Leader Selection for Optimal Network Coherence

Stacy Patterson and Bassam Bamieh

Abstract— We consider the problem of leader-based distributed coordination in networks where agents are subject to stochastic disturbances, but where certain designated leaders are immune to those disturbances. Specifically, we address the effect of leader selection on the coherence of the network, defined in terms of an H_2 norm of the system. This quantity captures the level of agreement of the nodes in the face of the external disturbances. We show that network coherence depends on the eigenvalues of a principal submatrix of the Laplacian matrix, and we formulate an optimization problem to select the set of leaders that results in the highest coherence. As this optimization problem is combinatorial in nature, we also present several greedy algorithms for leader selection that rely on more easily computable bounds of the H_2 norm and the eigenvalues of the system. Finally, we illustrate the effectiveness of these algorithms using several network examples.

I. INTRODUCTION

Distributed consensus is a fundamental problem in the context of multi-agent systems and distributed formation control [1], [2], [3]. In this setting, agents must reach agreement on values like direction, rate of travel, and inter-agent spacing in a decentralized manner, and they must maintain this agreement in the face of uncertainties such as external disturbances, unreliable communication, and agent failures. Therefore, it is important to consider the robustness of consensus algorithms as they are applied to this setting.

In this work, we study the robustness of a class of consensus algorithms that are particularly relevant to formation control, leader-follower consensus algorithms. In leader-follower consensus, one or more agents are *leaders*, and these agents determine the consensus value (or values) of the network. The remaining agents are *followers* that use local consensus dynamics and are unaware of the consensus value. For example, leaders can be controlled by sophisticated machinery, equipped with highly accurate sensors, and have access to global information and absolute measurements. In contrast, followers must operate solely based on relative information such as measurements relative to neighboring agents. It has been shown that if the consensus value (also called reference value) is available to a node that is the root of a spanning tree of the network and there are no external disturbances, then the system will converge to consensus at that consensus value [4]. Our interest is in analyzing the effects of external disturbances on the convergence behavior.

Leader-follower consensus can also be applied to problems outside the realm of formation control such as opinion dynamics in social networks. In this setting, the objective is to maintain agreement in the network, for example to maintain support for a cause or political candidate. User opinions (node states) are influenced by the opinions of neighbors as well as by externalities. Several nodes can be selected as leaders who can be trained offline. The leaders act as experts who all share the same opinion and do not waiver. In this setting, network coherence is a measure of how well these leaders "keep the peace" in the network. A question of interest is how to select the set of leaders that best maintains agreement among the social network users.

In order to focus on the robustness of the network with respect to a set of leaders, we consider a simple version of leader-follower consensus in an undirected network with a time-invariant communication structure. We assume that the leaders are not affected by external disturbances and that they share the same, fixed consensus value. The followers are subject to external disturbances, and the objective is for every node's value to remain as close to the consensus value as possible. We call the level of agreement in the network the *coherence* of the network.

In this work, we present a formal definition of the coherence of a leader-follower consensus system in terms of an H_2 norm, and we show that network coherence depends on the eigenvalues of a principal submatrix of the Laplacian matrix. Based on this formulation, we give an optimization problem for finding the set of leaders that maximizes network coherence. As this optimization problem is combinatorial in nature, we also propose two more scalable, heuristic-based algorithms for leader selection. Finally, we compare these algorithms using several network examples. To our knowledge, this work is the first to consider the H_2 norm as a measure of robustness for leader-follower consensus and the first to explore the question of selecting the best leaders for maintaining coherence.

A. Related Work

Several works have investigated necessary and sufficient conditions for convergence in leader-follower consensus [4], [5], [6]. These works consider directed networks with no external disturbances, and the leader reference value may be time-varying. Pasqualetti et al. [7] also study directed networks, and they find the convergence rate of the consensus algorithm for the case where leaders move with constant velocity.

