On first passage time problems in collective decision-making with heterogeneous agents

Vaibhav Srivastava Naomi Ehrich Leonard

***Abstract*— We study first passage time problems in collective decision-making using the context of two alternative choice tasks. The properties of the first passage time of a high dimensional stochastic process are hard to compute. For a class of stochastic processes governed by coupled linear stochastic differential equations, we develop reduced order models that are amenable to efficient computation of the properties of the first passage time. We use the proposed reduced model to study collective decision-making in heterogeneous cooperative networks and leader-follower networks.**

I. Introduction

Collective decision-making in animal and human groups has received significant interest in a broad scientific community [1], [2], [3]. In situations where individuals in groups are free to make decisions whenever they have collected enough information, the decision-making process often corresponds to the *first passage time* of a stochastic process. The first passage time is the earliest time when the trajectory of a stochastic process starting within a region of interest hits the boundary of the region.

The first passage time problems are ubiquitous in a wide range of areas including neuroscience [4], biology [5], [6], [7], finance [8], ecology [9], and engineering [10]. For stochastic processes governed by stochastic differential equations (SDEs), properties of first passage times, e.g., their expected values and probability distribution functions, can be characterized in terms of second order partial differential equations (PDEs). However, the number of variables in these PDEs is equal to the dimension of the state of the SDE, which makes the computation of the solution of these PDEs intractable for high-dimensional state-spaces.

In this paper, we study first passage time problems in collective decision-making using the context of two alternative choice tasks. The number of variables in such problems is of the order of the number of individuals in the group, and thus, the computation of the properties of first passage times becomes intractable for large groups. To ameliorate these issues, we develop a generic model reduction technique and apply it to collective decision-making problems.

We particularly focus on decision-making in human groups. The collective decision-making in human groups is typically studied under two extreme communication regimes: the so-called *ideal group* and the *Condorcet group*. In an ideal group, each decision-maker interacts with every other decision-maker in the group and the group arrives at a consensus decision. In a Condorcet group, decision-makers do not interact with one another; instead a majority rule is employed to reach a decision. Collective decision-making in ideal human groups and Condorcet human groups is studied in [3] using the classical signal detection model for human performance in two alternative choice tasks.

Human decision-making in two alternative choice tasks is extensively studied and well understood [4], [11], [12]. In particular, human decision-making in a two alternative choice task is modeled well by the *drift-diffusion model* (DDM). Human decision-making is typically studied under two paradigms, namely, *interrogation*, and *free response*. In the interrogation paradigm, the human has to make a decision at the end of a prescribed time duration, while in the free response paradigm, the human takes their time to make a decision. The free response paradigm in the DDM correspond to a first passage problem of the Wiener process.

Collective decision-making in Condorcet human groups using the DDM and free response paradigm is studied in [13], [14]. Collective decision-making in ideal cooperative networks and ideal leader-follower networks using the DDM and interrogation paradigm are studied in [15] and [16], respectively. Collective decision-making in ideal cooperative networks with identical agents using the DDM and free-response paradigm is studied in [17]. In this paper, we extend the results in [17] to cooperative networks with heterogeneous agents, and leader-follower networks.

The DDM is a continuous time approximation to the evidence aggregation process in a hypothesis testing problem. Moreover, the finite sample and the sequential hypothesis testing problems correspond to the interrogation and the free response paradigm in human decision-making, respectively. Consequently, the collective decision-making problem in human groups is similar to distributed hypothesis testing problems studied in the engineering literature [18], [19], [20]. In contrast to existing results in the distributed hypothesis testing literature, which focus on asymptotic performance, we study the finite time collective decision-making performance.

The major contributions of this paper are twofold. First, we consider a class of coupled linear stochastic differential equations and associated first passage time problems. We develop reduced order models for these coupled SDEs that are amenable to efficient computation of first passage time properties.

Second, we study collective decision-making in cooper-

This research has been supported in part by ONR grants N00014-09-1-1074, N00014-14-1-0635, ARO grant W911NF-14-1-0431 and NSF grant ECCS-1135724.

V. Srivastava and N. E. Leonard are with the Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, USA, {vaibhavs, naomi}@princeton.edu.

ative networks with heterogeneous agents and in leader-follower networks. In both contexts, we use the reduced order models to analyze the performance of collective decision-making in terms of speed and accuracy.

The remainder of the paper is organized in the following way. In Section II, we recall fundamentals of the first passage time problems. In Section III, we study a class of coupled linear SDEs and develop decoupled approximations to it. In Section IV and V, we apply the developed techniques to collective decision-making in cooperative networks, and leader-follower networks, respectively. Finally, we conclude in Section VI.

II. First Passage Time Problems

In this section, we briefly review the fundamentals of first passage time problems. We refer the interested reader to [21] for more details. Consider the following stochastic differential equation

$$\mathrm{d}\boldsymbol{x}(t) = \boldsymbol{a}(\boldsymbol{x})\mathrm{d}t + \boldsymbol{b}(\boldsymbol{x})\mathrm{d}\boldsymbol{W}_m(t), \tag{1}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$, $\boldsymbol{a} : \mathbb{R}^n \to \mathbb{R}^n$, $\boldsymbol{b} : \mathbb{R}^n \to \mathbb{R}^{n\times m}$, and $\boldsymbol{W}_m(t)$ is the m-dimensional Wiener process. Let $B(\boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x})\boldsymbol{b}(\boldsymbol{x})^\top \in \mathbb{R}^{n\times n}$.

