

# Distributed Optimization of Average Consensus Containment with Multiple Stationary Leaders

Sushobhan Chatterjee and Rachel Kalpana Kalaimani

**Abstract**—In this paper, we consider the problem of containment control of multi-agent systems with multiple stationary leaders, interacting over a directed network. While, containment control refers to just ensuring that the follower agents reach the convex hull of the leaders' states, we focus on the problem where the followers achieve a consensus to the average values of the leaders' states. We propose an algorithm that can be implemented in a distributed manner to achieve the above consensus among followers. Next we optimize the convergence rate of the followers to the average consensus by proper choice of weights for the interaction graph. This optimization is also performed in a distributed manner using Alternating Direction Method of Multipliers (ADMM). Finally, we complement our results by illustrating them with numerical examples.

**Index Terms**—Containment control, distributed control, leader-follower, average consensus, distributed optimization, ADMM, directed graphs

## I. INTRODUCTION

The idea of achieving a common objective in a system with multiple agents using distributed and co-operative control has garnered a lot of interest among research community for a long time. A fundamental aspect of the distributed co-operative control is that the agents achieve the collective objective via local interaction among neighbours, thereby making the framework robust to communication failures, highly adaptable and cost effective.

A well addressed problem in this multi-agent system framework is the consensus problem, where the agents are expected to agree on a common value [15], [13]. Consensus with a single leader is addressed in [4]. When there are multiple leaders the problem is referred to as containment control where the objective is to bring the remaining follower agents from any arbitrary state, within the convex hull spanned by the leaders' state [20]. This problem is also motivated from nature such as gravel ants forming a boundary over caterpillars and transporting along with them due to their sugary extract. Practical applications include shaping of social dynamics using a set of guiding agents, military applications, safe disposal of hazardous waste etc. [2]. Extensive research is done in this area such as analysing dynamic leaders using PDE in [5], assuming directed static topology [9], time-varying topologies [1], agents with higher order dynamics and impact of time delays [8].

\*This work has been partially supported by DST-INSPIRE Faculty Grant, Department of Science and Technology (DST), Govt. of India (ELE/16-17/333/DSTX/RACH)

The authors are with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India - 600036. (email: chatterjeesushobhan8@gmail.com; rachel@ee.iitm.ac.in).

While the above containment algorithms only focus on general containment of followers, i.e. each follower achieving an arbitrary state, different from each other, within the convex hull of the leaders, the primary objective of this paper is to achieve a consensus among the followers regarding the final destination within the convex hull of the leaders. The idea of achieving consensus in a multiple leader containment scenario is motivated by various applications. One is opinion consensus of a group, driven by a couple of external agents/factors, in a direction that gets aligned with their own priors. Another application is a cooperative attack on a static or, moving target, where the target is encircled by UAVs (leaders) such that the target's live location coincides with the average of coordinates of UAVs. Then the missiles (followers) can easily home in to the target without the need for any individual radar tracking.

Motivated by the above applications for consensus containment, we provide an algorithm where the followers reach a consensus to the average of the leaders' state. [7] investigates consensus of followers inside the convex hull of stationary leaders, where a distributed control law has been designed by formulating a centralized multi-objective optimization (MOP) with a trade-off between performance error and convergence rate. This results in a significant asymptotic drift in the final consensus value due to non-convergence to the exact solution on account of trade-offs. We propose a distributed control law which ensures a fast convergence of the followers to the consensus within the convex hull spanned by the leaders over a directed network, without any aforementioned asymptotic drift. The fast convergence is designed using a distributed optimization algorithm unlike the centralized approach in [7]. Distributed optimization of convergence rate to consensus without leaders in a multi-agent, undirected framework is addressed in [18].

We summarize our contribution as follows.

- 1) We consider a multi-agent system with multiple stationary leaders and followers communicating with each other over a directed network topology. We propose a distributed algorithm which ensures consensus of followers to the average of leaders' states.
- 2) We formulate an optimization problem to maximize the rate of convergence of follower agents to the consensus value. We propose an algorithm based on Alternating Direction Method of Multipliers (ADMM) that solves the above optimization problem in a distributed manner.
- 3) We illustrate the results using numerical examples. We compare the results of the distributed optimization to the centralized optimization.

The outline of the paper is laid out as follows. Following the introduction, Section II focuses on basic mathematical preliminaries. In Section III, the two problems are formally discussed. Algorithms for solving these problems are proposed in Section IV. Section V provides a numerical example to elucidate the validity of the proposed algorithms. Finally, we give concluding remarks in Section VI. The proofs of some of the results are in Appendix.

## II. PRELIMINARIES

In this section, we briefly introduce some preliminaries pertaining to directed graphs and convex analysis.

### A. Notations

Let  $\mathbb{R}^n$  be the  $n$ -dimensional real vector and  $\mathbb{R}^{n \times n}$  be the real square matrix of order  $n$ . Let  $I_n$  be the  $n \times n$  identity matrix,  $0_n$  be the  $n \times n$  matrix with all entries as 0 and the vector  $1_n = [1, \dots, 1]^T$  with dimension  $n$ . For any vector  $x : x_i$  or  $[x]_i$  denotes its  $i^{\text{th}}$  element,  $\bar{x}$  denotes its average value, and  $\|x\|_2$  denotes its 2-norm. For a set  $S$ ,  $|S|$  denotes its cardinality. For a matrix  $P : \Lambda(P)$  and  $\rho(P)$  denote its eigenvalue set and spectral radius, respectively,  $P^{ij}$  or  $(P)^{ij}$  denotes its  $(i, j)^{\text{th}}$  element and  $\|P\|_F$  denotes its Frobenius-norm.

