

Node Certainty in Collective Decision Making

Ioannis Poulakakis *Member, IEEE*, Luca Scardovi *Member, IEEE*, and Naomi E. Leonard *Fellow, IEEE*

Abstract—This paper brings into focus the relationship between the location of a decision-making unit in a network of decision makers and its certainty in integrating information toward a decision. A collection of units, each represented by a Drift-Diffusion Model (DDM), accrues evidence in continuous time by observing a (noisy) stimulus. Their task is to make a decision that depends on accurately identifying the stimulus observed. It is shown that common structural centrality measures based on nodal degree or geodesic distance cannot be used to rank the units according to their certainty in integrating information. Instead, the variance associated with the state of a decision-making unit depends on the communication topology in a way that incorporates *all* possible paths connecting that unit with the rest.

I. INTRODUCTION

In collective implementations of decision-making tasks that depend on noisy measurements of a stimulus, information sharing can significantly impact the certainty of each unit as it accumulates evidence toward a decision. Depending on the structural properties of the underlying communication architecture, units that are more “central” than others may emerge as potentially more powerful decision makers. The goal of this work is to identify such units in a restricted, albeit common, decision-making scenario.

The complexity of the mechanisms governing the capacity of humans or animals to make accurate and fast decisions in multi-scale, time-varying, highly uncertain environments can be daunting. Nonetheless, certain decision-making scenarios can be understood on a phenomenological level through the introduction of archetypical reductive mathematical models [1], [2]. An example of such a scenario is the *Two-Alternative Forced-Choice (TAFC)* task; see [2] for an extensive account. The TAFC task represents a canonical behavioral experiment, in which each trial involves correctly identifying a noisy stimulus drawn at random between two possibilities. Both behavioral data in humans [2], [3], and direct recordings of neural firing rates in sensorimotor brain areas of primates [4], [5], performing simple TAFC tasks provide evidence supporting the *Drift-Diffusion Model (DDM)*, as a plausible model for formally investigating such tasks.

This work was supported in part by AFOSR grant FA9550-07-1-0-0528, ONR grant N00014-09-1-1074, NSF grant ECCS-1135724 and ARO grant W911NG-11-1-0385. L. Scardovi was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

I. Poulakakis is with the Department of Mechanical Engineering, University of Delaware, Newark, DE 19716, USA; e-mail: poulakas@udel.edu.

L. Scardovi is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4, Canada; e-mail: scardovi@scg.utoronto.ca.

N. E. Leonard is with the Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA; e-mail: naomi@princeton.edu.

We turn our attention to exploring the role of information sharing in a network of interconnected drift-diffusion decision makers. Assessing the consequences of the overall network structure on each individual unit naturally leads to the notion of centrality [6]. Centrality measures typically assign to each node a quantity that reflects its location in the network. Based on shortest paths, a family of such measures has been introduced in [7]. This choice, however, does not incorporate information transmitted between pairs of nodes through non-geodesic pathways. A centrality index that makes use of all possible paths connecting any pair of nodes has been proposed by Stephenson and Zelen in [8], where it was applied to interpret the spread of an infectious disease in a network of interconnected individuals.

In this work, we examine the relationship between the location of a node in the underlying communication topology and its certainty as it collects and exchanges information with its neighbors. Assuming that evidence is accumulated in continuous time according to the drift-diffusion paradigm, we provide a formal connection between the variance of the state of each unit and the notion of information centrality proposed by Stephenson and Zelen in [8]. Our result demonstrates that collective evidence accumulation in the context of coupled DDMs is a total network process: the entirety of paths—including those that are not geodesic—connecting a unit with the rest of the network affect the unit’s certainty as it integrates noisy information in the pursuit of a decision. This result provides a classification of nodes by their certainty as a function of the network topology. Such classification can be used to improve collective decision-making performance, e.g. by suitably weighting the information supplied—or the decision made [9]—by each unit in the network.

The structure of this paper is as follows. In Section II, the DDM is introduced and extended to a collective evidence accumulation setting. In Section III a node certainty index is defined. Our main result that interprets the node certainty index in terms of information centrality is presented in Section IV. Section V concludes the paper.

II. MODEL AND PROBLEM STATEMENT

A. Decision Making and Drift-Diffusion Models

The DDM describes the integration of information based on observations of a noisy stimulus. It has been used extensively as a model of evidence accumulation in decision-making tasks, including the TAFC task that is commonly employed in human and animal behavioral experiments [2]. Under the assumption that the difference between the amounts of evidence supporting each of the two choices is

integrated over each trial, and for unbiased initial conditions, the standard DDM accumulates evidence according to

$$dx = \beta dt + \sigma dW, \quad x(0) = 0, \quad (1)$$

where $x(t)$ denotes the accumulated value at time t of the difference in the information favoring one choice over the other; $x = 0$ means that the amounts of integrated evidence are equal. In (1), the constant drift β represents an increase in the evidence supporting the correct decision and σdW are increments drawn from a Wiener process with standard deviation σ . The probability density of solutions of (1) at t is normally distributed with mean $\mathbb{E}[x(t)] = \beta t$ and variance $\text{Var}(x(t)) = \sigma^2 t$, i.e. $p(x, t) = N(\beta t, \sigma^2 t)$; see [10].