The first passage time [21] is the earliest time at which a trajectory $\boldsymbol{x}(t)$, initially inside a region $\mathcal{R}$ with boundary $\mathcal{S}$, leaves the region. Let $T \in \mathbb{R}_{>0}$ be the first passage time. Let $G(\boldsymbol{x}_0, t)$ be the probability that a trajectory starting from $\boldsymbol{x}_0$ stays within $\mathcal{R}$ until time t, i.e., $G(\boldsymbol{x}_0, t) = \mathbb{P}(T \geq t|\boldsymbol{x}(0) = \boldsymbol{x}_0)$. Then, G obeys the backward Fokker-Planck equation

$$\frac{\partial G}{\partial t} = \sum_{i=1}^{n} a_i(\boldsymbol{x}_0)\frac{\partial G}{\partial x_{0i}} + \frac{1}{2}\sum_{i,j=1}^{n} B_{ij}(\boldsymbol{x}_0)\frac{\partial^2 G}{\partial x_{0i}\partial x_{0j}},$$

with the initial condition

$$G(\boldsymbol{x}_0, 0) = \begin{cases} 1, & \boldsymbol{x}_0 \in \mathcal{R}, \\ 0, & \text{elsewhere}, \end{cases}$$

and the boundary condition $G(\boldsymbol{x}_0, t) = 0$, for each $\boldsymbol{x}_0 \in \mathcal{S}$.

The mean first passage time $\mathrm{ET}(\boldsymbol{x}_0)$ satisfies

$$\sum_{i=1}^{n} a_i(\boldsymbol{x}_0)\frac{\partial \mathrm{ET}}{\partial x_{0i}} + \frac{1}{2}\sum_{i,j=1}^{n} B_{ij}(\boldsymbol{x}_0)\frac{\partial^2 \mathrm{ET}}{\partial x_{0i}\partial x_{0j}} = -1,$$

with boundary condition $\mathrm{ET}(\boldsymbol{x}_0) = 0$, for each $\boldsymbol{x}_0 \in \mathcal{S}$.

Let $\pi(\boldsymbol{a}, \boldsymbol{x}_0)$ be the probability of the exit through an element $d\mathcal{S}(\boldsymbol{a})$ centered at point $\boldsymbol{a}$ on the boundary $\mathcal{S}$. Then, $\pi(\boldsymbol{a}, \boldsymbol{x}_0)$ satisfies

$$\sum_{i=1}^{n} a_i(\boldsymbol{x}_0)\frac{\partial \pi}{\partial x_{0i}} + \frac{1}{2}\sum_{i,j=1}^{n} B_{ij}(\boldsymbol{x}_0)\frac{\partial^2 \pi}{\partial x_{0i}\partial x_{0j}} = 0,$$

with boundary condition $\pi(\boldsymbol{a}, \boldsymbol{x}_0) = \delta_s(\boldsymbol{a} - \boldsymbol{x})$, where δ_s is the surface delta function for the boundary $\mathcal{S}$.

The above PDEs are of second order in an n-dimensional space. For large n, such PDEs are hard to solve even numerically. In the following, we consider a class of coupled linear SDEs and semi-infinite rectangular cuboid regions $\mathcal{R}$, and develop tractable approximations to the above PDEs.

III. Coupled Drift Diffusion Equations

Consider the following coupled drift-diffusion equation (DDE)

$$\mathrm{d}\boldsymbol{x}(t) = -A\boldsymbol{x}(t)\mathrm{d}t + \boldsymbol{\beta}\mathrm{d}t + \Sigma\mathrm{d}\boldsymbol{W}_n(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0. \tag{2}$$

Suppose that the matrix A is diagonalizable and has eigenvalues with non-negative real parts. Suppose the eigenvalue with the smallest real part is purely real. Let $\boldsymbol{u}_p$ and $\boldsymbol{u}_p^\dagger$ be right and left eigenvectors of A associated with eigenvalue λ_p, $p \in \{1, \ldots, n\}$ such that $A = \sum_{p=1}^{n} \lambda_p \boldsymbol{u}_p \boldsymbol{u}_p^\dagger$. We assume that eigenvalues are ordered in increasing order of their real parts.

We are interested in the properties of the first passage time of the stochastic process $\{\boldsymbol{x}(t)\}_{t\in\mathbb{R}_{\geq 0}}$ from regions of the form $x_k(t) \in [-\eta_k, \eta_k]$, for some $k \in \{1, \ldots, n\}$, and some $\eta_k \in \mathbb{R}_{>0}$. In the following, we first develop decoupled approximations to (2), and then use them to determine properties of the first passage time.

A. Decomposition of coupled DDE

We now decompose the coupled DDE (2) into principal and residual components. We define the principal component of $\boldsymbol{x}(t)$ as $\boldsymbol{x}_{\text{prin}}(t) = x_p\boldsymbol{u}_1$, where $x_p = \boldsymbol{u}_1^\dagger\boldsymbol{x}(t)$. We define the residual component as $\boldsymbol{\epsilon}(t) = \boldsymbol{x}(t) - \boldsymbol{x}_{\text{prin}}(t)$.

It follows that the principal component satisfies

$$\begin{aligned}\mathrm{d}x_p(t) &= -\boldsymbol{u}_1^\dagger A\boldsymbol{x}(t) + \boldsymbol{u}_1^\dagger\boldsymbol{\beta}\mathrm{d}t + \boldsymbol{u}_1^\dagger\Sigma\mathrm{d}\boldsymbol{W}_n(t) \\ &= -\lambda_1 x_p(t)\mathrm{d}t + \beta_p\mathrm{d}t + \sigma_p\mathrm{d}W_1(t),\end{aligned} \tag{3}$$

where $\beta_p = \boldsymbol{u}_1^\dagger\boldsymbol{\beta}$, $\sigma_p = \|\boldsymbol{u}_1^\dagger\Sigma\|_2$, and $W_1(t)$ is the standard one dimensional Weiner process. Thus, the principal component evolves according to an Ornstein-Uhlenbeck (O-U) process [21].