### B. Graph Theory

Let  $\mathcal{G} := (\mathcal{V}, \mathcal{E}, A)$  denote a directed graph, where  $\mathcal{V} = [v_1, v_2, \dots, v_n]$  represents the vertex set,  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  denotes the directed edge set and  $A = [A^{ij}] \in \mathbb{R}^{n \times n}$  denotes the weighted adjacency matrix defined as follows. A directed edge is an ordered pair of distinct vertices  $(v_j, v_i)$  such that,  $i^{\text{th}}$  node (child) can access state information of  $j^{\text{th}}$  node (parent). The weight associated with the edge  $(v_j, v_i)$  is the entry  $A^{ij} \geq 0$ , in  $A$ . Also we assume  $A^{ii} > 0 \forall i = \{1, 2, \dots, n\}$ .

$L = [L^{ij}] \in \mathbb{R}^{n \times n}$  denotes the laplacian matrix defined as:

$$L^{ij} = \sum_{k=1, k \neq i}^{k=n} A^{ik} \forall i = j, \text{ and } L^{ij} = -A^{ij} \forall i \neq j.$$

For a vertex  $v_i$ , the set of in-neighbours is defined by  $\mathcal{N}_i^{\text{in}} = \{v_j \mid (v_j, v_i) \in \mathcal{E}\}$  and out-neighbours by  $\mathcal{N}_i^{\text{out}} = \{v_j \mid (v_i, v_j) \in \mathcal{E}\}$ . For a graph  $\mathcal{G} := (\mathcal{V}, \mathcal{E}, A)$ , the subgraph  $\mathcal{G}^1$  is the graph induced by a vertex set  $\mathcal{V}^1 \subseteq \mathcal{V}$  and edge set  $\mathcal{E}^1 = \{(v_i, v_j) \in \mathcal{E} \mid v_i, v_j \in \mathcal{V}^1\}$ .

A directed graph  $\mathcal{G}$  is said to be strongly connected if there exists a sequence of directed paths between any two distinct pair of vertices  $\{v_j, v_i\} \in \mathcal{V}$ , starting at  $v_j$  and ending at  $v_i$ .

### C. Convex Analysis

A set  $D \subset \mathbb{R}^q$  is said to be convex if  $\exists x, y \in D$  s.t.  $(1 - \theta)x + \theta y \in D \forall \theta \in [0, 1]$ .

**Definition 1:** [16] The convex hull of a finite set of points  $\{p_1, \dots, p_n\} \in \mathbb{R}^q$  is the smallest convex set containing all points  $p_i, i = 1, 2, \dots, n$  denoted by  $\mathbf{conv}\{p_1, \dots, p_n\} = \left\{ \sum_{i=1}^n \theta_i p_i \mid \theta_i \in \mathbb{R}, \theta_i \geq 0, \sum_{i=1}^n \theta_i = 1 \right\}$ .

## III. PROBLEM FORMULATION

We first explain the framework of a multi-agent system with more than one leader. Consider a group of  $n$  agents communicating with each other. The communication pattern is depicted by a directed graph  $\mathcal{G}$ . The vertices of the graph correspond to the agents and the state of an agent  $j$  is accessible by agent  $i$  only if there is a directed edge  $(v_j, v_i)$  in the graph. An agent is designated to be a leader if its in-neighbour set is empty. The rest of the agents are termed as followers. Let  $\mathcal{L}$  and  $\mathcal{F}$  denote respectively the index set of leaders and followers. Let  $\mathcal{G}^{\mathcal{F}}$  be the subgraph induced by the vertex set of the followers,  $v_i \forall i \in \mathcal{F}$  and  $\mathcal{G}^{\mathcal{F}\mathcal{L}}$  is a subgraph with vertex set  $\mathcal{V}$  and edge set given by  $\mathcal{E}^{\mathcal{F}\mathcal{L}} = \{(v_j, v_i) \in \mathcal{E} \mid i \in \mathcal{F}, j \in \mathcal{L}\}$ . Let  $x_i \in \mathbb{R}$  denote the state of each agent and assume that they follow the following discrete time dynamics.

$$x_i(k+1) = \alpha_i x_i(k) + u_i(k), \quad (1)$$

where  $u_i(k)$  is the control input to agent  $i$  at  $k^{\text{th}}$  iteration. The following are our assumptions on the agents.

- Assumption 1:** (a) The leaders are stationary, i.e.,  $\forall k \geq 0 \ \& \ i \in \mathcal{L}, u_i(k) = 0 \ \& \ \alpha_i = 1$ , and hence  $x_i(k+1) = x_i(k)$ .  
 (b) Every leader has at least one follower connected to it.  
 (c)  $\mathcal{G}^{\mathcal{F}}$  is strongly connected.

### A. Average Consensus Containment

Containment control refers to the problem of designing an input for the followers such that followers reach the convex hull spanned by the stationary leaders [1], [9]. Containment control does not guarantee a consensus within the convex hull. We focus on the problem where followers reach a consensus to average of leaders' states. We term this problem as *average consensus containment control*. We explain later how the existing containment protocol can be modelled to achieve average consensus containment using a centralized approach. We propose to address this problem using a distributed approach. Our problem is formulated below.

