Proceedings of the IEEE Conf. on Decision and Control, **Florence, Italy, 2013**

Information Centrality and Optimal Leader Selection in Noisy Networks

Katherine Fitch and Naomi Ehrich Leonard

Abstract— **We consider the leader selection problem in which a system of networked agents, subject to stochastic disturbances, uses a decentralized coordinated feedback law to track an unknown external signal, and only a limited number of agents, known as leaders, can measure the signal directly. The optimal leader selection minimizes the total system error by minimizing the steady-state variance about the external signal, equivalent to an H_2 norm of the linear stochastic network dynamics. Efficient greedy algorithms have been proposed in the literature for similar optimal leader selection problems. In contrast, we seek systematic solutions. We prove that the single optimal leader is the node in the network graph with maximal information centrality. In the case of two leaders, we prove that the optimal pair maximizes a joint centrality, which depends on the information centrality of each leader and how well the pair covers the graph. We apply these results to solve explicitly for the optimal single leader and the optimal pair of leaders in special classes of network graphs. To generalize we compute joint centrality for m leaders.**

I. Introduction

Analysis and design of distributed coordination dynamics in multi-agent systems has gained considerable attention in recent years [1]–[3]. The dynamics are rich and the possibilities for application are abundant, both in engineering, e.g., vehicle networks [4], and in nature, e.g., bird flocks [5] and social networks [6]. When such systems interact with the external environment, e.g., to collectively learn or track an environmental signal, performance will depend on the information that agents have about the environment, which can vary across the group, particularly when such information is costly to acquire. For example, for a robotic vehicle network tracking a chemical plume, it may be most efficient for only some subset of vehicles to do the sampling. Likewise, in a herd of migrating animals, it is likely that only some subset of the animals measure the migration route [7].

Indeed, it was shown in [7] that the evolutionary specialization of a migratory population into leaders who invest in noisy measurement of the migration route and followers who only use the available noisy social cues is an evolutionarily stable solution when the cost of investment is sufficiently high. In [8], the location of emergent leaders as a function of the network topology was examined for adaptive dynamics modeled after [7]. In this model each agent adjusts its investment strategy to minimize its steady-state variance about the reference value (migration route). This bottom-up approach has the advantage of being distributed; however, it does not necessarily give rise to leader selections that optimize group performance. Systematic methods to determine optimal leader sets could prove useful in the design of high-performing distributed adaptive leadership dynamics.

In the present paper we address the global optimal leader selection problem for leader-follower dynamics defined such that some agents (leaders) invest in the costly measurement of a reference (consensus) value and the rest (followers) use local consensus dynamics, relying soley on measurements of the relative state of others in the system. We are interested in the setting in which the dynamics are subject to stochastic perturbations [9], [10]. We seek the leader set, as a function of the fixed, undirected network graph, that minimizes total system error defined as the steady-state variance of the system state about the reference value, an H_2 norm of the stochastic linear system dynamics [10], [11].

Current literature focuses on the development of optimization-based algorithms to find the set of leaders that minimizes total system error in the same or similar leader-follower stochastic dynamics as we consider here [10], [12]–[15]. The algorithms are designed to be efficient but do not in general have guarantees on optimality. Since many of the algorithms are iterative (one leader at a time), they do not accommodate that a node in the optimal set of l leaders may often not appear in the optimal set of m leaders, $m > l$. In [15] the authors derive algorithms to obtain lower and upper bounds on the global optimal error in the case of noise-corrupted leaders and they consider a "swap" algorithm to improve upon an iterative greedy algorithm.

Our contributions complement the existing literature: we prove the dependence of the globally optimal one and two leader sets on centrality measures of the network graph for both noise-corrupted and noise-free leaders, and we derive exact solutions for some cases. We leverage a recent result [16], which shows for a network of stochastic evidence accumulating decision-makers that the ordering of nodes with respect to information centrality [17] predicts the ordering of nodes with respect to certainty. We prove that the optimal single leader is the node with maximal information centrality. And we prove that the optimal two leader set maximizes a joint centrality, which depends on the information centrality of each leader and how well the pair "covers" the graph. We apply these results for cyclic graphs and path graphs.

The paper is organized as follows. In Section II, we present the model dynamics and define the problem. In Section III we review information centrality, node certainty and the results of [16]. We prove our main results for the optimal single leader set in Section IV and for the optimal two leader set in Section V. We generalize by computing joint centrality

Supported in part by NSF Graduate Research Fellowship DGE 1148900, ONR grant N00014-09-1-1074, and NSF grant ECCS-1135724.

K. Fitch and N. E. Leonard are with the Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08540, USA.{ kfitch, naomi}@princeton.edu

for m leaders in Section VI. We conclude in Section VII.

II. Model and Problem Statement

We consider a model of n agents with system state denoted by $\mathbf{x} = [x_1, x_2, ..., x_n] \in \mathbb{R}^n$, where x_i is the state of agent i. For every agent i we let the set of neighbors $\mathcal{N}_i$ be the set of agents communicating information to agent i.