The robustness of leader-follower consensus has been previously studied using different performance measures. In [8], the authors define and analyze robustness in terms of the effects that disturbances entering at the leaders have

on the other nodes in the network. Wang et al. [9] study the robustness of leader-follower consensus algorithms to disturbances on the communication links and quantify this robustness in terms of the L_2 gain.

We also note that the H_2 norm has been used as a robustness measure in networks with leaderless consensus dynamics. In the case of undirected networks, Xiao et al. [10] propose an optimization algorithm to select the edge weights that minimize the H_2 norm. The work by Bamieh et al. [11] considers the platooning problem and presents asymptotic scalings of the H_2 norm for formations of different dimensions. Finally, the work by Young et al. [12] studies the H_2 norm in directed networks with noisy consensus dynamics and gives analytical results for several types of directed graphs.

B. Outline

This paper is organized as follows. In Section II, we describe our leader-follower consensus model and give the definition of network coherence. Section III gives analysis of the stability of the consensus algorithm and the performance with respect to the number of leaders. In Section IV, we present an optimization algorithm for selecting the leaders that maintain the best network coherence, and we propose several more scalable heuristic-based algorithms for leader selection. In Section V, we illustrate the performance of the optimization algorithm and the scalable alternatives for various network examples. We conclude in Section VI with a discussion of topics for future investigation.

II. PRELIMINARIES

We consider a connected, undirected network of n identical agents, with a time-invariant communication structure. The network is modeled by an undirected graph $G = (V, E)$ where V is the set of nodes (with $|V| = n$) and E is the set of edges (with $|E| = m$). The Laplacian matrix of G is

$$L = D - A$$

where D is the diagonal matrix of node degrees and A is the adjacency matrix of G.

In the leader-follower consensus problem, the objective is for all nodes to follow an identical, constant trajectory $\overline{x} \in \mathbb{R}$. A subset of k nodes are the *leaders* of the system. We assume that these leaders follow the trajectory exactly. The remaining $n - k$ nodes are *followers*. Each follower updates its state based only on its own state and that of its neighbors using a local consensus algorithm and is also subject to stochastic disturbances. Let h be an n-vector with $h_i = 1$ if i is a leader and equal to 0 otherwise. The dynamics of the entire system are given by

$$\dot{x}(t) = -(I - diag(h)) \, L \, x(t) + (I - diag(h))w(t) \quad (1)$$

where $w(t)$ is a vector of zero-mean white noise processes. Let $\tilde{x}(t)$ be the vector of deviations from the desired trajectory,

$$\tilde{x}(t) := x(t) - \overline{x}\mathbf{1}.$$

Here $\mathbf{1}$ denotes the vector of all ones. We note that the system dynamics for $\tilde{x}$ are identical to the dynamics for x,

$$
\begin{aligned}
\frac{d}{dt}(\tilde{x} + \overline{x}\mathbf{1}) &= -(I - diag(h)) \, L(\tilde{x}(t) + \overline{x}\mathbf{1}) \\
&\quad + (I - diag(h))w(t) \\
\frac{d}{dt}\tilde{x} &= -(I - diag(h))L\tilde{x}(t) \\
&\quad + (I - diag(h))w(t).
\end{aligned}
$$

Therefore, without loss of generality, we assume that $\overline{x} = 0$.

The expression (1) is equivalent, up to a permutation in state, to the following,

$$
\begin{bmatrix} \dot{x}_l \\ \dot{x}_f \end{bmatrix} = -\begin{bmatrix} 0 & 0 \\ * & L[h] \end{bmatrix} \begin{bmatrix} x_l \\ x_f \end{bmatrix} + \begin{bmatrix} 0 \\ w_f \end{bmatrix}
$$

where $L[h]$ denotes the $(n-k) \times (n-k)$ principal submatrix of L obtained by removing the columns and rows of L that correspond to the leaders, *i.e.* the components of h that are equal to 1. x_l and x_f are the states of the leader and follower nodes respectively, and w_f is an $n - k$ vector of zero mean white noise processes. Clearly, for each leader node i, if $x_i(0) = 0$, then $x_i(t) = 0$ for $t > 0$. The dynamics of the follower nodes are

$$\dot{x}_f = -L[h]x_f + w_f. \quad (2)$$

In this work, we study the relationship between the choice of leader nodes and the *coherence* of the network. By coherence, we mean how closely each node follows the specified trajectory $\overline{x}$. Formally, we define the *steady-state variance of the deviation from consensus* of the system by

$$V := \lim_{t \to \infty} \sum_{i \in V} \mathbf{E}\left\{ (x_i(t) - \overline{x})^2 \right\}. \quad (3)$$

The smaller this variance, the higher the network coherence.

