

# Coupled Stochastic Differential Equations and Collective Decision Making in the Two-Alternative Forced-Choice Task

Ioannis Poulakakis, Luca Scardovi and Naomi Ehrlich Leonard

**Abstract**—This paper investigates the effect of coupling in a collective decision-making scenario, in which the task is to correctly identify a (noisy) stimulus between two known alternatives. Multiple interconnected decision-making units, each represented by a Drift-Diffusion Model (DDM), accumulate evidence toward a decision. A number of different graph topologies among the DDM's are considered, and their effect on the accuracy of the decision is investigated. It is deduced that, for the same stimuli, the average of the collected evidence increases linearly with time toward the correct decision regardless of the communication topology. However, the uncertainty associated with the process is affected by the interconnection graph, implying that certain topologies are better than others.

## I. INTRODUCTION

Choosing between two alternatives represents a large class of real-world decision-making problems faced by humans and animals in their natural environments. *Two-Alternative Forced-Choice* (T AFC) tasks offer the prospect of a principled understanding of the dynamics of such decision-making behaviors. This can be achieved through the introduction of mathematical models amenable to tractable analysis, which, under reasonable assumptions, can faithfully describe and predict key aspects of T AFC tasks.

A wealth of behavioral data is available in humans performing simple T AFC tasks; see [9], [2] and references therein. Recently, direct recordings of neural firing patterns in visual and motor areas in primates performing such tasks permitted relating task performance to neuronal activity; see for instance [5], [10]. Both behavioral and neural data provide evidence supporting the *Drift-Diffusion Model* (DDM), and variants of it, as a plausible model for formally investigating the mechanisms governing simple T AFC tasks. More specifically, variations of the DDM have been employed to fit accuracy and reaction time in a variety of behavioral data in [9]. In addition, the DDM has been used successfully to describe neural firing rates in sensorimotor brain areas during T AFC tasks in [5] and [10], for example.

The DDM emerges in the relevant literature in a variety of ways. The work in [2] offers a unified framework, in which the DDM is put into perspective relative to behavioral and neurophysiological data obtained in simple T AFC tasks. It is shown that the DDM is equivalent—in an appropriate sense—to a continuum limit of the discrete Sequential

Probability Ratio Test (SPRT) widely employed in decision making. 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 neuronal populations, whose activities provide evidence for the two alternatives in a T AFC task.

In this paper, we depart from the pure DDM representing a single decision-making unit, to consider the more general setting of multiple such units interconnected according to particular communication topologies. Our purpose is to investigate the effect of coupling in enhancing the accuracy of decisions in simple T AFC tasks. The results presented here provide a first step toward a unified framework for studying collective decision making in biological and engineered systems. The framework adopted in this article differs from [8], which also investigates collective decision making, in that here coupling is introduced at the level of the DDM's. Finally, similar models to the ones used in this work are investigated in [7], albeit in multiple alternatives.

The structure of this paper is as follows. In Section II, the DDM as a model for individual decision making in T AFC tasks is first reviewed and then extended in a collective decision-making setting. Sections III and IV include results for a number of communication topologies, which are then compared in Section V in terms of their performance in the decision-making task considered.

## II. DRIFT-DIFFUSION MODELS FOR T AFC TASKS

The T AFC task—see [2] for an extensive account—is a canonical behavioral experiment, in which each trial involves correctly identifying a noisy stimulus drawn at random between two possibilities. In this section, the Drift-Diffusion Model (DDM), commonly used as a mathematical description of the phenomenology of such decision-making tasks, is reviewed and extended in a scenario where multiple decision-making units participate.

### A. Individual Decision Making in the T AFC Paradigm

The dynamics of decision making in the T AFC paradigm can be captured by a DDM under the assumption that the difference between the amounts of evidence supporting each alternative is integrated over each trial. A decision is then reached when, either a threshold is crossed by the accumulated evidence, or a fixed time has passed after stimulus onset. The latter corresponds to the *forced-response protocol*, in which the subjects are instructed to respond when a cue is presented; the former corresponds to the *free-response protocol*, in which the subjects respond in their own time.

This research was supported by AFOSR grant FA9550-07-1-0-0528.

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

L. Scardovi is with the Institute for Automatic Control Engineering, Technische Universität München, D-80290 Munich, Germany; e-mail: scardovi@tum.de.

In the pure DDM, the process starts with unbiased initial conditions and 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  $\beta$  represents increase in the evidence supporting the correct decision and  $\sigma dW$  are increments drawn from a Wiener process with standard deviation  $\sigma$ . 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 [1] and [2].