The communication topology can be represented by a graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$, where each agent is a node in the set $\mathcal{V} = \{1, 2, ...n\}$, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges, and $A \in \mathbb{R}^{n \times n}$ is the adjacency matrix with nonnegative entries $a_{i,j}$ corresponding to the weight on edge (i, j). The graph contains the edge (i, j) when $j \in \mathcal{N}_i$. In this paper we consider unweighted, undirected graphs so that if the graph contains edge (i, j), then $a_{i,j} = a_{j,i} = 1$ and otherwise $a_{i,j} = 0$. The degree matrix D is a diagonal matrix with entries $d_{i,i} = \sum_{j=1}^{n} a_{i,j}$. The Laplacian matrix associated with the graph $\mathcal{G}$ is defined as $L = D - A$.

Let the reference value from the environment be $\mu \in \mathbb{R}$. Let agent i invest in the measurement of the reference value with gain k_i. Let the set of leaders be S with cardinality m. If agent $i \in S$ then $k_i = k > 0$, otherwise $k_i = 0$. The dynamics are modeled as a system of interconnected Ornstein-Uhlenback stochastic processes of the form

$$dx_i = -k_i(x_i - \mu)dt - L_i\mathbf{x}dt + \sigma dW_i, \tag{1}$$

where L_i is the ith row of L and σdW_i represents increments drawn from independent Wiener processes with standard deviation σ.

It was shown in [15] that the noise-free leader formulation with $k < \infty$ is equivalent to the noise-corrupted leader formulation (1) when all leaders have arbitrarily high feedback gains on their states, i.e., $k \to \infty$. In this paper we derive results for both finite and infinite k.

Letting $K = \mathrm{diag}(k_i)$ and without loss of generality letting $\mu = 0$ allows the above multivariate process to be written in vector form as

$$d\mathbf{x} = -M\mathbf{x}dt + \sigma d\mathbf{W}, \tag{2}$$

where $M = K+L$. We assume that the graph $\mathcal{G}$ is connected. Then, if $k_i = k > 0$ for some agent i, $-M$ is Hurwitz. It follows that the steady-state covariance matrix Σ of $\mathbf{x}$ is the solution to the Lyapunov equation

$$M\Sigma + \Sigma M^T = \sigma^2 I. \tag{3}$$

The diagonal element $\Sigma_{i,i}$ is the steady state variance of x_i about the reference value. The total system error is defined as $\mathrm{tr}(\Sigma) = \sum_{i=1}^{n} \Sigma_{i,i}$ as in [10], [14]. We define group performance to be the inverse of total system error.

Let λ_i denote an eigenvalue of M and let $\nu^{(i)}$ be the corresponding normalized eigenvector. The Moore Penrose pseudo-inverse of a matrix P is indicated by P^+ and the conjugate transpose of P by P^*.

From [18], the covariance matrix of (2) is given by

$$\mathrm{Cov}(\mathbf{x}(t), \mathbf{x}(t)) = \sigma^2 \int_0^t e^{-M(t-\tau)} e^{-M^T(t-\tau)} d\tau. \tag{4}$$

Since the Laplacian of an undirected graph is a symmetric matrix, it follows that M will also be symmetric, and therefore normal. Furthermore, there exists a unitary matrix, U, such that $U^*MU = \Lambda$, where Λ is a diagonal matrix containing the eigenvalues of M. Eq (4) becomes

$$\mathrm{Cov}(\mathbf{x}(t), \mathbf{x}(t)) = \sigma^2 (UR(t)U^*), \tag{5}$$

with

$$R(t) := \int_0^t e^{-(\Lambda + \bar{\Lambda})(t-\tau)} d\tau. \tag{6}$$

Following [19], this gives

$$[\mathrm{Cov}(\mathbf{x}(t), \mathbf{x}(t))]_{ij} = \sigma^2 \sum_{p=1}^{n} \frac{1 - e^{-2\mathrm{Re}(\lambda_p)t}}{2\mathrm{Re}(\lambda_p)} \nu_i^{(p)} \bar{\nu}_j^{(p)}. \tag{7}$$

Since M is Hermitian, all eigenvalues of M will be real. Thus, in steady-state, the variance of the state of each node can be written as

$$\mathrm{Var}(x_i)_{ss} = \Sigma_{i,i} = \sigma^2 \sum_{p=1}^{n} \frac{1}{2\lambda_p} |\nu_i^{(p)}|^2. \tag{8}$$

By (8) the total system error is then

$$\sum_{i=1}^{n} \Sigma_{i,i} = \sigma^2 \sum_{i=1}^{n} \frac{1}{2\lambda_i} = \frac{\sigma^2}{2} \sum_{i=1}^{n} M_{i,i}^{-1}. \tag{9}$$

This is equivalent to an H_2 norm of the system, which characterizes coherence of the network [10], [11].

