Graph Distances and Controllability of Networks

A. Yasin Yazıcıoğlu, Waseem Abbas, and Magnus Egerstedt

Abstract—In this technical note, we study the controllability of diffusively coupled networks from a graph theoretic perspective. We consider leader-follower networks, where the external control inputs are injected to only some of the agents, namely the leaders. Our main result relates the controllability of such systems to the graph distances between the agents. More specifically, we present a graph topological lower bound on the rank of the controllability matrix. This lower bound is tight, and it is applicable to systems with arbitrary network topologies, coupling weights, and number of leaders. An algorithm for computing the lower bound is also provided. Furthermore, as a prominent application, we present how the proposed bound can be utilized to select a minimal set of leaders for achieving controllability, even when the coupling weights are unknown.

I. INTRODUCTION

Networks of diffusively coupled agents appear in numerous systems such as sensor networks (e.g., [1]), distributed robotics (e.g., [2]), power grids (e.g., [3]), social networks (e.g., [4]), and biological systems (e.g., [5]). A central question regarding such networks is whether a desired global behavior can be induced by directly manipulating only a small subset of the agents, referred to as the "leaders" in the network. This question has motivated numerous studies on the controllability of networks. In particular, there has been a large interest in relating the *network controllability* to the structure of the inter-action graph. In this technical note, we relate the controllability of diffusively coupled agents with single integrator dynamics to the distances between the nodes on the interaction graph.

Various graph theoretic tools have recently been utilized to provide some topology based characterizations of network controllability. Some of the graph theoretic constructs that are widely employed for this purpose include *equitable partitions* (e.g., [6], [7]), *maximum matching* (e.g., [8], [11]), *centrality based measures* (e.g., [13], [14]), and *dominating sets* (e.g., [15]). Recently, the *graph distances* have been used to acquire further insight on how the network structure and the locations of the leaders influence the network controllability [16], [17], [18]. The graph distances, which rely purely on the shortest paths between the nodes, provide a computationally tractable and perceptible characterization of the graph structure. Thus, the distance-based relationships reveal some innate connections between the network topology and the network controllability.

The main contribution of this technical note is a distance-based tight lower bound on the dimension of the controllable

A. Yasin Yazıcıoğlu is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, yasiny@mit.edu.

Waseem Abbas is with the Institute for Software Integrated Systems, Vanderbilt University, waseem.abbas@vanderbilt.edu

Magnus Egerstedt is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, magnus@gatech.edu.

subspace (Theroem 3.2). The bound is generic in the sense that it is applicable to systems with arbitrary network topologies, coupling weights, and number of leaders. Based on the distances between the leaders and the followers, we first define the distance-to-leaders (DL) vectors. Then, we define a certain ordering rule and derive the lower bound as the maximum length of the sequences of the DL vectors that satisfy the rule (Section III). An algorithm to compute the proposed bound is also presented in Section IV.

A prominent attribute of the proposed bound is that, unlike the dimension of the controllable subspace, it does not depend on the coupling weights. Thus, the bound is useful in many applications, particularly when the information about the network is incomplete. As an example, in Section V, we present how the bound can be used in leader selection for achieving the controllability of any given network, even when the coupling weights are unknown. Finally, some conclusions are provided in Section VI.

II. PRELIMINARIES

A. Graph Theory

A graph $\mathcal{G} = (V, E)$ consists of a node set $V = \{1, 2, \ldots, n\}$ and an edge set $E \subseteq V \times V$. For an undirected graph, each edge is represented as an unordered pair of nodes. For each $i \in V$, let $\mathcal{N}_i$ denote the *neighborhood* of i, i.e.,

$$\mathcal{N}_i = \{j \in V \mid (i,j) \in E\} \tag{1}$$

A path between a pair of nodes $i, j \in V$ is a sequence of nodes $\{i, \ldots, j\}$ such that each pair of consecutive nodes are linked by an edge. The *distance* between the nodes, $dist(i, j)$, is equal to the number of edges that belong to the shortest path between the nodes. A graph is *connected* if there exists a path between every pair of nodes. A graph is weighted if there is a corresponding weighting function $w : E \mapsto \mathbb{R}^+$. The *adjacency matrix*, $\mathcal{A}$, of a weighted graph is defined as

$$\mathcal{A}_{ij} = \begin{cases} w(i,j) & \text{if } (i,j) \in E \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

For any adjacency matrix $\mathcal{A}$, the corresponding *degree matrix*, Δ, is defined as

$$\Delta_{ij} = \begin{cases} \sum_{k \in \mathcal{N}_i} \mathcal{A}_{ik} & \text{if } i = j \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

The *graph Laplacian*, L, is defined as the difference of the degree matrix and the adjacency matrix, i.e.,

$$L = \Delta - \mathcal{A}. \tag{4}$$

B. Leader-Follower Networks

A network of diffusively coupled agents can be represented as a graph, where the nodes correspond to the agents, and the weighted edges exist between the coupled agents. For such a network $\mathcal{G} = (V, E)$, let the dynamics of each agent $i \in V$ be

$$\dot{x}_i = \sum_{j \in \mathcal{N}_i} w(i,j)(x_j - x_i), \tag{5}$$

where x_i denotes the state of i, and $w(i,j) \in \mathbb{R}^+$ represents the strength of the coupling between i and j.