In this work, we focus on the TAFC task administered under the forced-response protocol, in which the process evolves until a pre-specified cue time,  $T_{\text{dec}}$ , is reached. The sign of  $x(T_{\text{dec}})$  determines the response. If  $\beta > 0$  (resp.  $\beta < 0$ ) and  $x(T_{\text{dec}}) > 0$  (resp.  $x(T_{\text{dec}}) < 0$ ) the correct choice is made. The opposite case corresponds to an error; see Fig. 1. The quality of the decision is measured by the *error rate* (ER), [2]; i.e., the probability that, at time  $T_{\text{dec}}$ , the individual picks the wrong decision. For  $\beta > 0$ ,

$$\text{ER} := \mathbb{P}[x(T_{\text{dec}}) < 0] = \int_{-\infty}^0 p(x, T_{\text{dec}}) dx. \quad (2)$$

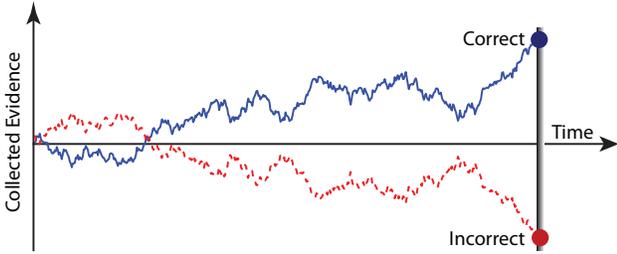


Fig. 1. The TAFC task under the forced-response protocol;  $\beta > 0$ . The vertical line corresponds to the cue time  $T_{\text{dec}}$ . One sample path (continuous line) results in the correct choice, the other (dashed line) does not.

### B. Collective Decision Making in the TAFC Paradigm

The objective of this study is to analyze the effect of interconnection in improving decision making in the TAFC task administered under the forced-response protocol. To this effect, we consider a generalized scenario in which  $n$  decision-making units are presented with the same stimulus, partially buried in noise, and are asked to correctly identify it between two alternatives. The units collect evidence according to the DDM (1) modified to include coupling among them according to particular communication topologies. We investigate a scenario in which the decision is made by the node having the least uncertainty in accumulating evidence.

Consider  $n$  copies of (1) coupled as follows

$$dx_k = \left[ \beta + \sum_{j=1}^n \alpha_{kj} (x_j - x_k) \right] dt + \sigma dW_k, \quad (3)$$

$k = 1, 2, \dots, n$ . In (3),  $\beta$  denotes the drift term (bias) leading to the correct decision, which is constant for all  $k$

to reflect the fact that all units are presented with the same stimulus, and  $\sigma dW_k$  are increments of independent Wiener processes with standard deviation  $\sigma$ . Finally,  $\alpha_{kj} \geq 0$  are the elements of an adjacency matrix corresponding to the coupling topology. The notation used here corresponds to a “sensing” convention, i.e. given a pair of nodes  $k, j$ ,  $\alpha_{kj} > 0$  implies the existence of a directed edge from  $k$  to  $j$  in the corresponding interconnection graph. In addition, we assume that there are no self-loops, i.e.  $\alpha_{kk} = 0$  for every  $k$ .

In matrix form (3) can be rewritten as

$$dx = [b - Lx]dt + C dW, \quad (4)$$

where  $x := \text{col}(x_1, \dots, x_n)$ ,  $b := \beta \mathbf{1}$ ,  $C := \sigma I$  and  $dW := \text{col}(dW_1, \dots, dW_n)$ . In (4),  $L$  is the graph Laplacian:

$$l_{kj} := \begin{cases} \sum_{i=1, i \neq k}^n \alpha_{ki}, & k = j, \\ -\alpha_{kj}, & k \neq j. \end{cases} \quad (5)$$

The following theorem characterizes the statistics of the random process  $\{x(t) : t \geq 0\}$  produced by (4) given deterministically zero initial conditions, i.e.  $\text{Cov}(x_0, x_0) = 0$  and  $\mathbb{E}[x_0] = 0$ , corresponding to unbiased decision making.

*Theorem 1:* Let  $x(0) = 0$  with probability 1. Then, the general solution of (4) is

$$x(t) = \int_0^t e^{-L(t-\tau)} b d\tau + \int_0^t e^{-L(t-\tau)} C dW, \quad (6)$$

in which the stochastic integral is interpreted in the Itô sense. In addition,

- 1) the mean and covariance of (6) are given by

$$\mathbb{E}[x(t)] = \int_0^t e^{-L(t-\tau)} b d\tau = (\beta t) \mathbf{1} \quad (7)$$

and

$$\text{Cov}(x(t), x(t)) = \sigma^2 \int_0^t e^{-L(t-\tau)} e^{-L^T(t-\tau)} d\tau, \quad (8)$$

respectively;

- 2) the stochastic process  $\{x(t) : t \geq 0\}$  is Gaussian.