The objective of this work is two-fold.

1) Characterize the coherence of the network as a function of the set nodes that are acting as leaders.
2) Determine a method to identify the optimal leaders, *i.e.* the set of nodes that minimize the steady-state variance (3).

We address the first goal in Section III, where we present analysis of the variance of the deviation from consensus of a leader-follower system in terms of the eigenvalues of the $L[h]$ matrix. The second goal is addressed in Section IV, which gives a formulation for selecting the optimal leader set as well as several more efficient heuristics for leader selection.

III. NETWORK COHERENCE ANALYSIS

The variance of the deviation from consensus is the square of the H_2 norm of the system (2) and is given by

$$V = \mathbf{tr}\left(\int_0^\infty e^{-L[h]^* t} e^{-L[h]t} dt \right). \quad (4)$$

We first show that for any leader set of size k, where $k > 0$, this variance is bounded. We do this by proving that the operator $L[h]$ is positive definite.

Lemma 3.1: If L is the Laplacian of an undirected, connected graph and $L[h]$ is the principal submatrix of L specified by $h \in \{0,1\}^n$, $L[h]$ is positive definite.

Proof: This is a proof by induction.

Base Case: It is well known that for an undirected connected graph,

$$0 = \lambda_1(L) < \lambda_2(L) \leq \lambda_3(L) \ldots \leq \lambda_n(L).$$

Let L_1 be a principal submatrix of L formed by the removal of any one row and column. By the Cauchy Interlacing Theorem (see A.1), it holds that $\lambda_i(L_1) \geq \lambda_i(L)$ for $i = 1 \ldots n-1$. Also, by the Kirchoff Matrix Tree Theorem [13],

$$\det(L_1) = \kappa,$$

where κ is the number of spanning trees of the graph G. Since G is connected, κ must be greater than 0. Therefore, L_1 is non-singular. Combing these two facts, it must be the case that $\lambda_i(L_1) > 0$ for $i = 1 \ldots n-1$.

Inductive step: Let L_j be a principal submatrix of L formed by the removal of j rows and j columns. If $\lambda_i(L_j) > 0$ for $i = 1 \ldots n - j$, then again, by the Cauchy Interlacing Theorem, for any L_{j+1} formed by the removal of a row and column from L_j, $\lambda_i(L_{j+1}) \geq \lambda_i(L_j) > 0$ for $j = 1 \ldots n - (j+1)$. ∎

This lemma leads to the following theorem on the value of the variance of the deviation from consensus.

Theorem 3.2: The steady-state variance of the deviation from consensus for a network with leaders specified by $h \in \{0,1\}^n$ is

$$V = \frac{1}{2}\mathbf{tr}\left(L[h]^{-1}\right) = \frac{1}{2}\sum_{i=1}^{n-k}\frac{1}{\lambda_i(L[h])}. \tag{5}$$

where $L[h]$ is the principal submatrix of the Laplacian matrix corresponding to h.