**Problem 1:** Consider a set of  $n$  agents with  $m$  leaders and  $n - m$  followers interacting over a directed graph satisfying *Assumption 1* with dynamics given in (1). Find a control input that can be designed and implemented in a distributed manner, such that the follower agents state converges to the average of the stationary leaders state.

$$\lim_{k \rightarrow \infty} x_i(k) \rightarrow \frac{1}{m} \sum_{j \in \mathcal{L}} x_j(0) \forall i \in \mathcal{F} \quad (2)$$

### B. Optimizing the convergence rate

Our next objective is to maximize the rate of convergence of the followers to the consensus mentioned in Problem 1. This can be achieved by suitably choosing the edge weights of the network over which the agents are communicating. Based on the results that we obtain for Problem 1, we first formulate the fastest convergence rate problem as an optimization problem. Then we attempt to solve that in a distributed manner. This is formulated as follows.

*Problem 2:* Consider Problem 1 with the mentioned assumptions and network model. We solve the following:

- 1) Formulate an optimization problem that maximizes the rate of convergence given in (2).
- 2) Propose to solve this optimization problem in a distributed manner.

#### IV. MAIN RESULTS

##### A. Average Consensus Containment

In this section, we first briefly explain the assumptions and protocol that ensures containment control. Then we introduce our algorithm which achieves average consensus containment that can be implemented in a distributed manner.

Consider a multi-agent network of  $n$  agents with  $m$  stationary leaders and  $n - m$  followers, having dynamics as given in (1) interacting over a network depicted by a graph  $\mathcal{G}$ . Assume that the agents use the following update protocol,

$$u_i(k) = \begin{cases} \sum_{j \in \mathcal{N}_i^{in}} \alpha_{ij} x_j(k), & i \in \mathcal{F} \\ 0, & i \in \mathcal{L} \end{cases} \quad (3)$$

for some suitable choice of weights  $\alpha_{ij}$ . Let  $x_L(k) \in \mathbb{R}^m$  and  $x_F(k) \in \mathbb{R}^{n-m}$  refer to the vector of all leader and follower states respectively at  $k^{\text{th}}$  iteration. Based on the weights in (3) and the system dynamics in (1), we define a matrix  $A \in \mathbb{R}^{n \times n}$ , where  $A^{ij} = \alpha_{ij}$ , for  $i \neq j$  and  $A^{ii} = \alpha_i$  which represents a weighted adjacency matrix corresponding to graph  $\mathcal{G}$ . The agent dynamics can be written in matrix form as follows.

$$\begin{bmatrix} x_F(k+1) \\ x_L(k+1) \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ 0 & I_m \end{bmatrix} \begin{bmatrix} x_F(k) \\ x_L(k) \end{bmatrix}, A = \begin{bmatrix} A_1 & A_2 \\ 0 & I_m \end{bmatrix} \quad (4)$$

The following result states the conditions for containment of followers within the convex hull of leader's states for a suitable choice of  $A_1$  and  $A_2$ .

*Proposition 1:* ([19], Theorem 9) Consider a set of multi-agents satisfying Assumption 1 following the protocol in (4). Assume  $A_1 = I_{n-m} - \alpha L_1$ ,  $A_2 = -\alpha L_2$ , where  $L_1$  and  $L_2$  are the weighted laplacian matrices associated with subgraphs  $\mathcal{G}^{\mathcal{F}}$  and  $\mathcal{G}^{\mathcal{FL}}$  respectively. The followers will converge to the convex hull formed by stationary leaders if and only if step-size satisfies;  $\alpha < \min_{\lambda_i \in \Lambda(L_1)} \frac{2\text{Re}(\lambda_i)}{\text{Re}^2(\lambda_i) + \text{Im}^2(\lambda_i)}$ , and the final state of all the followers are given by  $-L_1^{-1} L_2 x_L(0)$ . ■

From the above proposition, we observe that the follower states converge to some value within the convex hull formed the stationary leader's states and do not achieve any consensus. We next proceed to design an update protocol for the average consensus containment problem. We propose Algorithm 1 for average consensus containment under suitable assumptions on the matrix  $A$ . The existing Push sum algorithm [6] is adapted to the current problem. The convergence of the algorithm is proved in Theorem 1.

*Remark 1:* (Stopping Criterion for Algorithm 1) Each of the agent  $i \in \mathcal{F}$  fixes a small, arbitrary tolerance value,  $\gamma >$

0, and calculates its successive iterate error  $\{e_i(k)\}$ , at each iteration  $k$  as

$$e_i(k) = x_i(k) - x_i(k-1) \quad (5)$$

which indicates the extent of asymptotic convergence of the agents trajectory. The stopping criterion for Algorithm 1 is when

$$E_i(k) = |e_i(k)| \leq \gamma \quad \forall i \in \mathcal{F} \quad (6)$$

---

##### Algorithm 1 Distributed Average Consensus Containment

---

**Given :**  $A$ , Initial condition  $x(0)$  of all agents.

**Assumptions :**  $A_1$  and  $A_2$  defined in (4) are column stochastic and agents satisfy Assumption 1.

**Initialize :**  $s(0) = x(0)$ ,  $w(0) = \mathbf{1}_n$

**Iterate :**

**for**  $k \geq 0$  **do**

**Step 1 : Exchange Values :**  $\forall i \in \mathcal{F}$ ,

Receive  $s_j(k)$ ,  $w_j(k)$  from all in-neighbours  $j \in \mathcal{N}_i^{in}$

**Step 2 : Update sum and weight vectors :**

$s(k+1) = As(k)$

$w(k+1) = Aw(k)$

**Step 3 : Update state vector :**  $\forall i \in \mathcal{F}$

$x_i(k+1) = \frac{s_i(k+1)}{w_i(k+1)}$

**end for**

---

*Remark 2:* (Algorithm 1 implemented in a distributed and parallel setting) Each agent  $i$  maintains three variables,  $s_i(k)$ ,  $w_i(k)$  and  $x_i(k)$  at each iteration  $k$ . At every iteration, each agent updates its variables  $s_i$  and  $w_i$  by exchanging data only with its neighbours as indicated in Step 1 and 2. Furthermore, every agent updates the variable  $x_i$  as indicated in Step 3, parallelly.