The proof of Theorem 1 is a straightforward consequence of the results in [1, pp. 131–132]. It is only mentioned that (7) is a consequence of the fact that the exponential of the graph Laplacian is a stochastic matrix.

Before continuing with specializing to various graphs, the following important remarks are in order.

*Remark 1:* Theorem 1 implies that, for the simplified setting discussed here, the mean of the solution increases linearly with time *regardless* of the communication topology. However, the covariance, which represents the uncertainty of the process, *does* depend on the interconnection.

*Remark 2:* In general, not every node in the graph has the same uncertainty in accumulating evidence. This raises the question of how the decision will be made. In the setting adopted here, it is assumed that the decision made when the cue time is reached will be determined by a single

<sup>1</sup>Notation:  $\mathbf{1}$  is the column vector  $[1, \dots, 1]^T$  of appropriate dimension.

node, namely the one with the least uncertainty associated with its state. One of the objectives of the analysis in the following sections is to identify such node. This assumption, combined with the fact that  $x$  is a Gaussian random vector and  $\mathbb{E}[x_k(t)] = \beta t$  for all  $k$ , results in

$$\text{ER} := \frac{1}{2} \left[ 1 - \text{erf} \left( \frac{\beta t}{\sqrt{2 \min_k [\text{Var}(x_k(t))]} } \right) \right], \quad (9)$$

which corresponds to the (marginal) probability that the node with the least integrated variance selects the wrong choice. In (9),  $\text{erf}(\cdot)$  is the error function integral,

$$\text{erf}(x) := \frac{2}{\pi} \int_0^x e^{-u^2} du. \quad (10)$$

Defining the ER according to (9) implies that only the variance influences the error rate. Accordingly, we adopt the variance as a measure of performance in terms of accuracy.

### III. NORMAL GRAPHS

An important class of graphs, including all the undirected graphs, can be described by Laplacians that are normal; i.e., matrices that commute with their transpose, [6, Sec. 2.5]. In such cases, a simplified expression for the covariance matrix can be derived. Since  $L$  is normal, there exists a unitary matrix  $S$ , such that  $S^\dagger L S = \Lambda$ , where  $S^\dagger$  is the Hermitian transpose of  $S$  and  $\Lambda$  is a diagonal matrix containing the eigenvalues of  $L$ . Substitution in (8) results in

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

where

$$G(t) := \int_0^t \exp[-(\Lambda + \bar{\Lambda})(t - \tau)] d\tau. \quad (12)$$

Equation (11) can be used to derive an expression for the covariance matrix as the following lemma shows.

*Lemma 1:* Consider (4). Suppose  $L$  is normal and the underlying graph is connected<sup>2</sup>. Let  $v^{(p)}$  be the normalized eigenvector corresponding to the  $p$ -th eigenvalue  $\lambda_p$  of  $L$ . Then, the elements of the covariance matrix are given by

$$[\text{Cov}(x(t), x(t))]_{kj} = \sigma^2 \frac{t}{n} + \sigma^2 \sum_{p=2}^n \frac{1 - e^{-2\text{Re}(\lambda_p)t}}{2\text{Re}(\lambda_p)} v_k^{(p)} \bar{v}_j^{(p)}, \quad (13)$$

where  $\text{Re}(\lambda_p)$  denotes the real part of the eigenvalue  $\lambda_p$  and  $\bar{v}_k^{(p)}$  is the complex conjugate of the  $k$ -th component  $v_k^{(p)}$  of the  $p$ -th eigenvector.

*Proof:* Let  $S = [v^{(1)} | \dots | v^{(n)}]$ . Then, (11) gives

$$[\text{Cov}(x(t), x(t))]_{kj} = \sigma^2 \sum_{p=1}^n g_{pp}(t) v_k^{(p)} \bar{v}_j^{(p)}, \quad (14)$$

in which  $g_{pp}(t)$  denotes the  $p$ -th element of the diagonal matrix  $G(t)$  and is computed by (12) as

$$g_{pp}(t) = \frac{1 - e^{-2\text{Re}(\lambda_p)t}}{2\text{Re}(\lambda_p)}. \quad (15)$$

<sup>2</sup>Definition: The graph  $\mathcal{G}$  is *connected* if it contains a globally reachable node  $k$ , i.e., there exists a node  $k$  such that, for every node  $j$ , there exists a path in  $\mathcal{G}$  from  $j$  to  $k$ .