Given m, the optimal leader selection problem is to find the set of leaders S^* over all possible sets S of m leaders that minimizes the total system error as given by (9):

$$S^* = \arg\min_S \sigma^2 \sum_{i=1}^{n} \frac{1}{2\lambda_i} = \arg\min_S \frac{\sigma^2}{2} \sum_{i=1}^{n} M_{i,i}^{-1}. \tag{10}$$

III. Information Centrality and Certainty

The notion of information centrality was first proposed by Stephenson and Zelen in [17]. The authors define the information contained in a path between two nodes i, j in a graph as the inverse of the length of the path. Summing the information in all paths between the nodes i, j then gives the total information $I_{i,j}^{\mathrm{tot}}$. Information centrality c_i for node i is defined using the harmonic average of the total information between node i and every other node j:

$$c_i = \left(\frac{1}{n} \sum_{j=1}^{n} \frac{1}{I_{i,j}^{\mathrm{tot}}} \right)^{-1}. \tag{11}$$

As shown in [17], $I_{i,j}^{\mathrm{tot}}$ can be computed from the Laplacian matrix, L, without the need for path enumeration. Define $B = (L + \mathbf{1}_n \mathbf{1}_n^T)^{-1}$, then

$$I_{i,j}^{\mathrm{tot}} = (b_{ii} + b_{jj} - 2b_{ij})^{-1}. \tag{12}$$

Poulakakis et al. [16] apply the notion of information centrality to directly interpret the certainty level of each node in a network of stochastic evidence accumulating decision-makers in terms of the structural properties of the underlying communication graph. The certainty of node i, μ_i, is defined as the inverse of the difference between the variance of the state x_i about the reference signal and the minimum achievable variance as $t \to \infty$. It is shown that

$$\frac{1}{\mu_i} = \frac{\sigma^2}{2} L^+_{i,i} = \frac{\sigma^2}{2}\left(\frac{1}{c_i} - \frac{K_f}{n^2}\right), \qquad (13)$$

where K_f is the Kirchhoff index of $\mathcal{G}$. It follows from (13) that the ordering of nodes by information centrality is equivalent to the ordering of nodes by certainty. It is shown that the ordering by certainty is *not* predicted by a degree distribution or geodesic paths between nodes; indeed *all* paths between nodes need to be accounted for to determine the relative certainty of nodes in the network.

We show in the following sections that information centrality is fundamental to the solution of the optimal leader selection problem.

IV. Optimal Single Leader Selection

In this section we derive an explicit expression for total system error in terms of properties of the underlying network topology for the case of single leader selection.

A. Noise-Corrupted and Noise-Free Leader

Theorem 1: Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ be an undirected connected graph of order n. Let the cardinality of the leader set, S, be $m = 1$, and let the leader node be noise-corrupted ($k < \infty$) and indexed by p with information centrality c_p. Then total system error (9) for the system dynamics (2) is

$$\sum_{i=1}^{n} \Sigma_{i,i} = \frac{n\sigma^2}{2}\left(\frac{1}{k} + \frac{1}{c_p}\right), \qquad (14)$$

and the optimal leader set $S^* = \{p^*\} = \arg\max_p c_p$, the node with maximal information centrality.

Before proving Theorem 1, we state a lemma from [20] that we use in the proof.

Lemma 1: [20] Let $\mathbf{d}, \mathbf{e} \in \mathbb{R}^n$. A rank-1 update $\mathbf{e}\mathbf{d}^T$ for the Moore-Penrose pseudo-inverse of a real valued matrix, $F \in \mathbb{R}^{n\times n}$, is given by

$$(F + \mathbf{e}\mathbf{d}^T)^+ = F^+ + G \qquad (15)$$

where

$$G = -\frac{1}{\|\mathbf{w}\|^2}\mathbf{v}\mathbf{w}^T - \frac{1}{\|\mathbf{m}\|^2}\mathbf{m}\mathbf{h}^T + \frac{\beta}{\|\mathbf{m}\|^2\|\mathbf{w}\|^2}\mathbf{m}\mathbf{w}^T \qquad (16)$$

and $\beta = 1 + \mathbf{d}^T F^+ \mathbf{e}$, $\mathbf{v} = F^+\mathbf{e}$, $\mathbf{h} = (F^+)^T\mathbf{d}$, $\mathbf{w} = (I - FF^+)\mathbf{e}$, and $\mathbf{m} = (I - F^+F)^T\mathbf{d}$.

Proof: (Theorem 1) Let $\mathbf{1_n}$ be the vector of n ones. The following are properties of L^+ (see [16] for details):

$$LL^+ = L^+L = I_n - \frac{1}{n}\mathbf{1_n}\mathbf{1_n}^T \qquad (17)$$

$$\mathbf{1_n}^T L^+ = L^+\mathbf{1_n} = 0 \qquad (18)$$

$$Tr(L^+) = \frac{K_f}{n}. \qquad (19)$$

Let $\mathbf{e} = \mathbf{d}$ be vectors of length n with $\sqrt{k}$ in the p^{th} entry and zeros elsewhere, i.e. $\mathbf{e} = \mathbf{d} = [0, \quad 0, ..., \sqrt{k}, ..., 0]^T$. Then $K = \mathbf{e}\mathbf{d}^T$. Since $M = L + K$, Lemma 1 can be applied to compute M^{-1} and thus its trace as

$$\sum_{j=1}^{n} M^{-1}_{j,j} = \sum_{j=1}^{n}(L + \mathbf{e}\mathbf{d}^T)^{-1}_{j,j} = \sum_{j=1}^{n}(L^+ + G)_{j,j}. \qquad (20)$$

