2013 American Control Conference (ACC)
Washington, DC, USA, June 17-19, 2013

Optimal Distribution of Heterogeneous Agents under Delays

Juan M. Nogales and Jorge Finke
Department of Electrical Engineering and Computer Science
Pontificia Universidad Javeriana
jmnogales@javerianacali.edu.co, finke@ieee.org

Abstract— **An analytical framework for the study of a generic distribution problem is introduced in which a group of agents with different capabilities intend to maximize total utility by dividing themselves into various subgroups without any form of global information-sharing or centralized decision-making. The marginal utility of belonging to a particular subgroup rests on the well-known concept in economic theory of the law of diminishing returns. For a class of discrete event systems, we identify a set of conditions that define local information and cooperation requirements, and prove that if the proposed conditions are satisfied a stable agent distribution representing a Pareto optimum is achieved even under random but bounded decision and transition delays.**

I. INTRODUCTION

One of the most profound trends in modern society is the movement from centralized to distributed applications. It is envisioned that large-scale networks of agents replace more costly and vulnerable centralized systems. Trying to maximize total utility, these agents must work together and coordinate their actions to complete a set of tasks [1]-[3]. Because agents often make decisions based on outdated sensing about their surroundings and their actions affect the state of the network only at some distant point in the future, delays become a major obstacle in the development of network-based applications. The question of how to share information depends in general on the effect of delays on both the agents' perceptions and interactions. Identifying what information over the network is of value is important to overcome these limitations and design interaction strategies that guarantee the performance of the group as a whole [4]-[7].

The work in [4] presents an epidemic model that compares the effects of different intervention strategies in social networks. It identifies conditions under which the process of treating wait-listed households gives the least effective outcome (i.e., under constant intervention delays). Moreover, it shows that allowing for the immediate treatment of 33 percent of the infected households is sufficient to prevents outbreaks. The work in [5] proposes various strategies to control epidemic outbreaks based on the structure of contact networks. The authors compare the effects of delays both in implementing mass and targeted vaccination strategies and in keeping infected patients away from other individuals for some period of time. They show that if the delays in implementation are short, a targeted vaccination can be as effective as a mass vaccination. Moreover, if quarantine periods are short, outbreaks can be contained with a targeted strategy, avoiding the full cost of mass vaccination.

In technological networks, delays are often introduced due to resource and information flow constraints [6]-[7]. The work in [6] characterizes the consensus problem for a group of agents under constant communication delays. Taking advantage of a regular network topology, the authors propose a distributed algorithm based on a proportional-derivative control law for second-order dynamics. They identify regions of stability that characterize the response of the system to variations in delays and controller gains. The work in [7] considers the load balancing problem on a network based on outdated estimates (due to network bandwidth constraints). Delays limit the range of controller gains and as consequence the rate of convergence to an optimal distribution. As in [6], the results in [7] suggest a trade-off between the optimal parameters of the control strategies (gains leading to a better system response), and the risk for instability due to delays.

This paper considers the effects of delays on a generic distribution problem of a group of heterogeneous agents. Heterogeneous agents have two or more distinctive features that define the various types of agents. Like [8], we present a cooperation strategy that coordinates the actions between agents and takes into account to what extent each type of agent contributes toward the group's common goal. We use a discrete event systems framework to capture how the group achieves an equilibrium distribution that represents a Pareto optimum in the sense that no subgroup of agents can benefit from reallocating agents without making the group as a whole worse off. We associate to each subgroup mathematical functions that satisfy the well-known concept in economic theory of the law of diminishing returns. The proposed framework allows the heterogeneous agents to achieve the optimal distribution despite the effect of delays on both the agents' perceptions and interactions. Adding random but bounded delays extends the work in [3] allowing us to propose a cooperation scheme based on local, outdated information. In particular, the proposed framework quantifies the trade-off between the optimal degree of cooperation for faster convergence and the size of the delays in decision-making and transitioning (which is of interest in various contexts of multi-agent networks [9]-[11]).

The remaining sections are organized as follows: Section II introduces the notion of utility functions and formalizes the generic distribution problem. Section III introduces the model and presents sufficient conditions for a group of heterogeneous agents to reach the optimal distribution under both types of delays. Our analytical results in Section IV characterize the stability properties of the Pareto optimum equilibrium point of the network. Section V provides Monte Carlo simulations that explore the link between cooperation, settling time, information-sharing, and the size of the delays.

II. THE NOTATION AND BASIC PROBLEM

Let a node represent an activity, task, or subgroup of agents; nodes belong to a set N, indexed from 1 to n. An agent represents a resource or unit of supply and may be of various types; agent types belong to a set M, indexed from 1 to m. The interconnection among nodes is described by a network $G = (N, A)$, where A represents the set of edges. If $\{j, k\} \in A$, agents at node j can sense (outdated) information about node k and can move to k, and agents at node k can also sense (or move) to node j. Let $N_j = \{k : \{j, k\} \in A\}$ denote the set of neighboring nodes of node j. We assume that the number of agents of each type is large enough to be appropriately represented by a continuous variable (as in [12]). Let $\mathcal{R} = [0, \infty)^m$ be the space of all combinations of available agents and let $\hat{r}_j^\ell$ represent the number of agents of type ℓ at node j and $\hat{r}_j = [\hat{r}_j^1, \ldots, \hat{r}_j^m]^\top \in \mathcal{R}$ be the distribution of agents at node j (e.g., the number of agents of each type assigned to a particular subgroup). Let $\Delta_c \subset \mathcal{R}^n$ denote the $m(n-1)$ dimensional simplex defined by the equality constraint $\sum_{i=1}^n \hat{r}_i = c$, where c is a vector $[c^1, \ldots, c^m]^\top \in \mathcal{R}$ and c^ℓ the available number of agents of type ℓ. The utility of having a distribution of agents $\hat{r}_j$ at node j is given by the function $f_j : \mathcal{R} \to [0, \infty)$ and the total utility function is defined by $f : \mathcal{R}^n \to [0, \infty)$, $f(\hat{r}) = \sum_{i=1}^n f_i(\hat{r}_i)$, where $\hat{r} = [\hat{r}_1^\top, \ldots, \hat{r}_n^\top]^\top$. The objective is to identify local requirements that allow a group of heterogeneous agents to solve the following optimization problem

