Detection of Homophilic Communities and Coordination of Interacting Meta-Agents: A Game-Theoretic Viewpoint

Omid Namvar Gharehshiran, William Hoiles, *Student Member, IEEE*, and Vikram Krishnamurthy, *Fellow, IEEE*

***Abstract*—This paper studies two important signal processing aspects of homophilic behavior namely, *detection* of homophilic communities and the distributed *coordination* of meta-agents, which interact with the detected homophilic communities. First, the theory of revealed preferences from microeconomics is used to construct a nonparametric decision test for homophilic behavior using only the time series of external influences and associated agents' responses. These tests rely on rationalizing the dataset of agents' actions as the play from the Nash equilibrium of a concave potential game. A stochastic gradient algorithm is given to optimize the external influence signal in real time to minimize the Type-II error probabilities of the detection test subject to specified Type-I error probability. Using the decision test, methods are provided to detect for homophilic communities. Subsequently, a nonparametric algorithm is presented that uses the constructed potential function for the potential game to predict the preferences of the detected homophilic communities. Second, we present a non-cooperative game model for interaction of meta-agents that interact with the communities and propose an algorithm that prescribes meta-agents how to take actions based on the preference of the communities and past interaction information with other meta-agents. The proposed algorithm has two timescales: the *slow* timescale is the nonparametric preference learning presented in the first part, and the *fast* timescale is a regret-matching stochastic approximation algorithm. It is shown that, if all meta-agents follow the proposed algorithm, their collective behavior is attracted to the correlated equilibria set of the game. This means that meta-agents can co-ordinate their strategies in a distributed fashion as if there exists a centralized coordinating device that they all trust to follow. We provide a real-world example using the energy market, and a numerical example to detect malicious agents in an online social network.**

***Index Terms*—Multiagent signal processing, nonco-operative games, correlated equilibrium, homophily, revealed preferences, Afriat's theorem, stochastic approximation algorithm.**

Manuscript received December 09, 2014; revised May 27, 2015 and November 12, 2015; accepted January 12, 2016. Date of publication January 19, 2016; date of current version February 05, 2016. This research was supported in part by the NSERC Strategic Grant, in part by the SSHRC Insight Grant, and in part by the Canada Research Chairs program. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alejandro R. Ribeiro.

O. N. Gharehshiran is with the Department of Statistical Sciences, University of Toronto, Toronto ON M5S 3G3, Canada (e-mail: omidn@utstat.toronto.edu).

W. Hoiles and V. Krishnamurthy are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: whoiles@ece.ubc.ca; vikramk@ece.ubc.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSIPN.2016.2519766

I. Introduction

HOMOPHILY[1] refers to a tendency of various types of individuals to associate with others who are similar to themselves, and makes individuals' preferences become localized in the sociodemographic space [1], [2]. Consider a social network comprising of several homophilic communities as depicted in Fig. 1. Each such community is characterized by self-interested agents sharing similar preferences over a discrete set of objects (e.g. videos in YouTube, electricity consumption during day vs. night, etc.) resulting from an external influence which affects attributes of objects (e.g. quality of videos, price of electricity during day vs. night, etc.). There also exists a set of meta-agents that make decisions based on the preferences of the detected homophilic communities and other meta-agents. An example is where the homophilic communities represent the power demand of different regions in the power grid, and the meta-agents are network operators that must decide which generators to turn on to meet the demands of the homophilic communities. This paper answers the following questions:

i) Can one detect homophilic communities and compute community preferences using a dataset comprising of a time series of agents' actions in response to a time series of external influences?
ii) Can selfish meta-agents coordinate their decision strategies (in their interaction with other meta-agents) to reach a global coordination that adapts to the evolution of preferences within the detected communities?

Rationalizability, learning, and equilibrium in games are of central importance in answering the above questions. On the one hand, rationalizing a dataset of agents' actions in their interaction with each other via a potential game leads to conclusions about their homophilic behavior. On the other hand, the game-theoretic notion of equilibrium describes a condition of global coordination where all meta-agents are content with their social welfare. Reaching an equilibrium, however, involves a complex process of guessing what others will do. Game-theoretic learning explains how such coordination might arise as a consequence of a long-run process of learning from interactions and adapting behavior [3].

[1] In [1] the following illustrative example is provided for homophily behavior in social networks: "If your friend jumped off a bridge, would you jump too?" A possible reasons for answering "yes" is that you are friends as a result of your fondness for jumping off bridges. Notice that this is different to contagion behavior where "your friend inspired you to jump off the bridge." Due to space restrictions we do not consider contagion behavior in this paper.

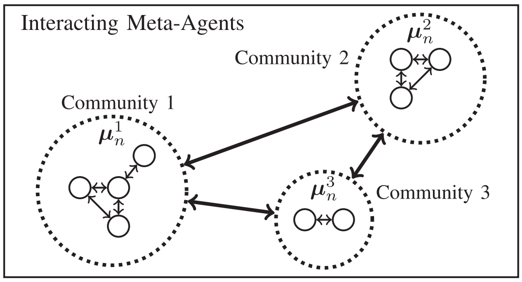

Fig. 1. Schematic of a social network comprising of three interacting meta-agents and three homophilic communities. Agents within each community share similar preferences, denoted by $\boldsymbol{\mu}_n^c$, $c \in \{1, 2, 3\}$, over a discrete set of objects as the result of an external influence, denoted by $\mathbf{p}_n$, which affects attributes of objects.

A. Main Contributions and Organization

The main results in this paper are summarized below:

1) Detection of Homophilic Communities and Preference Estimation: In Sec. II, the theory of revealed preference [4]–[6] from microeconomics is used to construct both a decision test and a statistical test to detect for homophilic communities. The proposed tests only require the time series of agents' actions and external influences that affect agents' actions. The statistical test for homophilic communities has a guaranteed error bound on Type-I errors. Additionally, we provide a stochastic gradient algorithm to reduce the probability of Type-II errors in the statistical test for homophilic behavior. Having detected homophilic communities, Sec. III presents a non-parametric algorithm that can be used to predict the preferences of such communities. Two types of communities are considered: *i) homophilic community*, in which all agents maximize the same utility function and, hence, their actions are identical (if their resource budgets are identical); *ii) generalized homophilic community*, in which agent actions are consistent with play from the Nash equilibrium of a concave potential game. In potential games, the incentive of all agents to change their strategy can be explained using a single global function. Therefore, agents playing a concave potential game form a homophilic community.

The proposed scheme is fundamentally different to the *model-based* theme that is widely used in the signal processing literature in which an objective function (typically convex) is proposed and then algorithms are constructed to compute the minimum. In contrast, the revealed preference approach is *data-centric*; that is, we wish to determine whether the dataset is obtained from the interaction of agents with similar preferences.

2) Preference-Based Coordination Among Meta-Agents: In Sec. IV, we present a non-cooperative game model for preference-based interaction of meta-agents that interact with homophilic communities, and elaborate on correlated equilibrium [7] as the solution concept. As the agents' preferences change over time, so does the utility function of the meta-agents; therefore, we practically deal with a regime-switching game [8].

Section V then combines the results from Sec. II and III with regret-matching game-theoretic learning [9]–[11]. An adaptive two timescale algorithm is described that prescribes meta-agents how to adjust their strategies based on the preferences of the detected homophilic communities and their interaction with other meta-agents. As found in [12], [13], agents preferences typically change on a slow timescale spanning several months. Therefore, the slow timescale corresponds to detection of homophilic communities and computing their preferences. The fast timescale is a stochastic approximation algorithm that relies on the regret-matching learning procedure [9], [10] and prescribe meta-agents which action to take so as to minimize their experienced regrets. We show that, if each meta-agent follows the proposed algorithm[2] independently, its experienced regret can be made arbitrarily small after sufficient interaction with other meta-agents. Moreover, if all meta-agents follow the proposed algorithm, their collective behavior across the social network will asymptotically reach the set of correlated equilibria, and track its evolution as the preferences of communities change over time.

An interesting aspect of both parts described above is the ordinal nature of decision making. In the dataset parsing of Sec. II, the utility function obtained is *ordinal*; in the learning dynamics of Sec. V, the decision strategy is an ordinal function of the experienced regrets. Humans make ordinal decisions[3] since they tend to think in symbolic ordinal terms.

3) Illustration of Results on Real Datasets: We use real datasets of aggregate power consumption from the Ontario energy market social network, obtained from the Independent Electricity System Operator (IESO) website, and show that the power consumption behavior of agents is consistent with players engaged in a concave potential game (i.e. generalized homophilic behavior). We further estimate the community preference for using power, which can be utilized to improve demand side management strategies. We also provide an example on detection of malicious agents in an online social network illustrating the detection of homophilic communities, performance of the statistical test for homophilic communities, estimating community preferences, and the coordination of meta-agents.

B. Literature

In classical revealed preference theory, Afriat's theorem gives a non-parametric finite sample test to detect if an agent's response to an external influence is consistent with utility maximization [16], [17]. Afriat's test is used in [18] for detection of non-interacting malicious agents. For multiple interacting agents (i.e., players in a game), single agent tests are not suitable. Typically, the study of players in a potential game involves parametric assumptions on the utility function of the players. A notable exception is [19] in which a non-parametric detection test for players engaged in a concave potential game is developed for intra-household consumption data based on Varian's

[2]Although we use the term "algorithm," the learning procedure is aimed towards mimicking human behavior; it involves minimizing a moving average regret in the presence of inertia in changing decisions [14], [15].

[3]Humans typically convert numerical attributes to ordinal scales before making decisions. For example, it does not matter if the cost of a meal at a restaurant is \$200 or \$205; an individual would classify this cost as "high." Also, credit rating agencies use ordinal symbols such as AAA, AA, A.

and Afriat's work [16], [17]. If agents' actions are observed in noise, then the detection test [19] cannot be used.

Game theoretic models for social networks have been studied widely [20]–[22]. Regret-matching [9], [11], [23] is known to guarantee convergence to the set of correlated equilibria [7] under no structural assumptions on the game model. The correlated equilibrium arguably provides a natural way to capture conformity to social norms [24]. It can be interpreted as a mediator [25] instructing people to take actions according to some commonly known probability distribution. The regret-based adaptive procedure in [9] assumes a fully connected network topology, whereas the regret-based reinforcement learning algorithm in [23] assumes a set of isolated agents who neither know the game, nor share information of their past decisions with others. In [8], a regret-matching algorithm is developed when agents exchange information over a non-degenerate network connectivity graph.

II. Detection Test for Homophilic Communities

This section uses the principle of revealed preferences to detect whether a group of agents exhibit homophilic behavior and how to cluster agents into homophilic communities. Homophily refers to a tendency of various types of individuals to associate with others who are similar to themselves. Two types of homophilic communities are considered, namely, *homophilic community* and *generalized homophilic community*. A homophilic community is one in which the utility function of all agents are identical and is defined below.

Definition 2.1: A group of agents, indexed by $k = 1, 2, \ldots, K$, form a *homophilic community* if, for all external influence vectors $\mathbf{p} \in \mathbb{R}_+^s$, the associated action vectors of each agent k, denoted by $\mathbf{x}^k \in \mathbb{R}_+^s$, satisfy

$$\mathbf{x} = \frac{\mathbf{x}^1}{\mathbf{p}'\mathbf{x}^1} = \frac{\mathbf{x}^2}{\mathbf{p}'\mathbf{x}^2} = \cdots = \frac{\mathbf{x}^K}{\mathbf{p}'\mathbf{x}^K} \tag{1}$$

with each $\mathbf{x}^k$ resulting from the maximization of the associated agent's utility function:

$$\mathbf{x}^k \in \underset{\{\mathbf{x}^k : \mathbf{p}'\mathbf{x}^k \le B^k\}}{\arg\max} \; u^k(\mathbf{x}^k). \tag{2}$$

In (2), $u^k(\cdot)$ is the utility function of agent k and B^k denotes the resource budget of agent k.

The above definition corresponds to classical revealed preferences where the utility functions of agents are independent of actions taken by others. In this paper, we consider the more general setting, where the utility functions of agents are interdependent. That is, the utility function of each agent depends on the actions of other agents (e.g., agents in a game), as defined below.