Similarly, the residual component satisfies

$$\begin{aligned}\mathrm{d}\boldsymbol{\epsilon}(t) &= -A\boldsymbol{x}(t)\mathrm{d}t + \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^\dagger\boldsymbol{x}(t) + \bar{\boldsymbol{\beta}}\mathrm{d}t + \bar{\Sigma}\mathrm{d}\boldsymbol{W}_n(t) \\ &= -A\boldsymbol{x}(t)\mathrm{d}t + A\boldsymbol{u}_1\boldsymbol{u}_1^\dagger\boldsymbol{x}(t) + \bar{\boldsymbol{\beta}}\mathrm{d}t + \bar{\Sigma}\mathrm{d}\boldsymbol{W}_n(t) \\ &= -A\boldsymbol{\epsilon}(t)\mathrm{d}t + \bar{\boldsymbol{\beta}}\mathrm{d}t + \bar{\Sigma}\mathrm{d}\boldsymbol{W}_n(t),\end{aligned} \tag{4}$$

where $\bar{\boldsymbol{\beta}} = \boldsymbol{\beta} - \boldsymbol{u}_1^\dagger\boldsymbol{\beta}\boldsymbol{u}_1$ and $\bar{\Sigma} = (I_n - \boldsymbol{u}_1\boldsymbol{u}_1^\dagger)\Sigma$.

B. Properties of components of coupled DDE

We now characterize properties of two components of the coupled DDE, namely, the principal component and the residual component.

*Proposition 1 (**Properties of components**):* The following statements hold for the principal component $x_p(t)$ and the residual component $\boldsymbol{\epsilon}(t)$ defined in equations (3) and (4):

(i) the expected value and variance of principal component x_p satisfy

$$\mathbb{E}[x_p(t)] = \lim_{\lambda\to\lambda_1^+} \frac{\beta_p}{\lambda}(1 - e^{-\lambda t}), \text{ and}$$

$$\mathrm{Var}[x_p(t)] = \lim_{\lambda\to\lambda_1^+} \frac{\sigma_p^2}{2\lambda}(1 - e^{-2\lambda t});$$

(ii) the expected value and covariance of residual component $\boldsymbol{\epsilon}(t)$ satisfy

$$\lim_{t\to+\infty}\mathbb{E}[\boldsymbol{\epsilon}(t)] = \sum_{p=2}^{n}\frac{1}{\lambda_p}\boldsymbol{u}_p\boldsymbol{u}_p^\dagger\boldsymbol{\beta}, \text{ and}$$
$$\lim_{t\to+\infty}\mathrm{Cov}[\boldsymbol{\epsilon}(t)] = \sum_{q=2}^{n}\sum_{r=2}^{n}\frac{\boldsymbol{u}_q^\dagger\Sigma\Sigma^\top\boldsymbol{u}_r^\ddagger}{\lambda_q+\lambda_r}\boldsymbol{u}_q\boldsymbol{u}_r^\top,$$

where $\boldsymbol{u}_q^\ddagger$ is the transpose of $\boldsymbol{u}_q^\dagger$;

(iii) the covariance between the principal component and the residual component satisfies

$$\lim_{t\to+\infty}\mathrm{Cov}(x_p(t),\boldsymbol{\epsilon}(t)) = \sum_{q=2}^{n}\frac{\boldsymbol{u}_q^\dagger\Sigma\Sigma^\top\boldsymbol{u}_1^\ddagger}{\lambda_1+\lambda_q}\boldsymbol{u}_q.$$

Proof: Consider the following augmented coupled dynamics:

$$\begin{bmatrix}\mathrm{d}x_p(t)\\ \mathrm{d}\boldsymbol{\epsilon}(t)\end{bmatrix} = \begin{bmatrix}-\lambda_1 & 0\\ 0 & -A\end{bmatrix}\begin{bmatrix}x_p(t)\\ \boldsymbol{\epsilon}_k(t)\end{bmatrix}\mathrm{d}t + \begin{bmatrix}\beta_p\\ \boldsymbol{\beta}\end{bmatrix}\mathrm{d}t + \begin{bmatrix}\boldsymbol{u}_1^\dagger\Sigma\\ (I_n-\boldsymbol{u}_1\boldsymbol{u}_1^\dagger)\Sigma\end{bmatrix}\mathrm{d}\boldsymbol{W}_n(t). \quad (5)$$

Since, equation (5) is a linear SDE, it follows

$$\mathbb{E}\begin{bmatrix}x_p(t)\\ \boldsymbol{\epsilon}(t)\end{bmatrix} = \begin{bmatrix}\int_{\tau=0}^{t}e^{-\lambda_1\tau}\beta_p\mathrm{d}\tau\\ \int_{\tau=0}^{t}e^{-A\tau}(I_n-\boldsymbol{u}_1\boldsymbol{u}_1^\dagger)\boldsymbol{\beta}\mathrm{d}\tau\end{bmatrix},$$

and the expressions for the expected values follow immediately.