By symmetry of the Laplacian and its pseudo-inverse, $\mathbf{m} = \mathbf{w} = \frac{\sqrt{k}}{n}\mathbf{1_n}$. By (16), G can be written as

$$G = -L^+_p\mathbf{1_n}^T - \mathbf{1_n}L^{+T}_p + \frac{(1 + kL^+_{p,p})}{k}\mathbf{1_n}^T\mathbf{1_n}.$$

Therefore (20) becomes

$$\sum_{j=1}^{n} M^{-1}_{j,j} = \sum_{j=1}^{n} L^+_{j,j} - 2\sum_{j=1}^{n}(L^+_{j,p}) + \frac{n(1 + kL^+_{p,p})}{k}.$$

Applying (13), (18), (19),

$$\sum_{j=1}^{n} M^{-1}_{j,j} = \frac{K_f}{n} + \frac{n}{k} + n\left(\frac{1}{c_p} - \frac{K_f}{n^2}\right) = \frac{n}{k} + \frac{n}{c_p}. \qquad (21)$$

Substituting into (9) gives the total system error as

$$\sum_{i=1}^{n} \Sigma_{i,i} = \frac{n\sigma^2}{2}\left(\frac{1}{k} + \frac{1}{c_p}\right). \qquad (22)$$

Total system error (22) is minimized when the leader is node p corresponding to maximal information centrality, c_p. ■

Corollary 1: Consider the conditions of Theorem 1 where the leader is noise-free. Then total system error (9) is

$$\sum_{i=1}^{n} \Sigma_{i,i} = \frac{n\sigma^2}{2}\left(\frac{1}{c_p}\right), \qquad (23)$$

and the optimal leader set $S^* = \{p^*\} = \arg\max_p c_p$, the node with maximal information centrality.

Proof: A noise-free leader is equivalent to a leader with arbitrarily large k. Taking the limit of total system error as $k \to \infty$ yields

$$\lim_{k\to\infty}\sum_{i=1}^{n} \Sigma_{i,i} = \lim_{k\to\infty}\frac{n\sigma^2}{2}\left(\frac{1}{k} + \frac{1}{c_p}\right) = \frac{n\sigma^2}{2}\left(\frac{1}{c_p}\right). \qquad (24)$$

■

B. Examples

Consider the unweighted, undirected network of Fig. 1. In Table I, each node is ranked according to system performance in steady-state if that node were the leader. As predicted by (14), the node with highest information centrality (shown in blue) is the optimal choice of leader for performance. Further, the steady-state performance ranking is equivalent to ranking by information centrality. The inclusion of closeness centrality in Table I serves to demonstrate that this measure, which depends only on geodesic paths between nodes, does not help to determine the optimal leader since it does not distinguish among nodes b, d, and f. This highlights the importance of non-geodesic paths in information transfer within the network and therefore on system performance.

Consider next the cycle graph and the path graph. By symmetry every node in the cycle graph has the same information centrality; thus, any node is the optimal single leader. The optimal single leader in the path graph is the node at the midpoint of the path.

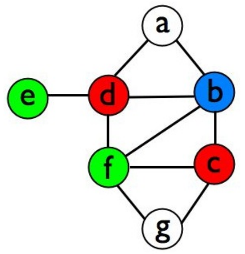

Fig. 1. The connected, undirected graph used in Example 1. Node b with highest information centrality is blue. The optimal noise-corrupted leader pair is red and the optimal noise-free leader pair is green.

Node	Perf. Rank	Infor. Centrality	Closeness Centrality
a	5	1.132	0.091
b	1	1.565	0.125
c	4	1.305	0.100
d	2	1.540	0.125
e	7	0.733	0.077
f	3	1.534	0.125
g	6	1.055	0.083

TABLE I
TABLE RANKING EACH NODE'S PERFORMANCE AS LEADER AND CORRESPONDING CENTRALITY MEASURES

V. OPTIMAL TWO LEADER SELECTION

In this section we derive an explicit expression for total system error in terms of properties of the underlying network topology for the case of two leader selection. In the previous section it was shown that total system error depends upon the information centrality of the single leader. We show here, for two leaders, that total system error depends on a notion that we define as "joint centrality" of the leaders. We employ the notion of resistance distance $r_{i,j}$ between two nodes i, j on an undirected graph $\mathcal{G}$ defined as [21]

$$r_{i,j} = L^+_{i,i} + L^+_{j,j} - 2L^+_{i,j}. \tag{25}$$

The resistance distances with respect to node j and the information centrality of node j are related by

$$\sum_{i=1}^{n} r_{i,j} = \frac{n}{c_j}. \tag{26}$$

A. Noise-corrupted and Noise-free Leaders

Theorem 2: Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ be an undirected connected graph of order n. Let the cardinality of the leader set, S, be $m = 2$, and let the leader nodes be noise-corrupted ($k < \infty$) and indexed by p, s, with information centrality c_p, c_s, respectively. Then total system error (9) for the system dynamics (2) is