Definition 2.2: A group of agents, indexed by $k = 1, 2, \ldots, K$, form a *generalized homophilic community* if, for all external influences $\mathbf{p} \in \mathbb{R}_+^s$, the following conditions hold:

1) Each agent's action, denoted by $\mathbf{x}^k \in \mathbb{R}^s$, satisfies:

$$\mathbf{x}^k = \mathbf{x}^{k^*}(\mathbf{p}, \mathbf{x}^{-k}) \in \underset{\{\mathbf{x}^k : \mathbf{p}'\mathbf{x}^k \le B^k\}}{\arg\max} \; u^k(\mathbf{x}^k, \mathbf{x}^{-k}), \tag{3}$$

where $\mathbf{x}^{-k}$ denotes the actions of all agents excluding agent k and B^k denotes the resource budget of agent k.

2) For all $\mathbf{x}^k, \widehat{\mathbf{x}}^k$ that satisfy $\mathbf{p}'\mathbf{x}^k \le B^k$ and $\mathbf{p}'\widehat{\mathbf{x}}^k \le B^k$, there exists a concave potential function V that satisfies

$$\begin{aligned} &u^k(\mathbf{x}^k, \mathbf{x}^{-k}) - u^k(\widehat{\mathbf{x}}^k, \mathbf{x}^{-k}) > 0 \\ &\quad \text{iff } V(\mathbf{x}^k, \mathbf{x}^{-k}) - V(\widehat{\mathbf{x}}^k, \mathbf{x}^{-k}) > 0 \end{aligned} \tag{4}$$

for all utility functions $u^k(\cdot)$, $k = 1, 2, \ldots, K$.

Notice that, if all agents seek to maximize their associated utility functions via (3), then the agents are engaged in a concave potential game and the action profile of agents satisfies a unique Nash equilibrium of the game.

A. Decision Test for Generalized Homophilic Community

Given a set of K agents, this section provides a decision test to detect if they form a generalized homophilic community. Specifically, we consider a version of Afriat's theorem [16] for deciding if a dataset of responses from a group of agents in a social network can be rationalized as play from the equilibrium of a potential game. In a potential game, the incentive of all agents to change their strategy can be expressed using a single global function. Therefore, agents playing a potential game form a homophilic community. Potential games have been used in telecommunication networking for tasks such as routing, congestion control, power control in wireless networks, and peer-to-peer file sharing [26], and in social networks to study the diffusion of technologies, advertisements, and influence [27].

Consider a group of interacting agents, indexed by $k = 1, 2, \ldots, K$, as depicted in Fig. 2. The question is: Given the following time-series of data from agents

$$\mathcal{D} = \left\{ \left(\mathbf{p}_n, \mathbf{x}_n^1, \ldots, \mathbf{x}_n^K\right) : n \in \{1, 2, \ldots, N\} \right\}, \tag{5}$$

where $\mathbf{p}_n \in \mathbb{R}_+^s$ denotes the external influence signal and $\mathbf{x}_n^k \in \mathbb{R}_+^s$ represents the action of agent k at time n, is it possible to detect if agents are playing a potential game and maximizing their individual utilities? As depicted in Fig. 2, the actions of agents depend both on the external influence $\mathbf{p}_n$ and the actions of the other agents with whom they interact in the social network. The utility function of each agent k has the form $u^k(\mathbf{x}^k, \mathbf{x}^{-k})$, where $\mathbf{x}^k$ denotes the action of agent k, and $\mathbf{x}^{-k}$ represents the joint action profile of all agents excluding agent k. Deb, following Varian's and Afriat's work, shows in [28] that refutable restrictions exist for the dataset $\mathcal{D}$ to satisfy Nash equilibrium; however, these conditions are satisfied by most datasets. Agents engaging in a concave potential game and generating actions that satisfy Nash equilibrium impose stronger restrictions on the dataset $\mathcal{D}$. We refer to this behavior as *Nash rationality*, and formally define it below.

Definition 2.3: The dataset $\mathcal{D}$, defined in (5), is consistent with *Nash equilibrium* play if there exist utility functions $u^k(\mathbf{x}^k, \mathbf{x}^{-k})$ such that (3) holds for each agent k and all data pairs $\left(\mathbf{p}_n, \mathbf{x}_n^1, \ldots, \mathbf{x}_n^K\right)$, $n = 1, 2, \ldots, N$, where B_n^k in (3) denotes the resource budget of agent k at time n, $\mathbf{p}_n \in \mathbb{R}_+^s$, and $u^k(\mathbf{x}^k, \mathbf{x}^{-k})$ is a *non-satiated* utility function in $\mathbf{x}^k$: For

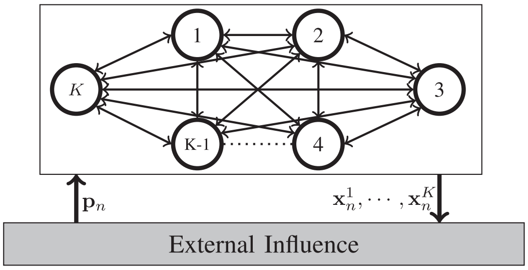

Fig. 2. Schematic of a network comprising of K interacting agents. $\mathbf{p}_n$ and $\mathbf{x}_n^k$ denote the external influence and the action of agent k in response to the external influence and other agents at time n, respectively. The aim is to determine if the dataset $\mathcal{D} = \{(\mathbf{p}_n, \mathbf{x}_n^1, \mathbf{x}_n^2, \ldots, \mathbf{x}_n^K) : n \in \{1, 2, \ldots, N\}\}$ is consistent with play from a Nash equilibrium of players engaged in a concave potential game.

any $\epsilon > 0$, there exists an $\mathbf{x}^k$ with $\|\mathbf{x}^k - \mathbf{x}_n^k\|_2 < \epsilon$ such that $u^k(\mathbf{x}^k, \mathbf{x}_n^{-k}) > u^k(\mathbf{x}_n^k, \mathbf{x}_n^{-k})$.

If for all $\mathbf{x}^k, \widehat{\mathbf{x}}^k$ that satisfy $\mathbf{p}_n' \mathbf{x}^k \leq B_n^k$ and $\mathbf{p}_n' \widehat{\mathbf{x}}^k \leq B_n^k$, there exists a concave potential function V that satisfies (4) for all utility functions $u^k(\cdot)$, $k = 1, 2, \ldots, K$, and for all $\mathbf{p}_n', B_n^k$, $n = 1, 2, \ldots, N$, then the dataset $\mathcal{D}$ satisfies *Nash rationality*.

The following theorem provides necessary and sufficient conditions for a dataset to be consistent with Nash rationality. The proof is analogous to Afriat's Theorem when the concave potential function of the game is differentiable [28], [29].

Theorem 2.1: Given a dataset $\mathcal{D}$, defined in (5), the following statements are equivalent:

1) $\mathcal{D}$ is consistent with Nash rationality (see Definition 2.3) for a K-player concave potential game.
2) Given scalars v_n and $\lambda_n^k > 0$, the set of inequalities

$$v_m - v_n - \sum_{k=1}^{K} \lambda_n^k \mathbf{p}_n' (\mathbf{x}_m^k - \mathbf{x}_n^k) \leq 0. \qquad (6)$$

has a feasible solution for $n, m \in \{1, \ldots, N\}$.
3) A concave potential function that satisfies (3) is given by:

$$\widehat{V}(\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^K) = \min_{n \in \{1,2,\ldots,N\}} \left\{ v_n + \sum_{k=1}^{K} \lambda_n^k \mathbf{p}_n' (\mathbf{x}^k - \mathbf{x}_n^k) \right\}. \qquad (7)$$

4) The dataset $\mathcal{D}$ satisfies the Potential Generalized Axiom of Revealed Preference (PGARP) if the following two conditions are satisfied:
   a) For every dataset $\mathcal{D}_m^k = \{(\mathbf{p}_n, \mathbf{x}_n^k) : n \in \{1, 2, \ldots, m\}\}$, for all $k \in \{1, \ldots, K\}$ and all $m \in \{1, \ldots, N\}$, $\mathcal{D}_m^k$ satisfies GARP (defined in [18]).
   b) The actions $\mathbf{x}_n^k$ originated from agents in a concave potential game. ■

Note that, if only a single agent is considered (i.e. $K = 1$), then Theorem 2.1 is identical to Afriat's Theorem. As pointed out in [17], a remarkable feature of Afriat's Theorem is that, if the dataset can be rationalized by a non-trivial utility function, then it can be rationalized by a continuous, concave, monotonic utility function. "Put another way, violations of continuity, concavity, or monotonicity cannot be detected with only a finite number of demand observations." Similar to Afriat's Theorem, the constructed concave potential function in (7) is ordinal; that is, it is unique up to positive monotone transformations. Therefore, there exist several options for $\widehat{V}(\cdot)$ that would produce identical preference relations to the actual potential function $V(\cdot)$. In 4 of Theorem 2.1, the first condition only provides necessary and sufficient conditions for the dataset $\mathcal{D}$ to be consistent with a Nash equilibrium of a game. Therefore, the second condition is required to ensure consistency with other statements in Theorem 2.1. The intuition that connects 1 and 3 in Theorem 2.1 is provided by the following result from [30]: For any smooth potential game that admits a concave potential function V, a sequence of responses $\{\mathbf{x}_n^1, \mathbf{x}_n^2, \ldots, \mathbf{x}_n^K\}$ are generated by a pure-strategy Nash equilibrium if and only if it is a maximizer of the potential function,

$$\begin{aligned} \mathbf{x}_n &= \left\{\mathbf{x}_n^1, \mathbf{x}_n^2, \ldots, \mathbf{x}_n^K\right\} \\ &\in \underset{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^K}{\arg\max}\ V\left(\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^K\right) \\ &\quad \text{s.t.} \qquad \mathbf{p}_n' \mathbf{x}^k \leq B_n^k \qquad \forall k \in \{1, 2, \ldots, K\} \end{aligned} \qquad (8)$$

for each probe vector $\mathbf{p}_n \in \mathbb{R}_+^s$.

The non-parametric test for Nash rationality involves determining if (6) has a feasible solution. Computing parameters v_n and $\lambda_n^k > 0$ in (6) involves solving a linear program with N^2 linear constraints in $(K+1)N$ variables, which has polynomial time complexity [31]. In the special case of one agent, the constraint set in (6) is the dual of the *shortest path problem* in network flows. The parameters v_n and λ_n in (6) can then be computed using Warshall's algorithm with $O(N^3)$ [17].

B. Statistical Test for Generalized Homophilic Behavior

A real world dataset can fail the Nash rationality test (6) as a result of agents' joint action profile $\mathbf{x}_n$ being corrupted by measurement noise. This section provides a statistical test to detect Nash rationality in the presence of such noise.