In a *leader-follower* setting, the objective is to drive the overall system by injecting external control inputs to some of the nodes, which are called the *leaders*. The set of leaders can be represented as $\mathcal{L} = \{l_1, \ldots, l_m\} \subseteq V$, where, without loss of generality, the leaders are labeled such that $l_j < l_{j+1}$. For any leader-follower network, a global state vector x can be obtained by stacking the states of all the nodes. Without loss of generality, let $x \in \mathbb{R}^n$, and let $u \in \mathbb{R}^m$ be the control input injected to the leaders. Then, the overall dynamics of the system can be expressed as

$$\dot{x} = -Lx + Bu, \tag{6}$$

where B is an $n \times m$ matrix with the following entries

$$B_{ij} = \begin{cases} 1 & \text{if } i = l_j. \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

For the system in (6), the *controllable subspace* consists of the states that can be reached from $x(0) = \mathbf{0}$ in any finite time via an appropriate choice of $u(t)$. The controllable subspace is the range space of the *controllability matrix*, i.e.,

$$\Gamma = \begin{bmatrix} B & (-L)B & (-L)^2 B & \ldots & (-L)^{n-1} B \end{bmatrix}. \tag{8}$$

III. LEADER-FOLLOWER DISTANCES AND CONTROLLABILITY

In this section, we present a connection between the controllability of networks and the distances of the nodes to the leaders on the interaction graph. More specifically, we utilize such distances to define a tight lower bound on the dimension of the controllable subspace, i.e. the rank of the controllability matrix. First, we provide some definitions prior to our analysis.

Definition 3.1 *(Distance-to-Leaders (DL) Vector)*: For each node i in a network with m leaders, the DL vector $d_i \in \mathbb{R}^m$ is defined as

$$d_{i,j} = dist(i, l_j), \tag{9}$$

where $d_{i,j}$ denotes the jth entry of d_i, and l_j denotes the jth leader for $j \in \{1, 2, \ldots, m\}$.

In our analysis, we utilize a specific sequence, which we define as a *pseudo-monotonically increasing sequence*, of the DL vectors. For any vector sequence D, let D_i be the ith vector in the sequence, and let $D_{i,j}$ denote the jth entry of D_i.

Definition 3.2 *(Pseudo-Monotonically Increasing (PMI) Sequence)*: A sequence D of vectors, where each vector is in $\mathbb{R}^m$, is a PMI sequence if for every D_i there exists some $\alpha(i) \in \{1, 2, \cdots, m\}$ such that

$$D_{i,\alpha(i)} < D_{j,\alpha(i)}, \ \forall i < j. \tag{10}$$

Condition (10) simply means that if D_i is the ith vector in the sequence with $D_{i,\alpha(i)}$ being its $\alpha(i)^{th}$ entry, then the corresponding (i.e. $\alpha(i)^{th}$) entries of all the subsequent vectors in the sequence should be greater than $D_{i,\alpha(i)}$.

Example – Consider the network shown in Fig. 1. For this network, one can build a PMI sequence of five DL vectors as

$$D = (D_1, D_2, D_3, D_4, D_5) = (d_1, d_6, d_5, d_3, d_4)$$
$$= \left(\begin{bmatrix} \textcircled{0} \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ \textcircled{0} \end{bmatrix}, \begin{bmatrix} 2 \\ \textcircled{1} \end{bmatrix}, \begin{bmatrix} \textcircled{1} \\ 2 \end{bmatrix}, \begin{bmatrix} \textcircled{2} \\ 2 \end{bmatrix} \right),$$

where, for each D_i, the $\alpha(i)^{th}$ entry that satisfies (10) is encircled. For instance, consider D_1, for which $\alpha(1) = 1$ and $D_{1,1} = 0$. Note that the first entries of all the subsequent vectors are greater than 0. Similarly, for the second vector D_2, $\alpha(2) = 2$ and $D_{2,2} = 0$. Note that $D_{j,2} > 0$ for $j > 2$, and so on. For this example, another PMI sequence of five DL vectors could be build as $(d_1, d_3, d_6, d_5, d_4)$, where $\alpha(1) = \alpha(2) = 1$ and $\alpha(3) = \alpha(4) = \alpha(5) = 2$.
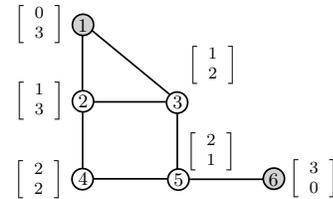

Fig. 1. A leader-follower network with two leaders (shown in gray), $l_1 = 1$ and $l_2 = 6$. Each node i has its DL vector d_i given next to itself.