$$\sum_{i=1}^{n} \Sigma_{i,i} = \frac{n\sigma^2}{4}\left(\frac{1}{\rho_{k,p,s}}\right), \tag{27}$$

where $\rho_{k,p,s}$ is a k-dependent joint centrality of nodes p, s given by

$$\rho_{k,p,s} = \Big(\frac{1}{k} + \frac{1}{c_p} + \frac{1}{c_s} - \frac{r_{s,p}}{2} - \frac{k\sum_{i=1}^{n}(r_{i,p} - r_{i,s})^2}{2n(2+kr_{s,p})}\Big)^{-1}. \tag{28}$$

The optimal leader set will be $S^* = \{p^*, s^*\} = \arg\max_{p,s} \rho_{k,p,s}$, the two nodes with the maximal k-dependent joint centrality.

Before proving Theorem 2, we state a lemma from [22] that we use in the proof.

Lemma 2: [22] For rank one square matrix H and non-singular X and $X + H$, $(X + H)^{-1}$ can be written as

$$(X + H)^{-1} = X^{-1} - \frac{1}{1+g}X^{-1}HX^{-1}, \tag{29}$$

where $g = \mathrm{tr}(HX^{-1})$.

Proof: (Theorem 2) Let K_p, K_s be rank one matrices with $K_{p_{p,p}} = k$, $K_{s_{s,s}} = k$ where $k > 0$ and all other elements of K_p, K_s are zero. Let $K = K_p + K_s$, let $N = L + K_p$ and let $M = L + K = N + K_s$.

By (29) of Lemma 2 we can compute

$$\begin{aligned} M^{-1} &= (N + K_s)^{-1} \\ &= N^{-1} - \frac{1}{1 + \mathrm{tr}(K_s N^{-1})} N^{-1} K_s N^{-1}. \end{aligned} \tag{30}$$

Since $N = L + K_p$, by Lemma 1 it holds that

$$N^{-1} = L^+ - L^+_p \mathbf{1}_n^T - \mathbf{1}_n L_p^{+T} + \frac{1 + kL^+_{p,p}}{k}\mathbf{1}_n\mathbf{1}_n^T. \tag{31}$$

Thus, by (31)

$$\begin{aligned} \mathrm{tr}(K_s N^{-1}) = kN^{-1}_{s,s} &= 1 + kL^+_{s,s} - 2kL^+_{s,p} + kL^+_{p,p} \\ &= 1 + k\, r_{s,p}. \end{aligned} \tag{32}$$

Using (32) in (30), total system error (9) is

$$\begin{aligned} \sum_{i=1}^{n} \Sigma_{i,i} &= \frac{\sigma^2}{2}\sum_{i=1}^{n} M^{-1}_{i,i} \\ &= \frac{\sigma^2}{2}\sum_{i=1}^{n}\left(N^{-1}_{i,i} - \frac{1}{2 + kr_{s,p}}(N^{-1}K_sN^{-1})_{i,i}\right). \end{aligned} \tag{33}$$

Expanding in terms of L^+ and using (18) and (19) yields

$$\begin{aligned} \sum_{i=1}^{n} M^{-1}_{i,i} =& \frac{n}{k} + \frac{n}{c_p} - \frac{1}{2 + k\, r_{s,p}}\Big(k\sum_{i=1}^{n}(L^+_{i,s} - L^+_{i,p})^2 \\ &+ nk(L^+_{p,s})^2 - 2nL^+_{p,s} - 2nkL^+_{p,p}L^+_{s,p} \\ &+ 2nL^+_{p,p} + nk(L^+_{p,p})^2 + \frac{n}{k}\Big), \end{aligned} \tag{34}$$

where a significant amount of algebra has been omitted due to space constraints. Using (13) and (25) in (34) gives

$$\begin{aligned} \sum_{i=1}^{n} M^{-1}_{i,i} =& \frac{1}{2 + k\, r_{s,p}}\Big(- k\sum_{i=1}^{n}\Big(\frac{1}{2c_s} - \frac{r_{i,s}}{2} - \frac{1}{2c_p} + \frac{r_{i,p}}{2}\Big)^2 \\ &+ \frac{n}{k} + \frac{n}{c_p} + \frac{n}{c_s} - \frac{kn}{4c_p^2} - \frac{kn}{4c_s^2} + \frac{kn}{2c_pc_s} \\ &+ \frac{knr_{s,p}}{2c_p} + \frac{knr_{s,p}}{2c_s} - \frac{knr^2_{s,p}}{4}\Big). \end{aligned} \tag{35}$$