Consider the dataset

$$\mathcal{D}_{\text{obs}} = \left\{\left(\mathbf{p}_n, \mathbf{y}_n^1, \mathbf{y}_n^2, \ldots, \mathbf{y}_n^K\right) : n \in \{1, 2, \ldots, N\}\right\}, \qquad (9)$$

consisting of external influences $\mathbf{p}_n$, and noisy observations of agents' actions $\mathbf{y}_n^k = \mathbf{x}_n^k + \mathbf{w}_n^k$, where $\mathbf{w}_n^k$ represents additive measurement noise. Given $\mathcal{D}_{\text{obs}}$, a feasibility test is required to check if the clean dataset $\mathcal{D}$ satisfies Nash rationality. Let H_0 and H_1 denote the null hypothesis that the clean dataset $\mathcal{D}$ satisfies Nash rationality, and the alternative hypothesis that $\mathcal{D}$ does not satisfy Nash rationality, respectively. In devising a statistical test for H_0 versus H_1, there are two possible sources of error:

Type-I error: Reject H_0 when H_0 is valid,
Type-II error: Failure to reject H_0 when H_0 is invalid. (10)

Given the noisy dataset $\mathcal{D}_{\text{obs}}$, the following statistical test can be used to detect if a group of agents select actions that satisfy Nash equilibrium (3) of a concave potential game and, hence, exhibit homophilic behavior:

$$\int_{\Phi^*(\mathbf{y})}^{+\infty} f_M(\psi) \mathrm{d}\psi \underset{H_1}{\overset{H_0}{\gtrless}} \gamma. \qquad (11)$$

Here, $\mathbf{y} = \{\mathbf{y}_n^1, \ldots, \mathbf{y}_n^K\}_{n=1}^N$ is the set of noisy observations of agents' actions to external influences $P = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N]$, γ is the *significance level* of the statistical test, and the *test statistic* $\Phi^*(\mathbf{y})$ is the solution of the following constrained optimization problem:

$$\begin{aligned} &\min \ \Phi \\ &\text{s.t. } v_m - v_n - \sum_{k=1}^K \lambda_n^k \mathbf{p}_n'(\mathbf{y}_m^k - \mathbf{y}_n^k) - \sum_{k=1}^K \lambda_n^k \Phi \le 0, \\ &\quad \Phi \ge 0, \quad \lambda_n^k > 0 \quad \text{for} \quad n, m \in \{1, 2, \ldots, N\} \end{aligned} \tag{12}$$

where the minimization is performed over the variables $\{v_n, \lambda_n^1, \ldots, \lambda_n^K, \xi_n^1, \ldots, \xi_n^K\}_{n=1}^N$. Further, f_M is the probability density function of the random variable M, defined as

$$M \equiv \max_{n,m} \left[\sum_{k=1}^K |\mathbf{p}_n'(\mathbf{w}_n^k - \mathbf{w}_m^k)|\right]. \tag{13}$$

Note that (12) is non-convex due to $\sum_k \lambda_n^k \Phi$; however, since the objective function is given by the scalar Φ, for any fixed value of Φ, (12) becomes a set of linear inequalities allowing feasibility to be straightforwardly determined [18]. Further, the noise distribution of $\mathbf{w}$ must have bounded support and be known to apply the statistical test (11). Given a sample of actions, the noise distribution can be estimated using methods such as the Kolmogorov-Smirnov test, Shapiro-Wilk test, and Anderson-Darling test. Non-parametric density estimators can also be used to estimate the noise distribution from sample data; the reader is referred to [32] for further details. In the latter case, (11) becomes a *non-parametric* statistical test for homophilic behavior, however the Type-I error probability of the test cannot be guaranteed as the upper bound for M (13) may be underestimated.

The main idea of the statistical test (11) is to first compute how the actions of agents must be adjusted to satisfy homophilic behavior (6), then test if the adjustment is "statistically significant". In (11), the adjustment parameter is denoted by Φ. Given a dataset $\mathcal{D}_{\text{obs}}$, if the solution of (12) is $\Phi = 0$, then $\mathcal{D}_{\text{obs}}$ satisfies Nash rationality (6) and, hence, homophilic behavior directly. For $\Phi > 0$, the random variable M, defined in (13), upper bounds the adjustment Φ for $\mathcal{D}_{\text{obs}}$ to satisfy homophilic behavior. The integral in (11) computes the probability that $\Phi \le M$. For the dataset $\mathcal{D}_{\text{obs}}$ to satisfy the statistical test (11), the probability of $\Phi \le M$ has to be greater than the significance level γ of the test.

The following theorem characterizes the performance of the above statistical test.

Theorem 2.2: Consider the noisy dataset $\mathcal{D}_{\text{obs}}$, defined in (9). The probability that the statistical test (11) yields a Type-I error (rejects H_0 when it is true) is less then γ.

Proof: See Appendix A.

A weakness with (11) is that there is no guaranteed upper bound on Type-II errors (accept H_0 when it is false) for a general external influence $\mathbf{p}_n$. However, as K increases, the probability of Type-II errors decreases. As a guideline for single agents, if the additive noise satisfies a normal distribution, then the magnitude of the measurement error should be on the scale of 5% of the maximum possible actions of the agents [33].

The statistical test (11) can be enhanced by adaptively optimizing the external influences $P = [\mathbf{p}_1, \ldots, \mathbf{p}_N]$ to reduce Type-II errors via

$$P^* \in \operatorname*{arg\,min}_{P \in \mathbb{R}_+^{s \times N}} J(P)$$

$$J(P) = \underbrace{\mathbb{P}\left(\int_{\Phi^*(\mathbf{y})}^{+\infty} f_M(\beta)\mathrm{d}\beta > \gamma \,\middle|\, \{P, \mathbf{x}(P)\} \in \mathcal{P}\right)}_{\text{Probability of Type-II error}}. \tag{14}$$

Here,

$$\mathbf{x} = \left\{\left(\mathbf{x}_n^1(\mathbf{p}_n), \ldots, \mathbf{x}_n^K(\mathbf{p}_n)\right)\right\}_{n=1}^N$$

denotes the set of agents' responses to external influences P, f_M is the probability density function of the random variable M defined in (13), and $J(\cdot)$ is the conditional probability that the test (11) accepts H_0 for all agents given that H_0 is false. Further, the set $\mathcal{P}$ contains all pairs $\{P, \mathbf{x}(P)\}$ that do not satisfy Nash rationality (see Definition 2.3). The objective function in (14) cannot be evaluated analytically as the set $\{P, \mathbf{x}(P)\} \in \mathcal{P}$ has no closed form solution. Therefore, only noisy estimates can be made of the objective function in (11). In real-world applications, there must exist a source of noisy observations in which the underlying actions of agents do not satisfy utility maximization; see [18] for a single agent example.

Remark 2.1: For a known discrete and finite action space, the set $\mathcal{P}$ can be constructed using the results from [34]. In this case, the optimal probe signal P^* can be selected using geometric methods such as the Chebyshev center that selects P^* such that $\{P^*, \mathbf{x}(P^*)\}$ is as far as possible from the boundary of $\mathcal{P}$.

Computing (14) requires a stochastic optimization algorithm as the probability density function f_M is unknown. We use the simultaneous perturbation stochastic gradient (SPSA) method [35]. Let $I(Y)$ denote the indicator function: $I(Y) = 1$ if Y is true, and 0 otherwise. Let further $F_M(\cdot)$ denote an estimate of the cumulative distribution function of M constructed by generating random samples according to (13). The SPSA algorithm for optimization of external influences is then summarized in Algorithm 1. The benefit of using the SPSA algorithm is that the estimated gradient $\widehat{\nabla}_P J_q(P_q)$ in (16) can be computed using only two measurements of the function (15) per iteration; see [35] for tutorial exposition and the convergence properties of the SPSA algorithm. For constant step-size ε, it is shown in [18] that the algorithm converges weakly (in probability) to a local stationary point.

C. Constructing Homophilic and Generalized Homophilic Communities

Here we consider methods for detecting and constructing homophilic and generalized homophilic communities for a given dataset $\mathcal{D}$, defined in (5). The methods presented in this section rely on the ability to detect if the actions of agents are a result of the maximization of their associated utility functions, (2) for homophilic communities and (3) for generalized homophilic communities.

To detect the homophilic communities in $\mathcal{D}$, we directly apply the utility maximization condition (6) given in Definition 2.1, where only one agent is considered at a time. If a feasible solution exists for the agents, then the associated

Algorithm 1. SPSA for Optimization of External Influences

Initialization: Choose initial probe

$$P_0 = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N] \in \mathbb{R}_+^{s\times N},$$

and the significance level γ in (11). Set further L, which controls the accuracy of the empirical probability of Type-II errors (14), the gradient step-size $\sigma > 0$, and the step-size $0 < \rho < 1$ for updating the probe vector.

Step 1: Estimate the cost (probability of Type-II errors) via

$$\widehat{J}_q(P_q) = \frac{1}{L}\sum_{l=1}^{L} I\left(F_M(\Phi^*(\mathbf{y}_l)) \le 1-\gamma\right), \tag{15}$$

where $\Phi^*(\mathbf{y}_l)$ is computed using (12) with the noisy observations $\mathbf{y}_l = \mathbf{x}(P_q) + \mathbf{w}_l$. Note that $\mathbf{w}_l$ is a fixed realization of the additive measurement noise $\mathbf{w} \in \mathbb{R}^{s\times N\times K}$.

Step 2: Compute the gradient estimate $\widehat{\nabla}_P J_q(P_q)$:

$$\widehat{\nabla}_P J_q = \frac{\widehat{J}_q\left(P_q + \Delta_q\sigma\right) - \widehat{J}_q\left(P_q - \Delta_q\sigma\right)}{2\sigma\Delta_{qr}}, \tag{16}$$

where Δ_{qr} is the realization of the Bernoulli random variable $\Delta \in \mathbb{R}^{s\times N}$.

Step 3: Update the probe vector P_q:

$$P_{q+1} = P_q - \rho\widehat{\nabla}_P J_q(P_q). \tag{17}$$

Recursion. Set $q \leftarrow q+1$, and go Step 1.

homophilic communities can be constructed using standard cluster analysis methods such as *density-based spatial clustering of applications with noise* (DBSCAN) [36] given the normalized actions of agents (i.e. $\mathbf{x}^k/\mathbf{p}'\mathbf{x}^k$).

To detect if agents form a generalized homophilic community, it is required to check if agents actions can be rationalized by the play from a concave potential game. Using an extension of Afriat's theorem [16], Sec. II-A provided a test (6) to detect if such a concave potential function exists given the dataset $\mathcal{D}$. What if the agents $\{1,2,\ldots,K\}$ fail to satisfy (3)? This may result as one or more of the agents are not maximizing their associated utility functions. In this case, detecting if agents form a generalized homophilic community is computationally demanding as we have to apply the decision test (6) to detect which partitions of $\{1,2,\ldots,K\}$, in which the cardinality of each disjoint subset is strictly greater than one, satisfy (3). However, for detecting the generalized homophilic communities with the largest number of agents, this can be formulated using a series of convex mixed-integer nonlinear programming (MINLP) problems as described in Algorithm 2. Convex MINLPs can be solved using a variety of methods in the combinatorial optimization literature [37]–[39]. Convex MINLP problems are NP-hard; however, a suboptimal solution can be found by dropping the integrity constraints of the convex MINLP converting it into a convex nonlinear programming problem that can be solved efficiently [37]–[39]. There are also numerical methods which can both detect if

Algorithm 2. Generalized Homophilic Communities

Initialization: Given the dataset $\mathcal{D}$, defined in (5), remove all agents that are independent utility maximizers, i.e., their actions satisfy (2). Index the resulting reduced set by $k' \in \mathcal{K}' = \{1,2,\ldots,K'\}$.

Step 1: Detect for the largest community in $\mathcal{K}'$ by solving the following convex mixed-integer nonlinear program:

$$\mathbf{b}^c \in \underset{\mathbf{b}=[b^1,\cdots,b^{K'}]}{\arg\max} \left\{\textstyle\sum_{k=1}^{K'} b^k\right\} \tag{19}$$

$$\text{s.t.} \quad v_m - v_n - \textstyle\sum_{k=1}^{K'} b^k\lambda_n^k\mathbf{p}_n'(\mathbf{x}_m^k - \mathbf{x}_n^k) \le 0,$$
$$\lambda_n^k > 0, \quad \mathbf{b} \in \{0,1\}^{K'}.$$

If $|\mathbf{b}^c| < 2$, then no generalized homophilic community exists in $\mathcal{K}'$. Else, a generalized homophilic community does exist with the agents given by the non-zero entries in $\mathbf{b}^c$.