$$\text{maximize } f(\hat{r}), \text{ subject to } \hat{r} \in \Delta_c \tag{1}$$

under decision delays (agents have a delayed perception of the marginal utility function of the neighboring nodes) and transition delays (agents require some finite time to move from one node to another), both random but bounded.

We suppose that each utility function f_j satisfies the following three assumptions, common in economic theory [13]. First, each function f_j is continuously differentiable on $\mathcal{R}$. Second, an increase in utility must satisfy the law of diminishing returns, i.e.,

$$\frac{f_j(\hat{r}_j + u^\ell h^\ell) - f_j(\hat{r}_j)}{u^\ell} > \frac{f_j(\hat{r}_j + w^\ell h^\ell) - f_j(\hat{r}_j)}{w^\ell} \tag{2}$$

where $\hat{r}_j \in \mathcal{R}$, $w^\ell > u^\ell > 0$ are some number of agents of type $\ell \in M$, and $h^\ell \in \mathcal{R}$ is a vector with one in the ℓth row and zeros otherwise. In other words, we assume a decreasing average return with respect to (w.r.t.) increasing magnitudes of agent additions. Third, an increase of agents at a node must always increase the utility associated to that node and satisfy the bounds

$$0 < \frac{f_j(\hat{r}_j + v^\ell h^\ell) - f_j(\hat{r}_j)}{v^\ell} < \infty \tag{3}$$

where $v^\ell > 0$, $\ell \in M$. Note that (2) and (3) can be viewed as a particular case of the law of diminishing returns. Under the above assumptions, the partial derivative of f_j w.r.t. agents of type $\ell \in M$, that is the marginal utility w.r.t. agents of type ℓ, denoted by s_j^ℓ, satisfies

$$-a^\ell \leq \frac{s_j^\ell(x_j) - s_j^\ell(y_j)}{x_j^\ell - y_j^\ell} \leq -b^\ell \tag{4}$$

for any $x_j, y_j \in \mathcal{R}$, $x_j^\ell \neq y_j^\ell$, and some constants $0 < b^\ell \leq a^\ell$ defined for each type of agent $\ell \in M$. Moreover, because of the assumptions on f_j, the functions s_j^ℓ are continuous on $\mathcal{R}$, strictly decreasing, and non-negative (see [3] for details).

III. THE MODEL

Agent transitions across the network are driven by the asynchronous occurrence of discrete events at time index $t = 0, 1, 2, \ldots$. At time t, the number of agents of type $\ell \in M$ transitioning from node j to a neighboring node k is defined as $r_{j\to k}^\ell(t)$. The total number of agents of type ℓ transitioning from node j to its neighboring nodes is defined as $\vec{r}_j^{\,\ell}(t) = \sum_{k \in N_j} r_{j\to k}^\ell(t)$. Let $\vec{r}_j(t) = [\vec{r}_j^{\,1}(t), \ldots, \vec{r}_j^{\,m}(t)]^\top$. To take into account delays in transitioning between nodes let

$$\vec{r}(t) = \begin{bmatrix} \vec{r}_1(t) & \ldots & \vec{r}_1(t - B_t + 1) \\ \vdots & \ddots & \vdots \\ \vec{r}_n(t) & \ldots & \vec{r}_n(t - B_t + 1) \end{bmatrix}$$

where $\vec{r} \in \mathcal{R}^{n \times B_t}$, $B_t > 1$. Similarly, to capture delays in decision-making let

$$\hat{r}(t) = \begin{bmatrix} \hat{r}_1(t) & \ldots & \hat{r}_1(t - B_s + 1) \\ \vdots & \ddots & \vdots \\ \hat{r}_n(t) & \ldots & \hat{r}_n(t - B_s + 1) \end{bmatrix}$$

where $\hat{r} \in \mathcal{R}^{n \times B_s}$, $B_s > 1$. The state of the system is defined as $r = [\vec{r}, \hat{r}]$, $r \in \mathcal{R}^{n \times (B_t + B_s)}$.