For any connected $\mathcal{G} = (V, E)$ and any set of leaders $\mathcal{L} \subseteq V$, let $\mathcal{D}_{\mathcal{L}}$ denote the set of all PMI sequences of the corresponding DL vectors. Furthermore, let $\delta_{\mathcal{L}}$ denote the length of the longest sequence in $\mathcal{D}_{\mathcal{L}}$, i.e.

$$\delta_{\mathcal{L}} = \max_{D \in \mathcal{D}_{\mathcal{L}}} |D|. \tag{11}$$

In the following analysis, we show that, for any weighting function $w : E \mapsto \mathbb{R}^+$, the rank of the resulting controllability matrix is lower bounded by $\delta_{\mathcal{L}}$.

Lemma 3.1. *Let $\mathcal{G} = (V, E)$ be a connected graph. Then, for any weighting function $w : E \mapsto \mathbb{R}^+$,*

$$[(-L)^r]_{ij} \begin{cases} = 0 & \text{if } 0 \leq r < dist(i,j), \\ \neq 0 & \text{if } r = dist(i,j), \end{cases} \tag{12}$$

where $dist(i,j)$ is the distance between i and j on $\mathcal{G}$.

Proof. Using (4), $(-L)^r$ can be expanded as

$$(-L)^r = (\mathcal{A} - \Delta)^r = \mathcal{A}^r + \sum_{m=0}^{r-1} (-1)^{r-m} \mathcal{S}_m, \tag{13}$$

where $\mathcal{S}_m$ denotes the sum of all matrices that can be expressed as a multiplication of m copies of $\mathcal{A}$ and $r - m$ copies of Δ. For instance, if $r = 2$, then $\mathcal{S}_0 = \Delta^2$ and $\mathcal{S}_1 = \mathcal{A}\Delta + \Delta\mathcal{A}$. Note that, for $w : E \mapsto \mathbb{R}^+$, any matrix

that can be represented as such a multiplication has only non-negative entries since Δ and $\mathcal{A}$ have only non-negative entries. Moreover, Δ is a diagonal matrix and, for any connected graph, it has only positive entries on the main diagonal. As such, it does not alter the signs of entries when multiplied by a matrix. Hence,

$$[\mathcal{S}_m]_{ij} = 0 \Leftrightarrow [\mathcal{A}^m]_{ij} = 0. \tag{14}$$

Using (13),

$$[(-L)^r]_{ij} = [\mathcal{A}^r]_{ij} + \sum_{m=0}^{r-1} (-1)^{r-m} [\mathcal{S}_m]_{ij}. \tag{15}$$

Since $\mathcal{A}$ is the adjacency matrix of a weighted graph with positive edge weights, $[\mathcal{A}^k]_{ij}$ is equal to a positive scalar times the number of walks of length k from node i to node j. Hence, $[\mathcal{A}^k]_{ij} = 0$ for all $0 \le k < dist(i,j)$, and $[\mathcal{A}^{dist(i,j)}]_{ij} \ne 0$ for any connected graph. Consequently, (14) and (15) together imply (12). $\qquad\square$

Theorem 3.2. *Let $\mathcal{G} = (V, E)$ be a connected graph, and let $\mathcal{L} \subseteq V$ be the set of leaders. Then, for any weighting function $w : E \mapsto \mathbb{R}^+$, the controllability matrix Γ satisfies*

$$rank(\Gamma) \ge \delta_{\mathcal{L}}, \tag{16}$$

where $\delta_{\mathcal{L}}$ is defined in (11).

Proof. For any connected $\mathcal{G} = (V, E)$ and any set of leaders $\mathcal{L}$, let $\{d_1, d_2, \cdots, d_n\}$ be the corresponding DL vectors. Let $D = (D_1, D_2, \cdots, D_{\delta_{\mathcal{L}}})$ be a PMI sequence of maximum length. Now, consider vectors of the form

$$(-L)^{r_p} b_{\alpha(p)}, \tag{17}$$

where $\alpha(p)$ is the index of D_p as per the definition of a PMI sequence (Definition 3.2), $r_p = D_{p,\alpha(p)}$, and $b_{\alpha(p)}$ denotes the $\alpha(p)^{th}$ column of the input matrix B. If D_p corresponds to the DL vector of node i, i.e. $D_p = d_i$, then $r_p = d_{i,\alpha(p)}$. As a result of Lemma 3.1, i^{th} entry of the vector in (17) is non-zero. Also, for any node j with $d_{j,\alpha(i)} > d_{i,\alpha(i)}$, the jth entry of the vector in (17) equal to zero. Using this along with the definition of PMI sequences, we conclude that the $n \times \delta_{\mathcal{L}}$ matrix

$$\begin{bmatrix} (-L)^{r_1} b_{\alpha(1)} & (-L)^{r_2} b_{\alpha(2)} & \dots & (-L)^{r_{\delta_{\mathcal{L}}}} b_{\alpha(\delta_{\mathcal{L}})} \end{bmatrix}, \tag{18}$$

has a full column rank since each column contains the leftmost non-zero entry in some rows. Note that for every $p \in \{1, 2, \dots, \delta_{\mathcal{L}}\}$, we have $r_p = D_{p,\alpha(p)} \le n - 1$ since the distance between any two nodes is always smaller than or equal to $n - 1$. Hence, each column of the matrix in (18) is also a column of Γ. Consequently, $rank(\Gamma) \ge \delta_{\mathcal{L}}$. $\qquad\square$

