposable, aperiodic, stochastic matrices*, Proceedings of the American Mathematical Society **14** (1963), no. 4, 733–737.