Let $(\hat{r})_{pq} \in \mathcal{R}^{n \times B_s}$ denote the vector of the distribution of all types of agents across node p in column q of its matrix argument (i.e., the distribution at time $t + 1 - q$). Let $\mathcal{S} = \{1, \ldots, B_s\}$. Note that each utility function f_j is strictly concave on $\mathcal{R}$ and f is strictly concave on $\mathcal{R}^n$. Thus, the optimal point $\bar{r}$ that satisfies (1) is unique and belongs to the set $\Delta_c^\star$ defined as

$$\{r \in \Delta_c | \, \forall p \in N, \forall p' \in N_p, \forall \ell \in M, \forall q, q' \in \mathcal{S}, \vec{r}_p^{\,\ell} = 0, \\ s_p^\ell((\hat{r})_{pq}) < s_{p'}^\ell((\hat{r})_{p'q'}) \Rightarrow \hat{r}_p^{\,\ell} = 0\} \tag{5}$$

In other words, when $r \in \Delta_c^\star$, it must be that for any type of agents $\ell \in M$ there are no agents transitioning between nodes (because they would not represent a gain in terms of utility) and if a particular node has a lower marginal utility than a neighboring node, then the optimal distribution has no agents of type ℓ at that node [14].

Let $e^{\ell}_{j\rightarrow k}(t)$ denote the departure of $u^{\ell}_{j\rightarrow k}(t)$ agents of type $\ell \in M$ from node j to node k at time t. Let $e_{j\rightarrow k}(t)$ denote the set of all possible departures from node j to node k for any type of agent. Similarly, let $e^{\ell}_{j\leftarrow i}(t)$ denote the arrival of $w^{\ell}_{j\leftarrow i}(t)$ agents of type $\ell \in M$ at node j from node i at time t and $e_{j\leftarrow i}(t)$ denote the set of all possible arrivals for any type of agents. Finally, let $\mathcal{E} = \mathcal{P}(\{e_{j\rightarrow k}(t)\} \cup \mathcal{P}(\{e_{j\leftarrow i}(t)\}) - \{\emptyset\}$ be the set of events of all simultaneous transitions (departures and arrivals) between nodes. An event $e(t) \in \mathcal{E}$ is defined as a set with each element representing a transition of an number of agents of type $\ell \in M$ between two neighboring nodes.

If an event $e(t) \in \mathcal{E}$ occurs at time t, the update of the state of the system is given by $r(t+1) = g(r(t))$. For agents of type $\ell \in M$, the number of agents at node $j \in N$ at time $t+1$ is given by

$$\begin{aligned} \hat{r}_j^{\ell}(t+1) = \hat{r}_j^{\ell}(t) - &\sum_{\{k:\, e^{\ell}_{j\rightarrow k}(t) \,\in\, e(t)\}} u^{\ell}_{j\rightarrow k}(t) \\ + &\sum_{\{i:\, e^{\ell}_{j\leftarrow i}(t) \,\in\, e(t)\}} w^{\ell}_{j\leftarrow i}(t) \end{aligned} \quad (6)$$

In other words, the number of agents of type ℓ at node $j \in N$ at time $t+1$ is the actual amount, minus the number of agents that departures toward its neighboring nodes, plus the number of agents that arrives from some neighboring nodes. The number of agents that are in transition from node $j \in N$ to its neighboring nodes at time $t+1$ is given by

$$\begin{aligned} \vec{r}_j^{\,\ell}(t+1) = \vec{r}_j^{\,\ell}(t) + &\sum_{\{k:\, e^{\ell}_{j\rightarrow k}(t) \,\in\, e(t)\}} u^{\ell}_{j\rightarrow k}(t) \\ - &\sum_{\{k: e^{\ell}_{k\leftarrow j}(t) \,\in\, e(t)\}} w^{\ell}_{k\leftarrow j}(t) \end{aligned} \quad (7)$$

Note that the number of agents of type ℓ that are in transition depends on the number of agents that departs from node j and the number of agents that arrives at any of the destination nodes. We do not differentiate between agents transitioning to different destination nodes. Moreover, solving (1) requires that the agents on a network G satisfy the following three assumptions.

Assumption 1 (on the network): The network $G = (N, A)$ is connected, i.e., $\forall j, k \in N$ there exists a path from node j to node k.

The requirement that G has one network component places minimum conditions on the sensing constraints and the possible transitions of the agents.

Assumption 2 (on the various types of agents): The total number of agents of each type $\ell \in M$, c^{ℓ}, is large enough to have some number of agents of type ℓ at each node when $r \in \Delta_c^{\star}$.

Following a similar argument as in [3], we can show that when Assumption 1 and 2 hold, $\Delta_c^{\star}$ in (5) can be written as

$$\begin{aligned} \{r \in \Delta_c |\, \forall p \in N, \forall p' \in N_p, \forall \ell \in M, \forall q, q' \in \mathcal{S}, \vec{r}_p^{\,\ell} = 0, \\ s_p^{\ell}((\hat{r})_{pq}) = s_{p'}^{\ell}((\hat{r})_{p'q'})\} \end{aligned} \quad (8)$$

Assumption 3 (on the delays): At time t, agents of type ℓ at node j decide to transition to neighboring nodes based on a delayed value of the marginal utility of node k w.r.t. agents of type ℓ. This perception of the marginal utility of node $k \in N_j$ satisfies

$$\rho^{\ell}_{jk}(t) \in \{s^{\ell}_k(\hat{r}_k(t')) : t' \in [t - B_s + 1, t]\}$$

where $B_s > 1$. Moreover, $\rho^{\ell}_{jj}(t) = s^{\ell}_j(\hat{r}_j(t))$ and agents that start transitioning from node j at time t must arrive at their destination node within $B_t > 1$.