Similarly,

$$\mathrm{Cov}\begin{bmatrix}x_p(t)\\ \boldsymbol{\epsilon}(t)\end{bmatrix} = \int_{\tau=0}^{t}\begin{bmatrix}e^{-\lambda_1\tau} & 0\\ 0 & e^{-A\tau}\end{bmatrix}\begin{bmatrix}\boldsymbol{u}_1^\dagger\Sigma\\ (I_n-\boldsymbol{u}_1\boldsymbol{u}_1^\dagger)\Sigma\end{bmatrix}\begin{bmatrix}\Sigma^\top\boldsymbol{u}_1^\ddagger & \Sigma^\top(I_n-\boldsymbol{u}_1^\ddagger\boldsymbol{u}_1^\top)\end{bmatrix}\begin{bmatrix}e^{-\lambda_1\tau} & 0\\ 0 & e^{-A^\top\tau}\end{bmatrix}\mathrm{d}\tau.$$

Simplifying the above expression, we obtain

$$\mathrm{Var}(x_p(t)) = \int_{\tau=0}^{t}e^{-2\lambda_1\tau}\sigma_p^2\mathrm{d}\tau = \lim_{\lambda\to\lambda_1^+}\frac{\sigma_p^2}{2\lambda}(1-e^{-2\lambda t}),$$

and

$$\begin{aligned}&\mathrm{Cov}(\boldsymbol{\epsilon}(t))\\ &= \int_{\tau=0}^{t}e^{-A\tau}(I_n-\boldsymbol{u}_1\boldsymbol{u}_1^\dagger)\Sigma\Sigma^\top(I_n-\boldsymbol{u}_1^\ddagger\boldsymbol{u}_1^\top)e^{-A^\top\tau}\mathrm{d}\tau\\ &= \int_{\tau=0}^{t}\Big(\sum_{q=2}^{n}e^{-\lambda_q\tau}\boldsymbol{u}_q\boldsymbol{u}_q^\dagger\Big)\Sigma\Sigma^\top\Big(\sum_{r=2}^{n}e^{-\lambda_r\tau}\boldsymbol{u}_r^\ddagger\boldsymbol{u}_r^\top\Big)\mathrm{d}\tau\\ &= \sum_{q=2}^{n}\sum_{r=2}^{n}\frac{1-e^{-(\lambda_q+\lambda_r)t}}{\lambda_q+\lambda_r}\boldsymbol{u}_q^\dagger\Sigma\Sigma^\top\boldsymbol{u}_r^\ddagger\boldsymbol{u}_q\boldsymbol{u}_r^\top.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{Cov}(x_p(t),\boldsymbol{\epsilon}(t)) &= \int_0^t e^{-A\tau}(I_n-\boldsymbol{u}_1\boldsymbol{u}_1^\dagger)\Sigma\Sigma^\top\boldsymbol{u}_1^\ddagger e^{-\lambda_1\tau}\mathrm{d}\tau\\ &= \sum_{q=2}^{n}\frac{1-e^{-(\lambda_1+\lambda_q)t}}{\lambda_1+\lambda_q}\boldsymbol{u}_q\boldsymbol{u}_q^\dagger\Sigma\Sigma^\top\boldsymbol{u}_1^\ddagger.\end{aligned}$$

This completes the proof. ∎

C. Decoupled approximation to coupled DDE

In this section, we develop a decoupled approximation to the coupled DDE (2). The principal component is already decoupled. In particular, the k-th principal component is $x_k^{\text{prin}}(t) = x_p(t)u_{1k}$, where u_{1k} is the k-th element of $\boldsymbol{u}_1$, and can be modeled by the O-U process (3).

We now focus on the residual component $\boldsymbol{\epsilon}(t)$. Since $\boldsymbol{\epsilon}(t)$ is the solution of the linear SDE (4) with constant coefficients and a constant input, it is a continuous Gaussian process. The asymptotic values of expected value and variance of $\boldsymbol{\epsilon}(t)$, derived in Proposition 1, are constants. A simple, scalar, and continuous Gaussian process with a constant asymptotic expected value and a constant asymptotic variance is the O-U process. In the following, we approximate the k-th element of the residual component $\epsilon_k(t)$ by an O-U process. Such an approximate O-U process must also capture the correlation between $x_p(t)$ and $\epsilon_k(t)$. In the following we propose such an approximation $\varepsilon_k(t)$ of $\epsilon_k(t)$.

We first introduce some notation. Let

$$\frac{1}{\mu_k} := \lim_{t\to+\infty}\mathrm{Var}(\epsilon_k(t)),\ \alpha_k := \lim_{t\to+\infty}\mathbb{E}[\epsilon_k(t)], \text{ and}$$
$$\gamma_k := \lim_{t\to+\infty}\mathrm{Cov}(x_p(t),\epsilon_k(t)).$$

We now propose the following coupled dynamics for $x_k^{\text{prin}}(t)$ and $\varepsilon_k(t)$

$$\begin{bmatrix}\mathrm{d}x_k^{\text{prin}}(t)\\ \mathrm{d}\varepsilon_k(t)\end{bmatrix} = \begin{bmatrix}-\lambda_1 & 0\\ 0 & -\frac{\mu_k}{2}\end{bmatrix}\begin{bmatrix}x_k^{\text{prin}}(t)\\ \varepsilon_k(t)\end{bmatrix}\mathrm{d}t + \begin{bmatrix}\beta_p u_{1k}\\ \frac{\alpha_k\mu_k}{2}\end{bmatrix}\mathrm{d}t + \begin{bmatrix}\sigma_p u_{1k} & 0\\ \frac{(2\lambda_1+\mu_k)\gamma_k}{2\sigma_p} & \sqrt{1-\frac{(2\lambda_1+\mu_k)^2\gamma_k^2}{4\sigma_p^2}}\end{bmatrix}\begin{bmatrix}\mathrm{d}W_1(t)\\ \mathrm{d}W_2(t)\end{bmatrix}. \quad (6)$$

Equation (6) can be solved to show that it asymptotically matches the metrics for residual component (4) established in Proposition 1. We now state this fact without proof.