Step 2: Construct the new set $\mathcal{K}'$ with all unassigned agents from Step 1. If $|\mathcal{K}'| \ge 2$, then return to Step 1 to detect for the existence of the next generalized homophilic community $c+1$. Else, if $|\mathcal{K}'| < 2$, no further generalized homophilic communities exist.

the convex MINLP has an optimal solution, and can guarantee termination at the optimal solution [39]. Prior to detecting any generalized homophilic communities, all agents that cannot be part of a community are removed in the Initialization step of Algorithm 2. This relies on the property that the utility functions of agents engaged in a concave potential game must satisfy the following symmetry relation [28]:

$$\frac{\partial^2 u^k}{\partial\mathbf{x}^k\partial\mathbf{x}^i} = \frac{\partial^2 u^i}{\partial\mathbf{x}^i\partial\mathbf{x}^k} \quad \forall k,i \in \{1,2,\ldots,K\}. \tag{18}$$

If an agent maximizes its utility function independent of the actions of other agents, then it cannot satisfy (18) and is not engaged in a concave potential game. In Step 1, a convex MINLP is formulated to detect the largest possible generalized homophilic community given the set of agents $\mathcal{K}'$. Note that the maximization of the MINLP is taken over the number of agents whose actions satisfy the decision test (6). Of course, if $|\mathcal{K}'|$ is sufficiently small, then brute force enumeration can be applied to detect for the largest generalized homophilic community. The process of constructing communities continues until the number of non-zero entries in $\mathbf{b}$ is less then two or the cardinality of $\mathcal{K}'$ is less then two—that is, until it is impossible to construct a generalized homophilic community. In Sec. VI-B we provide a numerical example illustrating how Algorithm 2 can be applied to detect for generalized homophilic communities given a time-series $\mathcal{D}$ (5) of external influences and responses.

III. Community Preferences

This section presents a method to compute the preferences of homophilic or generalized homophilic communities, detected using the methods presented in Sec. II. This allows us to detect which actions the agents in a community prefer. For example, in a power grid, a network operator can use knowledge of the power usage preferences of the community which they serve so as to select which power generation facilities need to be activated to meet the power demand of the community. In this section we provide a measure of the community preferences based on the constructed utility function, for homophilic communities, and concave potential function, for generalized homophilic communities.

Consider a set of (generalized) homophilic communities indexed by $c \in \mathcal{C} = \{1, \cdots, C\}$, and suppose each community c comprises of agents indexed by the set $\mathcal{K}_c = \{1, 2, \ldots, K_c\}$. The community preference is formally defined below.

Definition 3.4: Given the external influence $\mathbf{p}_n$, the community preference vector $\boldsymbol{\mu}_n^c = [\mu_n^c(1), \ldots, \mu_n^c(s)] \in \mathbb{R}_+^s$ for a discrete set of s objects is evaluated by:

$$\mu_n^c(i) = \frac{\sum_{k\in\mathcal{K}_c} x_n^k(i)}{\sum_{k\in\mathcal{K}_c}\sum_{w=1}^{s} x_n^k(w)}, \quad i = 1, 2, \ldots, s, \tag{20}$$

where $x_n^k(i)$ represents the action of agent k in community c associated with object i in response to the external influence $\mathbf{p}_n$.

Note that the denominator of (20) is the total resources used in community c in response to $\mathbf{p}_n$. To compute the community preferences $\boldsymbol{\mu}_n^c$, the actions $\mathbf{x}_n^k$ of all agents $k \in \mathcal{K}_c$ have to be jointly evaluated for each external influence $\mathbf{p}_n$ and social budget B_n^k. To this end, we use the concave potential function of the community formed by the agents. For a dataset from community c:

$$\mathcal{D}_{\text{obs}}^c = \left\{ \left(\mathbf{p}_n, \mathbf{y}_n^1, \ldots, \mathbf{y}_n^{K_c}\right) : n \in \{1, 2, \ldots, N\} \right\},$$

the following minimum distance criterion can be used to estimate $\{v_n, \lambda_n, \xi_n^k\}$ in order to construct a concave potential function:

$$\begin{aligned} &\min_{\xi_n^k} \sum_{n=1}^{N} \sum_{k\in\mathcal{K}_c} \left(\mathbf{y}_n^k - \xi_n^k\right)' \left(\mathbf{y}_n^k - \xi_n^k\right) \\ &\text{s.t. } v_m - v_n - \sum_{k\in\mathcal{K}_c} \lambda_n^k \mathbf{p}_n' \left(\mathbf{y}_m^k - \mathbf{y}_n^k + \xi_n^k - \xi_m^k\right) \le 0, \\ &\qquad \lambda_n^k > 0 \quad \text{for} \quad n, m \in \{1, 2, \ldots, N\}. \end{aligned} \tag{21}$$

Here, $\xi_n^k \in \mathbb{R}^s$ denotes the minimum deviation of the observed data $\mathbf{y}_n^k$ necessary for $\mathbf{x}_n^k$ to have originated from a homophilic community. Specifically, the solution of (21) provides an estimate of the agents' actions $\mathbf{x}_n^k = \mathbf{y}_n^k - \xi_n^k$ which satisfy homophilic behavior (6) and are unique almost everywhere [40]. A solution to (21) can be computed efficiently using numerical methods such as sequential quadratic programming [41]. Note that (21) will always have a feasible solution since the main assumption here is that agents in community c satisfy homophilic behavior, i.e., the dataset $\mathcal{D}_{\text{obs}}$ satisfies the statistical test (11). Eq. (12) is utilized to compute the minimum deviation $\Phi \in \mathbb{R}$ of all the agents' actions in $\mathcal{D}_{\text{obs}}$ to satisfy Nash rationality. If satisfied, then the quadratic maximization (21) computes the minimum deviation of each agent's action, denoted by $\xi_n^k \in \mathbb{R}^s$, required for $\mathcal{D}_{\text{obs}}$ to satisfy Nash rationality. A quadratic cost is used because the minimum deviation ξ_n^k in the sum of the objective function may be positive or negative. Recall that, if $\mathcal{D}_{\text{obs}}$ satisfies homophilic behavior with $\xi_n^k = 0$, then (6) can be used to evaluate the parameters $\{v_n, \lambda_n\}$ to construct a concave potential function.

Having computed the parameters $\{v_n, \lambda_n, \eta_n^k\}$, agents' actions can be evaluated from the following optimization problem for a given $\mathbf{p}_o \in \mathbb{R}_+^s$ and $B_o^k \in \mathbb{R}_+$:

$$\begin{aligned} \widehat{\mathbf{x}}_o^c =& \{\widehat{\mathbf{x}}_o^1, \widehat{\mathbf{x}}_o^2, \ldots, \widehat{\mathbf{x}}_o^{K_c}\} \\ &\in \underset{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^{K_c}}{\arg\max} \ \widehat{V}\left(\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^{K_c}\right) \\ &\quad \text{s.t.} \quad \mathbf{p}_o'\mathbf{x}^k \le B_o^k \qquad \forall k \in \{1, 2, \ldots, K_c\} \end{aligned} \tag{22}$$

where $\widehat{V}(\cdot)$ is given by (7) and $\mathbf{x}_n^k = \mathbf{y}_n^k - \xi_n^k$. Note that, if $\mathbf{p}_n$ and B_n^k are selected for the optimization (22), then the result will be $\widehat{\mathbf{x}}_n^c = \{\widehat{\mathbf{x}}_n^1, \widehat{\mathbf{x}}_n^2, \ldots, \widehat{\mathbf{x}}_n^{K_c}\}$. The optimization problem in (22) is a linear program with a piecewise linear objective and can be solved in polynomial time. Having estimated agents' actions via (22), the community preferences $\boldsymbol{\mu}_n^c$ can be evaluated using (20). The above procedure for computing community preferences is summarized below Algorithm 3.

Algorithm 3. Computing Community Preferences

Initialization: Select an external influence $\mathbf{p}_o \in \mathbb{R}_+^s$.

Step 1: For a dataset $\mathcal{D}_{\text{obs}}^c$ that satisfies homophilic behavior (11), use (21) to compute the parameters

$$\{v_n, \lambda_n^1, \ldots, \lambda_n^K, \xi_n^1, \ldots, \xi_n^K\}_{n=1}^N.$$

Step 2: Solve the following linear programming problem:

$$\begin{aligned} &\widehat{\mathbf{x}}_o^c \in \max_{\mathbf{x}_o^1, \mathbf{x}_o^2, \ldots, \mathbf{x}_o^{K_c}} z \\ &\text{s.t. } z \le v_n + \sum_{k\in\mathcal{K}_c} \lambda_n^k \mathbf{p}_n'(\mathbf{x}_o^k - \mathbf{y}_n^k + \xi_n^k) \quad \forall n \in \{1, \ldots, N\} \\ &\mathbf{p}_o'\mathbf{x}_o^k \le 1, \quad x_o^k(i) \ge 0 \quad \forall k \in \mathcal{K}_c, \forall i \in \{1, \ldots, m\} \end{aligned} \tag{23}$$

to estimate agents' actions $\widehat{\mathbf{x}}_o^c$ (22) with all agents having a resource budget $B^k = 1$.

Step 3: Compute the community preferences $\boldsymbol{\mu}^c$ by plugging $\widehat{\mathbf{x}}_o^c$ into (20).

IV. Game-Theoretic Model for Preference-Based Interaction of Meta-Agents

Having presented a statistical test to detect homophilic communities of agents and a nonparametric algorithm to predict their preferences, this section formulates the interaction among meta-agents whose actions are dependent on the preferences of the communities. As an example, consider a corporate network of operators in the power grid that must decide which power generation stations to switch on to meet the power demand of consumers. Each operator in the corporate network does not only consider minimizing the cost of switching on a generator or purchasing power from other operators, but also accounts for the preferences for power usage of agents in their associated community. Given that agent preferences change over a slow timescale relative to the power demands, and that each operator attempts to maximize its associated utility, the network of operators can be considered to be engaged in a non-cooperative game. Since the community preferences change over time, the associated utility function of the meta-agents also changes. Hence, we practically deal with a regime switching non-cooperative game. This section introduces the network-based non-cooperative game model, and a variant of correlated equilibrium adapted to such a time-varying environment as the solution concept for the formulated game.

A. Network-Level Non-Cooperative Game Model

The standard representation of a non-cooperative game, known as *normal form*, for a network incorporating several homophilic communities is comprised of three elements:

1. Set of meta-agents: Each meta-agent is associated with a (generalized) homophilic community detected using the tools in Sec. II-III. Therefore, we use the same indexing for meta-agents as that used for communities. That is, the set of meta-agents is denoted by $\mathcal{C} = \{1, \cdots, C\}$ and individual meta-agents are indexed by $c \in \mathcal{C}$

2. Set of actions: Each meta-agent c repeatedly takes actions a^c from the set $\mathcal{A}^c = \{1, \cdots, A^c\}$. A generic *action profile* of all meta-agents is denoted by

$$\mathbf{a} = \left(a^c, \mathbf{a}^{-c}\right), \quad \text{where} \quad \mathbf{a}^{-c} = \left(a^1, \cdots, a^{c-1}, a^{c+1}, \cdots, a^C\right)$$

denotes the action profile of all meta-agents excluding c. Note that, in contrast to individual agent's actions that are in the continuous domain, meta-agents' actions are typically discrete as they make high-level decisions based on the preferences of agents they represent.

3. Utility function: For each meta-agent c, the utility function takes the form $U^c(a^c, \mathbf{a}^{-c}, \boldsymbol{\mu}^c)$, where $\boldsymbol{\mu}^c \in \mathbb{R}^s$ represents the preference of (generalized) homophilic community c. The interpretation of such a utility function is the aggregated rewards and costs associated with the chosen action as the outcome of the interaction with other meta-agents. The utility function can be quite general: It could reflect reputation or privacy [42], [43], and benefits/costs associated with maintaining links in a social network [44], [45] or with production, download, and upload in content production and sharing over peer-to-peer networks [46].

Note that meta-agents' actions are interdependent as the utility realized by each meta-agent is a function of other meta-agents' actions. Each meta-agent's action also depends on the preferences of the (generalized) homophilic community with which it interacts.