In the relevant literature, (1) arises in a variety of ways. For example, it is shown in [2] that under the assumption of infinitesimal increments of information arriving at each moment in time, (1) is the (weak) limit of the logarithmic likelihood ratio in typical binary hypothesis tests. In a different context, it is also shown in [2] that the DDM can be derived through appropriate reductions in models of competing leaky accumulators representing two neural populations with activities that provide evidence for the two alternatives in a TAFC task. More recently, [11] established a connection between (1) and the long-time behavior of an integrated Ornstein-Uhlenbeck velocity process obtained as the (weak) limit of an excitatory-inhibitory shot noise pair at high intensities. The work in [11] provides a quantitative link from the microscopic, short-time statistics of neuronal representations to the macroscopic, long-time statistics of information accumulation processes.

B. Networks of Interconnected DDMs

The objective of this work is to analyze the role of exchanging information in reducing uncertainty in a network of agents observing a signal corrupted by noise. A collection of n agents coupled according to certain communication topologies is considered, and the drift-diffusion paradigm (1) is adopted and extended to model evidence accumulation at each agent in the presence of information received from its neighbors. Mathematically, for each $k = 1, \dots, n$, the state x_k of the k -th agent evolves according to

$$dx_k = \left[\beta + \sum_{j=1}^n \alpha_{kj} (x_j - x_k) \right] dt + \sigma dW_k, \quad (2)$$

where, in analogy with (1), β represents a constant drift term and σdW_k corresponds to increments drawn from independent Wiener processes with standard deviation σ . In (2), $\alpha_{kj} \geq 0$ denotes the attention paid by the agent k to the difference between its state x_k and the state x_j of the j -th agent; $\alpha_{kj} = 0$ implies that k and j do not communicate.

The model (2) can be directly associated with a collective decision-making scenario, in which n decision-making units are presented with a stimulus (e.g., a deterministic signal) partially buried in noise, and are asked to correctly identify it between two alternatives [12]. Continuous-time models described by (2) with $\beta = 0$ have been used to provide

sufficient and necessary conditions for mean-square average consensus under measurement noise [13], and to analyze robustness in linear consensus algorithms in the presence of (white) noise [14].

It is natural to identify the communication topology with a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ with vertex set \mathcal{V} containing the agents and edge set \mathcal{E} containing the communication links among them according to the adjacency matrix A with elements α_{kj} . Then, (2) takes the form

$$dx = (b - Lx) dt + HdW, \quad (3)$$

where¹ $x := \text{col}(x_1, \dots, x_n)$, $b := \beta \mathbf{1}_n$, $H := \sigma I_n$, $dW := \text{col}(dW_1, \dots, dW_n)$, and L is the Laplacian matrix of the interconnection graph, \mathcal{G} , defined by

$$L_{kj} := \begin{cases} \sum_{i=1, i \neq k}^n \alpha_{ki}, & k = j, \\ -\alpha_{kj}, & k \neq j. \end{cases} \quad (4)$$

C. Basic Properties of the Model

It is known that the solution $\{x(t) : t \geq 0\}$ of (3) is a Gaussian process [10]. Under deterministically zero initial conditions, i.e., $\text{Cov}(x_0, x_0) = 0$ and $\mathbb{E}[x_0] = 0$, the mean and covariance of $x(t)$ are given by

$$\mathbb{E}[x(t)] = \int_0^t e^{-L(t-\tau)} b d\tau \quad (5)$$

and

$$\text{Cov}(x(t), x(t)) = \sigma^2 \int_0^t e^{-L(t-\tau)} e^{-L^T(t-\tau)} d\tau, \quad (6)$$

respectively; see [10]. The following lemma provides lower and upper bounds for the variance associated with the state of each unit.