Agents at node j must know the actual value of $s^{\ell}_j(\hat{r}_j(t))$, but both decision and transition delays can be random but bounded. Transitions may be stochastic, but rest on the following conditions for $e(t) \in \mathcal{E}$ at time t.

Condition 1: If $\rho^{\ell}_{jk}(t) \le s^{\ell}_j(\hat{r}_j(t))\ \forall k \in N_j$ then $e^{\ell}_{j\rightarrow k}(t) \in e(t)$ has $u^{\ell}_{j\rightarrow k}(t) = 0$ (i.e., agents at node j cannot transition to any neighboring node).

Condition 2: If $\rho^{\ell}_{jk}(t) > s^{\ell}_j(\hat{r}_j(t))$ for some $k \in N_j$, then $e^{\ell}_{j\rightarrow k}(t) \in e(t)$ has $u^{\ell}_{j\rightarrow k}(t) > 0$ and node k is the unique destination node such that

$$\rho^{\ell}_{jk}(t) \ge \rho^{\ell}_{ji}(t) > s^{\ell}_j(\hat{r}_j(t)),\ \forall i, k \in N_j \quad (9)$$

The number of agents of type ℓ that starts to transition from node j to node k is

$$0 < u^{\ell}_{j\rightarrow k}(t) = \phi^{\ell}[\rho^{\ell}_{jk}(t) - s^{\ell}_j(\hat{r}_j(t))] \quad (10)$$

where $\phi^{\ell} \in (0, 1/2a^{\ell}]$.

Eq. (9) guarantees that agents transition toward a unique node perceived to have higher or equal marginal utility than all other neighboring nodes. Eq. (10) bounds the number of agents that can transition. The value of ϕ^{ℓ} in (10) captures the degree of cooperation between agents of type ℓ.

IV. RESULTS

By restraining the allowable events in the network, the assumptions and conditions in Section III define a distribution process that leads to a Pareto solution that captures the optimum distribution of the heterogeneous agents across the network.

Theorem 1: Suppose Assumptions 1-3 hold. Moreover, decision-making on the network G satisfies Conditions 1-2. Then, the point $r \in \Delta_c^{\star}$ is an equilibrium point of the model and has a region of asymptotic stability equal to Δ_c.

Theorem 1 implies that from any initial distribution of agents, the group of agents will reach the point $r \in \Delta_c^{\star}$ that maximizes the total utility, despite local requirements and decision and transition delays.

Proof:

Let the Lyapunov candidate function be defined as

$$\begin{aligned} V(r(t)) = \sum_{\ell=1}^{m} \Bigg[&\max_i \{s^{\ell}_i(\hat{r}_i(t')) : t' \in [t - B_s + 1, t]\} - s^{\ell}_i(\bar{r}_i) \\ &+ \frac{1}{n} \Bigg(\max_i \{s^{\ell}_i(\hat{r}_i(t')) : t' \in [t - B_s + 1, t]\} \\ &- \frac{1}{nB_t} \sum_{i=1}^{n} \sum_{t'=t-B_t+1}^{t} s^{\ell}_i(\hat{r}_i(t) + \vec{r}_i(t')) \Bigg) \Bigg] \end{aligned} \quad (11)$$

where $\bar{r}$ is the case when $r \in \Delta_c^\star$. The supplement to this paper (available at http://www.jfinke.org/public_html/publications.html) shows that there exist a metric $\varphi(r(t), \Delta_c^\star)$ and two constants $\eta_1, \eta_2 > 0$ such that $\eta_1 \varphi(r(t), \Delta_c^\star) \leq V(r(t)) \leq \eta_2 \varphi(r(t), \Delta_c^\star)$. Moreover, if the maximum marginal utility decreases, then the Lyapunov function $V(r(t)) = 0$ as $t \to \infty$. Here, we will show that the maximum marginal utility must decrease after $(n–1)(B_s+B_t)$ time steps.

Note that when $r \in \Delta_c^\star, \forall j \in N, \forall k \in N_j, \forall \ell \in M$, because $\vec{r}_j^{\,\ell} = 0$, we know that there are no agents transitioning between nodes and marginal utilities have not changed their value within the last $B_s - 1$ time steps. Because $\rho_{jk}^\ell(t) = s_k^\ell(\hat{r}_k(t)) = s_j^\ell(\hat{r}_j(t))$, $u_{j\to k}^\ell = 0$ (i.e., agents cannot transition toward any neighboring node). Therefore, $\Delta_c^\star$ is invariant.

Let $r \notin \Delta_c^\star$. Fix a time index $t_1 > B_s$. Fix a type of agent $\ell \in M$ such that for some node j either $\vec{r}_j^{\,\ell}(t_1) \neq 0$ (some agents are transitioning) or $\max_i\{s_i^\ell(\hat{r}_i(t))\} > s_j^\ell(\hat{r}_j(t))$ at some $t \in [t_1 - B_s + 1, t_1]$ (there exists a node with a marginal utility below the maximum value within the last $B_s - 1$ time steps). (Because $r \notin \Delta_c^\star$, at least one case must be true). Let $t_0 = \operatorname{argmin}_t \max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1 - B_s + 1, t_1]\}$ be the minimum time index at which some node has the maximum marginal utility. Let $\mathcal{K}^\ell(t) = \operatorname{argmax}_i\{s_i^\ell(\hat{r}_i(t))\}$ be the set of nodes with the highest marginal utility w.r.t. agents of type ℓ at time t. We now show that the transition of agents of type ℓ from any node j to its neighboring nodes will not cause the maximum marginal utility value to increase over time (i.e., there isn't any node that exceeds the value of the nodes in $\mathcal{K}^\ell(t_0)$).