*Theorem 1:* The average consensus containment protocol proposed in Algorithm 1 ensures that followers converge to a consensus which is the average of the leaders' state values.

The proof is given in Appendix. ■

##### B. Optimizing the convergence rate

In this section we address Problem 2, where we maximize the rate of convergence of the follower agents to consensus which is the average of stationary leaders' states. We first identify the parameter that influences the rate of convergence and then formulate the relevant optimization problem. The following lemma characterises the rate of convergence of the follower agents in terms of the weighted adjacency matrix.

*Lemma 1:* Consider a multi-agent system following dynamics given in (1) and satisfying Assumption 1. Assume that the agents are implementing Algorithm 1. The rate of convergence is characterised by  $\rho(A_1 - \frac{1}{n-m} \mathbf{1}\mathbf{1}^T)$ . ■

The proof of the lemma is given in Appendix.

From the above lemma, it is clear that we have to optimize the weights of the follower interaction, i.e., entries of  $A_1$ , so as to maximize the convergence rate. For this, we formulate the following optimization problem with appropriate

constraints.

$$\begin{aligned} & \underset{A_1}{\text{minimize}} && \rho\left(A_1 - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T\right) \\ & \text{subject to} && A_1\mathbf{1} = \mathbf{1}, A_1^T\mathbf{1} = \mathbf{1}, \\ & && (A_1)^{ij} = 0, (i, j) \notin \mathcal{E}^{\mathcal{F}} \end{aligned} \quad (A)$$

where  $\mathcal{E}^{\mathcal{F}}$  is the edge set of the subgraph  $\mathcal{G}^{\mathcal{F}}$  and  $\mathbf{1} = 1_{n-m}$ . Here,  $A_1$  should be column stochastic because of the assumption in Algorithm 1, and the requirement of row stochasticity has been explained later. The third equality is the topological constraint imposed by the communication pattern of the follower agents.

The above minimization problem is not convex due to  $\rho(A_1 - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T)$  being a non-convex function of  $A_1$  [14]. The problem is modified by using a convex function  $\|A_1 - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T\|_2$  in place of the objective function  $\rho(A_1 - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T)$ . The relaxed optimization problem is as follows.

$$\begin{aligned} & \underset{A_1}{\text{minimize}} && \left\|A_1 - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T\right\|_2 \\ & \text{subject to} && A_1\mathbf{1} = \mathbf{1}, A_1^T\mathbf{1} = \mathbf{1}, \\ & && (A_1)^{ij} = 0, (i, j) \notin \mathcal{E}^{\mathcal{F}} \end{aligned} \quad (B)$$

Since, for a directed graph,  $A_1 \neq A_1^T$ ,  $\rho(A_1 - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T) \leq \|A_1 - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T\|_2$ . Suppose  $A_1^*$  is the solution to above optimization problem (B), then with  $A_1^*$  being doubly stochastic it can be shown that  $\|A_1^* - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T\|_2 < 1$  ([18], Lemma 3), and hence  $\rho(A_1^* - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T) < 1$ .

Note that the above problem can be solved as a centralized optimization problem using any of the standard optimization solvers. Since our proposed Algorithm 1 is implemented in a distributed manner, we propose to solve the above problem (B) also in a distributed manner using ADMM over directed graphs. For this purpose, we modify the problem as follows.

$$\begin{aligned} & \underset{(A_1)_i \forall i \in \mathcal{F}}{\text{minimize}} && \sum_{i=1}^{n-m} \frac{\|(A_1)_i - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T\|_2}{n-m} \\ & \text{subject to} && (A_1)_i = (A_1)_j, (i, j) \in \mathcal{E}^{\mathcal{F}}, \\ & && (A_1)_i\mathbf{1} = \mathbf{1}, (A_1)_i^T\mathbf{1} = \mathbf{1}, \\ & && (A_1)_i^{ij} = 0, (i, j) \notin \mathcal{E}^{\mathcal{F}} \end{aligned} \quad (C)$$

In problem (C), each of the matrices  $(A_1)_i$  are copies of the optimization variable  $A_1$  maintained by each agent  $i \in \mathcal{F}$ , and are their estimates of the centralized optimal solution  $A_1^*$ . Now, since  $\mathcal{G}^{\mathcal{F}}$  is strongly connected,  $(A_1)_i = (A_1)_j$ ,  $(i, j) \in \mathcal{E}^{\mathcal{F}}$  implies that all  $(A_1)_i$ 's are equal. Hence, we conclude that (B) and (C) are equivalent. Since we want the constraints to be decoupled for a distributed approach, we ensure that each  $(A_1)_i$  has the graph topological constraints only on its  $i^{\text{th}}$  row and not the entire matrix. This is ensured in the last constraint of the above problem.