Re-arrangement of terms and application of (26) gives

$$\begin{aligned} \sum_{i=1}^{n} \Sigma_{i,i} =& \frac{n\sigma^2}{4}\Big(\frac{1}{k} + \frac{1}{c_p} + \frac{1}{c_s} - \frac{r_{s,p}}{2} \\ &- \frac{k\sum_{i=1}^{n}(r_{i,p} - r_{i,s})^2}{2n(2 + kr_{s,p})}\Big) = \frac{n\sigma^2}{4\rho_{k,p,s}}, \end{aligned} \tag{36}$$

where $\rho_{k,p,s}$ is defined by (28). Total system error (22) is minimized when the leaders are nodes p, s corresponding to maximal k-dependent joint centrality, $\rho_{k,p,s}$. ■

Comparing total system error for two noise-corrupted leaders (27) and for a single noise-corrupted leader (14), the k-dependent joint centrality $\rho_{k,p,s}$ (28) can be seen to play an analogous role to the k-dependent individual information centrality defined as $c_{k,p} = (\frac{1}{k} + \frac{1}{c_p})^{-1}$. Leaders that maximize these centrality terms optimize performance by minimizing total system error. In both cases, higher k yields better performance. High information centrality of the leader or leaders also contributes to performance. However, while in the case of a single leader the node with maximal information central is the optimal leader, in the case of two leaders, there is a tradeoff between centrality of each of the two leaders which should be high, resistance $r_{s,p}$ between the two leaders which should be large, and the k-dependent distribution term $\sum_{i=1}^{n}(r_{i,p}-r_{i,s})^2/(2+kr_{s,p})$ which should be large. Since resistance is a metric, the last term is a covering term that should be optimized, i.e., the sum over nodes of the squares of the differences between the resistance distance between the node and each leader normalized by a k-dependent function of the resistance distance between leaders. Referring back to Example 1 shown in Fig. 1, the optimal two noise-corrupted leaders with state feedback gain $k = 1$ are nodes c and d, which are highlighted in red. These two nodes clearly do not have the first and second highest information centralities; however they represent the optimal tradeoff of the terms in $\rho_{k,p,s}$ (28). We note that an iterative algorithm that first (correctly) chooses node b as the optimal leader and then seeks a second optimal leader would be unable to solve for the optimal pair c and d. This is likely true in general (see, for instance, the example in [10]).

Corollary 2: Consider the conditions of Theorem 2 where the leaders are noise-free. Then total system error (9) is

$$\sum_{i=1}^{n} \Sigma_{i,i} = \frac{n\sigma^2}{4}\left(\frac{1}{\rho_{p,s}}\right), \tag{37}$$

where $\rho_{p,s}$ is a joint centrality of nodes p, s given by

$$\rho_{p,s} = \Big(\frac{1}{c_p} + \frac{1}{c_s} - \frac{r_{s,p}}{2} - \sum_{i=1}^{n}\frac{(r_{i,p}-r_{i,s})^2}{2nr_{s,p}}\Big)^{-1}. \tag{38}$$

The optimal leader set $S^* = \{p^*, s^*\} = \arg\max_{p,s}\rho_{p,s}$, the two nodes with the maximal joint centrality.

Proof: Assigning noise-free leaders is equivalent to leaders with arbitrarily large k. Taking the limit of total system error as $k \to \infty$ yields

$$\lim_{k\to\infty}\sum_{i=1}^{n}\Sigma_{i,i} = \lim_{k\to\infty}\frac{n\sigma^2}{4\rho_{k,p,s}} = \frac{n\sigma^2}{4\rho_{p,s}} \tag{39}$$

where $\rho_{p,s}$ is given by (38). ■

Now comparing (37) with (23), the k-independent joint centrality $\rho_{p,s}$ (38) in the two leader case can be seen to play an analogous role to the k-independent information centrality c_p in the single leader case. Leaders that maximize these centralities optimize performance by minimizing total system error. The case of the single optimal leader is unchanged as compared to the noise-corrupted case, whereas the optimal two leaders may be different in the noise-free versus the noise-corrupted cases because the covering term (last term in (38)) no longer depends on k. Indeed, in Example 1 of Fig. 1, the optimal two noise-free leaders are nodes e and f, which are highlighted in green.

The following lemma provides an upper bound on $\rho_{p,s}$.

Lemma 3:

$$\rho_{p,s} \le \Big(\frac{1}{c_s} + \frac{1}{c_p} - r_{s,p}\Big)^{-1}$$

Proof: Resistance is a metric, $r_{i,p} \le r_{i,s} + r_{s,p}$. So,

$$\sum_{i=1}^{n}\frac{(r_{i,p}-r_{i,s})^2}{2nr_{s,p}} \le \sum_{i=1}^{n}\frac{r_{s,p}^2}{2nr_{s,p}} = \frac{nr_{s,p}^2}{2nr_{s,p}} = \frac{r_{s,p}}{2}.$$

■

B. Two Leaders on an Undirected Cycle

Theorem 3: Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ be an undirected cycle graph of order n where n is even. Let the cardinality of the leader set, S, be $m = 2$, and let the leader nodes be noise-corrupted and indexed by p, s. The optimal leader set S^* is any two nodes with maximal resistance distance $r_{s,p} = \frac{n}{4}$, which corresponds to geodesic distance $d_{s,p} = \frac{n}{2}$ (antipodal nodes).