Since the graph is assumed to be connected,  $\mathbf{1}$  spans the kernel of  $L$ , i.e.  $L\mathbf{1} = 0$ , implying that  $v^{(1)} = (1/\sqrt{n})\mathbf{1}$  is the normalized eigenvector corresponding to the zero eigenvalue  $\lambda_1$ , thus resulting in (13). ■

*Remark 3:* According to Lemma 1, the variance associated with the state of each node is given by

$$\text{Var}(x_k(t)) = \sigma^2 \frac{t}{n} + \sigma^2 \sum_{p=2}^n \frac{1 - e^{-2\text{Re}(\lambda_p)t}}{2\text{Re}(\lambda_p)} |v_k^{(p)}|^2, \quad (16)$$

obtained from (13) for  $k = j$ . In view of the fact that  $\text{Re}(\lambda_p) > 0$ ,  $p \in \{2, \dots, n\}$ , for connected graphs, (16) implies that the uncertainty associated with the evidence collected by each node cannot be smaller than  $\sigma^2 t/n$ .

*Remark 4:* An interesting limiting case is obtained from (13) if, for finite  $t > 0$  as is the case in the forced-response protocol,  $\text{Re}(\lambda_p)$ ,  $p \in \{2, \dots, n\}$ , is very large, in the limit infinite. In this case, which corresponds to strong coupling among the nodes of the graph, the covariance matrix becomes

$$\mathbf{K}(t) = \sigma^2 \frac{t}{n} \mathbf{1}\mathbf{1}^\top. \quad (17)$$

It is noted that, as was mentioned in Remark 3,  $\sigma^2 t/n$  corresponds to the smallest achievable variance. Since  $\text{rank}\{\mathbf{1}\mathbf{1}^\top\} = 1$ , the dimension of the nullspace  $\mathcal{N}(\mathbf{K}(t))$  of the limiting covariance matrix is  $n - 1$  implying that the solution  $x$  is a singular random vector. Defining  $n - 1$  new variables by

$$y_k := x_k - \frac{1}{n-1} \sum_{j=1, j \neq k}^n x_j, \quad (18)$$

results in  $\mathbb{E}[y_k] = 0$  and  $\text{Var}(y_k) = 0$ , i.e.,  $y_k$  is a deterministic variable. In words, in this limiting case, the evidence collected by any arbitrary node of the graph is equal with probability 1 to the average of the evidence collected by all other nodes.

In the remainder of this section, (16) is particularized to special classes of connected normal graphs. This allows to compare the performance in terms of accuracy of the decision-making task described in Section II-B as a function of the communication topology.

#### A. Circulant Graphs

A particular class of normal graphs can be represented by Laplacian matrices that are circulant. Examples of circulant graphs include the complete graph (all-to-all communication), and the directed and undirected ring topologies. Let  $L_0$  be a Laplacian matrix that is also circulant and define  $L$  in (4) by

$$L = \alpha L_0, \quad (19)$$

where  $\alpha > 0$  is a parameter representing the strength of the communication links, and  $L_0$  is defined by its row elements  $\{d_0, d_1, \dots, d_{n-1}\}$  satisfying  $\sum_{j=0}^{n-1} d_j = 0$  through circulant shifts, see [4].

Many properties of circulant matrices can be derived in closed form; most notably, exact expressions for the eigenvalues and eigenvectors of such matrices are known, see

for instance [4]. Assuming a connected graph and applying Lemma 1 in view of results in [4], the following expression is derived for the diagonal elements of the covariance matrix

$$\text{Var}(x_k(t)) = \sigma^2 \frac{t}{n} + \sigma^2 \frac{1}{n} \sum_{p=2}^n \frac{1 - e^{-2\text{Re}(\lambda_p)t}}{2\text{Re}(\lambda_p)}, \quad (20)$$

where  $\text{Re}(\lambda_p)$  denotes the real part of the  $p$ -th eigenvalue

$$\lambda_p = \alpha \sum_{\ell=0}^{n-1} d_\ell e^{-i\frac{2\pi\ell(n+1-p)}{n}}. \quad (21)$$

of the Laplacian  $L$ .