Proof: This result follows from Lemma 3.1 and the fact that each $L[h]$ is normal. We refer the reader to [11] for details. ∎

The Effect of Adding Leaders

The result in Theorem 3.2 is similar to the results presented for the robustness of distributed average consensus algorithms in systems with random additive noise [12], [11]. In such systems, there is no leader, and every node is subject to random external disturbances. The dynamics are given by

$$\dot{x}(t) = -Lx(t) + w(t),$$

where, as in our formulation, L is the Laplacian matrix of an undirected, connected graph. The mean of the node values drifts in a random walk around the origin, and unlike in the leader-follower consensus case, the variance of the deviation from $\overline{x} = 0$ is unbounded. However, the variance of the deviation from the average of the current values, $\frac{1}{n}\sum_{i=1}^n x_i(t)$, is bounded, this variance is given by

$$V_{dav} = \frac{1}{2}\sum_{i=2}^n \frac{1}{\lambda_i(L)}. \tag{6}$$

Both V and V_{dav} are measures of deviation from consensus. In a system with a leader, the consensus value is fixed and equal to the state of the leaders, while, in leaderless systems, the value changes over time. An interesting question is how the variance of the deviation from consensus in a leaderless system compares to the variance in a system with a leader. Using equations (5) and (6), it is possible to analyze the relationship between these two variances.

Let G be an undirected, connected graph. Let L be the Laplacian matrix of G and let let L_1 be a principal submatrix of L formed by the removal of one row and column (L_1 is the $L[h]$ matrix in Eq. (2)). By the Cauchy Interlacing Theorem, we obtain the following inequality,

$$\frac{1}{\lambda_{i+1}(L)} \leq \frac{1}{\lambda_i(L_1)}.$$

Therefore, the variance of the deviation from consensus in the system with one leader relates to the variance in the leaderless system as follows

$$V_{dav} = \frac{1}{2}\sum_{i=1}^{n-1}\frac{1}{\lambda_{i+1}(L)} \leq \frac{1}{2}\sum_{i=1}^{n-1}\frac{1}{\lambda_i(L_1)} = V.$$

This inequality shows that the coherence, in terms of the deviation from consensus, of a network with one leader is *worse* than the coherence of that same network when no leader is present.

Another interesting question is what is the effect of adding additional leaders to a leader-follower consensus system. Let L_j denote a principal submatrix of L formed by the removal of j rows and columns; L_j corresponds to a system with j leaders. Let V_j denote the variance of the deviation from consensus for this system. The Cauchy Interlacing Theorem establishes the following relationship between the eigenvalues of L_j and L_{j+1}, where L_{j+1} is a principal submatrix of L_j obtained by the removal of a single row and column.

$$\frac{1}{\lambda_i(L_j)} \geq \frac{1}{\lambda_i(L_{j+1})}.$$

Therefore,

$$V_j = \frac{1}{2}\sum_{i=1}^{n-j}\frac{1}{\lambda_i(L_j)} \geq \frac{1}{2}\sum_{i=1}^{n-(j+1)}\frac{1}{\lambda_i(L_{j+1})} = V_{j+1}.$$

The above inequality shows that once a network has a leader, adding additional leaders can only improve coherence (or maintain the current level of coherence).

We note that the above result does not imply that the coherence of a given network with $j+1$ leaders will always be better than (or at least the same as) the coherence of the same network with j leaders. In fact, the variance of the deviation from consensus depends not only on the number of leaders but also on the particular choice of leader nodes. In the next section, we explore the problem of selecting the leaders for a given network that will yield the smallest variance.

IV. OPTIMIZING LEADER SELECTION

In the leader selection problem, we are given a network $G = (V, E)$ and a budget k, and the goal is to select the k leaders that will maximize the coherence of the network. The leader selection problem can be formulated as the following optimization problem,

$$\text{minimize} \quad \frac{1}{2}\mathbf{tr}\left(L[h]^{-1}\right)$$
$$\text{subject to} \quad h_j \in \{0,1\} \quad j = 1 \ldots n$$
$$\|h\|_1 = k.$$

where h is the $\{0,1\}$ vector that indicates the leader nodes (as defined in the previous section).

The optimal solution can be obtained through an exhaustive search of the $\binom{n}{k}$ candidate leader sets to find the set that minimizes V. This approach is combinatorial in n and k, and so for large networks, it may be infeasible to find the optimal solution even if k is small. Therefore, we explore several alternatives that are more computationally attractive.