Since (16) holds for any weighting function $w : E \mapsto \mathbb{R}^+$, the proposed lower bound is closely related to the notion of *strong structural controllability*. A network is said to be structurally controllable if and only if there exists $w : E \mapsto \mathbb{R}^+$ such that the resulting controllability matrix is of full rank [9]. Furthermore, a network is said to be strongly structurally controllable if and only if for any $w : E \mapsto \mathbb{R}^+$, the resulting

controllability matrix is of full rank [10]. In this regard, the dimension of controllable subspace in the sense of strong structural controllability can be defined as the minimum of the ranks of all controllability matrices obtained by using arbitrary weighting functions $w : E \mapsto \mathbb{R}^+$ [12]. Accordingly, $\delta_{\mathcal{L}}$ is essentially a lower bound on the dimension of controllable subspace in the sense of strong structural controllability.

We also emphasize that the lower bound in Theorem 3.2 is tight, i.e., there exist $\mathcal{G} = (V, E)$, $\mathcal{L} \subseteq V$, and $w : E \mapsto \mathbb{R}^+$ such that $rank(\Gamma) = \delta_{\mathcal{L}}$. Any cycle graph with two adjacent nodes being the leaders, or any path graph with a leaf node being the leader are some of the examples that satisfy (16) with equality. In these two examples, $rank(\Gamma) = \delta_{\mathcal{L}}$ follows from the fact that both cases lead to $\delta_{\mathcal{L}} = n$, where n is the total number of nodes in the graph. In the remainder of this section, we present the connections between the proposed lower bound and some closely related distance-based measures, namely the maximum distance to the leaders and the number of unique DL vectors.

For networks with a single leader, any PMI sequence of the DL vectors consists of one dimensional vectors with monotonically increasing entries. Hence, for such networks, $\delta_{\mathcal{L}}$ is equal to one plus the maximum distance to the leader, which was indeed proposed in [16] as a lower bound of the controllability matrix for single-leader networks. An extension to the case of multiple leaders was later presented in [17] by taking the maximum of this value among all the leaders, i.e.

$$\mu_{\mathcal{L}} = \max_{i \in V, j \in \mathcal{L}} dist(i,j) + 1. \tag{19}$$

The relationship between $\mu_{\mathcal{L}}$ and $\delta_{\mathcal{L}}$ can be seen through the following observation: If one considers only the PMI sequences that satisfy (10) for some fixed entry, i.e. $\alpha(1) = \dots = \alpha(|D|)$, then the longest PMI sequence in this constrained set, $\mathcal{D}_{\mathcal{L}}^* \subseteq \mathcal{D}_{\mathcal{L}}$, has length $\mu_{\mathcal{L}}$. Hence, one can conclude that the following inequality holds for any leader-follower network:

$$rank(\Gamma) \ge \delta_{\mathcal{L}} \ge \mu_{\mathcal{L}}. \tag{20}$$

In light of (20), while the two quantities are equal for single-leader networks, the proposed lower bound $\delta_{\mathcal{L}}$ is richer than $\mu_{\mathcal{L}}$ in capturing the controllability of networks with multiple leaders. In fact, since the maximum distance between any two nodes in a graph is the diameter of the graph by definition, $\mu_{\mathcal{L}}$ is always less than or equal to one plus the diameter of the graph, even when every node is a leader. In general, the difference between $\delta_{\mathcal{L}}$ and $\mu_{\mathcal{L}}$ depends on the graph topology and the leader assignment. For instance, the example in Fig. 1 yields $\delta_{\mathcal{L}} = 5$ and $\mu_{\mathcal{L}} = 3$. Numerical comparisons of the bounds for Erdös-Renyi and Barabási-Albert graphs with two leaders are illustrated in Fig. 2. In this figure, each point on the plot corresponds to the average value for 50 randomly generated cases (each case is a randomly generated graph and a pair of randomly assigned leaders). The results indicate that $\delta_{\mathcal{L}}$ provides a significantly better utilization of the graph distances in the controllability analysis of multi-leader networks, even when the network has only a pair of leaders.
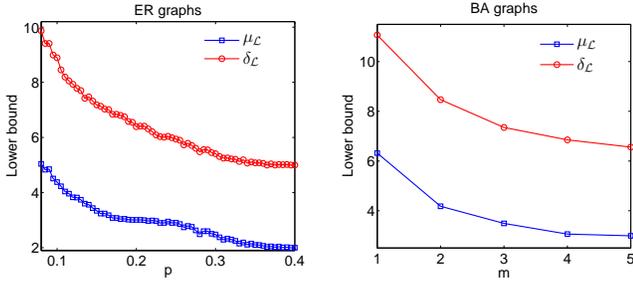
Fig. 2. Comparison of the lower bounds for two randomly selected leaders on Erdös-Renyi random graphs with 50 nodes in which any two nodes are adjacent with the probability p; and Barabási-Albert graph with 50 nodes in which each new node is connected to m existing nodes through a preferential attachment strategy.