B. Correlated Equilibrium

The game-theoretic concept of equilibrium describes a condition of global coordination where all players are content with the outcome. Here, we focus on the correlated equilibrium [7], defined as follows:

Definition 4.5 (Correlated Equilibrium): Let $\boldsymbol{\pi}$ denote a joint distribution on the joint action space $\mathcal{A}^{\mathcal{C}} = \times_{c=1}^{C} \mathcal{A}^c$ of all meta-agents, i.e.,

$$\pi(\mathbf{a}) \geq 0, \ \forall \mathbf{a} \in \mathcal{A}^{\mathcal{C}}, \ \text{and} \ \textstyle\sum_{\mathbf{a} \in \mathcal{A}^{\mathcal{C}}} \pi(\mathbf{a}) = 1.$$

The set of *correlated equilibria* $\mathcal{Q}(\boldsymbol{\mu})$ for each preference vector $\boldsymbol{\mu} = [\boldsymbol{\mu}^1, \ldots, \boldsymbol{\mu}^C]$ is the convex polytope: [see (24) at the bottom the page], where $\pi^k(i, \mathbf{a}^{-k})$ denotes the probability that agent k picks action i and the rest $\mathbf{a}^{-k}$.

Note in (24) that $\mathcal{Q}(\boldsymbol{\mu})$ changes with time as the preferences of communities $\boldsymbol{\mu}^c$ evolve over time. Several reasons motivate adopting the correlated equilibrium in large-scale networks. It is structurally and computationally simpler than the Nash equilibrium. The coordination among agents in the correlated equilibrium can further lead to potentially higher utilities than if agents take their actions independently (as required by Nash equilibrium) [7]. Finally, it is more realistic as the observation of the common history of actions (or realized utilities) naturally correlates agents future decisions [23].

An intuitive interpretation of correlated equilibrium is "coordination in decision-making." Suppose a mediator is observing a repeated interactive decision making process among multiple selfish meta-agents. The mediator, at each period, gives private recommendations as what action to take to each meta-agent. The recommendations are correlated as the mediator draws them from a joint probability distribution on the action profile of all meta-agents; however, each meta-agent is only given recommendations about its own decision. Each meta-agent can freely interpret the recommendations and decide if to follow. A correlated equilibrium results if neither of meta-agents wants to deviate from the provided recommendation. That is, in correlated equilibrium, meta-agents' decisions are coordinated as if there exists a global coordinating device that all meta-agents trust to follow.

V. Adaptive Preference-Based Decision Making

This section presents a two timescale stochastic approximation algorithm that combines the detection test and preference

$$\mathcal{Q}(\boldsymbol{\mu}) = \{\boldsymbol{\pi} : \textstyle\sum_{\mathbf{a}^{-c}} \pi^c\left(i, \mathbf{a}^{-c}\right)\left[U^c\left(j, \mathbf{a}^{-c}, \boldsymbol{\mu}^c\right) - U^c\left(i, \mathbf{a}^{-c}, \boldsymbol{\mu}^c\right)\right] \leq 0, \ \forall i, j \in \mathcal{A}^c, c \in \mathcal{C}\} \tag{24}$$

$$\begin{aligned} r_n^c(i,j) = &(1-\varepsilon)^{n-1}\left[U^c\left(j, \mathbf{a}_1^{-c}, \boldsymbol{\mu}_1\right) - U^c\left(i, \mathbf{a}_1^{-c}, \boldsymbol{\mu}_1\right)\right] \cdot I(a_1^c = i) \\ &+ \varepsilon \textstyle\sum_{2 \leq \tau \leq n} (1-\varepsilon)^{n-\tau}\left[U^c\left(j, \mathbf{a}_\tau^{-c}, \boldsymbol{\mu}_\tau\right) - U^c\left(i, \mathbf{a}_\tau^{-c}, \boldsymbol{\mu}_\tau\right)\right] \cdot I(a_\tau^c = i) \end{aligned} \tag{25}$$

$$r_{n+1}^c(i,j) = r_n^c(i,j) + \varepsilon\left(\left[U^c\left(j, \mathbf{a}_n^{-c}, \boldsymbol{\mu}_n\right) - U^c\left(i, \mathbf{a}_n^{-c}, \boldsymbol{\mu}_n\right)\right] \cdot I\left(a_n^c = i\right) - r_n^c(i,j)\right) \tag{26}$$

prediction procedure of Sec. II and III with a regret-matching learning functionality [9], [23]. This algorithm enables meta-agents to adapt their decisions to preferences of the communities they interact with, yet obtain a sophisticated and rational global behavior at the network level, namely, correlated equilibrium.

A. Two Timescale Adaptive Learning Algorithm

Time is discrete $n = 1, 2, \ldots$. At each time n, the meta-agent makes a decision a_n^c according to a randomized strategy $\boldsymbol{\phi}^c(R_n^c)$, a probability distribution over the action space $\mathcal{A}^c$, which is a function of a *regret* matrix $R_n^c = [r_n^c(i,j)]$. Each element $r_n^c(i,j)$ records the discounted time-averaged regrets—losses in utilities—had the meta-agent selected action j every time it played action i in the past. More precisely, $r_n^c(i,j)$ is defined by: [see (25) at the bottom of the previous page], and is updated via the recursive expression: [see (26) at the bottom of the previous page]. Here, $0 < \varepsilon \ll 1$ is a small parameter that represents the adaptation rate of the strategy update procedure, and is required when meta-agents face repeated decision making in a non-stationary environment (due to changes in community preference over time) [8]. Further, $I(Y)$ denotes the indicator operator: $I(Y) = 1$ if statement Y is true, and 0 otherwise. Positive $r_n^c(i,j)$ implies the opportunity to gain by switching from action i to j in future. Therefore, the regret-matching learning procedure [9] assigns positive probabilities to all actions j for which $r_n^c(i,j) > 0$. In fact, the probabilities of switching to different actions are proportional to their regrets relative to the current action, hence the name "regret-matching". This procedure is viewed as particularly simple and intuitive as no sophisticated updating, prediction, or fully rational behavior is required by the meta-agents.

The proposed stochastic approximation algorithm is summarized in the following protocol that mimics human's learning process [14], [15]:

Decision Protocol

Step 1: Chooses action a_n^c randomly from a weight vector (probabilities) $\boldsymbol{\phi}^c(R_n^c)$. This weight vector is an ordinal function[4] of regret due to its previous actions.

Step 2: Update regrets R_n^c based on the chosen action, observed decisions of other meta-agents and the homophilic community preferences $\boldsymbol{\mu}^c$, according to (25). (The exponential discounting places more importance on recent actions.)

Let $|x|^+ = \max\{0, x\}$. Below we abstract the above decision protocol into Algorithm 4 so as to facilitate analysis of the global behavior. The two timescales in Algorithm 4 can be explained as follows: Since the preferences $\boldsymbol{\mu}^c$ typically change on a slow timescale spanning several months [12], [13], each meta-agent c collects responses of agents within each community over a longer period (e.g. a month) and runs Algorithm 3 to obtain the community preference vector $\boldsymbol{\mu}^c$. In contrast, the high-level decisions of meta-agents a_n^c made by Algorithm 4 can be updated more frequently (even on an hourly basis). For instance, the edge servers in a content distribution network can update their caching decisions on a faster time as compared with the preferences of the users in their locale. The subscript n in Algorithm 4 represents the fast timescale, wherein $\boldsymbol{\mu}_n^c$ is a slow jump changing process. That is, it remains constant until N data points of agents' actions $\mathbf{y}_n^c$ and probe signal $\mathbf{p}_n$ are collected, at which point it jumps into its new value computed via Algorithm 3.

Algorithm 4. Regret-Matching With Diffusion Cooperation

Initialization: Set

$$\nu^c > A^c \left| U_{\max}^c - U_{\min}^c \right|,$$

where $U_{\max}^c$ and $U_{\min}^c$ denote the upper and lower bounds on the utility function, respectively. Set the step-size $0 < \varepsilon \ll 1$, sample a_0^c uniformly from $\mathcal{A}^c$, and initialize $R_0^c = \mathbf{0}$.

Step 1: Choose Action Based on Past Regret.

$$a_n^c \sim \boldsymbol{\phi}^c(R_n^c) = \left(\phi_1^c(R_n^c), \ldots, \phi_A^c(R_n^c)\right),$$

where

$$\phi_i^c(R_n^c) = \begin{cases} \frac{1}{\nu^c} \left| r_n^c\left(a_{n-1}^c, i\right) \right|^+, & i \neq a_{n-1}^c, \\ 1 - \sum_{j \neq i} \phi_j^c(R_n^c), & i = a_{n-1}^c. \end{cases} \tag{27}$$

Step 2: Update Individual Regret.

$$R_{n+1}^c = R_n^c + \varepsilon \left[F^c\left(a_n^c, \mathbf{a}_n^{-c}, \boldsymbol{\mu}_n^c\right) - R_n^c \right] \tag{28}$$

where $F^c = [f_{ij}^c]$ is an $A^c \times A^c$ matrix with elements

$$f_{ij}^c = \left[U^c\left(j, \mathbf{a}_n^{-c}, \boldsymbol{\mu}_n^c\right) - U^c\left(i, \mathbf{a}_n^{-c}, \boldsymbol{\mu}_n^c\right) \right] \cdot I\left(a_n^c = i\right).$$

Step 3: Compute Community Preferences.
If $|\mathcal{D}_{\text{obs}}^c| = N$, compute $\boldsymbol{\mu}_n^c$ via Algorithm 3, and set $\mathcal{D}_{\text{obs}}^c = \emptyset$. Otherwise, set $\boldsymbol{\mu}_{n+1}^c = \boldsymbol{\mu}_n^c$, collect new probe signal $\mathbf{p}_n$ and user responses $\mathbf{y}_n^c$, and update $\mathcal{D}_{\text{obs}}^c$:

$$\mathcal{D}_{\text{obs}}^c \leftarrow \mathcal{D}_{\text{obs}}^c \cup \left(\mathbf{p}_n, \mathbf{y}_n^c\right). \tag{29}$$

Recursion. Set $n \leftarrow n + 1$, and go Step 1.

B. Discussion and Intuition

Distinct properties of Algorithm 4 are as follows:

1) Inertia: The choice of ν^c in the decision strategy (27) guarantees that there is always a positive probability of picking the same action as the last period. Therefore, ν^c can be viewed as an *inertia* parameter. It mimics humans' decision making process and plays a significant role in breaking away from the so-called "bad cycles." This inertia is, in fact, the very factor that makes convergence to the correlated equilibria set possible under (almost) no structural assumptions on the underlying game [9].

2) Ordinal Choice of Actions: The decision strategy (27) is an ordinal function of the experienced regrets. Actions are

[4]An ordinal function orders pairs of alternatives such that one is considered to be worse than, equal to, or better than the other.

ordered based on the regret values with the exception that all actions with negative regret are considered to be equally desirable—see footnote 3.

3) Nonlinear Adaptive Filtering: Let us rearrange the elements of the regret matrix R^c as a vector in $(A^c)^2$-dimensional space and, with slight abuse of notation, still denote it by R^c. The distance of any vector R^c to the negative orthant is defined by

$$D(R^c) = \frac{1}{2}\sum_{i,j\in\mathcal{A}^c}\left(|r^c(i,j)|^+\right)^2.$$

The decision strategy (27) can then be expressed by

$$\phi^c(R_n^c) = \left.\frac{\nabla_{ij}D}{\nu^c}\right|_{R^c=R_n^c} = \frac{|r_n^c(i,j)|^+}{\nu^c}.$$

Therefore, Steps 1 and 2 of Algorithm 4 can be interpreted as a nonlinear adaptive filtering algorithm that aims to minimize the loss function $D(R_n^c)$ that characterizes how far each meta-agent is from the ideal behavior, namely, experiencing no regrets due to previous decisions, by properly choosing future actions. The adaptation process is based on learning from the sequence of decisions made by other meta-agents, which affect the realized utilities and, hence, the experienced regrets.