Lemma 1: Consider (3). For any interconnection graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ and any node $v_k \in \mathcal{V}$,

$$\frac{\sigma^2}{n} t \leq \text{Var}(x_k(t)) \leq \sigma^2 t. \quad (7)$$

Proof: Let \mathbf{q}_k denote the $n \times 1$ vector with a 1 in the k -th entry and zeros in the remaining entries; clearly, $\sum_{k=1}^n \mathbf{q}_k = \mathbf{1}_n$. Then,

$$\begin{aligned} \text{Var}(x_k(t)) &= \mathbf{q}_k^T \text{Cov}(x(t), x(t)) \mathbf{q}_k \\ &= \sigma^2 \int_0^t \|e^{-L^T(t-\tau)} \mathbf{q}_k\|^2 d\tau \\ &= \sigma^2 \int_0^t \sum_{\ell=1}^n \left(\mathbf{q}_\ell^T e^{-L^T(t-\tau)} \mathbf{q}_k \right)^2 d\tau, \end{aligned} \quad (8)$$

where (6) has been used. The lower bound in (7) is obtained through Jensen's inequality²,

$$\sum_{\ell=1}^n \left(\mathbf{q}_\ell^T e^{-L^T(t-\tau)} \mathbf{q}_k \right)^2 \geq \frac{1}{n} \left(\sum_{\ell=1}^n \mathbf{q}_\ell^T e^{-L^T(t-\tau)} \mathbf{q}_k \right)^2, \quad (9)$$

¹Notation: $\mathbf{1}_n$ is the n -dimensional column vector with entries all equal to one and I_n is the $n \times n$ identity matrix.

²Jensen's inequality: Let f be a convex function on an interval \mathcal{X} and $x_j \in \mathcal{X}$ for $j \in \{1, \dots, n\}$. Then, $f\left(\frac{1}{n} \sum_{j=1}^n x_j\right) \leq \frac{1}{n} \left(\sum_{j=1}^n f(x_j)\right)$.

by observing that $e^{-L^T(t-\tau)}\mathbf{q}_k$ corresponds to the k -th column of $e^{-L^T(t-\tau)}$, that is, the k -th row of $e^{-L(t-\tau)}$, and by noticing that $e^{-L(t-\tau)}$ is row-stochastic³. Finally, the upper bound follows from

$$\sum_{\ell=1}^n \left(\mathbf{q}_\ell^T e^{-L^T(t-\tau)} \mathbf{q}_k \right)^2 \leq \left(\sum_{\ell=1}^n \left| \mathbf{q}_\ell^T e^{-L^T(t-\tau)} \mathbf{q}_k \right| \right)^2 \quad (10)$$

in a similar fashion. ■

Remark 1: In view of the fact that $\sigma^2 t$ corresponds to the variance of the state of an *isolated* DDM (1), the upper bound of inequality (7) implies that the uncertainty associated with the evidence collected by any of the units in a graph cannot exceed that of a unit collecting evidence in isolation.

Remark 2: When t is sufficiently small, (6) results in $\text{Cov}(x(t), x(t)) \approx (\sigma^2 t) I_n$, implying that all the units behave like isolated DDMs at the beginning of the process. However, as time evolves and the units collect and communicate their accumulated evidence to their neighbors, their certainty improves with respect to that of an isolated unit in a way that depends on the topology of the communication. An objective of this work is to identify units with variance that evolves more closely to the lower bound in (7).

III. A NODE CERTAINTY INDEX

This section introduces an index that characterizes the certainty of each unit in connected undirected graphs. We begin with an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ with Laplacian matrix L . Since L is symmetric, there exists an orthogonal matrix U such that $U^T L U = \Lambda$, where Λ is the diagonal matrix containing the eigenvalues of L . For such graphs, the covariance matrix (6) can be written as

$$\text{Cov}(x(t), x(t)) = \sigma^2 (U G(t) U^T), \quad (11)$$

where

$$G(t) := \int_0^t e^{-2\Lambda(t-\tau)} d\tau. \quad (12)$$

Under the assumption that \mathcal{G} is connected⁴, [12, Lemma 1] provides the following expression for the variance $\text{Var}(x_k(t))$ of the state of the k -th unit:

$$\text{Var}(x_k(t)) = \sigma^2 \frac{t}{n} + \sigma^2 \sum_{p=2}^n \frac{1 - e^{-2\lambda_p t}}{2\lambda_p} |u_k^{(p)}|^2, \quad (13)$$

where $u_k^{(p)}$ is the k -th component of the normalized eigenvector that corresponds to the nonzero⁵ eigenvalue λ_p , $p = 2, \dots, n$, of L .

Equation (13) can be used to define an index, based on which the nodes can be classified in terms of their certainty as they accrue evidence. First, note that, by [15, Theorem

³This is a consequence of the zero row sums of the Laplacian matrix L .

⁴Definition: \mathcal{G} is connected if there exist a path between any two vertices.

⁵By [15, Theorem 1.37] connectivity of \mathcal{G} implies that $\mathbf{1}_n$ spans the kernel of L , so that $u^{(1)} = (1/\sqrt{n}) \mathbf{1}_n$ is the normalized eigenvector corresponding to the zero eigenvalue $\lambda_1 = 0$. Equation (13) follows.