First, consider an event $e(t) \in \mathcal{E}$ at time t with a transition $e_{j\to k'}^\ell(t)$ where $u_{j\to k'}^\ell(t) > 0$ from a node $j \in N$ to some node $k' \in N_j$. Because agents of type ℓ at node j can only transition toward a unique destination node, using (6), the marginal utility of node j at time $t + 1$ is overbounded by the number of agents that transition toward node k', so that

$$s_j^\ell(\hat{r}_j(t+1)) \leq s_j^\ell(\hat{r}_j(t) - u_{j\to k'}^\ell(t) h^\ell) \tag{12}$$

Note that agents that arrive at a node can only decrease its marginal utility. Moreover, using (4) for node j with $x_j = \hat{r}_j(t)$ and $y_j = \hat{r}_j(t) - u_{j\to k'}^\ell(t) h^\ell$ yields

$$s_j^\ell(\hat{r}_j(t) - u_{j\to k'}^\ell(t) h^\ell) \leq s_j^\ell(\hat{r}_j(t)) + a^\ell u_{j\to k'}^\ell(t) \tag{13}$$

According to (10) and because $0 < \phi^\ell \leq 1/2a^\ell$, we know that

$$\begin{aligned} s_j^\ell(\hat{r}_j(t)) + a^\ell u_{j\to k'}^\ell(t) &\leq \rho_{jk'}^\ell(t) - a^\ell u_{j\to k'}^\ell(t) \\ &= (1 - a^\ell \phi^\ell) \rho_{jk'}^\ell(t) + a^\ell \phi^\ell s_j^\ell(\hat{r}_j(t)) \end{aligned} \tag{14}$$

Combining (12), (13), and (14), we have that

$$s_j^\ell(\hat{r}_j(t+1)) \leq (1 - a^\ell \phi^\ell) \rho_{jk'}^\ell(t) + a^\ell \phi^\ell s_j^\ell(\hat{r}_j(t)) \tag{15}$$

Because $\rho_{jk'}^\ell(t) \in \{s_{k'}^\ell(\hat{r}_{k'}(t')) : t' \in [t - B_s + 1, t]\}$ and $a^\ell \geq b^\ell > 0$, we can overbound (15) by

$$\begin{aligned} s_j^\ell(\hat{r}_j(t+1)) \leq &\max_i\{s_i^\ell(\hat{r}_i(t')) : t' \in [t - B_s + 1, t]\} \\ &- b^\ell \phi^\ell \Big[\max_i\{s_i^\ell(\hat{r}_i(t')) : t' \in [t - B_s + 1, t]\} \\ &- s_j^\ell(\hat{r}_j(t))\Big] \end{aligned} \tag{16}$$

Second, consider an event $e(t) \in \mathcal{E}$ at time t in which no agents depart from node j, $e_{j\to k'}^\ell(t)$ with $u_{j\to k'}^\ell(t) = 0$ for all $k' \in N_j$. According to (6), the marginal utility of node j cannot increase at time $t + 1$. (Eq. (16) also holds for nodes from which no agents depart). Thus, for a fix time t_1, the marginal utility at any node $j \in N$ does not exceed the maximum value within the last $B_s - 1$ time steps.

Next, using (16) with $t = t'$, for $t' \in [t_1, t_1 + B_s - 1]$, we know that the occurrence of $e_{j\to k'}^\ell(t') \in e(t')$ with $u_{j\to k'}^\ell(t') \geq 0$, cannot increase the marginal utility of any node $j \in N$ above $\max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1 - B_s + 1, t_1]\}$. In other words, for all $t'' \in (t_1, t_1 + B_s]$, it must be that $s_j^\ell(r_j(t'')) \leq \max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1 - B_s + 1, t_1]\}$ for all node $j \in N$. Thus, we know that

$$\begin{aligned} &\max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1 - B_s + 1, t_1]\} \\ &\geq \max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1, t_1 + B_s]\} \end{aligned} \tag{17}$$

To show that after a certain number of time steps the maximum marginal utility in fact decreases, consider the following two cases.

First, if for some $t' \in [t_1, t_1 + B_s]$, $\max_i\{s_i^\ell(\hat{r}_i(t'))\} < \max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1 - B_s + 1, t_1]\}$, because $s_j^\ell(\hat{r}_j(t')) \leq \max_i\{s_i^\ell(\hat{r}_i(t'))\}$ for all $j \in N$, using (16) with $t = t'$, we get that $s_j^\ell(\hat{r}_j(t'+1)) < \max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1 - B_s + 1, t_1]\}$. Using a similar argument as above, the marginal utility values cannot reach $\max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1 - B_s + 1, t_1]\}$ for all $t'' > t'$. Thus, we can conclude that (17) holds with strict inequality.