Algorithm 1 Greedy Leader Selection Algorithm.

1: $h \leftarrow [0 \ 0 \ \cdots \ 0]^*$
2: $V_{min} \leftarrow \infty$
3: **for** $i = 1$ to k **do**
4:    **for** $j \in$ V $-$ {nodes already selected as leaders} **do**
5:       $g \leftarrow h$
6:       $g_j \leftarrow 1$
7:       $V \leftarrow \frac{1}{2}\mathbf{tr}\left(L[g]^{-1}\right)$
8:       **if** $V \leq V_{min}$ **then**
9:          $V_{min} \leftarrow V$
10:         $min\_id \leftarrow j$
11:       **end if**
12:    **end for**
13:    $h_{min} \leftarrow 1$
14: **end for**

The first alternative is to use a greedy approach, choosing the leaders one at a time, each time selecting the node that gives the smallest variance. The greedy leader selection algorithm is shown in Algorithm 1. The algorithm produces the vector h where the selected leaders correspond to the components of h that are equal to 1. The inner loop of this algorithm, beginning on line 4, checks what the variance would be if each node were added to the current leader set. The node that results in the smallest variance is then added to the leader set, and the process repeats until all k leaders have been identified. This algorithm is more efficient than solving the original optimization problem. However, in order to select each leader, it is necessary to compute the trace of the inverse of an $O(n) \times O(n)$ matrix (line 7) for each iteration of the inner loop.

The greedy algorithm may still be too computationally intensive for very large networks. This motivates the development of algorithms that use bounds on V that are easier to compute. For the selection of leaders 2 through k, we can use an approximation of V based on a result on eigenvalues of inverses of principal submatrices (see A.2). Let $\bar{L}$ be an $(n - j) \times (n - j)$ principal submatrix of L, with $j > 0$,

---

Algorithm 2 Greedy Leader Selection Algorithm with Approximated Variances.

1: $h \leftarrow [0 \ 0 \ \cdots \ 0]^*$
2: $min\_id \leftarrow$ id of node corresponding to *smallest*
3:          diagonal entry of $((L + J)^{-1} - J)$
4: $h_{min} \leftarrow 1$
5: **for** $i = 2$ to k **do** {Find leaders 2 to k.}
6:    $min\_id \leftarrow$ id of node corresponding to *largest*
7:          diagonal entry of $L[h]^{-1}$
8:    $h_{min} \leftarrow 1$
9: **end for**

---

and let $\bar{L}[g]$ be an $(n - (j + 1)) \times (n - (j + 1))$ principal submatrix of L formed by removing a column and row from $\bar{L}$. The following inequality provides a upper bound on the eigenvalues of $\bar{L}[g]^{-1}$

$$\lambda_i\left(\bar{L}^{-1}[g]\right) \geq \lambda_i\left(\bar{L}[g]^{-1}\right), \quad (7)$$

for $i = 1 \ldots n - (j + 1)$. It follows that

$$\mathbf{tr}\left(\bar{L}^{-1}[g]\right) \geq \mathbf{tr}\left(\bar{L}[g]^{-1}\right).$$

Therefore, given $\bar{L}$, where j leaders have already been chosen, to choose the $(j + 1)^{th}$ leader, we select the node that corresponds to the maximum diagonal entry of $\bar{L}^{-1}$. This choice give the g that minimizes $\lambda_i\left(\bar{L}^{-1}[g]\right)$ for $i = 1 \ldots n - (j + 1)$.