1.37], connectivity of \mathcal{G} implies $\lambda_p > 0$ for $p = 2, \dots, n$, so that by (13) the limit

$$\lim_{t \rightarrow +\infty} \left(\text{Var}(x_k(t)) - \sigma^2 \frac{t}{n} \right) =: \frac{1}{\mu(v_k)} \quad (14)$$

is well defined. Then, for each node v_k in a connected undirected graph \mathcal{G} , we define the *node certainty index* $\mu(v_k) > 0$ as the inverse of the difference between the variance of v_k and the minimum achievable variance $\sigma^2 t/n$ as $t \rightarrow +\infty$. For a node v_k , a high value of $\mu(v_k)$ corresponds to small uncertainty associated with the state x_k of that node. Equivalently, the variance of the state x_k of v_k evolves closely to the minimum possible variance $\sigma^2 t/n$, after transients decay. By convention, $\mu(v_k) = \infty$ corresponds to nodes that achieve the highest possible certainty.

In view of (13), equation (14) implies

$$\frac{1}{\mu(v_k)} = \sigma^2 \sum_{p=2}^n \frac{1}{2\lambda_p} |u_k^{(p)}|^2. \quad (15)$$

It is evident from (15) that the node certainty index μ depends on the eigenstructure of the graph Laplacian, reflecting the fact that the certainty of a node is a function of its location in the underlying interconnection graph. Classifying the nodes of a graph based on their certainty and interpreting this classification in terms of the structural properties of the interconnection graph is at the core of this work.

Before continuing with clarifying the relation between the node certainty index μ and suitable notions of network centrality, a few remarks are in order.

Remark 3: From (15) it is easy to see that

$$\sum_{v_k \in \mathcal{V}(\mathcal{G})} \frac{1}{\mu(v_k)} = \sigma^2 \sum_{p=2}^n \frac{1}{2\lambda_p}. \quad (16)$$

As was discussed in [14], the sum in the right hand side of (16) corresponds to the expected steady-state *dispersion* around the consensus subspace of the evidence collected by the decision-making units. Hence, the inverse of $\mu(v_k)$ represents the individual contribution of node v_k to the dispersion of the evidence; the higher $\mu(v_k)$, the smaller the contribution of the node v_k . Note that the sum in the right-hand side of (16) is related to the effective resistance K_f of the graph,

$$K_f := n \sum_{p=2}^n \frac{1}{\lambda_p}, \quad (17)$$

also known as the Kirchhoff index; see [16]. Clearly,

$$\sum_{v_k \in \mathcal{V}(\mathcal{G})} \frac{1}{\mu(v_k)} = \sigma^2 \left(\frac{K_f}{2n} \right). \quad (18)$$

Remark 4: The definition of the node certainty index can be extended to strongly connected digraphs under the assumption that the associated Laplacian matrix is normal. Indeed, the normality condition implies that L is unitarily diagonalizable, so that

$$\text{Var}(x_k(t)) = \sigma^2 \frac{t}{n} + \sigma^2 \sum_{p=2}^n \frac{1 - e^{-2\text{Re}(\lambda_p)t}}{2\text{Re}(\lambda_p)} |u_k^{(p)}|^2; \quad (19)$$

see [12, Lemma 1] for details. Since by [15, Theorem 1.37] strong connectivity implies $\text{Re}(\lambda_p) > 0$ for $p = 2, \dots, n$ the limit (14) is well defined, and the node certainty index can be computed by the following rule

$$\frac{1}{\mu(v_k)} = \sigma^2 \sum_{p=2}^n \frac{1}{2\text{Re}(\lambda_p)} |u_k^{(p)}|^2. \quad (20)$$

IV. NODE CERTAINTY AS A CENTRALITY MEASURE

This section provides an interpretation of the node certainty index μ , and the associated node classification, through a suitable centrality measure. It is intuitively discussed and rigorously formalized that the impact of network architecture on node certainty depends on the *totality* of paths—and not just the geodesic paths—in the network.

A. Motivating Examples

To motivate the discussion, we begin with two examples of undirected graphs. For each node v_k , we compute the certainty index $\mu(v_k)$ using (15). In addition, we provide its degree—that is, the number of edges attached to v_k —and the corresponding *closeness* centrality, $\kappa_{\text{close}}(v_k)$, computed as the inverse of the mean geodesic distance $d(v_k, v_j)$ from v_k to v_j , averaged over all nodes v_j in the graph, i.e.

$$\kappa_{\text{close}}(v_k) = \left(\frac{1}{n} \sum_{j=1}^n d(v_k, v_j) \right)^{-1}; \quad (21)$$

see [6, Section 7.6]. It will be shown that, for general undirected graphs, node certainty cannot be captured by centrality measures based on degrees or geodesic paths. This is a consequence of the fact that the evidence accumulated by each unit is transmitted through the network and reaches the rest of the units via circuitous, non-geodesic pathways.