Now (C) can be implemented in a distributed manner, but parallel operation across all the agents is not possible due to the fact that  $(A_1)_i$ 's of different agents all come up in the same constraint,  $(A_1)_i = (A_1)_j$ ,  $(i, j) \in \mathcal{E}^{\mathcal{F}}$  in (C). In order to circumvent this issue, we formulate it using Fenchel duality ([12], Section 3.1.2). We introduce an auxiliary primal

variable  $Z \in \mathbb{R}^{(n-m) \times (n-m)}$  as follows, which decouples the consensus constraint by enforcing cohesion among all copies.

$$\begin{aligned} & \underset{(A_1)_i \forall i \in \mathcal{F}}{\text{minimize}} && \sum_{i=1}^{n-m} \frac{\|(A_1)_i - \frac{1}{n-m}\mathbf{1}\mathbf{1}^T\|_2}{n-m} \\ & \text{subject to} && (A_1)_i = Z, i \in \mathcal{F}, \\ & && (A_1)_i\mathbf{1} = \mathbf{1}, (A_1)_i^T\mathbf{1} = \mathbf{1}, \\ & && (A_1)_i^{ij} = 0, (i, j) \notin \mathcal{E}^{\mathcal{F}} \end{aligned} \quad (D)$$

We propose Algorithm 2 that solves the above optimization problem (D) in a distributed manner using ADMM for directed graphs. The steps involved in the derivation are given in the Appendix. ■

---

**Algorithm 2** Local estimation of optimal weight matrix  $A_1^*$  by agent  $i \in \mathcal{F}$  over a directed network using ADMM

---

**Initialize :**  $\rho > 0$ ,  $(A_1)_i(0) = 0_{n-m}$ ,  $Z_i(0) = 0_{n-m}$ ,  $C_i(0) = 0_{n-m}$ ,  $v_i(0, 0) > 0$  to a small value,  $\forall i \in \mathcal{F}$

**Iterate :**  $\forall i \in \mathcal{F}$

**for**  $k \geq 0$  **do**

**Primal Update :**

Evaluate  $(A_1)_i(k+1)$  as per (17).

Initialize  $M_i(0, k) = (A_1)_i(k+1)$

**for**  $t = 0, 1, \dots, H-1$  **do**

Send  $v_i(t, k)$  and  $M_i(t, k)$  to out-neighbours.

Compute  $v_i(t+1, k)$  using (20)

Compute  $M_i(t+1, k)$  using (19)

**end for**

Fix  $v_i(0, k+1) = v_i(H, k)$

Fix  $Z_i(k+1) = M_i(H, k)$

**Dual Update :**

Compute  $C_i(k+1)$  as per (18).

**end for**

---

*Remark 3:* The update step for the primal variable  $Z$  in optimization problem (D) involves computation of an average across all agents over a directed network, as shown in (16). Hence for each iteration  $k$  in Algorithm 2, every agent  $i \in \mathcal{F}$  computes an estimate of  $Z$ , denoted as  $Z_i$ , by performing  $H$  rounds of communication in the inner loop. The performance of the algorithm improves as  $H$  increases. A lower bound on  $H$  for the convergence of the algorithm is discussed in [17].

*Remark 4:* (Stopping Criterion for Algorithm 2) Each of the agent  $i \in \mathcal{F}$  fixes a small, arbitrary tolerance value,  $\epsilon > 0$ , and calculates its *residual* values at each iteration  $k$  as

$$r_{ij}^1(k) = \frac{1}{n-m} \|(A_1)_i(k+1) - (A_1)_j(k+1)\|_F, j \in \mathcal{N}_i^{\text{in}} \quad (7)$$

$$r_{ij}^2(k) = |(A_1)_i^{ij}(k)|, j \notin \mathcal{N}_i^{\text{in}} \quad (8)$$

which indicates the extent of violations of the constraints involved. The stopping criterion for Algorithm 2 is when

$$R_i(k) = \max\{r_{ij}^1(k), r_{ij}^2(k); \forall j\} \leq \epsilon \forall i \in \mathcal{F} \quad (9)$$

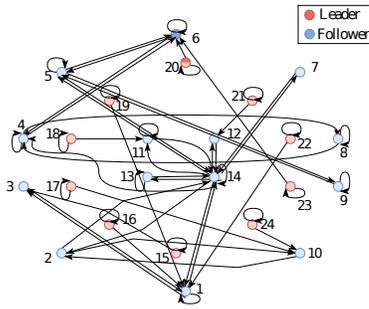


Fig. 1. A directed topology of  $n = 24$  agents

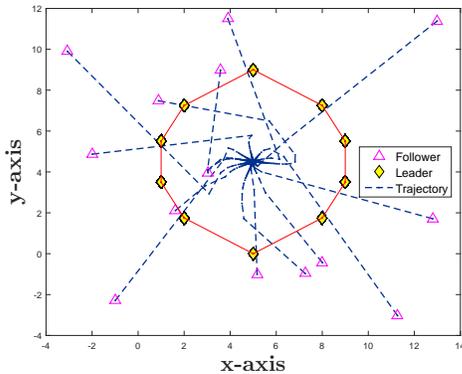


Fig. 2. State trajectories of all agents

## V. EXAMPLE

In this section, we present a numerical example that illustrates our result when all the agents are in 2-D plane.

We consider a set of 24 agents interacting over a directed network topology, i.e.,  $n = 24$ . Assume that there are  $m = 10$  leaders, and the rest are followers interacting as shown in Fig. 1. Let  $x^1$  and  $x^2$  denote two sets of state variable that indicates the position of each agent in 2-D plane. The state values of the stationary leader agents are given by  $x_L^1 = [5 \ 3 \ 2 \ 2 \ 3 \ 5 \ 7 \ 8 \ 8 \ 7]^T$  and  $x_L^2 = [1 \ 2 \ 4 \ 5 \ 7 \ 8 \ 7 \ 5 \ 4 \ 2]^T$ . Algorithm 2 is employed to compute an optimal  $A_1$  matrix in a distributed manner. Then using that  $A_1$  and any random choice of  $A_2$  following Assumption 1, Algorithm 1 is employed to ensure that follower agents reach a consensus to the average of the stationary leader states which is  $(\bar{x}_L^1, \bar{x}_L^2) = (5, 4.5)$ , is illustrated in Fig. 2.