The inequality (7) holds only when $\bar{L}$ is positive definite. When selecting the first leader, $\bar{L} = L$, and L has an eigenvalue of 0. So, a different approximation is needed to select the first leader. A natural choice is to consider the pseudoinverse of L,

$$L^{\dagger} := (L + J)^{-1} - J.$$

Here, $J = \frac{1}{n}\mathbf{1}\mathbf{1}^*$. $L^{\dagger}$ is such that its trace is equal to the total variance of the deviation from average of the leaderless system. However, in experiments, we observed that choosing the node that corresponds to the maximal diagonal entry of $L^{\dagger}$ actually produces the worst leader rather than the best. This indicates a complicated relationship between the variance of the deviation from average and the variance of the deviation from $\overline{x}$. We note that in a system with one leader, the two variances are related as follows

$$
\begin{aligned}
V_{dav} &= V - \frac{n+1}{n^2}\mathbf{1}^*X\mathbf{1} \\
&= V - \frac{n+1}{n^2}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\text{cov}(x_i, x_j),
\end{aligned}
$$

where X is the covariance matrix of x_f. The details of this derivation have been omitted due to space constraints. This expression shows that the variance of the deviation from average is maximized when the sum of covariances of the node states is maximized. We conjecture that the best leader is actually the one that will minimize the sum of these covariances and thus maximize the total variance of the deviation from average. Therefore, in the approximation-based greedy leader selection algorithm, we choose the
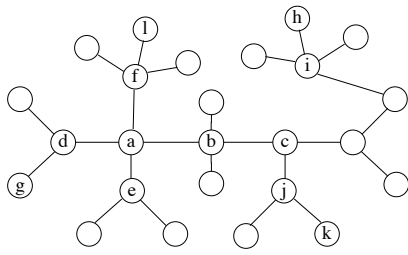
Fig. 1. Undirected graph with 25 nodes and 24 edges, used to generate the results in Fig. 2 and Table I.

TABLE I
LEADER SETS SELECTED BY DIFFERENT LEADER SELECTION
ALGORITHMS FOR GRAPH IN FIG. 1

k	Optimal	Greedy	Approx. Greedy	Max. Degree
1	b	b	b	a
2	a,i	b,i	b,h	a,b
3	a,i,j	b,f,i	b,g,h	a,b,e
4	e,f,i,j	b,f,i,j	b,g,h,k,l	a,b,e,i
5	d,e,f,i,j	b,f,e,i,j	b,g,h,k,l	a,b,c,e,i

first leader to be the node that corresponds to the smallest diagonal entry of $L^\dagger$. This heuristic has been shown to perform well in our experiments, as demonstrated in the next section.

The leader selection algorithm that relies on these two approximations of the variance of the deviation from consensus is given in Algorithm 2. The performance improvement comes from the fact that selecting each leader requires computing the inverse of only one $O(n) \times O(n)$ matrix as opposed to the n matrices required by Algorithm 1.

V. EXAMPLES

In this section, we present results of the various leader selection algorithms for several example networks. We first give the results for the graph shown in Fig. 1. The graph has 25 nodes and 24 edges. For $k = 1 \ldots 7$, we use each algorithm to find the "best" leader set of size k. Figure 2 shows the variance of the network that results from each leader set. Since this is a small graph, we are able to compute the optimal leader set and variance. We
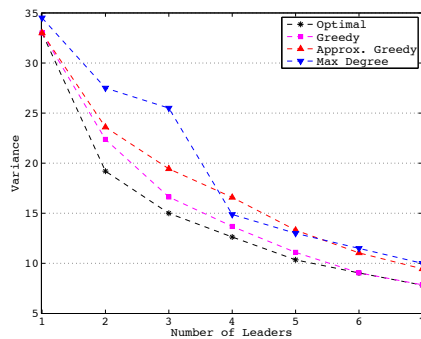

Fig. 2. Variance of deviation from consensus in the graph shown in Fig. 1 for leader sets selected by different leader selection algorithms.
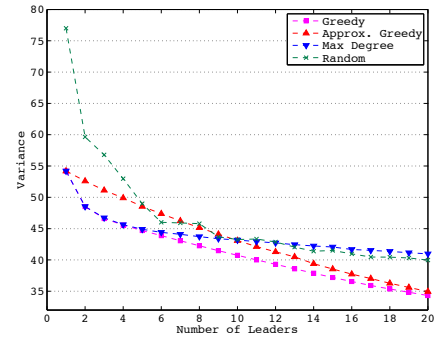