Example 1: Consider the unweighted undirected tree used in Fig. 1. According to the table in Fig. 1, the node v_5 is the most certain, in the sense that the variance of its state evolves closest to the minimum achievable variance $\sigma^2 t/n$. Clearly, v_5 is not the maximum degree node; the degree of v_5 is equal to three, i.e. less than the degree of v_6 , which is four. Thus, the local nature of a centrality measure based on nodal degrees precludes its use to distinguish the nodes of the given tree in terms of their certainty. Instead, compared to any other node in the tree, v_5 is the node that minimizes the sum of the lengths of the paths that start from it. Hence, this example suggests that closeness centrality (21) is a suitable measure that captures the classification of the nodes based on the certainty index μ ; see also the table in Fig. 1. While this observation is true for trees, the following example shows that it does not hold when the graph contains cyclic paths.

Example 2: Consider the graph of Fig. 2. It is clear from the table in Fig. 2 that $\mu(v_3) = \mu(v_4) > \mu(v_5)$. In contrast to Example 1, this distinction between vertex v_5 and the vertices v_3, v_4 cannot be captured by closeness centrality. In particular, any of the vertices v_3, v_4 and v_5 in the graph of Fig. 2 is connected to the rest of the nodes through two geodesic paths of length 2 and two geodesic paths of length 1, resulting

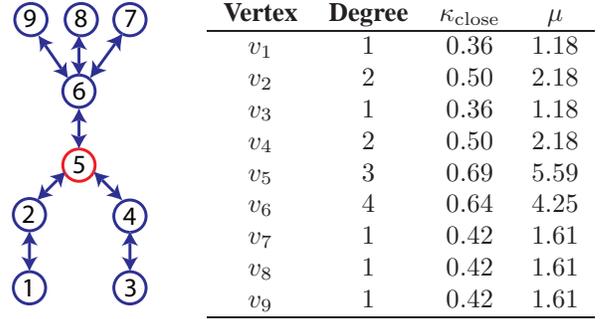


Fig. 1. Left: The unweighted undirected tree used in Example 1. The node that maximizes the certainty index is v_5 . Right: The table summarizes properties of the nodes, such as degree, κ_{close} and μ .

in $\kappa_{\text{close}}(v_3) = \kappa_{\text{close}}(v_4) = \kappa_{\text{close}}(v_5)$. The reason for the ambiguity is that the definition (21) of closeness centrality is based on geodesic paths, and does not incorporate non-geodesic pathways. To provide some intuition, consider the pairs $\{v_1, v_4\}$ and $\{v_1, v_5\}$, and enumerate all possible paths connecting them. For $\{v_1, v_4\}$ we have the paths $v_1 - v_4$, $v_1 - v_2 - v_3 - v_4$ and $v_1 - v_5 - v_2 - v_3 - v_4$; for $\{v_1, v_5\}$ we have the paths $v_1 - v_5$, $v_1 - v_2 - v_5$ and $v_1 - v_4 - v_3 - v_2 - v_5$. Thus, the evidence transmitted by v_1 reaches v_4 via three paths of length 1, 3 and 4, respectively. On the other hand, it reaches v_5 via three paths of lengths 1, 2 and 4. This difference is reflected in the node certainty index, and cannot be captured by shortest paths, based on which v_4 and v_5 are indistinguishable. This observation reveals the non-geodesic nature underlying information transmission, which can be captured only if all possible paths between any pair of nodes are taken into account.

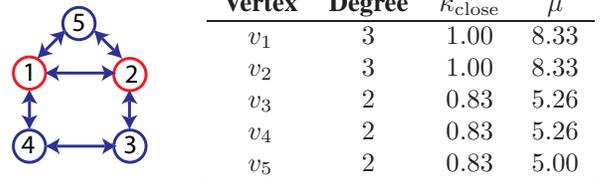


Fig. 2. Left: The connected undirected graph used in Example 2. The nodes that maximize the certainty index are v_1 and v_2 . Right: The table summarizes properties of the nodes, such as degree, κ_{close} and μ .