*Proposition 2 (**Asymptotic Matching**):* The following statements hold for the dynamics (6):

(i) the asymptotic expected value of $\varepsilon_k(t)$ is

$$\lim_{t\to+\infty}\mathbb{E}[\varepsilon_k(t)] = \alpha_k;$$

(ii) the asymptotic variance of $\varepsilon_k(t)$ is

$$\lim_{t\to+\infty}\mathrm{Var}(\varepsilon_k(t)) = \frac{1}{\mu_k};$$

(iii) the asymptotic covariance between $x_k^{\text{prin}}(t)$ and $\varepsilon_k(t)$ is

$$\lim_{t\to+\infty}\mathrm{Cov}(x_k^{\text{prin}}(t),\varepsilon_k(t)) = u_{1k}\gamma_k.$$

Notice that if A is a symmetric matrix and Σ is an identity matrix, then the decomposition in (3) and (4) corresponds to principal components of the covariance of $\boldsymbol{x}(t)$. In such a case, $x_k^{\text{prin}}(t)$ and $\varepsilon_k(t)$ are uncorrelated and the diffusion matrix in (6) is a diagonal matrix.

The approximate dynamics (6) can be used to obtain the dynamics of the approximate state $y_k(t) = x_k^{\text{prin}}(t) + \varepsilon_k(t)$.

$$\begin{bmatrix} \mathrm{d}y_k(t) \\ \mathrm{d}\varepsilon_k(t) \end{bmatrix} = \begin{bmatrix} -\lambda_1 & \lambda_1 - \frac{\mu_k}{2} \\ 0 & -\frac{\mu_k}{2} \end{bmatrix} \begin{bmatrix} y_k(t) \\ \varepsilon_k(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \beta_p u_{1k} + \frac{\alpha_k \mu_k}{2} \\ \frac{\alpha_k \mu_k}{2} \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \sigma_p u_{1k} + \frac{(2\lambda_1 + \mu_k)\gamma_k}{2\sigma_p} & \sqrt{1 - \frac{(2\lambda_1+\mu_k)^2\gamma_k^2}{4\sigma_p^2}} \\ \frac{(2\lambda_1+\mu_k)\gamma_k}{2\sigma_p} & \sqrt{1 - \frac{(2\lambda_1+\mu_k)^2\gamma_k^2}{4\sigma_p^2}} \end{bmatrix} \begin{bmatrix} \mathrm{d}W_1(t) \\ \mathrm{d}W_2(t) \end{bmatrix}. \quad (7)$$

Now, the PDEs in Section II can be used with (7) to determine approximate first passage time properties. Note that the PDEs corresponding to (7) evolve in a two dimensional space, and can be solved numerically in an efficient manner.

IV. Decision-making in cooperative networks

In this section, we study collective decision-making in cooperative networks using the context of two-alternative choice tasks. In cooperative networks, agents exchange information with one another, and each agent makes its own decision from its collected information. In the following, we first describe the model for evidence aggregation in collective decision-making, and then develop a decoupled approximation to this model. Finally, we numerically investigate the accuracy of the decoupled approximation.

A. Cooperative coupled DDM

Consider a set of n decision-makers performing a two alternative choice task and let their interaction topology be modeled by a connected undirected graph $\mathcal{G}$ with Laplacian matrix $L \in \mathbb{R}^{n\times n}$. The evidence aggregation in collective decision-making is modeled in the following way. At each time $t \in \mathbb{R}_{\geq 0}$, every decision-maker $k \in \{1, \ldots, n\}$ (i) computes a convex combination of her evidence with her neighbor's evidence; (ii) collects new evidence; and (iii) adds the new evidence to the convex combination. This collective evidence aggregation process is mathematically described by the following cooperative coupled drift diffusion model [15]

$$\mathrm{d}\boldsymbol{x}(t) = \boldsymbol{\beta}\mathrm{d}t - L\boldsymbol{x}(t)\mathrm{d}t + \mathrm{d}\boldsymbol{W}_n(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \quad (8)$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the evidence at time t, $\boldsymbol{\beta} \in \mathbb{R}^n_{\geq 0}$ is the vector of drift rates, L is the Laplacian matrix associated with the interaction graph of agents, and $\boldsymbol{W}_n$ is the n-dimensional Weiner process. Agent k makes a decision in favor of the correct (incorrect) alternative, whenever its evidence $x_k(t)$ crosses the threshold $+\eta_k(-\eta_k)$. The decision-making process for each decision-maker corresponds to a first passage time problem associated with (8).

B. Reduced model for cooperative coupled DDM

We now specialize the decoupled approximation (6) to (8). The approximate evidence at node k is $y_k(t) := \frac{1}{n}\mathbf{1}_n^\top \boldsymbol{x}(t) + \varepsilon_k(t)$, where $\mathbf{1}_n$ is the column n-vector of all ones. Here, the principal component corresponds to evidence averaged across the network, and the residual component corresponds to the deviation from this average. The evidence aggregation dynamics for the reduced model are

$$\begin{bmatrix} \mathrm{d}y_k(t) \\ \mathrm{d}\varepsilon_k(t) \end{bmatrix} = \begin{bmatrix} \hat{\beta} + \frac{\mu_k}{2}(\alpha_k - \varepsilon_k(t)) \\ \frac{\mu_k}{2}(\alpha_k - \varepsilon_k(t)) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \frac{1}{\sqrt{n}} & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{d}W_1(t) \\ \mathrm{d}W_2(t) \end{bmatrix},$$

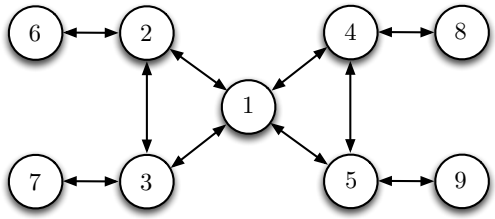


Fig. 1. Interaction topology of the agents.

with $y_k(0) = \bar{x}_0$, $\varepsilon_k(0) = 0$, and $\hat{\beta} = \frac{1}{n}\mathbf{1}_n^\top \boldsymbol{\beta}$. We refer to this reduced model as the *reduced DDM*.