Fig. 3. Variance of deviation from consensus for leader sets selected by different leader selection algorithms in a 250 node Erdös-Rényi random graph where nodes are connected with probability 0.02.
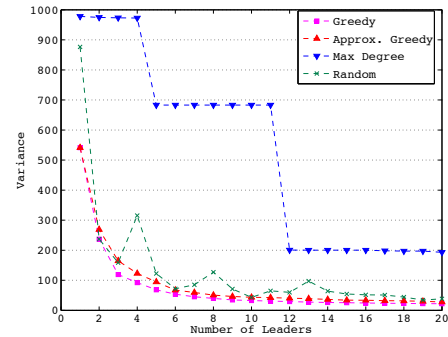

Fig. 4. Variance of deviation from consensus in a 250 node geometric graph for leader sets selected by different leader selection algorithms.

also show the variance of the leader sets generated by the greedy algorithm (Algorithm 1), the greedy algorithm with approximate variances (Algorithm 2), and a nïve leader selection scheme, Max Degree. In this scheme, the leaders are simply the k nodes with maximal vertex degree. We note that our three schemes perform well in this network, selecting the optimal leader for $k = 1$, while the maximum degree scheme performs erratically. The leader nodes chosen by each algorithm are given in Table I. An interesting point to note is that for $k = 1$, the optimal leader is node b, but for $k = 2$, b is no longer in the optimal set. This demonstrates a shortcoming of the greedy schemes, since when leaders are chosen incrementally, b will also be in the leader set. However, even with this shortcoming, the greedy schemes both outperform the nïve approach.

Figure 3 shows results for a 250 node Erdös-Rényi random graph, where an edge connects each pair of vertices with probability 0.02. The mean vertex degree of the graph is 5, and the graph has 1,259 edges. We compare the variance of the network for the leaders selected by the greedy algorithm and the greedy algorithm with approximate variances to the variance when leaders are generated using the max degree heuristic and when leaders are selected at random. Although a single random trial does not give an indication of the expected performance of a random selection, it provides some indication of the low coherence that may result from a random leader set. We note that, in this example, the greedy

algorithm, the approximate variance greedy algorithm, and the max degree algorithm all select the same first leader. The max degree heuristic performs well when the number of leaders is small, and it is outperformed by the approximate variance greedy algorithm for larger leader set sizes. This graph exhibits high degree skew, which may be a reason for the good performance of the max degree scheme. However, the results for larger leader sets indicate a complex relationship between the node degree of the leaders and the variance of the network.

Finally, Figure 4 shows results for a 250 node geometric graph, where nodes are placed randomly on a unit square, and an edge is drawn between nodes i and j if the Euclidean distance between i and j is less than 0.0163. The graph has 1,178 edges and a mean vertex degree of 9. As in the previous example, we compare the greedy algorithm and greedy algorithm with approximate variances to the max degree heuristic and random leader selection. In this example, the approximation-based greedy algorithm performs nearly as well as the greedy algorithm, and both perform better than the two other schemes. Unlike the previous examples, the max-degree heuristic leads to particularly bad leader sets in this network. This may be due to the fact that in a geometric graph, the degree skew is low, and so the node degree does not play a significant role in the "importance" of a node in reducing variance.

VI. CONCLUSION AND FUTURE WORK

We have presented a characterization of the robustness of leader-follower consensus algorithms in undirected networks in terms of the H_2 norm of the system. This value captures the *coherence* of the network in the face of external disturbances for a particular choice of leaders. We have shown that network coherence depends on the eigenvalues of a principal submatrix of the Laplacian and presented the coherence maximization problem as an optimization problem. We have also proposed two more scalable, heuristic-based leader selection algorithms, and we have compared the performance of these algorithms on several network examples.

The generalization of coherence analysis and leader selection to systems with dynamics based on a weighted Laplacian is straightforward and will be included in the full version of this paper. We anticipate that this work can also be easily extended to double integrator consensus dynamics in undirected networks. With respect to directed networks, the analysis of the H_2 norm can be applied to certain directed graphs where the Laplacian matrix is normal. However, the heuristics for leader selection depend on the symmetry of the Laplacian, and therefore alternate heuristics must be considered for this case.