It is readily seen from (20) that the variance associated with the state of each node is the same for all nodes. This observation implies that, for circulant graphs, all the nodes have the same uncertainty in collecting evidence toward the correct decision. Hence, any one of them could be used to make a decision. Furthermore, for finite values of  $t > 0$  and  $\alpha > 0$ , this uncertainty decreases as the cardinality of the graph increases. In fact, if the number of nodes  $n \geq 2$  is very large, the process becomes nearly deterministic.

*Remark 5:* The fact that the variance associated with the state of each node decreases with the number of nodes can be used to improve the precision of decision making. Given any specification  $\epsilon > 0$  for the precision and a fixed decision time  $t = T_{\text{dec}} > 0$ , (20) can be used to find the smallest number of nodes  $n$  that is required in order to have  $\text{ER} < \epsilon$ , with the error rate defined in Remark 2.

### B. Undirected Path Topology

In this section we consider the case of an undirected path graph, see Fig. 2. The corresponding Laplacian is

$$L = \begin{bmatrix} \alpha & -\alpha & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\alpha & 2\alpha & -\alpha & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\alpha & 2\alpha & -\alpha & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\alpha & 2\alpha & -\alpha \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\alpha & \alpha \end{bmatrix}, \quad (22)$$

which has the structure of a symmetric tridiagonal matrix.

Since symmetric matrices are normal, the results in Lemma 1 apply to connected undirected paths. As in the case of circulant matrices, closed-form expressions for the eigensystems of special classes of tridiagonal matrices, such as (22), can be found in the literature; see for instance [3]. As a result, the following expression for the diagonal elements of the covariance matrix can be derived

$$\text{Var}(x_k(t)) = \sigma^2 \frac{t}{n} + \sigma^2 \frac{2}{n} \sum_{p=2}^n \frac{1 - e^{-2\lambda_p t}}{2\lambda_p} \cos^2 \left[ \frac{\pi}{n} (p-1) \left( k - \frac{1}{2} \right) \right], \quad (23)$$

where

$$\lambda_p = 2\alpha \left( 1 - \cos \left[ \frac{\pi}{n} (p-1) \right] \right) \quad (24)$$

is the  $p$ -th eigenvalue of the Laplacian.

By way of contrast to the circulant graphs in Section III-A, the variances associated with the states of the nodes are *not* equal. From (23) it can be seen that nodes symmetrically located with respect to the mid-point of the path, i.e., the pairs  $(k, n-k+1)$  for  $k = 1, 2, \dots, n$  have the same variance. Moreover, the closer a node is to the midpoint of the path, the smallest is its variance. This situation is depicted in Fig. 2, which shows the variances of the nodes in an undirected path graph for  $n = 3$  and  $n = 6$ . It is therefore natural to restrict attention to nodes that are the closest to the midpoint of the path, i.e. node 2 for  $n = 3$  and either of the nodes 3 and 4 for  $n = 6$  as Fig. 2 suggests.

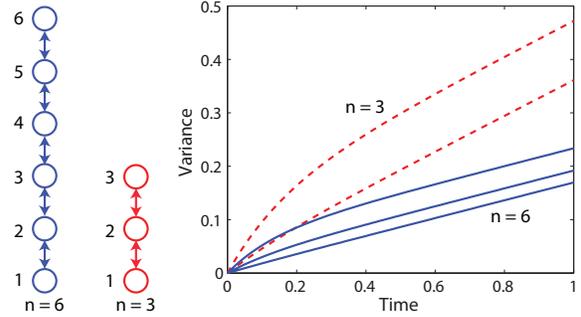


Fig. 2. Diagonal elements of the covariance matrix for the undirected path graphs on the left;  $n = 3$  (dashed) and  $n = 6$  (continuous);  $T_{\text{dec}} = 1$  s,  $\alpha = 1$  and  $\sigma = 1$ . The upper dashed line shows  $\text{Var}(x_1(t)) = \text{Var}(x_3(t))$ , the lower shows  $\text{Var}(x_2(t))$  (mid-point of the path). The upper continuous line shows  $\text{Var}(x_1(t)) = \text{Var}(x_6(t))$ , the middle shows  $\text{Var}(x_2(t)) = \text{Var}(x_5(t))$  and the lower shows  $\text{Var}(x_3(t)) = \text{Var}(x_4(t))$ .