Second, if for all $t' \in [t_1, t_1 + B_s]$, $\max_i\{s_i^\ell(\hat{r}_i(t'))\} = \max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1 - B_s + 1, t_1]\}$, i.e., the maximum marginal utility value at time t_0 does not decrease between t_1 and $t_1 + B_s$, then there exist a set of (possibly a few) nodes $k', \ldots, k'' \in \mathcal{K}^\ell(t')$ for all $t' \in [t_1, t_1 + B_s]$ with the same marginal utility value as node $k \in \mathcal{K}^\ell(t_0)$. From (16) with $t = t'$, we know that the marginal utility of node $j \notin \mathcal{K}_\ell(t')$ at some $t' \in [t_1, t_1 + B_s]$ cannot reach the same value as node $k \in \mathcal{K}^\ell(t_0)$. Therefore, we can conclude that $|\mathcal{K}^\ell(t'+1)| \leq |\mathcal{K}^\ell(t')|$ for all $t' \in [t_1, t_1 + B_s]$.

Furthermore, as long as $r \notin \Delta_c^\star$, if the maximum marginal utility value does not decrease between t_1 and $t_1 + B_s$, there exists some node $k' \in \mathcal{K}^\ell(t')$ for all $t' \in [t_1, t_1 + B_s]$ with a neighboring node $j \notin \mathcal{K}_\ell(t_1 + B_s - 1)$ because the network G is connected (Assumption 1). Note that the perception of any neighboring node about the marginal utility of node k' must be equal to its actual value, $s_j^\ell(\hat{r}_j(t_1 + B_s - 1)) < \rho_{jk'}^\ell(t_1 + B_s - 1) = \max_i\{s_i^\ell(\hat{r}_i(t)) : t \in [t_1 - B_s + 1, t_1]\}$. Because $\rho_{jk'}^\ell(t_1 + B_s - 1) > s_j^\ell(\hat{r}_j(t_1 + B_s - 1))$, agents at node j must transition toward node k'. According to Assumption 3, those agents that transition at time $t_1 + B_s - 1$ must arrive at node k' within B_t time steps, i.e., $e_{k'\leftarrow j}^\ell(t'') \in e(t'')$ with $w_{k'\leftarrow j}^\ell(t'') > 0$ must occur at some $t'' \in [t_1 + B_s, t_1 + B_s + B_t - 1]$.

Let $\tau_{k'} = \operatorname{argmin}_{t''}\{w_{k'\leftarrow j}^\ell(t'') > 0 : j \in N_{k'}, t'' \in [t_1 + B_s, t_1 + B_s + B_t - 1]\}$ be the minimum time index at which transitioning agents arrive at node k'. Using (6) with $t = \tau_{k'}$, we have that

$$
\begin{aligned}
&s_{k'}^{\ell}(\hat{r}_{k'}(\tau_{k'}+1)) \\
&\le s_{k'}^{\ell}\left(\hat{r}_{k'}(\tau_{k'}) + \sum_{\{j:\, e_{k'\leftarrow j}^{\ell}(\tau_{k'})\in e(\tau_{k'})\}} w_{k'\leftarrow j}^{\ell}(\tau_{k'})h^{\ell}\right) \\
&< s_{k'}^{\ell}(\hat{r}_{k'}(\tau_{k'})) \\
&= \max_i\{s_i^{\ell}(\hat{r}_i(t)) : t\in[t_1-B_s+1, t_1]\} \qquad (18)
\end{aligned}
$$

Because $\tau_{k'}\in[t_1+B_s, t_1+B_s+B_t-1]$, if for all $t'\in[t_1, t_1+B_s]$, $\max_i\{s_i^{\ell}(\hat{r}_i(t'))\}=\max_i\{s_i^{\ell}(\hat{r}_i(t)):t\in[t_1-B_s+1,t_1]\}$, we know that after B_s+B_t time steps either $|\mathcal{K}^{\ell}(t_1+B_s+B_t)|<|\mathcal{K}^{\ell}(t_1)|$ or $\max_i\{s_i^{\ell}(\hat{r}_i(t_1+B_s+B_t))\}<\max_i\{s_i^{\ell}(\hat{r}_i(t)):t\in[t_1-B_s+1,t_1]\}$ (due to (18)). Furthermore, because the set $\mathcal{K}^{\ell}(t')$ for all $t'\in[t_1, t_1+B_s]$ may contain several nodes, after at most $(n-1)(B_s+B_t)$ time steps, some agents transitioning must have reached all the nodes with the maximum marginal utility within $[t_1-B_s+1, t_1]$, i.e.,

$$
\begin{aligned}
&\max_i\{s_i^{\ell}(\hat{r}_i(t)) : t\in[t_1-B_s+1,t_1]\} \\
&> \max_i\{s_i^{\ell}(\hat{r}_i(t)) : t\in[t_2-B_s+1,t_2]\}
\end{aligned}
$$

where $t_2=t_1+n(B_s+B_t)$. Note that if $r\notin\Delta_c^{\star}$, Condition 2 will be always satisfied and will cause that some agents start to transition toward a node with a higher perception. As consequence, the maximum will decrease and the variations in the Lyapunov function in (11) will be negatives. Thus, $V(r(t))=0$ as $t\to\infty$ and the optimal point $r\in\Delta_c^{\star}$ is asymptotically stable in the region Δ_c. More details can be found in the supplement to this paper.