The proposed lower bound is also closely related to the number of unique DL vectors. For any connected $\mathcal{G} = (V, E)$ and any leader set $\mathcal{L} \subseteq V$, let $\upsilon_{\mathcal{L}}$ be the number of unique DL vectors. Note that, due to (10), each vector in a PMI sequence has an entry that is strictly smaller than the corresponding entries of all the following vectors in that sequence. Hence, a PMI sequence cannot contain two identical vectors. Consequently, the proposed lower bound is always less than or equal to the number of unique DL vectors, i.e.

$$\upsilon_{\mathcal{L}} \geq \delta_{\mathcal{L}}. \tag{21}$$

The relationship in (21) facilitates a deeper understanding for potential applications of the proposed lower bound. For instance, for any $k \in \{1, \ldots, n\}$, having $\upsilon_{\mathcal{L}} = k$ is a necessary condition to have $\delta_{\mathcal{L}} = k$. Hence, it is possible to conclude that a network is completely controllable as per the proposed lower bound, only if each follower has a distinctive DL vector. The relationships in (20) and (21) may naturally yield the question of whether $\upsilon_{\mathcal{L}}$ can capture $rank(\Gamma)$ better than $\delta_{\mathcal{L}}$ and $\mu_{\mathcal{L}}$. In the following result, we show that there is no such universal relationship between $rank(\Gamma)$ and $\upsilon_{\mathcal{L}}$.

Proposition 3.3. *For leader-follower networks, the number of unique DL vectors, $\upsilon_{\mathcal{L}}$, is not a universal bound of $rank(\Gamma)$.*

Proof. We prove this statement by providing some examples both for $rank(\Gamma) > \upsilon_{\mathcal{L}}$ and for $rank(\Gamma) < \upsilon_{\mathcal{L}}$.

1) A path graph with uniform edge weights is controllable from any single node if and only if the the number of nodes is a power of two, i.e. $n = 2^k$ for some $k \in \mathbb{N}$ [19]. Note that, for a path graph with a single leader, $\delta_{\mathcal{L}} = n$ if and only if the leader is a leaf node. Hence, for any $n = 2^k$, if a non-leaf node is the only leader of a path graph with uniform edge weights, then $rank(\Gamma) > \upsilon_{\mathcal{L}}$.

2) A cycle graph with uniform edge weights is controllable from any pair of nodes if and only if the the number of nodes is a prime number [19]. Consider any cycle graph of n nodes such that n is an odd composite number. For such a graph with two leaders, the clockwise and counterclockwise paths between the leaders have different lengths. Hence, any pair of nodes that have equal distances from one of the leaders have different distances from the other leader, i.e. $\upsilon_{\mathcal{L}} = n$. Furthermore, in light of [19], there exists a pair of nodes

that render the system uncontrollable if they are assigned as the leaders. If such a pair is assigned as the leaders, then $rank(\Gamma) < \upsilon_{\mathcal{L}}$. $\qquad\square$

We would like to conclude this section with a remark regarding the application of the presented results to directed networks.

Remark 3.1 (*Directed Networks*) The formulation in (1)-(8) is applicable to directed networks with the interaction graph $\mathcal{G} = (V, E)$, where each $(i, j) \in E$ denotes that i is influenced by j as in (5). For such a network, powers of the adjacency matrix $\mathcal{A}$ have the property that $[\mathcal{A}^r]_{ij} = 0$ for all $0 \leq r < dist(i, j)$ and $[\mathcal{A}^{dist(i,j)}]_{ij} \neq 0$, where $dist(i, j)$ is the length of the shortest directed path from i to j. Hence, by using the corresponding DL vectors as in (9), the results in both Lemma 3.1 and Theorem 3.2 can be extended to strongly connected networks (i.e., there exists a directed path from every node to every other node).

IV. AN ALGORITHM FOR COMPUTING THE LOWER BOUND

In this section, we present an algorithm to compute the proposed lower bound, $\delta_{\mathcal{L}}$. Note that the main contribution of this work is the lower bound itself, and the algorithm in this section is provided to facilitate some practical use of our result.

Let $S = \{d_1, d_2, \ldots, d_n\}$ be the set of all DL vectors for a given leader-follower network. Given these vectors, we present an iterative way of generating the longest PMI sequences. Let C_p be the set of all DL vectors that can be assigned as the pth element of such a sequence D. According to these definitions, $C_1 = S$. Once a vector from C_p is assigned as the pth element of the sequence, D_p, and an index $\alpha(p)$ satisfying (10) is chosen, the resulting C_{p+1} can be obtained from C_p as

$$C_{p+1} = \{d_i \in C_p \mid d_{i,\alpha(p)} > D_{p,\alpha(p)}\}. \tag{22}$$

In order to obtain longer sequences, this iteration must be continued until $C_p = \emptyset$. However, in general there are too many possible sequences that can be obtained this way, and it is not feasible to find the maximum length for PMI sequences by searching among all these possibilities. Instead, we present a necessary condition for a PMI sequence to have the maximum possible length. This necessary condition significantly lowers the number of sequences to consider.