In [17], we investigated first passage properties of (2) for agents with the same drift rate. For such agents, it can be verified that the expected value of $\epsilon_k(t)$ is zero uniformly in time, and consequently, the drift term $\mu_k\alpha_k/2$ in $\varepsilon_k(t)$ dynamics is zero.

Suppose $T_k = \inf\{y_k(t) \in \{-\eta_k, +\eta_k\} \mid t \in \mathbb{R}_{>0}\}$. Let $\mathrm{ET}_k = \mathbb{E}[T_k]$ and error rate $\mathrm{ER}_k = \mathbb{P}(y_k(T_k) = -\eta_k)$. Then, using the theory presented in Section II, we obtain the following PDEs for ET_k and ER_k

$$\Big(\hat{\beta} + \frac{\mu_k(\alpha_k - \varepsilon_{0k})}{2}\Big)\frac{\partial \mathrm{ET}_k}{\partial y_{0k}} + \frac{\mu_k(\alpha_k - \varepsilon_{0k})}{2}\frac{\partial \mathrm{ET}_k}{\partial \varepsilon_{0k}} + \frac{1}{2}\Big(\frac{n+1}{n}\frac{\partial^2 \mathrm{ET}_k}{\partial {y_{0k}}^2} + 2\frac{\partial^2 \mathrm{ET}_k}{\partial y_{0k}\partial \varepsilon_{0k}} + \frac{\partial^2 \mathrm{ET}_k}{\partial {\varepsilon_{0k}}^2}\Big) = -1,$$

with boundary conditions $\mathrm{ET}_k(\cdot, \pm\eta_k) = 0$, $\mathrm{ET}_k(\pm\bar{\eta}_k, \cdot) = 0$, and $\bar{\eta}_k \to +\infty$. Similarly,

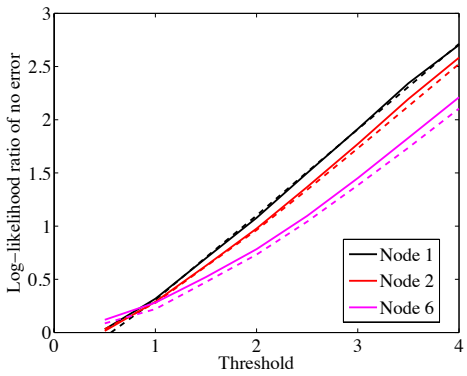
$$\Big(\hat{\beta} + \frac{\mu_k(\alpha_k - \varepsilon_{0k})}{2}\Big)\frac{\partial \mathrm{ER}_k}{\partial y_{0k}} + \frac{\mu_k(\alpha_k - \varepsilon_{0k})}{2}\frac{\partial \mathrm{ER}_k}{\partial \varepsilon_{0k}} + \frac{1}{2}\Big(\frac{n+1}{n}\frac{\partial^2 \mathrm{ER}_k}{\partial {y_{0k}}^2} + 2\frac{\partial^2 \mathrm{ER}_k}{\partial y_{0k}\partial \varepsilon_{0k}} + \frac{\partial^2 \mathrm{ER}_k}{\partial {\varepsilon_{0k}}^2}\Big) = 0,$$

with boundary conditions $\mathrm{ER}_k(\cdot, \eta_k) = 0$, $\mathrm{ER}_k(\cdot, -\eta_k) = 1$, $\mathrm{ER}_k(\bar{\eta}_k, \cdot) = 0$, $\mathrm{ER}_k(-\bar{\eta}_k, \cdot) = 1$, and $\bar{\eta}_k \to +\infty$.

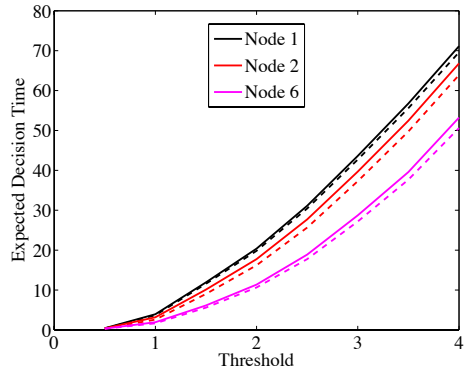
C. Numerical Illustration

Consider a set of nine agents with the interaction topology shown in Figure 1. Let the vector of drift rates be $\boldsymbol{\beta} = [\mathbf{0}_5^\top \ 0.1 \times \mathbf{1}_4^\top]^\top$. A comparison of the performance of the nodes as predicted by the coupled DDM and by the reduced DDM for different thresholds is shown in Figure 2. Note that for a better presentation $\log\big(\frac{1-\mathrm{ER}_k}{\mathrm{ER}_k}\big)$, which is a decreasing function of ER_k, is plotted in Figure 2. The performance metrics obtained using the reduced DDM have the same order as the performance metrics obtained using the coupled DDM, e.g., node 6 has the smallest expected decision time for a given threshold under both models.

It can be seen that for a sufficiently large threshold a more central node has a larger expected decision time and a smaller error rate as compared to a less central node. This is interesting because the most central node 1 has no direct access to the signal, i.e., its drift rate is zero, yet it is able to achieve better decision accuracy as compared to node 6 which has direct access to the signal.



(a) Log-likelihood ratio of no error



(b) Log-likelihood ratio of no error

Fig. 2. A comparison of the performance of the reduced DDM and the cooperative coupled DDM with heterogeneous agents. The solid lines and the dashed lines represent the coupled DDM and the reduced DDM, respectively.