C. Asymptotic Local and Global Behavior

In what follows, we present the main theorem that characterizes both local and global behavior emerging from meta-agents individually following Algorithm 4. The regret matrices R_n^c will be used as indicatives of local experience of meta-agents. The *global behavior* $\mathbf{z}_n$ at the network level at each time n is defined as the *discounted empirical frequency* of joint action profile of all meta-agents up to time n. Formally,

$$\mathbf{z}_n = (1-\varepsilon)^{n-1}\mathbf{e}_{\mathbf{a}_1} + \varepsilon\sum_{2\le\tau\le k}(1-\varepsilon)^{k-\tau}\mathbf{e}_{\mathbf{a}_\tau}, \quad (30)$$

where $\mathbf{e}_{\mathbf{a}_\tau}$ is a unit vector on the space of all possible joint action profiles $\mathcal{A}^{\mathcal{C}}$ (see Definition 4.1) with the element corresponding to the joint play $\mathbf{a}_\tau$ being equal to one. The small parameter $0 < \varepsilon \ll 1$ is the same as the adaptation rate in (28). It introduces an exponential forgetting of the past decision profiles to enable adaptivity of the network behavior to the evolution of the community preferences. Given $\mathbf{z}_n$, the average utility accrued by each meta-agent can be straightforwardly evaluated, hence the name global behavior. It is more convenient to define $\mathbf{z}_n$ via the stochastic approximation recursion

$$\mathbf{z}_n = \mathbf{z}_{n-1} + \varepsilon\left[\mathbf{e}_{\mathbf{a}_n} - \mathbf{z}_{n-1}\right]. \quad (31)$$

We use stochastic averaging [47] in order to characterize the asymptotic behavior of Algorithm 4. The basic idea is that, via a ‘local’ analysis, the noise effects in the stochastic algorithm is averaged out so that the asymptotic behavior is determined by that of a ‘mean’ dynamical system. To this end, in lieu of working with the discrete-time iterates directly, one works with continuous-time interpolations of the iterates. Accordingly, define the piecewise constant interpolated processes

$$R^{c,\varepsilon}(t) = R_n^c,\ \mathbf{z}^\varepsilon(t) = \mathbf{z}_n \ \text{ for } \ t\in[n\varepsilon,(n+1)\varepsilon). \quad (32)$$

Further, with slight abuse of notation, denote by R_n^c the regret matrix rearranged as vectors of length $(A^c)^2$ (rather than an $A^c \times A^c$ matrix) and let $R^{c,\varepsilon}(\cdot)$ represent the associated interpolated vector processes. Let further $\|\cdot\|$ denote the Euclidean norm, $\mathbb{R}_-$ represent the negative orthant in the Euclidean space of appropriate dimension, and define the distance to a set $\mathcal{R}$ by $\operatorname{dist}[\mathbf{r},\mathcal{R}] = \inf_{\widehat{\mathbf{r}}\in\mathcal{R}}\|\mathbf{r}-\widehat{\mathbf{r}}\|$. The following theorem characterizes the local and global behavior emerging by following Algorithm 4.

Theorem 5.1: Let t_ε be any sequence of real numbers satisfying $t_\varepsilon \to \infty$ as $\varepsilon \to 0$. The following results hold:

(i) If meta-agent c follows Algorithm 4, the asymptotic regret $R^{c,\varepsilon}(\cdot + t_\varepsilon)$ converges in probability to the negative orthant. That is, for any $\beta > 0$,

$$\lim_{\varepsilon\to 0} P\left(\operatorname{dist}\left[R^{c,\varepsilon}(\cdot+t_\varepsilon),\mathbb{R}_-\right] > \beta\right) = 0. \quad (33)$$

(ii) If all meta-agents follow Algorithm 4, the global behavior $\mathbf{z}^\varepsilon(\cdot + t_\varepsilon)$ converges in probability to the correlated equilibria set $\mathcal{Q}(\boldsymbol{\mu})$. That is, for any $\beta > 0$,

$$\lim_{\varepsilon\to 0} P\left(\operatorname{dist}\left[\mathbf{z}^\varepsilon(\cdot+t_\varepsilon),\mathcal{Q}(\boldsymbol{\mu})\right] > \beta\right) = 0. \quad (34)$$

Proof: See Appendix B for a sketch of the proof. ■

The first result in the above theorem simply asserts that, if a meta-agent individually follows Algorithm 4, it will asymptotically experience zero regret in its interaction with other meta-agents. Here, $R^{c,\varepsilon}(\cdot + t_\varepsilon)$ looks at the asymptotic behavior of R_n^c. From the technical point of view, the requirement $t_\varepsilon \to \infty$ as $\varepsilon \to 0$ means that we look at R_n^c for a small ε but large n with $\varepsilon n \to \infty$. For a small ε, R_n^c eventually spends nearly all of its time (with an arbitrarily high probability) in a β-neighborhood of the negative orthant, where $\beta \to 0$ as $\varepsilon \to 0$. Note that, when ε is small, R_n^c may escape from such a β-neighborhood. However, if such an escape ever occurs, it will be a rare event. The order of the escape time is often of the form $\exp(c/\varepsilon)$ for some $c > 0$; see [47, Section 7.2] for details. The second result in Theorem 5.1 states that, if now all meta-agents in the social network start following Algorithm 4 independently, their collective behavior for a small step-size ε is in a β-neighborhood of the correlated equilibria set. That is, meta-agents can coordinate their strategies in a distributed fashion so that the distribution of their joint behavior is close to that determined by a correlated equilibrium. From the game-theoretic point of view, it shows that simple and self-oriented local behavior of meta-agents can still lead to the manifestation of globally sophisticated and rational behavior at the network level.

VI. Examples of Homophilic Behavior: Energy Market and Detection Malicious Agents

In this section we provide two examples to illustrate how the decision test (6), statistical detection test (11), and stochastic optimization algorithm (16) from Sec. II can be applied to detect for homophilic (Defintion 2.1) and generalized homophilic (Defintion 2.2) communities. The first example uses real-world aggregate power consumption data from the

Ontario energy market social network. The second example concerns the detection of malicious agents in an online social network comprised of normal agents, malicious agents (i.e. homophilic and generalized homophilic communities), and an authentication agent. Finally, the game-theoretic model in Sec. V is applied to coordinate the actions of fictitious networks, controlled by a central authentication agent, to optimally detect communities of malicious agents.

A. Homophilic Community Detection and Community Preferences in Ontario Electrical Energy Market

In this seciton we apply the tools developed in Sec. II to detect for homophilic (Defintion 2.1) and generalized homophilic (Defintion 2.2) communities using the aggregate power consumption of different zones in the Ontario power grid. A sampling period of $N = 79$ days starting from January 2014 is used to generate the dataset $\mathcal{D}$ for the analysis. All price and power consumption data is available from the *Independent Electricity System Operator*[5] (IESO) website. Each zone is considered as an agent in the corporate network for power distribution. The study of corporate social networks was pioneered by Granovetter [48] which shows that the social structure of the network can have important economic outcomes. Examples include agents choice of alliance partners, assumption of rational behavior, self interest behavior, and the learning of other agents behavior. This analysis provides useful information for constructing demand side management (DSM) strategies for controlling power consumption in the electricity market.

The zone's power consumption is regulated by the associated price of electricity set by the operators in the corporate network. Since there is a finite amount of power in the grid, each officer must communicate with other officers in the network to set the price of electricity. Here, we utilize the aggregate power consumption from each of the $K = 10$ zones in the Ontario power grid. To perform the analysis, the external influence $\mathbf{p}_n$ and action of agents $\mathbf{x}_n$ must be defined. In the Ontario power grid, the wholesale price of electricity is dependent on several factors such as consumer behavior, weather, and economic conditions. Therefore, the external influence is defined as $\mathbf{p}_n = [p_n(1), p_n(2)]$, where $p_n(1)$ denotes the average electricity price between midnight and noon, and $p_n(2)$ the average between noon and midnight with n denoting day. The action of each zone correspond to the total aggregate power consumption in each respective time associated with $p_n(1)$ and $p_n(2)$, and is given by $\mathbf{x}_n^k = [x_n^k(1), x_n^k(2)]$, $k = 1, 2, \ldots, 10$. The budget B_n^k of each zone has units of dollars as $\mathbf{p}_n$ has units of \$/kWh and $\mathbf{x}_n^k$ has units of kWh. Note that the budget B_n^k for zones connected to other distribution stations outside the Ontario power grid are adjusted to account for the impact power import/export has on the budget—for import the budget increases, for export the budget decreases. Additionally, the average price of electricity $\mathbf{p}_n$ is given by the IESO historical data. Therefore, the stochastic gradient algorithm (Algorithm 1) cannot be applied to reduce Type-II errors in this real-world dataset.

[5]http://ieso-public.sharepoint.com/

Given the dataset $\mathcal{D}$, defined in (5), are there any homophilic communities (Defintion 2.2) present? First, we must test if the power demands of each zone satisfy the utility maximization requirement given in (2). Using $\mathcal{D}$ and Warshall's algorithm, each zone does not satisfy utility maximization. Therefore, there are no homophilic communities present in the Ontario power grid. This, however, may be the result of interdependence of the power demands of different zones. Next, we test if the overall power demand of the zones can be represented by one generalized homophilic community (Defintion 2.2). Recall from Sec. II this is equivalent to testing if the partition $\{1, 2, \ldots, 10\}$ is consistent with play from a concave potential game. The joint power consumption of all zones in the Ontario power gird is consistent with Nash rationality. This is an expected result as network congestion games have been shown to reduce peak power demand in distributed demand management schemes [49]. Using (6) and (7), a concave potential function for the Ontario power grid is constructed. The question is then: When do agents prefer to consume power? Using the constructed concave potential function, we evaluate the community preferences $\boldsymbol{\mu}_n$, defined in (20). Using Algorithm 3, we consider the external influence $\mathbf{p}_o = [0.47, 0.59]$ \$/kWh from which the community preferences are given by the vector $\boldsymbol{\mu} = [0.49, 0.51]$. Recall from Sec. III that community preferences depend on $\mathbf{p}_o$. This is confirmed by setting $\mathbf{p}_o = [0.57, 0.81]$ \$/kWh which results in $\boldsymbol{\mu} = [0.38, 0.62]$; that is, the community prefers to use power in the time period associated with $x_n(2)$ (between midnight and noon) as $\boldsymbol{\mu}(2) > \boldsymbol{\mu}(1)$.

B. Detecting Malicious Agents in Online Social Networks

Socialbots and spambots are autonomous programs which attempt to imitate human behavior and are prevalent on popular social networking sites such as Facebook and Twitter. In this section numerical examples are provided for the detection of malicious agents (i.e., socialbots) in an online social network using the detection tests for homophilic (Defintion 2.2) and generalized homophilic (Defintion 2.1) communities presented in Sec. II–III.

We consider the detection of malicious agents (i.e., socialbots) in an online social network comprised of normal agents, malicious agents, and a network authentication agent as depicted in Fig. 3. We further compute the community preferences to gain insight into the preferred avoidance strategy of the malicious network.

Recent techniques for detecting malicious agents in social networks (i.e., socialbots and spambots) use a method known as *behavioral blacklisting* which attempts to detect emails, tweets, friend and follower requests, and URLs which have originated from malicious agents [50]. Behavioral blacklisting use the fact that socialbots and spambots tend to have different behaviors than humans. For example, in Twitter, socialbots and spambots tend to re-tweet far more than normal (i.e., human) agents. In contrast, normal accounts tend to receive more replies, mentions, and re-tweets [51]. The goal of malicious agents is to increase their connectivity in the social network to deliver harmful content such as viruses, gaining followers and friends, marketing, and political campaigning. Consider the network

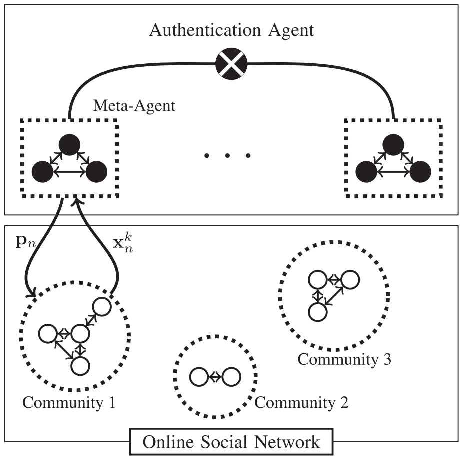

Fig. 3. Schematic of an online social network with a network authentication agent that is able to create meta-agents which interact with suspicious communities containing real agents (white) via a set of fictitious agents (black) in the social network. The goal of the authentication agent is to detect and eliminate malicious agents from the online social network. The parameters $\mathbf{p}_n$ and actions $\mathbf{x}_n^k$ are defined in Sec. VI-B.

topology depicted in Fig. 3. Each agent is capable of issuing commands which result in operations related to social structure (e.g., connect/disconnect from other agents) and social interaction (e.g., read/write messages). The authentication agent is designed to detect and eliminate malicious agents in the network. To this end, the authentication agent is able to construct fictitious accounts to study the actions of other agents in the network. Denoting by $\mathbf{p}_n \in \mathbb{R}_+^s$ the external influence representing m different classes of authentication queries at time n, the response of agent k is given by $\mathbf{x}_n^k \in \mathbb{R}_+^s$ and is the total number of successfully targeted followers and friends for each of the m authentication queries. Note that larger values of $\mathbf{p}_n$ indicate stronger authentication queries. Examples of queries for authentication can be found in [52]–[56]. For different authentication strategies, the malicious network community preferences $\boldsymbol{\mu}^c$, defined in (20), provides insight into the malicious networks preferred method of authentication.