A topic of primary interest is the scalability of the leader selection algorithms. The greedy heuristics presented in this work can be applied to networks with numbers of nodes in the thousands. The application of the leader selection problem to massive networks on the order of millions of nodes will require more scalable techniques, and we will address this issue in future work.

REFERENCES

[1] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, pp. 988–1001, June 2003.

[2] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, pp. 1520–1533, September 2004.

[3] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, pp. 169–182, February 2005.

[4] Z. Li, Z. Duan, and L. Huang, "Leader-follower consensus of multi-agent systems," in *Proc. American Control Conf.*, 2009, pp. 3256–3261.

[5] W. Ren, "Multi-vehicle consensus with a time-varying reference state," *Systems & Control Letters*, vol. 56, no. 7-8, pp. 474–483, 2007.

[6] M. Ji, A. Muhammad, and M. Egerstedt, "Leader-based multi-agent coordination: controllability and optimal control," in *Proc. American Control Conf.*, 2006, pp. 1358–1363.

[7] F. Pasqualetti, S. Martini, and A. Bicchi, "Steering a leader-follower team via linear consensus," in *Proc. 11th Int. Workshop on Hybrid Systems*, 2008, pp. 642–645.

[8] H. G. Tanner, G. J. Pappas, and V. Kumar, "Leader-to-formation stability," *IEEE Trans. Robotics and Automation*, vol. 20, no. 3, pp. 443–455, 2004.

[9] J. Wang, Y. Tan, and I. Mareels, "Robustness analysis of leader-follower consensus," *J. Syst. Sci. and Complexity*, vol. 22, no. 2, pp. 186–206, 2009.

[10] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *J. Parallel Distrib. Comput.*, vol. 67, no. 1, pp. 33–46, 2007.

[11] B. Bamieh, M. Jovanovic, P. Mitra, and S. Patterson, "Coherence in large-scale networks: Dimension dependent limitations of local feedback," *IEEE Trans. Autom. Control (accepted)*, 2010.

[12] G. F. Young, L. Scardovi, and N. E. Leonard, "Robustness of noisy consensus dynamics with directed communication," in *Proc. American Control Conf.*, 2010, pp. 6312–6317.

[13] G. Kirchhoff, "Über die auflösung der gleichungen, auf welche man bei der untersuchung der linearen verteilung galvanischer ströme geführt wird," *Ann. Phys. Chem.*, no. 72, pp. 497–508, 1847.

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.

APPENDIX

We list some useful theorems that relate the eigenvalues of a Hermitian matrix to the eigenvalues of the compression or perturbation of that matrix. See [14] for a general reference.

Cauchy Interlacing Theorem

Theorem A.1: Let A be an Hermitian $n \times n$ matrix, and let ω be an n-vector with each $\omega_i \in \{0, 1\}$. $A[\omega]$ denotes the $p \times p$ principal submatrix of A that is formed by removing the rows and columns corresponding to the components of ω that are equal to 1. The eigenvalues of A and $A[\omega]$ are related as follows,

$$\lambda_j(A) \leq \lambda_j(A[\omega]) \leq \lambda_{n-p+j}(A).$$

For example, if B is an $n - 1 \times n - 1$ principal submatrix of A, then

$$\alpha_1 \leq \beta_1 \leq \alpha_2 \leq \beta_2 \leq \ldots \leq \alpha_{n-1} \leq \beta_{n-1} \leq \alpha_n.$$

Eigenvalues of Inverses of Principal Submatrices

Theorem A.2: Let A be an Hermitian $n \times n$ matrix and let $A[\omega]$ be a principal submatrix of A. The following inequality holds for $i = 1 \ldots m$.

$$\lambda_i \left(A^{-1}[\omega] \right) \geq \lambda_i \left(A[\omega]^{-1} \right).$$