B. Main Result: Node Certainty and Information Centrality

This section clarifies the relation between node certainty, as measured by the index μ , and the location of a node in the underlying interconnection graph through the notion of *information centrality* [8]. To demonstrate the concept in an intuitive way, we first describe a path enumeration procedure. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ be a connected undirected graph; for simplicity assume an unweighted graph. Consider the pair of vertices $v_k, v_j \in \mathcal{V}(\mathcal{G})$. Suppose there are l_{kj} paths $\mathcal{P}_{kj}(q)$, $q = 1, \dots, l_{kj}$, connecting v_k and v_j , and that different paths $\mathcal{P}_{kj}(q_1)$ and $\mathcal{P}_{kj}(q_2)$, $q_1, q_2 = 1, \dots, l_{kj}$ with $q_1 \neq q_2$, do not have common edges. Denoting the length of each such path by $\ell(\mathcal{P}_{kj}(q))$, we define the information contained in it by

the inverse of its length, i.e.,

$$I_{kj}(q) = \frac{1}{\ell(\mathcal{P}_{kj}(q))}. \quad (22)$$

Then, the total information transmitted through all paths connecting the nodes v_k and v_j is given by

$$I_{kj}^{\text{tot}} = \sum_{q=1}^{l_{kj}} I_{kj}(q). \quad (23)$$

Intuitively, I_{kj}^{tot} represents the information contained in a ‘‘combined’’ path, which incorporates all paths $\mathcal{P}_{kj}(q)$ from v_k to v_j , by weighting each according to the factor $I_{kj}(q)/I_{kj}^{\text{tot}}$. The weight $I_{kj}(q)/I_{kj}^{\text{tot}}$ corresponds to the percentage of the total information transmitted from v_k to v_j that is contained in $\mathcal{P}_{kj}(q)$.

When common edges exist among different paths connecting a pair of vertices, the procedure for computing I_{kj}^{tot} via path enumeration needs to be modified so that the contribution of each common edge to the information of the combined path is weighted appropriately. Rather than describing the modified procedure, we provide the following lemma due to Stephenson and Zelen [8], which gives a simple way to compute I_{kj}^{tot} without path enumeration. Note that Lemma 2 incorporates the case where different paths between pairs of vertices share one or more of their edges, and can be used to compute I_{kj}^{tot} in weighted graphs as well.

Lemma 2 (Stephenson and Zelen, [8]): Let \mathcal{G} be an undirected connected graph of order n and let L be its Laplacian. Then, the total information I_{kj}^{tot} transmitted via all paths connecting $v_k, v_j \in \mathcal{V}(\mathcal{G})$ is

$$I_{kj}^{\text{tot}} = (c_{kk} + c_{jj} - 2c_{kj})^{-1}, \quad (24)$$

where c_{kj} , $k, j = 1, \dots, n$, are the entries of the matrix

$$C = (L + \mathbf{1}_n \mathbf{1}_n^T)^{-1}. \quad (25)$$

Then, according to [8], information centrality is defined for the vertex $v_k \in \mathcal{V}(\mathcal{G})$ by the harmonic average⁶ of I_{kj}^{tot} ,

$$\kappa_{\text{info}}(v_k) = \left(\frac{1}{n} \sum_{j=1}^n \frac{1}{I_{kj}^{\text{tot}}} \right)^{-1}. \quad (26)$$

where I_{kj}^{tot} is computed based on Lemma 2 for each pair.

The relation between node certainty and information centrality is now established by the following theorem.

Theorem 1: Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ be an undirected connected graph of order n . Then, for $v_k \in \mathcal{V}(\mathcal{G})$

$$\frac{1}{\mu(v_k)} = \frac{\sigma^2}{2} \left(\frac{1}{\kappa_{\text{info}}(v_k)} - \frac{K_f}{n^2} \right), \quad (27)$$

where $\mu(v_k)$ and $\kappa_{\text{info}}(v_k)$ are the certainty index and the information centrality of v_k given by (15) and (26), respectively; K_f is the Kirchhoff index of \mathcal{G} defined by (17). Hence, if k_1, k_2, \dots, k_n are indices such that

$$\mu(v_{k_1}) \geq \mu(v_{k_2}) \geq \dots \geq \mu(v_{k_n}), \quad (28)$$

⁶An alternative definition provided in [17] uses the arithmetic average.

then,

$$\kappa_{\text{info}}(v_{k_1}) \geq \kappa_{\text{info}}(v_{k_2}) \geq \dots \geq \kappa_{\text{info}}(v_{k_n}), \quad (29)$$

and vice versa.

Before continuing with the proof of Theorem 1, which is given in Section IV-C below, the following remark is made.

Remark 5: In the case where the graph \mathcal{G} is an unweighted tree, for every pair of nodes $v_k, v_j \in \mathcal{G}$ there exists a unique path \mathcal{P}_{kj} connecting them. Then, (22) and (23) imply that the total information transmitted between v_k and v_j is equal to the inverse of the length $\ell(\mathcal{P}_{kj})$ of \mathcal{P}_{kj} . Hence, (26) reduces to (21), which explains why closeness centrality can be used to discriminate the nodes of an undirected unweighted tree such as the one used in Example 1.