### C. Undirected Star Topology

In this section, the communication topology depicted in Fig. 3(a) is considered. It corresponds to an undirected star graph, whose Laplacian is given by

$$L = \begin{bmatrix} (n-1)\alpha & -\alpha \mathbf{1}^T \\ -\alpha \mathbf{1} & \alpha I \end{bmatrix}. \quad (25)$$

The following expression for the covariance matrix can be derived using (8) through the explicit computation of powers of the Laplacian and substitution in the exponential series,

$$\text{Cov}(x(t), x(t)) = \sigma^2 \begin{bmatrix} c_1 & c_2 \mathbf{1}^T \\ c_2 \mathbf{1} & c_3 I + c_4 \mathbf{1}\mathbf{1}^T \end{bmatrix}, \quad (26)$$

where

$$\begin{aligned} c_1(t, n, \alpha) &= \frac{t}{n} + \frac{n-1}{2\alpha n^2} (1 - e^{-2nat}), \\ c_2(t, n, \alpha) &= \frac{t}{n} - \frac{1 - e^{-2nat}}{2n^2\alpha}, \\ c_3(t, \alpha) &= \frac{1 - e^{-2at}}{2\alpha}, \\ c_4(t, n, \alpha) &= \frac{t}{n} - \frac{1 - e^{-2at}}{2(n-1)\alpha} + \frac{1 - e^{-2nat}}{2n^2(n-1)\alpha}. \end{aligned} \quad (27)$$

It can be shown based on (27) that the variance associated with the center node is smaller than that of the rest of the

nodes, suggesting that it is reasonable to decide based on its state. Furthermore, from (27) for fixed  $t$  and  $\alpha$ , increasing the number  $n$  of nodes will decrease the variance of the center node to an arbitrarily small number—thus increasing precision as Remark 5 suggests—while the variance of the other nodes cannot become smaller than  $c_3$ . Finally, it should be mentioned that as time increases, the difference between the variances of the center and of the other nodes approaches a constant, which can be computed as

$$\lim_{t \rightarrow \infty} [\text{Var}(x_1(t)) - \text{Var}(x_k(t))] = -\left(1 - \frac{2}{n}\right) \frac{1}{2\alpha}, \quad (28)$$

for  $k = 2, \dots, n$ . This is consistent with what was mentioned in Remark 4; the larger the coupling strength  $\alpha$  the closer the two variances become, both approaching  $\sigma^2 t/n$ .

#### IV. EXAMPLES OF NON-NORMAL GRAPHS

It is of interest to consider examples of non-normal graphs. In particular, the disconnected exploding star and the imploding star graphs are studied below.

##### A. Exploding Star Topology

In this section, the exploding star graph depicted in Fig. 3(b) is considered. It is an example of a disconnected directed graph, and the corresponding Laplacian is

$$L = \begin{bmatrix} (n-1)\alpha & -\alpha \mathbf{1}^T \\ 0 & 0 \end{bmatrix}. \quad (29)$$

The relatively simple form of (29) allows for explicit computation of the covariance matrix, which is found to be

$$\text{Cov}(x(t), x(t)) = \sigma^2 \begin{bmatrix} c_1 & c_2 \mathbf{1}^T \\ c_2 \mathbf{1} & t I \end{bmatrix}, \quad (30)$$

where

$$\begin{aligned} c_1(t, n, \alpha) &= -\frac{2}{(n-1)^2\alpha} + \frac{t}{n-1} + \frac{2}{(n-1)^2\alpha} e^{-(n-1)\alpha t} \\ &\quad + \frac{n}{2(n-1)^2\alpha} \left(1 - e^{-2(n-1)\alpha t}\right), \\ c_2(t, n, \alpha) &= -\frac{1}{(n-1)^2\alpha} + \frac{t}{n-1} + \frac{1}{(n-1)^2\alpha} e^{-(n-1)\alpha t}. \end{aligned} \quad (31)$$

It is apparent from (31) that the variances associated with the state of each node are not equal; the variance of the center node, perceived as the *informed* node, is smaller than that of the rest, perceived as the *uninformed* nodes. Furthermore, for finite  $t$  and  $\alpha$ , the larger the number  $n$  of nodes the smaller  $c_1$  will be, implying that the uncertainty associated with the state of the informed node can be made arbitrarily small by increasing  $n$ , as was the case in normal graphs; see Remark 5. It is evident from this discussion that, in such an “informed/uninformed” decision hierarchy, it makes sense to make a decision based on the state of the informed node.