Computation method	$\ A_1 - \frac{1}{n-m} \mathbf{1}\mathbf{1}^T\ _2$	$\rho(A_1 - \frac{1}{n-m} \mathbf{1}\mathbf{1}^T)$
Centralized $[A_1^*]$	<b>0.7071</b>	0.5010
WBA	0.9344	0.9317
Algorithm 2	<b>0.7086</b>	0.5146

TABLE I

CONVERGENCE FACTORS OF DIFFERENT WEIGHT MATRICES FOR THE NETWORK IN FIG.1

The convergence factors of the weight matrices ( $A_1$ ) generated by different methods have been given in Table I, for comparison purposes. A weight-balance algorithm (WBA) proposed in [10] generates balancing weights for a directed network, which, although not optimal, is computed

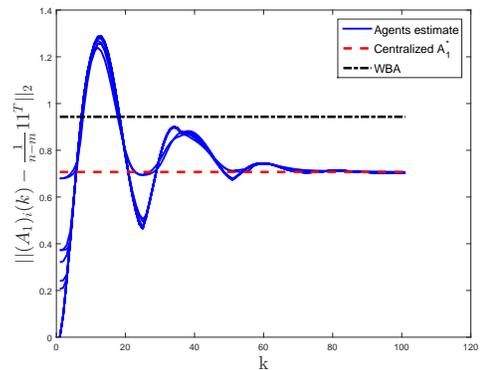


Fig. 3. Objective variation plot of the agents to the centralized optimal value  $\|A_1^* - \frac{1}{n-m} \mathbf{1}\mathbf{1}^T\|_2$

in a distributed manner.

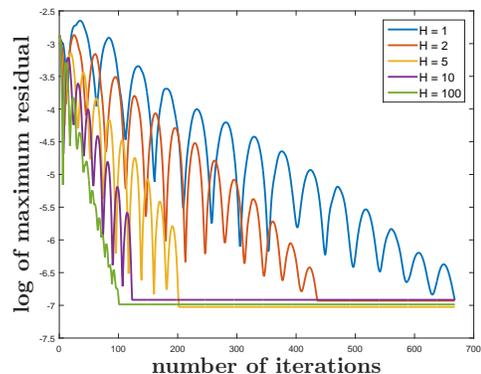


Fig. 4. Maximum log residual plot across all agents with  $\rho = 5$

It can be observed that  $A_1$  computed by Algorithm 2 is quite close to the optimal  $A_1^*$  given by centralized methods as depicted in Fig. 3, and performs much better than that of WBA. In Remark 3, we explained how the performance of Algorithm 2 improves with increase in  $H$ . This is illustrated in Fig. 4 where we observe that the *residual* values decreases faster for increase in values of  $H$ .

## VI. CONCLUSION

We considered a special case of the discrete-time containment control problem for multi-agent systems with multiple stationary leaders, interacting over a directed communication topology. We studied the problem of achieving faster convergence to a consensus among followers to the average of the leaders' state, terming it as the average consensus containment. First, we provided sufficient conditions that guarantee average consensus containment. A distributed algorithm was proposed that ensured average consensus containment. Using this analysis, an optimization problem was formulated to maximize the convergence rate which was also solved in a distributed manner using ADMM algorithm. Through numerical examples, we demonstrated that the optimal convergence rate obtained using locally computed weights in a distributed manner is very close to the optimal value obtained by centralized methods. Moving forward, it would be interesting

to investigate analogous results using dynamic leaders, and under a continuous-time setting.

## REFERENCES

- [1] Yongcan Cao and Wei Ren. Containment control with multiple stationary or dynamic leaders under a directed interaction graph. In *Proceedings of the 48th IEEE Conference on Decision and Control (CDC)*, pages 3014–3019. IEEE, 2009.
- [2] Yongcan Cao, Daniel Stuart, Wei Ren, and Ziyang Meng. Distributed containment control for multiple autonomous vehicles with double-integrator dynamics: algorithms and experiments. *IEEE Transactions on Control Systems Technology*, 19(4):929–938, 2010.
- [3] Jiu Ding and Noah H Rhee. On the equality of algebraic and geometric multiplicities of matrix eigenvalues. *Applied Mathematics Letters*, 24(12):2211–2215, 2011.
- [4] Ali Jadbabaie, Jie Lin, and A Stephen Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.
- [5] Meng Ji, Giancarlo Ferrari-Trecate, Magnus Egerstedt, and Annalisa Buffa. Containment control in mobile networks. *IEEE Transactions on Automatic Control*, 53(8):1972–1975, 2008.
- [6] David Kempe, Alin Dobra, and Johannes Gehrke. Gossip-based computation of aggregate information. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 482–491. IEEE, 2003.
- [7] Usman A Khan, Soumya Kar, and José MF Moura. Higher dimensional consensus: Learning in large-scale networks. *IEEE Transactions on Signal Processing*, 58(5):2836–2849, 2010.
- [8] Bo Li, Hong-Yong Yang, Zeng-Qiang Chen, and Zhong-Xin Liu. Distributed containment control of multi-agent systems with general linear dynamics and time-delays. *International Journal of Control, Automation and Systems*, 16(6):2718–2726, 2018.
- [9] Huiyang Liu, Guangming Xie, and Long Wang. Necessary and sufficient conditions for containment control of networked multi-agent systems. *Automatica*, 48(7):1415–1422, 2012.
- [10] Ali Makhdoumi and Asuman Ozdaglar. Graph balancing for distributed subgradient methods over directed graphs. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 1364–1371. IEEE, 2015.
- [11] Carl D Meyer. *Matrix Analysis and Applied Linear Algebra*, volume 71. Siam, 2000.
- [12] Giuseppe Notarstefano, Ivano Notarnicola, and Andrea Camisa. Distributed optimization for smart cyber-physical networks. *arXiv preprint arXiv:1906.10760*, 2019.
- [13] Reza Olfati-Saber, J Alex Fax, and Richard M Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.
- [14] Michael L Overton and Robert S Womersley. On minimizing the special radius of a nonsymmetric matrix function: Optimality conditions and duality theory. *SIAM Journal on Matrix Analysis and Applications*, 9(4):473–498, 1988.
- [15] Wei Ren and Randal W Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, 50(5):655–661, 2005.
- [16] R Tyrrell Rockafellar. *Convex analysis*, volume 36. Princeton university press, 1997.
- [17] Kiran Rokade and Rachel Kalpana Kalaimani. Distributed admm over directed graphs. *arXiv preprint arXiv:2010.10421*, 2020.
- [18] Kiran Rokade and Rachel Kalpana Kalaimani. Distributed computation of optimal weights for consensus using admm. *arXiv preprint arXiv:2002.08106*, 2020.
- [19] Dong Wang and Wei Wang. Necessary and sufficient conditions for containment control of multi-agent systems with time delay. *Automatica*, 103:418–423, 2019.
- [20] Feng Xiao and Long Wang. Consensus behavior of agents in networked systems under general communication topologies. In *2006 IEEE Conference on Computer Aided Control System Design*, pages 862–867. IEEE, 2006.