C. Proof of Main Result

In this section, Theorem 1 is proved through a sequence of lemmas. We begin with a result that relates the node certainty index μ with the diagonal elements of the group inverse of the Laplacian. To fix terminology, recall that the group inverse of an $n \times n$ matrix P , when it exists, is the unique matrix X that satisfies:

$$(i) PXP = P, (ii) XPX = X, (iii) PX = XP; \quad (30)$$

see [18, Section 4.4] for details. In what follows, the group inverse of a matrix P is denoted by $P^\#$. Note that, if P is Hermitian, its group inverse corresponds to the standard Moore-Penrose pseudoinverse of P .

Lemma 3: Let \mathcal{G} be a connected undirected graph of order n with Laplacian matrix L . Then, the group inverse $L^\#$ of L exists and is unique. In particular,

$$L^\# = U_r (U_r^T L U_r)^{-1} U_r^T. \quad (31)$$

where the columns of the $n \times (n-1)$ matrix U_r contain the normalized eigenvectors of L corresponding to the nonzero eigenvalues λ_p , $p = 2, \dots, n$. Moreover,

$$\frac{1}{\mu(v_k)} = \frac{\sigma^2}{2} L_{kk}^\#. \quad (32)$$

Proof: Existence and uniqueness of $L^\#$ follows from [18, Thm. 1, p. 162], by the fact that $\lambda_1 = 0$ is a simple eigenvalue of L . Since L is symmetric, let $U = [u^{(1)} \mid U_r]$ with $Lu^{(1)} = 0$ be the orthogonal matrix that diagonalizes L . Then, $LU_r = U_r \Lambda_r$, where Λ_r is the diagonal matrix containing the nonzero eigenvalues of L . Using this fact, and the properties

$$U_r U_r^T = I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T, \quad (33)$$

$$U_r^T U_r = I_{n-1}, \quad (34)$$

it is straightforward to show that the matrix

$$X = U_r (U_r^T L U_r)^{-1} U_r^T$$

satisfies the requirements (30) for the group inverse. Hence, $L^\# = X$, and the (k, j) -th entry of $L^\#$ is

$$L_{kj}^\# = \sum_{p=2}^n \frac{1}{\lambda_p} u_k^{(p)} u_j^{(p)}. \quad (35)$$

The result (32) follows for $k = j$ in view of (15). ■

The following lemma collects some useful facts about $L^\#$.

Lemma 4: Let $L^\#$ be the group inverse of the Laplacian L of a connected undirected graph \mathcal{G} of order n . Then,

$$LL^\# = L^\#L = I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top, \quad (36)$$

$$\mathbf{1}_n^\top L^\# = L^\# \mathbf{1}_n = 0, \quad (37)$$

$$\text{Tr}(L^\#) = \frac{K_f}{n}, \quad (38)$$

where K_f is the Kirchhoff index of \mathcal{G} .

Proof: Equations (36) and (37) follow from (31) in view of (33)-(34) and $\mathbf{1}_n^\top U_r = U_r^\top \mathbf{1}_n = 0$. Equation (38) is a consequence of (18) of Remark 3 in view of (32). ■

The following lemma establishes a correspondence between the group inverse of the graph Laplacian and the inverse C of the matrix $L + \mathbf{1}_n\mathbf{1}_n^\top$ in Lemma 2.

Lemma 5: Let \mathcal{G} be an undirected connected graph of order n with Laplacian matrix L . Then,

$$C = (L + \mathbf{1}_n\mathbf{1}_n^\top)^{-1} = L^\# + \frac{1}{n^2}\mathbf{1}_n\mathbf{1}_n^\top. \quad (39)$$

Proof: The result follows from

$$\begin{aligned} (L + \mathbf{1}_n\mathbf{1}_n^\top) \left(L^\# + \frac{1}{n^2}\mathbf{1}_n\mathbf{1}_n^\top \right) &= LL^\# + \frac{1}{n^2}(L\mathbf{1}_n)\mathbf{1}_n^\top \\ &\quad + \mathbf{1}_n(\mathbf{1}_n^\top L^\#) + \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top \\ &= I_n, \end{aligned}$$

where Lemma 4 and the fact $\mathbf{1}_n^\top \mathbf{1}_n = n$ have been used. ■

Theorem 1 is now proved by combining the lemmas above.