We consider the following utility function for malicious agents:

$$\begin{aligned} r^k(\mathbf{x}^k, \mathbf{x}^{-k}) &= \ln\left[\prod_{i=1}^{s} \frac{x^k(i)}{\sum_{w=1}^{K} x^w(i)}\right], \\ m^k(\mathbf{x}^k; \boldsymbol{\beta}, \boldsymbol{\alpha}^k) &= \ln\left[\prod_{i=1}^{s}\left(1+\frac{x^k(i)}{\beta(i)}\right)^{\alpha^k(i)}\right], \\ u^k(\mathbf{x}^k, \mathbf{x}^{-k}) &= \eta^k r^k(\mathbf{x}^k, \mathbf{x}^{-k}) + m^k(\mathbf{x}^k; \boldsymbol{\beta}, \boldsymbol{\alpha}), \end{aligned} \quad (35)$$

where $\mathbf{x}_n^{-k} \in \mathbb{R}_+^{s\times(K-1)}$ is the action profile of the other $(K-1)$ agents, $r(\cdot)$ captures the interdependence among targets, and $m(\cdot)$ expresses each agent's preference to avoid detection. As the number of captured friends and followers of agent k increases, so does its utility function $u^k(\cdot)$; however, as other agents capture friends and followers, $u^k(\cdot)$ decreases. The static accuracy (i.e., signal-to-noise ratio) of each authentication query (e.g., completely automatic public Turing test to tell computers and humans apart) is contained in elements of the vector $\boldsymbol{\beta} \in \mathbb{R}_+^s$. As the elements of $\boldsymbol{\beta}$ increase, the utility of the malicious agent decreases. The parameter $\boldsymbol{\alpha}^k \in \mathbb{R}_+^s$ is the agent's preference for capturing friends and followers when a specific authentication strategy is utilized. The positive scalar η^k provides a balance between how much the malicious agent weights its ability to capture friends and followers compared to other agents, and the possibility of being detected by authentication queries. If $\eta^k = 0$, the malicious agent acts independently of other malicious agents. The malicious social budget of each agent k is given by B_n^k. The total resources available to the authentication agent are limited such that, in any operating period n, the total resources available for queries for authentication is given by $\sum_{i=1}^{s} p_n(i)$. Consider the case with $s = 2$ authentication strategies. As the authentication agent commits larger resources to increase the authentication queries associated with $p_n(1)$, the associated number of friends and followers captured by the malicious agent $x_n(1)$ decreases. Given that the total resources available to the authentication agent is limited, as $p_n(1)$ increases, $p_n(2)$ must decrease. This causes an increase in the friends and followers captured by the malicious agent $x_n(2)$ for the authentication queries $p_n(2)$. Therefore, the malicious social budget is considered to satisfy the linear relation $B_n^k = \mathbf{p}_n' \mathbf{x}_n^k$. *Malicious agents* are those that attempt to maximize their respective utility function (35). In contrast, *normal agents* have no target preference and, therefore, select $\mathbf{x}_n^k$ in a uniform random fashion—that is, normal agents are not utility maximizers and take actions independent of $\mathbf{p}_n$. At each observation n, a noisy measurement $\mathbf{y}_n^k$ is made of the actions $\mathbf{x}_n^k$. Given the dataset $\mathcal{D}_{\text{obs}}$, defined in (9), a statistical test can be used to detect if malicious agents are present.

Consider a dataset $\mathcal{D}$ (5) composed of $C = 3$ communities each with $K = 3$ agents where community $c = 3$ is composed of normal agents. For $c = 1$, each malicious agent has $\eta^1 = \eta^2 = \eta^3 = 0$, and for $c = 2$ each malicious agent has $\eta^4 = \eta^5 = \eta^6 = 10$. Agents in community $c = 1$ are indexed by $\{1, 2, 3\}$, and those in community $c = 2$ are denoted by $\{4, 5, 6\}$. A total of $N = 20$ observations are made for all agents using the queries for authentication generated from the following normal distributions: $\mathbf{p}_n \sim \mathcal{N}(1, 5)$. The malicious agent resource budget for all agents are generated from the following normal distribution: $B_n \sim \mathcal{N}(500, 25)$. Two possible authentication strategies are considered with the static inaccuracy given by $\boldsymbol{\beta} = [3, 8]$ for all malicious agents. Additionally, the associated agent preferences are given by $\boldsymbol{\alpha}^1 = \boldsymbol{\alpha}^4 = [0.1, 0.9]$, $\boldsymbol{\alpha}^2 = \boldsymbol{\alpha}^5 = [0.9, 0.1]$, and $\boldsymbol{\alpha}^3 = \boldsymbol{\alpha}^6 = [0.4, 0.6]$. To construct the response $\mathbf{x}_n^k$ for agents in $c = 1$, the maximization (2) is performed and, to construct the responses $\mathbf{x}_n^k$ for agents in $c = 2$, the maximization (3) is performed. The response of normal agents in $c = 3$ are produced from generating samples from $\mathbf{x}_n^k \sim \mathcal{U}(200, 10)$. The datasets $\mathcal{D}_{\text{obs}}$, defined in (9), for normal and malicious agents are finally obtained using the clean dataset $\mathcal{D}$ and additive noise $\mathbf{w}^k \sim \mathcal{U}(0, \kappa)$, where κ represents the magnitude of the measurement error.

Given the dataset $\mathcal{D}_{\text{obs}}$ with $\kappa = 0.1$, we attempt to detect for the homophilic and generalized homophilic communities using the algorithms presented in Sec. II. Figure 4 provides the response data and the estimated communities. By design, we are aware that agents $\{1, 2, 3\}$ in community 1 are utility

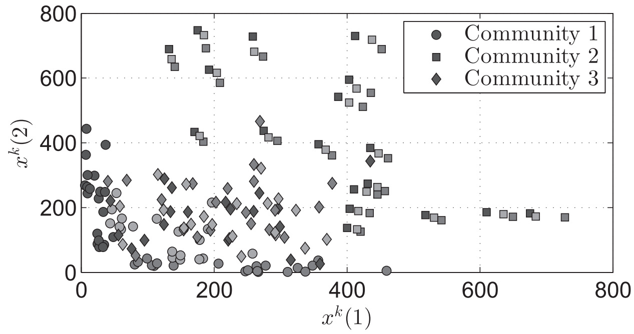

Fig. 4. Numerically computed response data for 9 agents distributed evenly in 3 communities as defined in Sec. VI-B.

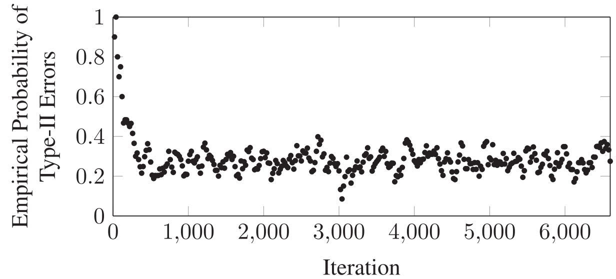

Fig. 5. Performance of the SPSA algorithm (Algorithm 1) for computing the locally optimal external influence P to reduce the probability of Type-II errors of the statistical test (11). The parameters are defined in Sec. VI-B.

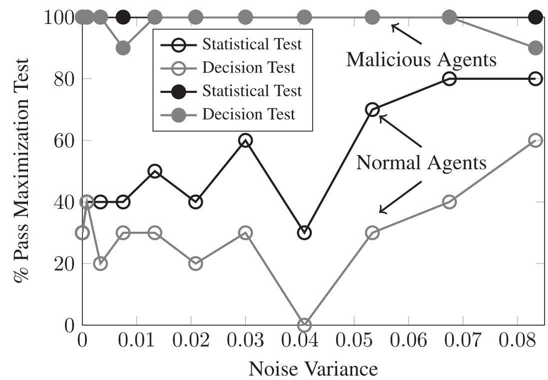

Fig. 6. Performance of the decision test (6) and statistical test (11) for the detection of malicious agents and normal agents. The parameters are defined in Sec. VI-B.

maximizers, community 2 is a generalized homophilic community, and agents $\{7,8,9\}$ in community 3 are normal agents. However, given only the data $\mathbf{x}_n^k$ in Fig. 4, how can we detect for these communities? Prior to applying the algorithms in Sec. II, consider that the responses of agents $\{4,5,6\}$ appear to be correlated. For example, $\mathbf{x}_1^4 = [298, 41]$, $\mathbf{x}_1^5 = [305, 41]$, and $\mathbf{x}_1^6 = [303, 41]$ suggesting that these agents are part of a generalized homophilic community (Definition 2.2), not a homophilic community (Definition 2.1) as the agent responses are not equal. Therefore, we consider the following test. First, let us detect if the agents $\{4,5,6\}$ satisfy the conditions of Definition 2.2. The responses from agents $\{4,5,6\}$ satisfy (6) and, therefore, these agents form a generalized homophilic community. Now, we consider detecting for utility maximization agents from the remaining agents $\{1,2,3,7,8,9\}$ using (6). Only agents $\{1,2,3\}$ satisfy utility maximization; however, each agent forms its own homophilic community as the responses from the agents are not equal. Finally, does any partition of the responses from agents $\{7,8,9\}$ satisfy the conditions of Defintion 2.2 for a generalized homophilic community? Using Algorithm 2, we find that no partition exists that satisfies (6) and, therefore, these are normal agents. As illustrated by this analysis, utilizing the tools developed in Sec. II allows us to detect the community structure present using a dataset $\mathcal{D}_{\text{obs}}$ of external influences and associated responses of agents.

We now consider how a set of external influences P can be designed to reduce the probability of Type-II errors from the statistical test (11) using Algorithm 1. Figure 5 plots the estimated cost (15) versus iterates generated by Algorithm 1 for $\sigma = 0.1$, $\epsilon = 0.2$, and $\kappa = 0.1$. Figure 5 illustrates that, by judiciously adapting the external influence via a stochastic gradient algorithm, the probability of Type-II errors can be decreased to approximately 30% allowing the statistical test (11) to adequately reject normal agents. Note that Type-II errors can be reduced further by increasing the number of observations N and then using Algorithm 1 to select the applied external influence P. For example, for $N = 50$, the estimated Type-II error probability is approximately 5%.

Let us consider the performance of the decision test (6) and non-parametric statistical test (11) for detecting agents engaged in a concave potential game (i.e. forming a generalized homophilic community). Two communities are considered: the first is a generalized homophilic community with 3 agents, and the other is a community of 3 normal agents as defined in Sec. VI-B. The additive noise introduced to the agents' actions is generated from $\mathbf{w}^k \sim \mathcal{U}(0,\kappa)$ with noise level κ. Prior to applying (11), we must first estimate the noise distribution using a non-parametric estimator as discussed in Sec. II-B. The noise distribution is estimated using the non-parametric piecewise linear estimator [57] using 40 samples of agent actions with the external influence held constant. This is performed for every noise level κ. Fig. 6 provides an estimate of the Type-I and Type-II error probabilities for the decision test and non-parametric statistical test where the percent that pass (6) and (11) are computed using $N = 20$ observed agent actions generated using 50 independent trials. The locally optimized external influence P was obtained from Algorithm 1, allowing the malicious and normal agents to be distinguished. As can be seen, the occurrence of Type-I errors in the statistical test is less then 5%. Though we have utilized a non-parametric noise density estimator, the results are in agreement with Theorem 2.2, which for a known noise density, guarantees that the statistical test (11) has less then a 5% Type-I error probability. Recall from Sec. II-B that the statistical test (11) is guaranteed to have a higher pass percentage compared to the decision test (6). Therefore, every dataset $\mathcal{D}_{\text{obs}}$ that satisfies (6) is guaranteed to satisfy (11) with $\Phi = 0$.