## APPENDIX

### PROOF OF THEOREM 1

Since  $A_1$  is a column stochastic matrix associated with a strongly-connected graph (Assumption 1), it is irreducible

[3]. Furthermore, [11, Sec. 8.3],  $A_1$  is primitive as its non-negative irreducible with positive diagonal entries due to each follower knowing its own state. Hence, 1 is a simple eigenvalue of  $A_1$ . Let  $u_r$  and  $u_l$  respectively denote the right and left eigenvector of  $A_1$  corresponding to eigenvalue 1. Note that  $u_l = \frac{1_{n-m}}{\sqrt{n-m}}$ . Using [11, Sec 7.10] we get

$$\lim_{k \rightarrow \infty} A_1^k = G, \quad (10)$$

where,  $G$  is the spectral projector onto nullspace of  $(I_{n-m} - A_1)$  along the column space of  $(I_{n-m} - A_1)$ , and is given by  $G = u_r u_l^T$ . Also, as per [11, Sec 8.4],  $A_1$  being column-stochastic, is Cesáro summable, with the Cesáro limit

$$\lim_{k \rightarrow \infty} \frac{A_1^{k-1} + \dots + A_1 + I_{n-m}}{k} = G, \quad (11)$$

From (10), (11),

$$\lim_{k \rightarrow \infty} \frac{A_1^{k-1} + \dots + A_1 + I_{n-m}}{k} = u_r u_l^T \quad (12)$$

In Algorithm 1, from Step 3 we get  $x_i(k) = \frac{[s(k)]_i}{[w(k)]_i}$  which when using the update law  $\forall i \in \mathcal{F}$ , becomes

$$x_i(k) = \frac{[A_1^k x_F(0) + \{A_1^{k-1} + A_1^{k-2} + \dots + A_1 + I_{n-m}\} A_2 x_L(0)]_i}{[A_1^k 1_{n-m} + \{A_1^{k-1} + A_1^{k-2} + \dots + A_1 + I_{n-m}\} A_2 1_m]_i}$$

Dividing numerator and denominator by  $k$  and taking limits with  $x_i^s := \lim_{k \rightarrow \infty} x_i(k)$ .

$$x_i^s = \frac{\left[ \lim_{k \rightarrow \infty} \frac{A_1^k x_F(0)}{k} + \lim_{k \rightarrow \infty} \left\{ \frac{A_1^{k-1} + A_1^{k-2} + \dots + A_1 + I_{n-m}}{k} \right\} A_2 x_L(0) \right]_i}{\left[ \lim_{k \rightarrow \infty} \frac{A_1^k 1_{n-m}}{k} + \lim_{k \rightarrow \infty} \left\{ \frac{A_1^{k-1} + A_1^{k-2} + \dots + A_1 + I_{n-m}}{k} \right\} A_2 1_m \right]_i}$$

From (10),  $\left[ \lim_{k \rightarrow \infty} \frac{A_1^k x_F(0)}{k} \right]_i \rightarrow 0$  and  $\left[ \lim_{k \rightarrow \infty} \frac{A_1^k 1_{n-m}}{k} \right]_i \rightarrow 0$ .