V. SIMULATIONS

This section presents the dynamics of a group of two types of agents across a network with a regular degree distribution and short, medium, and long size decision and transition delays. Short delays represent periods of time that are within 5% of the average settling time without delays (175 time steps). Medium and long delays correspond to periods within 5%-15% and 15-30%, respectively. Each node represents a subgroup that requires both types of agents but with a different marginal utility w.r.t. each type. Let the marginal utility functions of each node be characterized by the form $s_i^{\ell}(r_i)=c_i^{\ell}d_{i\ell}^{-c_i^{\ell}r_i^{\ell}}$ where c_i^{ℓ} and $d_{i\ell}$ are chosen from a uniform random distribution with support on $[0.5, 1.3]$ and $[2, 4]$, respectively. Figure 1 shows the average settling time, t_s, to reach the optimal distribution defined in (8) from an initial random distribution. Each plot considers 8 different networks with a particular density of links (i.e., proportion of links of G relative to the total number possible) and degrees of cooperation.

Figure 1a shows the impact of short decision and transition delays ($B_s=B_t=5$ time steps). Note that increasing the degree of cooperation leads to faster settling times with lower standard deviations, σ, for all densities. Moreover, networks with a high density have a lower σ for any fixed degree of cooperation.

Figure 1b considers medium decision delays and short transition delays($B_s=20$ and $B_t=5$). Note that increasing the degree of cooperation from 25% to 50% leads to a faster t_s and lower σ values for all density values, but more noticeable for networks with a low density (with values under .55). While increasing the degree of cooperation from 50% to 100% leads to faster t_s only for a low density (under .67), for networks with high densities (values between .89 and 1) t_s may increase. The results in Figure 1c are similar to Figure 1b. Note, however, that for degrees of cooperation between 50% and 100%, the effect of decision delays has a slightly stronger effect than transition delays on both t_s and σ for high density networks (values above .78).

In Figure 1d, decision and transition delays are medium ($B_s=B_t=20$). Note that increasing the degree of cooperation from 50% to 100% leads to slower settling times with higher σ values for networks with high density (values above .67). The higher the density values, the slower the settling times. For low density networks (with densities under .55), Figure 1d shows a faster t_s and lower σ values as the degree of cooperation increases. Note also that for degrees of cooperation under 75%, σ is lower for networks with high density than for low density values.

In Figure 1e, decision delays are long and transition delays are medium ($B_s=35$ and $B_t=20$). Note that increasing the degree of cooperation is useful (i.e., leads to a faster t_s and lower σ values) only for low densities (values under .44). Increasing the degree of cooperation for high density networks leads to slower t_s with higher σ values. The higher the density values are, the slower the settling times. As in Figure 1d, increasing the degree of cooperation from 25% to 50% always leads to faster t_s and lower σ values for all network densities. For degrees of cooperation under 75%, σ is lower for high density than for low density networks. The results in Figure 1f are similar to Figure 1e. In networks with high density, however, the effect of medium size decision delays and long transition delays is subtle compared with long decision delays and medium transition delays.

In general, Figure 1 shows that under short, medium, or long decision and transition delays, increasing the degree of cooperation leads to a faster settling times for networks with low densities. However, there exists a trade-off between the optimal degree of cooperation and the size of delays for high density values (i.e., less cooperation may lead to faster settling times as delays increase). Overall, when decision and transition delays are short, we can get a faster t_s as the degree of cooperation increases regardless the density of the network. When delays are medium or long, we can avoid slower settling times by letting the agents exploit the trade-off between the degree of cooperation and the size of the delays. Finally, the agents reach a Pareto optimum through local decision-making strategies despite the presence of both types delays with different sizes.

VI. CONCLUSIONS AND FUTURE WORK

The proposed mathematical framework generalizes models of generic team formation and should be of interest in be-