V. Decision-making in Leader-follower networks

In this section, we study collective decision-making in leader-follower networks using the context of two-alternative choice tasks. In leader-follower networks, there is a set of leaders who each have access to a noisy signal and may exchange information with others, and there is a set of followers who each receive information from leaders and exchange information with other followers. Then, each agent makes its own decision from its collected information. In the following, we first describe the model for evidence aggregation in collective decision-making, then develop a decoupled approximation to this model. Finally, we numerically investigate the accuracy of the decoupled approximation.

A. Leader-Follower Coupled DDM

We now consider the coupled DDM dynamics in leader-follower networks. For simplicity, we consider networks with a single leader. Without loss of generality, suppose the first agent is the leader. Let the Laplacian matrix associated with the interaction graph of agents be decomposed as follows

$$L = \begin{bmatrix} L_{11} & -\mathbf{1}_{n-1}^\top L_f \\ -L_f \mathbf{1}_{n-1} & L_f \end{bmatrix},$$

where $L_f \in \mathbb{R}^{(n-1)\times(n-1)}$ is the block of the Laplacian matrix associated with followers. Let

$$\bar{L} = \begin{bmatrix} 0 & \mathbf{0}_{n-1}^\top \\ -L_f \mathbf{1}_{n-1} & L_f \end{bmatrix}.$$

The coupled DDM for leader follower networks is

$$\begin{bmatrix} \mathrm{d}x^\ell(t) \\ \mathrm{d}\boldsymbol{x}^f(t) \end{bmatrix} = \begin{bmatrix} \beta_\ell \\ \mathbf{0}_{n-1} \end{bmatrix} \mathrm{d}t - \bar{L} \begin{bmatrix} x^\ell(t) \\ \boldsymbol{x}^f(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \sigma_\ell \mathrm{d}W_1(t) \\ \mathrm{d}\boldsymbol{W}_{n-1}(t) \end{bmatrix}, \tag{9}$$

with $x_1^\ell(0) = 0$ and $\boldsymbol{x}^f(0) = \mathbf{0}_{n-1}$, where $x_1^\ell(t)$ is the evidence of the leader at time t and $\boldsymbol{x}^f(t) \in \mathbb{R}^{n-1}$ is the vector of the evidence of followers at time t. The decision-making based on the evidence in (9) is modeled similarly to the cooperative coupled DDM in Section IV.

B. Reduced model for leader-follower coupled DDM

We now specialize the decoupled approximation (6) to (9). The approximate evidence at node k is $y_k(t) = x^\ell(t) + \varepsilon_k(t)$. Here, the principal component corresponds to the evidence of the leader, while the residual component corresponds to the deviation of a follower's evidence from the leader's evidence. The coupled dynamics of the evidence of the leader, and the approximate evidence for the k-th follower are

$$\begin{bmatrix} \mathrm{d}x^\ell(t) \\ \mathrm{d}y_k(t) \end{bmatrix} = \begin{bmatrix} \beta_\ell \\ \beta_\ell\big(1 - \frac{\alpha_k \mu_k}{2}\big) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 & 0 \\ \frac{\mu_k}{2} & -\frac{\mu_k}{2} \end{bmatrix} \begin{bmatrix} x^\ell(t) \\ y_k(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \sigma_\ell & 0 \\ \sigma_\ell\big(1 - \frac{\alpha_k \mu_k}{2}\big) & \sqrt{1 - \frac{\sigma_\ell^2 \alpha_k^2 \mu_k^2}{4}} \end{bmatrix} \begin{bmatrix} \mathrm{d}W_1(t) \\ \mathrm{d}W_2(t) \end{bmatrix},$$

with $x^\ell(0) = 0$ and $y_k(0) = 0$.

Suppose $T_k^f = \inf\{y_k(t) \in \{-\eta_k, +\eta_k\} \mid t \in \mathbb{R}_{>0}\}$. Let $\mathrm{ET}_k = \mathbb{E}[T_k^f]$ and error rate $\mathrm{ER}_k^f = \mathbb{P}(y_k(T_k^f) = -\eta_k)$. Then, using the theory presented in Section II, we can obtain the PDEs for the expected decision time and the error rate.

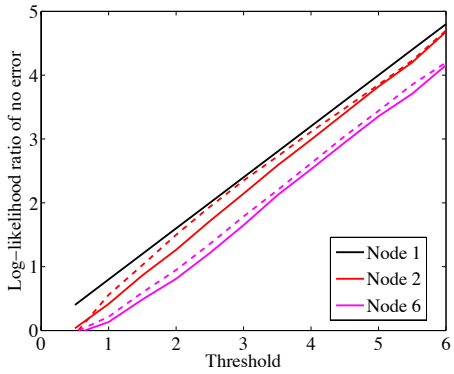
In the particular case when the leader has access to noise free observation, the approximate evidence for the k-th follower is

$$\mathrm{d}y_k(t) = \beta_\ell\big(1 - \frac{\alpha_k \mu_k}{2} + \frac{\mu_k t}{2}\big)\mathrm{d}t - \frac{\mu_k}{2} y_k \mathrm{d}t + \mathrm{d}W_2(t).$$

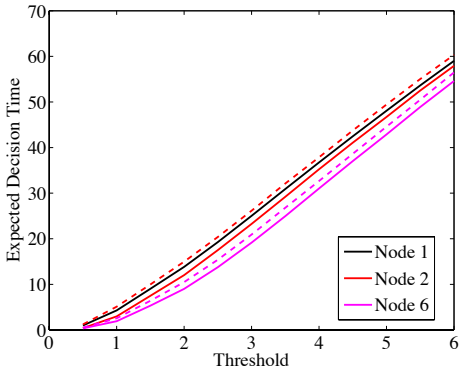
C. Numerical Illustration

Consider a set of nine agents with the interaction topology shown in Figure 1. Let node 1 be the leader. Let the drift rate and the diffusion rate for the leader be $\beta_\ell = 0.1$ and $\sigma_\ell = 0.5$, respectively. Let the diffusion rate for followers be unity. A comparison of the performance of the nodes as predicted by the coupled DDM and by the reduced DDM for different thresholds is shown in Figure 3. The performance metrics obtained using the reduced DDM have the same order as the performance metrics obtained using the coupled DDM.