Proof: For a circulant graph, $c_p = c_s = c$. By (28)

$$\rho_{k,p,s} = \Big(\frac{1}{k} + \frac{2}{c} - \frac{r_{s,p}}{2} - \frac{k\sum_{i=1}^{n}(r_{i,p}-r_{i,s})^2}{2n(2+kr_{s,p})}\Big)^{-1}. \tag{40}$$

By analyzing the graph as an electric circuit, the resistance distance between any two nodes in a cycle can be written as

$$\frac{1}{r_{j,p}} = \frac{1}{d_{j,p}} + \frac{1}{n-d_{j,p}} \tag{41}$$

where $d_{j,p}$ is the geodesic distance between nodes j and p. The maximum resistance distance is $r_{j,p} = \frac{n}{4}$, which is obtained between two nodes with $d_{j,p} = \frac{n}{2}$.

Using (41) gives

$$\begin{aligned}\sum_{i=1}^{n}(r_{i,s}-r_{i,p})^2 &= \sum_{i=1}^{n}\Big(d_{i,p} - d_{i,s} + \frac{d_{i,s}^2 - d_{i,p}^2}{n}\Big)^2 \\ &= \frac{d_{s,p}(d_{s,p}-n)(d_{s,p}^2 - nd_{s,p} - 2)}{3n}.\end{aligned} \tag{42}$$

Plugging (42) into (40) yields

$$\begin{aligned}\rho_{k,p,s} = \Big(&\frac{1}{k} + \frac{2}{c} - \frac{d_{s,p}(n-d_{s,p})}{2n} \\ &- \frac{kd_{s,p}(d_{s,p}-n)(d_{s,p}^2 - nd_{s,p} - 2)}{6n(2n + kd_{s,p}(n-d_{s,p}))}\Big)^{-1}.\end{aligned} \tag{43}$$

Then,

$$\begin{aligned}\frac{\partial\rho_{k.p,s}^{-1}}{\partial d_{s,p}} =& -\frac{1}{2n}(n - 2d_{s,p}) \\ &- \frac{2k[2(-d_{s,p}+d_{s,p}^3) + (1-3d_{s,p}^2)n + d_{s,p}n^2]}{3(2n + d_{s,p}k(-d_{s,p}+n))^2} \\ &- \frac{k^2[-2d_{s,p}^5 + 5d_{s,p}^4 n - 4d_{s,p}^3 n^2 + d_{s,p}^2 n^3]}{6n(2n + d_{s,p}k(-d_{s,p}+n))^2}.\end{aligned} \tag{44}$$

The first term of (44) will always be nonpositive. Furthermore, since $d_{s,p} \le \frac{n}{2}$ it can easily be shown that the two bracketed expressions in the second and third terms will be greater than zero for $n > 3$. Thus $\rho_{k,p,s}^{-1}$ decreases as $d_{s,p}$ increases and reaches its minimum when $d_{s,p}$ is maximal at $d_{s,p} = \frac{n}{2}$, corresponding to $r_{s,p} = \frac{n}{4}$. The result holds by Theorem 2. ■

Due to the symmetry of the cycle, the result of Theorem 3 holds in the noise-free two leader case as well. Because every node in the cycle has the same information centrality, the optimal leader set can be interpreted as the set that maximizes the distance between the two leaders plus the covering term, which simplifies further since the covering term is maximized for maximal distance between leaders. We conjecture that for a greater number of leaders $m > 2$, the optimal leader set should be any set of nodes that is uniformly distributed around the cycle.

C. Two Leaders on a Path

Theorem 4: Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ be an undirected, connected path graph of order n, which is the cycle graph with one link removed. Let the cardinality of the leader set, S, be $m = 2$, and let the leader nodes be noise-free and indexed by p, s where $s > p$. The optimal leader set S^* corresponds to $p^* = \text{rnd}(\frac{n}{5} + \frac{1}{2})$ and $s^* = \text{rnd}(\frac{4n}{5} + \frac{1}{2})$, where rnd is rounding to the closest integer.

Proof: For the path graph we have $r_{i,j} = \|i - j\|$ and

$$c_j = \frac{n}{\sum_{i=1}^{n} r_{i,j}} = \frac{2n}{(n-j)(1+n-j) - j + j^2}. \tag{45}$$

Indexing the leaders as s, p where $s > p$, substituting (45) into (38) and simplifying yields

$$\rho_{p,s}^{-1} = -\frac{1}{3n} + \frac{n + n^2 - p - s}{2n} + \frac{2(2p^2 + 2s^2 - s(3n + p))}{3n}. \tag{46}$$

Taking partial derivation of (46) with respect to s, p we find the minimum of $\rho_{p,s}^{-1}$ to be $p = \text{rnd}(\frac{n}{5} + \frac{1}{2})$ and $s = \text{rnd}(\frac{4n}{5} + \frac{1}{2})$. The level sets of $\rho_{p,s}^{-1}$ can be shown to be ellipses. A computation of the semi-axis lengths of the ellipses shows that the optimal leader set corresponds to rounding to the nearest integer the values that minimize $\rho_{p,s}^{-1}$. ■

VI. Joint centrality of m noise-free leaders

For the general case, we first perform a rank-1 update to find the inverse of M. We then apply the Sherman-Morrison-Woodbury formula [23] for a $m-1$ rank update to M^{-1}. Analogously to the previous sections, this allows us to determine the total system error for a set, S, of m leaders.