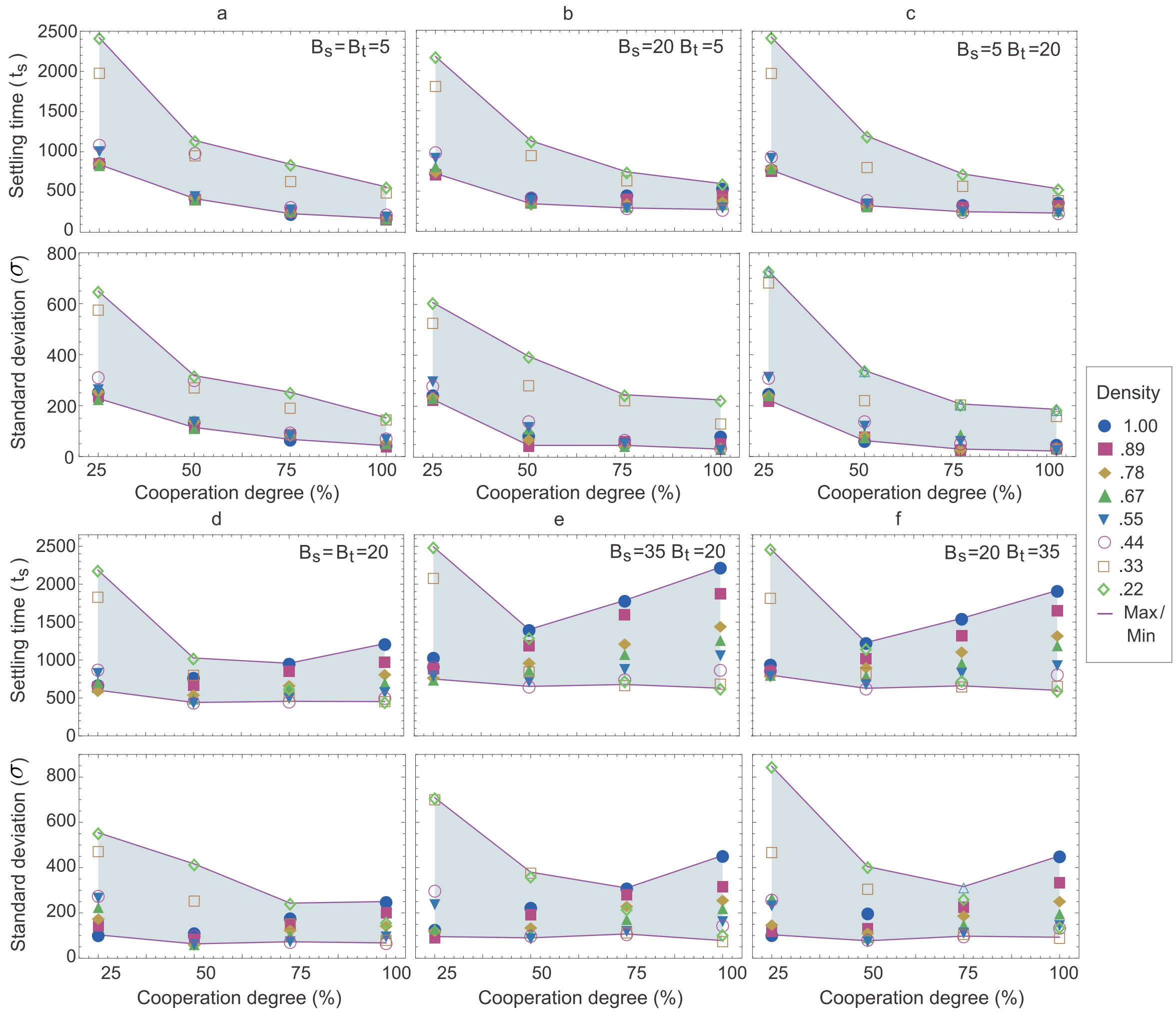

FIG. 1: Settling time for different density values under varying degrees of cooperation between agents and short, medium, and long delays. Each point represents 100 simulation runs. A density of .22 corresponds to a network with a ring topology, while a density of 1.0 corresponds to a fully connected network.

havioral sciences [4]-[8] and engineering applications [1]-[7] where collective outcomes are the result of coordinated actions by heterogeneous agents. Using discrete event systems theory, the model allows us to study how distributed agents reach a Pareto solution that maximizes the total utility despite local requirements and the presence of delays in both decision-making and transitioning.

REFERENCES

[1] Z. Meng et al. Leaderless and Leader-Following Consensus with Communication and Input Delays under a Directed Network Topology. IEEE Transactions on Systems, Man, and Cybernetics, vol. 41 (1), pp. 75-88, 2011.

[2] A. Seuret, D. V. Dimarogonas, and K. H. Johansson. Consensus under Communication Delays. In Proc. of 47th IEEE Conference on Decision and Control, Cancun, Mexico, pp. 4922-4927, Dec. 2008.

[3] J. Nogales and J. Finke. Local Requirements for Optimal Distribution of Heterogeneous Agents. In Proc. of the 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, pp. 3596-3601, Dec. 2011.

[4] S. Spencer and P. O'Neill. Assessing the Impact of Intervention Delays on Stochastic Epidemics. Methodology and Computing in Applied Probability, vol. 14 (1), pp. 1-18, 2012.

[5] S. Eubank et al. Modelling Disease Outbreaks in Realistic Urban Social Networks. Nature, vol. 429 (6988), pp. 180-184, 2004.

[6] R. Cepeda-Gomez and N. Olgac. Stability Analysis for the Group Dynamics Consensus with Time Delayed Communications. European Journal of Control, in print, 2012.

[7] J. Chiasson et al. The Effect of Time Delays on the Stability of Load Balancing Algorithms for Parallel Computations. IEEE Transactions on Control Systems Technology, vol. 13 (6), pp. 932-942, 2005.

[8] A. Arami, B. R. Rofoee, and C. Lucas. Multiple Heterogeneous Ant Colonies with Information Exchange. In Proc. of the IEEE World Congress on Computational Intelligence and Congress on Evolutionary Computation, Hong Kong, China, pp. 3298-3304, June 2008.

[9] Han-Lim Choi, A. K. Whitten, and J. P. How. Decentralized Task Allocation for Heterogeneous Teams with Cooperation Constraints. In Proc. of the American Control Conference, Baltimore, MD, pp. 3057-3062, June 2010.

[10] M. Franceschelli et al. Decentralized Stabilization of Heterogeneous Linear Multi-Agent Systems. In Proc. of the IEEE International Conference on Robotics and Automation, Anchorage, AL, pp. 3556-3561, May 2010.

[11] M. Yarvis et al. Exploiting Heterogeneity in Sensor Networks. In Proc. of the 24th IEEE Annual Joint Conference of the IEEE Computer and Communications Societies, Miami, FL, pp. 878-890, March 2005.

[12] B. J. Moore, J. Finke, and K. M. Passino. Optimal Allocation of Heterogeneous Resources in Cooperative Control Scenarios. Automatica, vol. 45 (3), pp. 711-715, 2009.

[13] R. Färe. Laws of Diminishing Returns. Lecture Notes in Economics and Mathematical Systems, vol. 176, Springer-Verlag, 1980.

[14] D. P. Bertsekas. Nonlinear Programming. Athena Scientific, 1999.