It can be seen that followers achieve a performance that is very close to the performance of the leader. However, there are individual differences in the performance, and these differences depend on the location of the individual in the network.



(a) Log-likelihood ratio of no error



(b) Log-likelihood ratio of no error

Fig. 3. A comparison of the performance of the reduced DDM and the coupled DDM for leader-follower network. The solid lines and the dashed lines represent the coupled DDM and the reduced DDM, respectively.

VI. Conclusions

In this paper we studied first passage time problems in collective decision-making using the context of two alternative choice tasks. We first considered a generic coupled linear SDE and developed a reduced order approximation to it. The reduced order model is amenable to efficient computation of the properties of first passage times. We applied these reduced order models to collective decision-making problems in heterogeneous cooperative networks and leader-follower networks, and characterized the performance of collective decision-making in terms of speed and accuracy.

There are several possible directions of future research. First, we developed methods for efficient computation of the error rate and the expected decision time. It is well known that small error rates and small expected decision times can not be simultaneously achieved. An interesting question is to characterize the tradeoff between the error rate and the expected decision time, i.e., the speed-accuracy tradeoff. In particular, it is of interest to understand the effect of the network structure and the distribution of heterogeneity in the network on the speed-accuracy tradeoff.

In this paper, we considered a given location of the leader. A designer has the liberty to choose where to place the leader in the network. This so-called leader selection problem has been studied in the literature, see e.g. [16]. An interesting direction is to study the leader selection problem in the context of optimizing the speed-accuracy tradeoff.

References

[1] I. D. Couzin. Collective cognition in animal groups. *Trends in Cognitive Sciences*, 13(1):36–43, 2009.

[2] L. Conradt and C. List. Group decisions in humans and animals: A survey. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1518):719–742, 2009.

[3] R. D. Sorkin, C. J. Hays, and R. West. Signal-detection analysis of group decision making. *Psychological Review*, 108(1):183, 2001.

[4] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen. The physics of optimal decision making: A formal analysis of performance in two-alternative forced choice tasks. *Psychological Review*, 113(4):700–765, 2006.

[5] H. C. Tuckwell and F. Y. M. Wan. First passage time to detection in stochastic population dynamical models for HIV-1. *Applied Mathematics Letters*, 13(5):79–83, 2000.

[6] G. Vahedi, B. Faryabi, J. Chamberland, A. Datta, and E. R. Dougherty. Intervention in gene regulatory networks via a stationary mean-first-passage-time control policy. *IEEE Transactions on Biomedical Engineering*, 55(10):2319–2331, 2008.

[7] S. Singh, D. J. Schneider, and C. R. Myers. Using multitype branching processes to quantify statistics of disease outbreaks in zoonotic epidemics. *Physical Review E*, 89(3):032702, 2014.

[8] S. Redner. *A Guide to First-Passage Processes*. Cambridge University Press, 2001.

[9] O Bénichou, M Coppey, M Moreau, PH Suet, and R Voituriez. Optimal search strategies for hidden targets. *Physical review letters*, 94(19):198101, 2005.

[10] W. Suwansantisuk, M. Z. Win, and L. A. Shepp. First passage time problems with applications to synchronization. In *IEEE International Conference on Communications*, pages 2580–2584, 2012.

[11] R. Ratcliff. A theory of memory retrieval. *Psychological Review*, 85(2):59–108, 1978.

[12] R. Ratcliff and G. McKoon. The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4):873–922, 2008.

[13] M. Kimura and J. Moehlis. Group decision-making models for sequential tasks. *SIAM Review*, 54(1):121–138, 2012.

[14] S. H. Dandach, R. Carli, and F. Bullo. Accuracy and decision time for sequential decision aggregation. *Proceedings of the IEEE*, 100(3):687–712, 2012.

[15] I. Poulakakis, G. F. Young, L. Scardovi, and N. E. Leonard. Node classification in networks of stochastic evidence accumulators. *arXiv preprint arXiv:1210.4235*, October 2012.

[16] K. E. Fitch and N. E. Leonard. Joint centrality distinguishes optimal leaders in noisy networks. *arXiv preprint arXiv:1407.1569*, 2014.

[17] V. Srivastava and N. E. Leonard. Collective decision-making in ideal networks: The speed-accuracy trade-off. *IEEE Transactions on Control of Network Systems*, 1(1):121–132, 2014.

[18] R. S. Blum, S. A. Kassam, and H. V. Poor. Distributed detection with multiple sensors II. Advanced topics. *Proceedings of the IEEE*, 85(1):64–79, 1997.

[19] P. Braca, S. Marano, V. Matta, and P. Willett. Asymptotic optimality of running consensus in testing binary hypotheses. *IEEE Transactions on Signal Processing*, 58(2):814–825, 2010.

[20] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma. Belief consensus and distributed hypothesis testing in sensor networks. In P. J. Antsaklis and P. Tabuada, editors, *Network Embedded Sensing and Control. (Proceedings of NESC'05 Worskhop)*, Lecture Notes in Control and Information Sciences, pages 169–182. Springer, 2006.

[21] C. Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer, fourth edition, 2009.