Proof: [Theorem 1] The definition of information centrality (26) combined with (24) in Lemma 2 gives

$$\frac{1}{\kappa_{\text{info}}(v_k)} = \frac{1}{n} \sum_{j=1}^n (c_{kk} + c_{jj} - 2c_{kj}). \quad (40)$$

By Lemma 5 in view of (32) in Lemma 3 we have

$$\frac{1}{n} \sum_{j=1}^n c_{kk} = c_{kk} = L_{kk}^\# + \frac{1}{n^2} = \frac{2}{\sigma^2} \frac{1}{\mu(v_k)} + \frac{1}{n^2}. \quad (41)$$

In addition, from Lemmas 5 and 4 and (17) we have

$$\frac{1}{n} \sum_{j=1}^n c_{jj} = \frac{1}{n} \text{Tr}(L^\#) + \frac{1}{n^2} = \frac{1}{n^2} K_f + \frac{1}{n^2} \quad (42)$$

and

$$\frac{1}{n} \sum_{j=1}^n 2c_{kj} = \frac{2}{n} \left(\sum_{j=1}^n L_{kj}^\# + \frac{1}{n} \right) = \frac{2}{n^2}, \quad (43)$$

The result follows. ■

Remark 6: Lemmas 3 and 4 can be extended to strongly connected digraphs with normal Laplacian matrices. The proofs are somewhat lengthier and for reasons of space are not given here. It is only mentioned that in this case the node certainty index is related to the diagonal elements of the Laplacian of the *mirror* graph of the interconnection digraph; see [19, Def. 2] for a definition of the mirror graph.

V. CONCLUSION

In this paper we studied a collection of units, each represented by a DDM, which accumulate evidence in continuous time by observing a deterministic signal partially corrupted by noise. We focused on the impact of the interconnection topology on the certainty of each unit as it integrates information in the pursuit of a decision. A node certainty index which captures how the location of a unit affects the variance of its state was defined based on the eigenstructure of the graph Laplacian. It was intuitively discussed and rigorously proved that ranking the units according to their certainty index can be captured by their information centrality, a centrality measure that incorporates *all* possible paths connecting each unit with the rest of the network. These results show that evidence accumulation in collective decision making is a total network process, and may be used to identify units that can potentially serve as more powerful decision makers.

REFERENCES

- [1] P. L. Smith and R. Ratcliff, "Psychology and neurobiology of simple decisions," *Trends in Neurosciences*, vol. 27, no. 3, pp. 161–168, 2004.
- [2] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, "The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks," *Psychological Review*, vol. 113, no. 4, pp. 700–765, 2006.
- [3] R. Ratcliff, T. Van Zandt, and G. McKoon, "Connectionist and diffusion models of reaction time," *Psychological Review*, vol. 106, no. 2, pp. 261–300, Apr. 1999.
- [4] I. Gold, J. and N. Shadlen, M., "Neural computations that underlie decisions about sensory stimuli," *Trends in Cognitive Sciences*, vol. 5, no. 1, pp. 10–16, 2001.
- [5] D. Schall, J., "Neural basis of deciding, choosing, and acting," *Nature Reviews: Neuroscience*, vol. 2, pp. 33–42, Jan. 2001.
- [6] M. E. J. Newman, *Networks: An Introduction*, 1st ed. Oxford University Press, 2010.
- [7] L. C. Freeman, "Centrality in social networks: Conceptual clarification," *Social Networks*, vol. 1, pp. 215–239, 1978/1979.
- [8] K. Stephenson and M. Zelen, "Rethinking centrality: Methods and examples," *Social Networks*, vol. 11, no. 1, pp. 1–37, 1989.
- [9] S. H. Dandach, R. Carli, and F. Bullo, "Accuracy and decision time for sequential decision aggregation," *Proceedings of the IEEE*, vol. 100, no. 3, pp. 687–712, 2012.
- [10] L. Arnold, *Stochastic Differential Equations*. Wiley, 1974.
- [11] P. L. Smith, "From Poisson shot noise to the integrated Ornstein-Uhlenbeck process: Neurally principled models of information accumulation in decision-making and response time," *Journal of Mathematical Psychology*, vol. 54, pp. 266–283, 2010.
- [12] I. Poulakakis, L. Scardovi, and N. E. Leonard, "Coupled stochastic differential equations and collective decision making in the two-alternative forced-choice task," in *Proceedings of the American Control Conference*, Baltimore, MD, 2010, pp. 69–74.
- [13] T. Li and J.-F. Zhang, "Mean square average-consensus under measurement noises and fixed topologies: Necessary and sufficient conditions," *Automatica*, vol. 45, pp. 1929–1936, 2009.
- [14] G. F. Young, L. Scardovi, and N. E. Leonard, "Robustness of noisy consensus dynamics with directed communication," in *Proceedings of the American Control Conference*, Baltimore, MD, 2010, pp. 6312–6317.
- [15] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotics Networks*. Princeton University Press, 2009.
- [16] A. Ghosh, S. Boyd, and A. Saberi, "Minimizing effective resistance of a graph," *SIAM Review*, vol. 50, no. 1, pp. 37–66, 2008.
- [17] M. Altmann, "Reinterpreting network measures for models of disease transmission," *Social Networks*, vol. 1, no. 1, pp. 1–17, March 1993.
- [18] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, 2nd ed. Springer-Verlag, 2003.
- [19] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, Sept. 2004.