Lemma 4.1. *Let D be a PMI sequence of DL vectors with the maximum possible length, then its pth entry, D_p, satisfies*

$$D_{p,\alpha(p)} = \min_{d_i \in C_p} d_{i,\alpha(p)}. \tag{23}$$

Proof. For the sake of contradiction, assume that this is not true for a PMI sequence D with the maximum length. Then, there exists a DL vector $d_j \in C_p$ such that $d_{j,\alpha(p)} < D_{p,\alpha(p)}$. By the construction of a PMI sequence, d_j can not be added to this sequence after D_p. However, d_j can be added right before D_p while keeping all the other parts of D the same since $d_{j,\alpha(p)}$ can be selected to satisfy (10) in the resulting sequence. Hence, it is possible to obtain a longer PMI sequence, which leads to the contradiction that D does not have the maximum possible length. $\qquad\square$

In light of (22), as far as the sequence length is concerned, the only important decision at each step p in building a PMI sequence satisfying (23) is the choice of $\alpha(p)$. Based on this observation, we propose Algorithm I for computing the lower bound.

Algorithm I

1: **initialize:** $\mathcal{C}_1 = \{\{d_1, d_2, \ldots, d_n\}\}$; $p = 1$
2: **while** $\mathcal{C}_{p,y} \neq \emptyset$ for some $y \in \{1, \cdots, |\mathcal{C}_p|\}$ **do**
3:      $q = 1$
4:      **for** $i = 1 : |\mathcal{C}_p|$ **do**
5:          **if** $\mathcal{C}_{p,i} \neq \emptyset$ **then**
6:              **for** $j = 1 : m$ **do**
7:                  $\mathcal{C}_{p+1,q} = \{d_t \in \mathcal{C}_{p,i} \mid d_{t,j} > \min_{d_s \in \mathcal{C}_{p,i}} d_{s,j}\}$
8:                  $q = q + 1$
9:              **end for**
10:          **end if**
11:      **end for**
12:      $p = p + 1$
13: **end while**
14: **return** $p - 1$

In Algorithm I, the variable $\mathcal{C}_p$ is the multiset, where each element $\mathcal{C}_{p,i}$ is the C_p resulting from (22) for specific choices of $\alpha(1), \ldots, \alpha(p-1)$, subject to the corresponding PMI sequences satisfying (23). The main while loop iterates as long as there exists a longer PMI sequence that satisfies (23). Note that for each $\mathcal{C}_{p,i} \neq \emptyset$, there are m (number of leaders) different $\mathcal{C}_{p+1,q}$, each corresponding to a particular choice of $\alpha(p)$. As such, Algorithm I computes $\delta_\mathcal{L}$ by generating no more than $m^{\delta_\mathcal{L}}$ elements $\mathcal{C}_{p,i}$.

Proposition 4.2. *Given the DL vectors for any connected leader-follower network, Algorithm I returns $\delta_\mathcal{L}$.*

Proof. By combining (22) and (23) (line 7 of Algorithm I), at each step p, Algorithm I builds all the possible C_{p+1} that correspond to the PMI sequences of p vectors satisfying (23). Hence, when $\mathcal{C}_{p,y} = \emptyset$ for every $y \in \{1, \cdots, |\mathcal{C}_p|\}$, there is not a longer PMI sequence that satisfies the necessary condition in Lemma 4.1. Consequently, Algorithm I always returns the length of the longest PMI sequence, $\delta_\mathcal{L}$. $\square$

Note that, in light of Remark 3.1, Algorithm I can also be used to compute the proposed lower bound for any directed network with a strongly connected interaction graph.

For a sample run of Algorithm I, consider the network in Fig. 1. For this example, the algorithm terminates after the fifth iteration of the while loop, and $\delta_\mathcal{L} = 5$. The resulting flow of Algorithm I can be represented via a tree diagram as illustrated in Fig. 3.
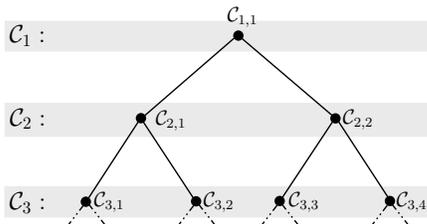

Fig. 3. An illustration of the flow of Algorithm I for the network in Fig. 1.