We determine the m-joint centrality of the noise-free leader set S to be

$$\rho_S = \Big(\frac{1}{2c_p} - \sum_{s,l \in S\setminus\{p\}} \frac{G_{s,l}}{2} \Big\{ \frac{r_{s,p} + r_{l,p}}{c_p} - \frac{r_{s,p}}{c_l} - \frac{r_{l,p}}{c_s} + r_{s,p} r_{l,p} + \sum_{i=1}^{n} \frac{1}{n} [r_{i,p}(r_{i,p} - r_{i,l} - r_{i,s}) + r_{i,s} r_{i,l}] \Big\} \Big)^{-1}, \tag{47}$$

where $G_{s,l}$ is the s, l element of the inverse of the $(m-1) \times (m-1)$ leader submatrix of M^{-1}.

VII. Final Remarks

In this paper we analyze and provide new insights on the optimal leader selection problem in a leader-follower multi-agent system subject to stochastic disturbances where performance is measured by coherence of the system. We prove that the optimal single leader maximizes information centrality. We prove that the optimal two leaders maximize a joint centrality that depends on the information centrality of each leader and on how the two leaders are distributed across the network. Future directions include exploring the generalization (47) to greater numbers of leaders and leveraging these results for the design of distributed adaptive leadership dynamics to yield high-performing dynamic networks.

References

[1] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, 2004.

[2] W. Ren, R. Beard, and E. Atkins, "A survey of consensus problems in multi-agent coordination," in *Proc. ACC*, 2005, pp. 1859–1864.

[3] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *Journal of Parallel and Distributed Computing*, vol. 67, pp. 33–46, 2007.

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

[5] G. F. Young, L. Scardovi, A. Cavagna, I. Giardina, and N. E. Leonard, "Starling flock networks manage uncertainty in consensus at low cost," *PLoS Computational Biology*, vol. 9, no. 1-7, 2013.

[6] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-bayesian social learning," *GEB*, vol. 76, pp. 210–225, 2012.

[7] V. Guttal and I. D. Couzin, "Social interactions, information use, and the evolution of collective migration," *PNAS*, vol. 107, p. 16172, 2010.

[8] D. Pais and N. E. Leonard, "Adaptive network dynamics and evolution of leadership in collective migration," *Physica D, http://dx.doi.org/10.1016/j.physd.2013.04.014*, 2013.

[9] J. Wang, Y. Tan, and I. Mareels, "Robustness analysis of leader-follower consensus," *J. Sys. Sci. Complex.*, vol. 22, pp. 186–206, 2009.

[10] S. Patterson and B. Bamieh, "Leader selection for optimal network coherence," in *Proc. IEEE CDC*, 2010, pp. 2693–2697.

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

[12] M. Fardad, F. Lin, and M. R. Jovanovic, "Algorithms for leader selection in large dynamical networks: noise-free leaders," in *Proc. IEEE CDC*, 2011, pp. 7188–7193.

[13] F. Lin, M. Fardad, and M. R. Jovanovic, "Algorithms for leader selection in large dynamical networks: noise-corrupted leaders," in *Proc. IEEE CDC*, 2011, pp. 2932–2937.

[14] A. Clark, L. Bushnell, and R. Poovendran, "A supermodular optimization framework for leader selection under link noise in linear multi-agent systems," *arXiv:1208.0946 [cs.SY]*, 2012.

[15] F. Lin, M. Fardad, and M. R. Jovanovic, "Algorithms for leader selection in stochastically forced consensus networks," *arXiv:1302.0450 [math.OC]*, 2013.

[16] I. Poulakakis, L. Scardovi, and N. E. Leonard, "Node classification in networks of stochastic evidence accumulators," *arXiv:1210.4235 [cs.SY]*, 2012.

[17] K. Stephenson and M. Zelen, "Rethinking centrality: Methods and examples," *Social Networks*, vol. 11, pp. 1–37, 1989.

[18] L. Arnold, *Stochastic Differential Equations: Theory and Applications*. Krieger Publishing Complany, 1992.

[19] I. Poulakakis, L. Scardovi, and N. E. Leonard, "Coupled stochastic differential equations and collective decision making in the two-alternative forced-choice task," in *Proc. ACC*, 2010, pp. 69–74.

[20] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, 2nd ed. Springer-Verlag, 2003.

[21] D. J. Klein and M. J. Randic, "Resistance distance," *J. Math. Chem.*, vol. 12, 1993.

[22] K. S. Miller, "On the inverse of the sum of matrices," *Mathematics Magazine*, vol. 54, no. 2, pp. 67–72, 1981.

[23] M. A. Woodbury, *Inverting modified matrices*. Statistical Research Group, Princeton Univ. Memorandum Rept. 42, 1950.