As a final remark, (31) implies that, for fixed  $t$  and  $n$ , in the limit as  $\alpha \rightarrow \infty$  the covariance matrix becomes

$$\mathbf{K}(t) = \sigma^2 \begin{bmatrix} \frac{t}{n-1} & \frac{t}{n-1} \mathbf{1}^T \\ \frac{t}{n-1} \mathbf{1} & t I \end{bmatrix}. \quad (32)$$

Notice in (32) that the relative role of “informed/uninformed” nodes is preserved as the coupling strength increases; the variance of the center node is  $\sigma^2 t/(n-1)$  and is smaller than the variance of all the other nodes. Note that the limiting covariance matrix is singular and  $\dim[\mathcal{N}(\mathbf{K}(t))] = 1$ . To provide insight into the singular nature of the random vector  $x(t)$ , consider the new variable

$$y = x_1 - \frac{1}{n-1} \sum_{j=2}^n x_j, \quad (33)$$

where  $x_1$  is the state of the informed node. Then,  $\mathbb{E}[y(t)] = 0$  and  $\text{Var}(y(t)) = 0$  meaning that  $y$  is deterministic. Hence, in this case, the evidence accumulated by the informed node is equal with probability 1 to the average of the evidence collected by all the other nodes. Essentially, the informed node “computes” the average of the collected evidence of the uninformed nodes.

##### B. Imploding Star Topology

In this section, a communication topology that is complementary to the exploding star graph is considered. This situation is depicted in Fig. 3(c) and corresponds to an imploding star directed graph. It is known, see for instance [11], that such graphs provide the fastest rate of convergence to the consensus state for linear consensus dynamics. The corresponding Laplacian matrix is

$$L = \begin{bmatrix} 0 & 0 \\ -\alpha \mathbf{1} & \alpha I \end{bmatrix}. \quad (34)$$

After some algebraic manipulations using (8) the elements of the covariance matrix are found to be

$$\text{Cov}(x(t), x(t)) = \sigma^2 \begin{bmatrix} t & c_1 \mathbf{1}^T \\ c_1 \mathbf{1} & c_2 \mathbf{1} \mathbf{1}^T + c_3 I \end{bmatrix}, \quad (35)$$

where

$$\begin{aligned} c_1(t, \alpha) &= t - \frac{1 - e^{-\alpha t}}{\alpha}, \\ c_2(t, \alpha) &= -\frac{3}{2\alpha} + t + \frac{2e^{-\alpha t}}{\alpha} - \frac{e^{-2\alpha t}}{2\alpha}, \\ c_3(t, \alpha) &= \frac{1}{2\alpha} - \frac{e^{-2\alpha t}}{2\alpha}. \end{aligned} \quad (36)$$

It is important to note that, contrary to all the graphs studied above, the coefficients  $c_1$ ,  $c_2$  and  $c_3$  in (36) are *independent* of the number  $n$  of the nodes. This should be expected based on the structure of the imploding star; since the information flows to the nodes from the center node but not in the other direction, the pairs formed by each node together with the center node are decoupled.

It is evident from (35) that the variance associated with the state of the center node corresponds to that of a single DDM, and is always larger than the variance associated with the state of each of the other nodes. The difference between  $\text{Var}(x_1(t))$  and  $\text{Var}(x_k(t))$  for  $k \in \{2, \dots, n\}$  as  $t$  grows eventually approaches the constant  $1/\alpha$ , which decreases as the coupling strength  $\alpha$  increases. Hence, for strong coupling, the variance associated with each node deteriorates, approaching  $\sigma^2 t$ , which is the variance of a single DDM.

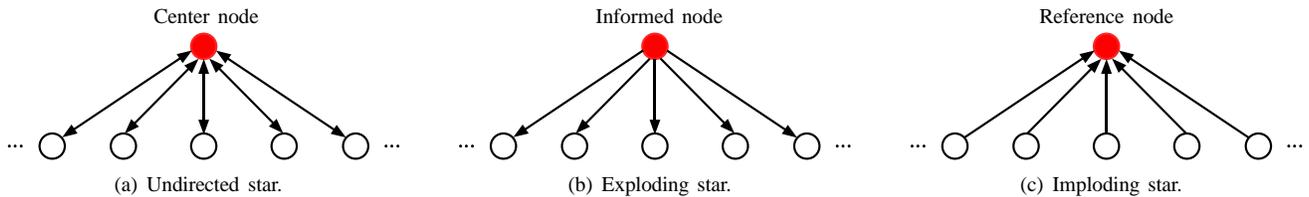


Fig. 3. Star graph topologies. The arrows follow the “sensing” convention: an arrow from  $k$  to  $j$  denotes that node  $k$  “sees” node  $j$ .

## V. A COMPARISON OF DIFFERENT GRAPHS

In the previous sections, the performance in terms of accuracy for a number of graphs in the decision-making scenario described in Section II-B has been analyzed based on the relative significance of the nodes in collecting evidence, their number and the coupling strength among them.