In Fig. 3., each node $\mathcal{C}_{p,i}$ at a given level $p > 1$ corresponds to an element of $\mathcal{C}_p$ that is computed in the $(p-1)^{st}$ iteration of the while loop. The left child of a node $\mathcal{C}_{p,i}$ is obtained from $\mathcal{C}_{p,i}$ by deleting all the DL vectors whose first entries are equal to the minimum value of the first entries among all the DL vectors in $\mathcal{C}_{p,i}$, i.e., obtained from $\mathcal{C}_{p,i}$ as per (23) and (22) for $\alpha(p) = 1$. Similarly, the right child of a node $\mathcal{C}_{p,i}$ is obtained by following the same procedure for $\alpha(p) = 2$. Accordingly, the elements in the first three levels are

$$\mathcal{C}_{1,1} = \left\{ \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix} \right\};$$

$$\mathcal{C}_{2,1} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix} \right\};$$

$$\mathcal{C}_{2,2} = \left\{ \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\};$$

$$\mathcal{C}_{3,1} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix} \right\};$$

$$\mathcal{C}_{3,2} = \mathcal{C}_{3,3} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\};$$

$$\mathcal{C}_{3,4} = \left\{ \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}.$$

V. DISTANCE-BASED LEADER SELECTION FOR CONTROLLABILITY

One of the main attributes of the proposed lower bound $\delta_\mathcal{L}$ is its independence of the edge weights. In contrast, $rank(\Gamma)$ depends on the values of the edge weights unless there is some constraint such as all the weights being equal. Consequently, computing $\delta_\mathcal{L}$ typically requires significantly less information about the overall system. This minimality of the required information makes the lower bound attractive in many applications such as leader selection for controllability.

Leader selection problems typically require finding a leader set $\mathcal{L}$ that optimizes a system objective such as robustness, mixing time, or controllability (e.g., [20], [21], [22], [23]). For instance, consider the problem of finding a minimum number of leaders that render a given network controllable under the resulting leader-follower dynamics. If the edge weights are known (or if they are known to be identical), then the rank of the controllability matrix can be computed for any set of leaders. Hence, a possible, yet not scalable, way to find a minimal set of leaders for controllability is to execute an exhaustive search. Note that, aside from the complexity issues, the rank computation is not applicable if the edge weights are unknown and arbitrary. In such cases, the leader selection problem needs to be solved by leveraging the structural properties of the interaction graph.

One approach to achieving controllability under arbitrary coupling weights is to choose the minimal $\mathcal{L}$ that achieves structural controllability (e.g., [8]). Structural controllability implies that the selected leaders provide complete controllability for some, not all, weighting functions $w : E \mapsto \mathbb{R}^+$. Hence, this approach may fall short in some applications, especially when there are constraints on the admissible edge weights. For instance, if all the edge weights in a network are equal by design, then it is known that a complete graph is not controllable by any single leader [6] whereas a single leader is

enough to achieve structural controllability [8]. Alternatively, the notion of strong structural controllability can be employed in the leader selection (e.g., [11]). In this regard, the proposed bound $\delta_{\mathcal{L}}$ can be used to ensure that the dimension of the controllable subspace is not smaller than some desired value, $k \in \{1, 2, \ldots, n\}$, for any $w : E \mapsto \mathbb{R}^+$ by formulating the leader selection problem as

$$\underset{\mathcal{L} \in 2^V}{\text{minimize}} \quad |\mathcal{L}|$$
$$\text{subject to} \quad \delta_{\mathcal{L}} \geq k, \tag{24}$$

In light of Theorem 3.2, any element in the feasible set of the problem in (24) renders $rank(\Gamma) \geq k$ for any weighting function $w : E \mapsto \mathbb{R}^+$. Note that the problem in (24) is always feasible, i.e., for any given $\mathcal{G} = (V, E)$ and $k \in \{1, 2, \ldots, n\}$, there always exists some $\mathcal{L} \subseteq V$ such that $\delta_{\mathcal{L}} \geq k$. Indeed, the feasibility can be shown by considering the trivial case, $\mathcal{L} = V$. In that case, for each DL vector, the entry that contains the distance of the corresponding node from itself satisfies (10) for any sequence of the corresponding DL vectors. Consequently, any sequence of the DL vectors is a PMI sequence if $\mathcal{L} = V$, and $\delta_V = n$. Note that the feasibility would not be guaranteed if the problem was posed by using $\mu_{\mathcal{L}}$ instead of $\delta_{\mathcal{L}}$ since $\mu_{\mathcal{L}}$ is always upper bounded by one plus the graph diameter. An example, where the leaders were assigned by solving (24) for $k = n$, is illustrated in Fig. 4. For this example, the leaders were selected through an exhaustive search by first looking for a single-leader solution and incrementing the number of leaders until a solution exists. Algorithm I was used to compute $\delta_{\mathcal{L}}$ for each candidate $\mathcal{L}$.
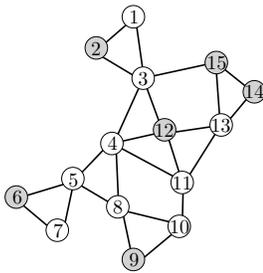

Fig. 4. A graph of 15 nodes, $\mathcal{G} = (V, E)$, and a minimal selection of leaders (shown in gray), $\mathcal{L} = \{2, 6, 9, 12, 14, 15\}$, such that $\delta_{\mathcal{L}} = 15$. The network is completely controllable via $\mathcal{L}$ for any weighting function $w : E \mapsto \mathbb{R}^+$.