From (10), (12), we get

$$x_i^s = \frac{\left[ \lim_{k \rightarrow \infty} A_1^k A_2 x_L(0) \right]_i}{\left[ \lim_{k \rightarrow \infty} A_1^k A_2 1_m \right]_i} = \frac{[u_r 1_{n-m}^T A_2 x_L(0)]_i}{[u_r 1_{n-m}^T A_2 1_m]_i} \quad (13)$$

Since,  $A_2$  is column stochastic,  $1_{n-m}^T A_2 = 1_m^T$ . Hence,

$$x_i^s = \frac{[u_r 1_m^T x_L(0)]_i}{[u_r 1_m^T 1_m]_i} = \frac{[u_r 1_m^T x_L(0)]_i}{[u_r m]_i} = \frac{1_m^T x_L(0)}{m} =: \bar{x}_L(0)$$

From the above equation, we observe that the followers are in consensus at the average of the leaders' state values. Hence the proof. ■

### PROOF OF LEMMA 1

From (13) we have  $\forall i \in \mathcal{F}$ ,

$$\lim_{k \rightarrow \infty} x_i(k) = \frac{\left[ \lim_{k \rightarrow \infty} A_1^k A_2 x_L(0) \right]_i}{\left[ \lim_{k \rightarrow \infty} A_1^k A_2 1_m \right]_i} \quad (14)$$

$A_1$  is primitive and as such, its eigenvalues satisfy  $\lambda_1 = 1 > \lambda_2 \geq \dots \geq \lambda_{n-m}$ . Therefore, the convergence rate of  $x_i(k)$  is determined by  $\lambda_2 = \rho(A_1 - \frac{1}{n-m} \mathbf{1}\mathbf{1}^T)$ . ■

DERIVATION OF ALGORITHM 2

The augmented Lagrangian ( $L_\rho$ ) for ( $D$ ) is given by

$$L_\rho = \frac{1}{n-m} \sum_{i=1}^{n-m} \left\| (A_1)_i - \frac{1}{n-m} \mathbf{1}\mathbf{1}^T \right\|_2 + \sum_{i=1}^{n-m} \text{trace} [((A_1)_i - Z)^T C_i] + \sum_{i=1}^{n-m} \frac{\rho}{2} \left\| (A_1)_i - Z \right\|_F^2 \quad (15)$$

where  $\rho > 0$  is the penalty parameter. Now, the standard 2-block ADMM algorithm for ( $D$ ) is given below. The primal variables  $[(A_1)_i(k) \forall i \in \mathcal{F}, Z(k)]$  are updated sequentially and dual variables  $[C_i(k) \forall i \in \mathcal{F}]$  are updated afterwards. The constraint set  $S := \{(A_1)_i \mid (A_1)_i \mathbf{1} = \mathbf{1}, (A_1)_i^T \mathbf{1} = \mathbf{1}, (A_1)_i^{ij} = 0, (i, j) \notin \mathcal{E}^{\mathcal{F}}\}$  has been incorporated into primal update steps rather than being dualized.

**$Z$  update step :**  $Z(k+1) = \underset{Z}{\operatorname{argmin}} \{L_\rho\}$

$$Z(k+1) = \frac{1}{n-m} \sum_{i=1}^{n-m} \left[ \frac{C_i(k)}{\rho} + (A_1)_i(k+1) \right] \quad (16)$$

The  $Z$  update step involves computing an average across all follower agents over a directed network. Hence we use an approach based on *dynamic average consensus*, proposed in [17], to get the  $Z$  update step, in a distributed manner. Each agent  $i \in \mathcal{F}$  maintains  $Z_i$  which is an estimate for  $Z$ . Before we provide a distributed update law for  $Z$ , we write the update law for the other variables.

**$(A_1)_i$  update step :**  $(A_1)_i(k+1) = \underset{S}{\operatorname{argmin}} \{L_\rho\}$

$$(A_1)_i(k+1) = \underset{S}{\operatorname{argmin}} \left\{ \frac{\| (A_1)_i - \frac{1}{n-m} \mathbf{1}\mathbf{1}^T \|_2}{n-m} + \text{trace} [((A_1)_i - Z_i(k))^T C_i(k)] + \frac{\rho}{2} \left\| (A_1)_i - Z_i(k) \right\|_F^2 \right\} \quad (17)$$

**$C_i$  update step :**

$$C_i(k+1) = C_i(k) + \rho [(A_1)_i(k+1) - Z_i(k+1)] \quad (18)$$

Note that in both (17) and (18), we use the estimate  $Z_i$  instead of  $Z$ . Next, we simplify the  $Z$  update step in (16). At every iteration  $k$ , each agent initializes a variable  $M_i(0, k) := Z_i(k) + (A_1)_i(k+1) - (A_1)_i(k) + \frac{C_i(k) - C_i(k-1)}{\rho}$ . From (18), this is further simplified as  $M_i(0, k) = (A_1)_i(k+1)$ . For some  $H \geq 1, 0 \leq t \leq H-1, i \in \mathcal{F}$  do

$$M_i(t+1, k) = [1 - d_i v_i(t, k)] M_i(t, k) + \sum_{j \in \mathcal{N}_i^{in}} v_j(t, k) M_j(t, k) \quad (19)$$

$$v_i(t+1, k) = \frac{1}{2} [v_i(t, k) + \frac{1}{d_i} \sum_{j \in \mathcal{N}_i^{in}} v_j(t, k)] \quad (20)$$

Then  $Z_i(k+1) = M_i(H, k)$ . In (19),  $d_i := |\mathcal{N}_i^{out}|$  is the out-degree of  $i^{th}$  follower agent. In (20),  $v_i(t, k) \in \mathbb{R}$  is the node-weight used by  $i^{th}$  agent to scale its outgoing information and evolves dynamically as shown. As defined in [10], the node-weights are initialized as  $v_i(0, 0) \leq (1/d^*)^{2D+1}$ , where  $D$  is the diameter of the subgraph  $\mathcal{G}^{\mathcal{F}}$  and  $d^* = \max_{i \in \mathcal{F}} d_i$ . Each of the agent  $i \in \mathcal{F}$  initialises a small value as its node-weight,  $v_i(0, 0) > 0$ . ■