It was observed that the variance associated with the state of each node is not constant among nodes, with the notable exception of circulant graphs. Furthermore, in the normal graphs examined in Section III, increasing the cardinality of the graph decreases the uncertainty associated with all the nodes. In the exploding star graph however, only the variance associated with the state of the center node is decreased, while, in the imploding star, the variances do not depend on the number of nodes. Finally, in the limiting case of strong communication, it was deduced that the variance associated with the states of all the nodes closely approximates the least achievable variance  $\sigma^2 t/n$  in all normal graphs. This is not the case for the non-normal graphs studied in Section IV: in the exploding star, only the variance of the center node approximates  $\sigma^2 t/(n-1)$ , the rest remain unchanged, while in the imploding star all the variances approach the variance of the solution of a single DDM, which is the worst case.

Let the “best” node in each graph be the one with the least variance integrated over time. Fig. 4 compares the various graphs studied above in terms of the time evolution of the variance associated with their best node for  $n = 10$ . As shown, every graph is better than a single DDM, even the imploding star, in which the covariance matrix is independent of the number of nodes. Interestingly, Fig. 4 shows that the disconnected exploding star nearly achieves the performance of the complete graph, albeit with fewer communication links. An issue arises, however, with respect to the robustness under node failure in this graph: if the center node fails, then the graph behaves as  $n - 1$  disconnected DDM’s. Such case would not be present, for instance, in complete graphs, suggesting the consideration of robustness measures together with accuracy in assessing the performance of each graph. This issue will be addressed in a future publication.

## VI. CONCLUSION

Motivated by a vast amount of behavioral and neurophysiological data suggesting the DDM as a model that captures the dynamics of decision making in T AFC tasks, we proposed and analyzed a collective scenario in which multiple units, coupled according to particular communication topologies, collect information toward the correct decision. The performance of each graph in enhancing the accuracy of

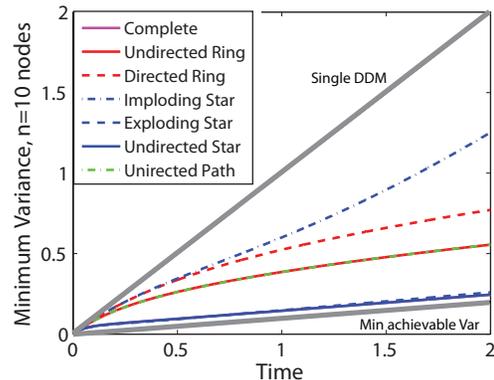


Fig. 4. Minimum variance for the graphs of Sections III and IV. The grey lines correspond to the variance of a single DDM ( $\sigma^2 t$ ) and the minimum achievable variance ( $\sigma^2 t/n$ ); all graphs are between these lines with the complete, undirected star and exploding star graphs approaching  $\sigma^2 t/n$ .

decision making was discussed. This paper should be viewed as a first step toward a common framework for studying collective decision making in T AFC tasks in groups involving biological and engineered systems.

## ACKNOWLEDGMENT

The authors wish to thank P. Holmes for useful discussions.

## REFERENCES

- [1] L. Arnold, *Stochastic Differential Equations*. Wiley, 1974.
- [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] H.-W. Chang, S.-E. Liu, and R. Burridge, “Exact eigensystems for some matrices arising from discretizations,” *Linear Algebra and its Applications*, vol. 430, pp. 999–1006, 2009.
- [4] P. J. Davis, *Circulant Matrices*. New York: Wiley, 1979.
- [5] J. I. Gold and M. N. Shadlen, “Neural computations that underlie decisions about sensory stimuli,” *Trends in Cognitive Sciences*, vol. 5, no. 1, pp. 10–16, 2001.
- [6] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
- [7] T. McMillen and P. Holmes, “The dynamics of choice among multiple alternatives,” *Journal of Mathematical Psychology*, vol. 50, pp. 30–57, 2006.
- [8] A. Nedic, D. Tomlin, P. Holmes, D. A. Prentice, and J. D. Cohen, “A simple decision task in a social context: Experiments, a model, and preliminary analyses of behavioral data,” in *Proceedings of the IEEE Conference on Decision and Control*, 2008.
- [9] 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.
- [10] J. D. Schall, “Neural basis of deciding, choosing, and acting,” *Nature Reviews: Neuroscience*, vol. 2, pp. 33–42, Jan. 2001.
- [11] C. W. Wu, *Synchronization in Complex Networks of Nonlinear Dynamical Systems*. World Scientific, 2008.