VI. CONCLUSION

In this technical note, we presented that the distances between the leaders and the followers on the interaction graph contain some fundamental information about the controllability of the leader-follower networks. In particular, we used the distance-to-leaders (DL) vectors to derive a tight lower bound on the dimension of the controllable subspace. The proposed bound is applicable to networks with arbitrary interaction graphs and weighting functions $w : E \mapsto \mathbb{R}^+$. We also provided some connections between the proposed lower bound and a pair of closely related distance-based measures, namely the maximum distance from the leaders and the number of distinct DL vectors. While the results were presented for undirected networks, we also showed how they can be

extended to directed networks. Furthermore, we presented an algorithm for computing the lower bound. The proposed bound may find its applications in various networked control problems, especially when the edge weights are unknown. As a prominent application, we presented how it can be utilized to find a minimal set of leaders that ensure the controllability of a leader-follower network with a given interaction graph under any weighting function $w : E \mapsto \mathbb{R}^+$.

REFERENCES

[1] A. Speranzon, C. Fischione, and K. H. Johansson, "Distributed and collaborative estimation over wireless sensor networks", *IEEE Conference Decision and Control*, pp. 1025–1030, 2006.

[2] A. Jadbabaie, J. Lin and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules", *IEEE Transactions on Automatic Control*, 48(6): 988–1001, 2003.

[3] F. Dorfler and F. Bullo, "Synchronization and transient stability in power networks and nonuniform Kuramoto oscillators", *SIAM Journal on Control and Optimization*, 50(3): 1616–1642, 2012.

[4] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks: A local interaction game with stubborn agents," *American Control Conference*, pp. 1982–1987, 2013.

[5] S. Gu *et al.*, "Controllability of structural brain networks," *Nature Communications*, 6, 2015.

[6] A. Rahmani, M. Ji and M. Egerstedt, "Controllability of multi-agent systems from a graph theoretic perspective", *SIAM J. Control Optim.*, 48(1): 162–186, 2009.

[7] M. Egerstedt, S. Martini, M. Cao, K. Camlibel, and A. Bicchi, "Interacting with networks: How does structure relate to controllability in single-leader consensus networks?" *IEEE Control Systems Magazine*, 32(4): 66–73, 2012.

[8] Y.-Y. Liu, J.-J Slotine, and A.-L.Barabási, "Controllability of complex networks," *Nature*, 473(7346): 167–173, 2011.

[9] C.-T Lin, "Structural controllability," *IEEE Transactions on Automatic Control*, 19(3): 201–208, 1974.

[10] H. Mayeda and T. Yamada, "Strong structural controllability," *SIAM Journal on Control and Optimization*, 17(1): 123–138, 1979.

[11] A. Chapman and M. Mesbahi, "On strong structural controllability of networked systems: a constrained matching approach," *American Control Conference*, pp. 6126–6131, 2013.

[12] J. Jarczyk, F. Svaricek, and B. Alt, "Strong structural controllability of linear systems revisited," *IEEE Conference on Decision and Control*, pp. 1213–1218, 2013.

[13] Y.-Y.Liu, J.-J Slotine, and A.-L.Barabási, "Control centrality and hierarchical structure in complex networks," *Plos One*, 7(9), 2012.

[14] Y. Pan and X. Li, "Structural controllability and controlling centrality of temporal networks," *PloS One*, 9(4), 2014.

[15] J. C. Nacher and T. Akutsu, "Analysis of critical and redundant nodes in controlling directed and undirected complex networks using dominating sets," *Journal of Complex Networks*, 2(4): 394–412, 2014.

[16] S. Zhang, M. K. Camlibel, and M. Cao, "Controllability of diffusively-coupled multi-agent systems with general and distance regular coupling topologies", *IEEE Conference on Decision and Control*, 2011.

[17] S. Zhang, M. Cao, and M. K. Camlibel, "Upper and lower bounds for controllable subspaces of networks of diffusively coupled agents," *IEEE Transactions on Automatic Control*, 59(3): 745–750, 2014.

[18] A. Y. Yazıcıoğlu, W. Abbas, and M. Egerstedt, "A tight lower bound on the controllability of networks with multiple leaders," *IEEE Conference on Decision and Control*, pp. 1978–1983, 2012.

[19] G. Parlangeli and G. Notarstefano, "On the reachability and observability of path and cycle graphs," *IEEE Transactions on Automatic Control*, 57(3): 743–748, 2012.

[20] F. Lin, M. Fardad, and M. R. Jovanovic, "Algorithms for leader selection in stochastically forced consensus networks," *IEEE Transactions on Automatic Control*, 59(7): 1789–1802, 2014.

[21] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran, "Leader selection in multi-agent systems for smooth convergence via fast mixing," *IEEE Conference on Decision and Control*, pp. 818–824, 2012.

[22] A. Y. Yazıcıoğlu and M. Egerstedt, "Leader selection and network assembly for controllability of leader-follower networks," *American Control Conference*, pp. 3802–3807, 2013.

[23] C. O. Aguilar and B. Gharesifard, "Graph controllability classes for the Laplacian leader-follower dynamics," *IEEE Transactions on Automatic Control*, 60(6): 1611–1623, 2015.