As discusses in Sec. IV, for meta-agents, it is useful to have a measure of the preferences of the community that they serve. Let us consider computing $\boldsymbol{\mu}^2$, as defined in (20), for the homophilic community $c = 2$ with the parameters

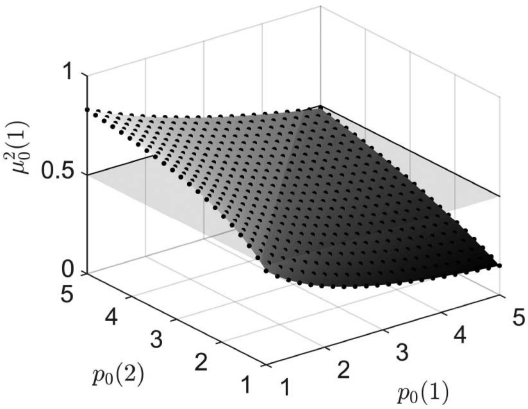

Fig. 7. Malicious community (i.e. homophilic community) preference $\mu^2(1)$ computed using Algorithm 3 with the parameters defined in Sec. VI-B. The black dots indicate the estimated $\mu^2(1)$. The translucent patch indicates the location where $\mu(1) = 0.5$.

defined in Sec. VI-B. Given that $s = 2$, and using (20), we have that the elements in $\boldsymbol{\mu}^2$ must satisfy $\mu^2(2) = 1 - \mu^1(1)$. Figure 7 provides the homophilic community preferences for $\mu_n^1(1)$ for different authentication queries $\mathbf{p}_o$. As can be seen, the homophilic community prefers the authentication queries deployed by $p_o(1)$; $\mathbf{p}_o = [1, 5]$ and $\mathbf{p}_o = [5, 1]$ yield $\mu^1(1) = 0.83$ and $\mu^1(1) = 0.15$, respectively. Therefore, the optimal detection strategy would be to use the authentication queries provided by $p_o(2)$—that is, the the homophilic community will capture more friends and followers when $p_o(1) < p_o(2)$ compared to when $p_o(1) > p_o(2)$.

C. Coordination of Meta-Agents for Detecting Malicious Agents in Online Social Networks

Here, we consider a network of fictitious agents interacting with a set of agents to detect if these agents belong to a homophilic community (i.e., they are malicious) as depicted in Fig. 3. The meta-agents coordinate their authentication strategies as a non-cooperative game defined in Sec. IV-A. The meta-agents account for the preference for authentication queries, denoted by $\boldsymbol{\mu}^c$, for each suspicious subgroup of agents indexed by c. Note that the detection of possible suspicious agents is equivalent to the detection of homophilic communities in which case $\boldsymbol{\mu}^c$ can be estimated using the tools presented in Sec. II–III. The results presented in this subsection provide insight into how authentication queries can be produced in a coordinated fashion by a set of meta-agents in order to detect malicious agents.

Let us define the non-cooperative game and utility function of the meta-agents. The set of meta-agents is denoted by $\mathcal{C} = \{1, 2, \ldots, C\}$, and the available authentication query strategies are denoted by $\mathcal{A}^c = \{0, 1, 2, \ldots, J\}$, for all $c \in \mathcal{C}$. The associated utility function for each meta-agent c is:

$$U^c(a^c, \mathbf{a}^{-c}, \boldsymbol{\mu}^c) = k_r^c \left[1 - \exp\left(\frac{S(a^c)}{\chi(a^c, \mathbf{a}^{-c})}\right)\right] - k_l^c \left[\exp\left(\frac{S(j)}{\boldsymbol{\mu}^c(j)s^c} \cdot \overline{d}(c, a^c)\right) - 1\right], \quad (36)$$

where

$$\chi(a^c, \mathbf{a}^{-c}) = 1 + \textstyle\sum_{c' \neq c} I\left(a^{c'} = a^c\right)$$

denotes the number of meta-agents sending authentication queries to the agents being served by meta-agents c. Here, $a^c = 0$ means that meta-agent c does not produce any authentication queries, and $a^c = j$, $0 \leq j \leq J$ corresponds to meta-agent c sending authentication queries to the suspicious subgroup of agents indexed by j. Further, s^c denotes the size of meta-agent c, $S(j)$ represents the size of the j-th suspicious subgroup of agents, and $\overline{d}(c, j)$ is a distance metric (e.g. minimum distance) between the meta-agent c and suspicious network j. The term $\frac{S(j)}{\boldsymbol{\mu}^c(j)s^c} \cdot \overline{d}(c, a^c)$ can be interpreted as the delay in sending authentication queries by fictitious network c to the j-th suspicious subgroup of agents, which delays the detection of malicious agents and increases the speed by which their friend and followers grow. In (36), $0 \leq k_r^c, k_l^c \leq 1$ are the relative weights of the reward and the loss in rewards due to delay in the utility of fictitious network c, respectively, and satisfy $k_r^c + k_l^c = 1$.

We set $C = 10$, and assume that each meta-agent is composed of 20 agents that can produce authentication queries and record the response of the queried agents. These agents are located in a 10×10 km^2 area in a random fashion as described next. First, a center point is uniformly picked, and then agents are uniformly distributed in a circle with radius 2 km. We further assume $S(j) \sim \mathcal{U}(1, \overline{S})$ for each suspicious subgroup of agents $j = 1, 2, \ldots, J$, and use the scheme described above for their placement. The distance $\overline{d}(c, j)$ in (36) is set to the distance between the center of the smallest circles within which agents in meta-agent c and suspicious subgroup j of agents. Finally, we assume $\boldsymbol{\mu}^c \sim \mathcal{U}(0, 1)$, for $c = 1, 2, \ldots, C$, and set $k_r^c = 0.6$, $k_l^c = 0.4$ in (36). Recall that $\boldsymbol{\mu}^c$ can be estimated for each suspicious subgroup of agents using the tools presented in Sec. II.

The following heuristics are used as benchmarks to compare the performance of Algorithm 4:

1) Location-Based: Each meta-agent c chooses the suspicious network $j^* = \min_{j \in \mathcal{A}^c} \overline{d}(c, j)$. This scheme aims at minimizing the delay costs associated with sending authentication queries.

2) Myopic Greedy: Each meta-agent c chooses the suspicious network $j^* = \max_{j \in \mathcal{A}^c} S(j)/\overline{d}(c, j)$. This scheme aims at maximizing the reward by sending authentication queries to more populated suspicious subgroup of suspicious agents located closer to the meta-agent.

We conduct a semi-analytical study to evaluate performance of Algortihm 4. Recall from Theorem 5.1 that $\mathbf{z}_n$ converges to the set $\mathcal{Q}(\boldsymbol{\mu}_n)$ rather than a particular point in that set. In fact, even if $\boldsymbol{\mu}_n = \boldsymbol{\mu}$ is fixed, $\mathbf{z}_n$ can generally move around in the polytope $\mathcal{Q}(\boldsymbol{\mu})$. Since the game-theoretic regret-matching is run on the faster timescale, its behavior in the slower timescale can be modeled by arbitrarily picking $\mathbf{z}_n$ from the polytope $\mathcal{Q}(\boldsymbol{\mu}_n)$. In our simulations, we use MATLAB's "cprnd" function, which draws samples $\boldsymbol{\pi}_n^c$ using a uniform distribution over the interior of the polytope $\mathcal{Q}(\boldsymbol{\mu}_n)$ defined by a system of linear

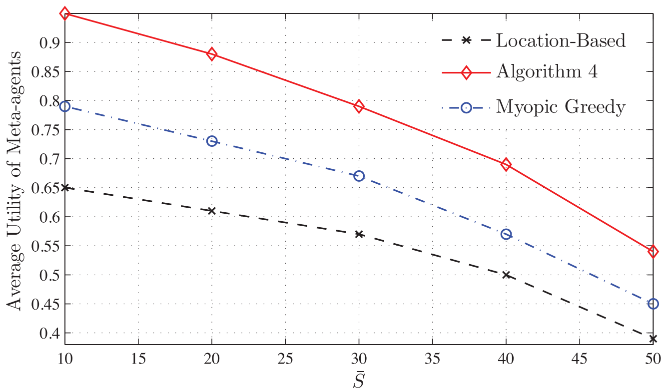

Fig. 8. Average utility of meta-agents versus the maximum size of the number of agents in a suspicious subgroup.

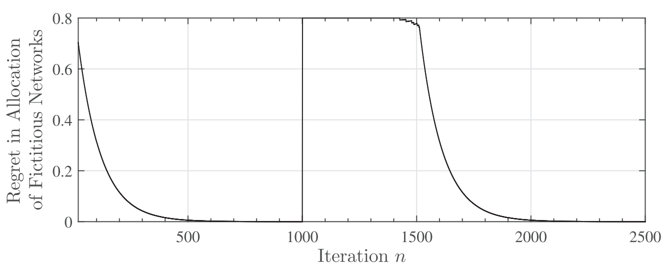

Fig. 9. Adaptive behavior of meta-agents. The work load changes at $n = 10^3$ due to the removal of a detected malicious network. The meta-agent adapts its decision strategy by learning from interaction with other meta-agents.

inequalities[6]. Using a uniform distribution is no loss of generality since it assumes no prior on the interior points of $\mathcal{Q}(\boldsymbol{\mu}_n)$, and matches the essence of convergence to a set. Once the sample $\boldsymbol{\pi}_n^c$ is taken, it is marginalized to obtain decision strategy of individual fictitious networks.

Figure 8 shows the average individual meta-agent's utility (36) versus the maximum size of the number of agents in a suspicious subgroup, denoted by $\overline{S}$ above. Each point on the graph is an average over 100 independent runs of the algorithms for different realizations of the simulation setup explained above. As the size of the suspicious subgroup increases, the utility of meta-agents decrease since the delay in sending authentication queries increases. As the results in Fig. 8 illustrate, the average behavior of meta-agents determined by the correlated equilibria set of the game (Algorithm 4) outperforms the heuristic methods (i.e., *Location-Based* and *Myopic Greedy*) explained above.

Figure 9 illustrates the average regret that each meta-agent experiences along the way to reach a correlated equilibrium, and further demonstrates adaptivity of the decision strategy implemented by Algorithm 4. At $n = 10^3$, a detected malicious network is eliminated and, therefore, the workload $\boldsymbol{\mu}^c$ of the fictitious network c that performed authentication queries drops to 0. As a result, the set of correlated equilibria of the underlying game undergoes a regime switching. As seen in Fig. 9, the meta-agent representing fictitious network c adjusts its strategy so as to minimize its experienced regret once it realizes the change.

[6]Theoretically, finding a correlated equilibrium in a game is equivalent to solving a linear feasibility problem of the form $A\boldsymbol{\pi} < \mathbf{0}$; see (24) for details.

VII. Summary

This paper studied the detection of homophilic communities and the distributed coordination of meta-agents each interacting with a homophilic community. First, we studied how to parse a dataset into different homophilic communities. A non-parametric detection test was constructed to detect homophilic communities which only requires a time series of external influences and actions of the agents. Given the set of detected homophilic communities, a non-parametric algorithm was provided that can be used to estimate the preferences of the communities which can be used by the meta-agent to coordinate their actions. Second, a non-cooperative game formulation was presented for the interaction among meta-agents which, in turn, interact with the communities. A two timescale stochastic approximation algorithm was presented that prescribed the actions meta-agents should take based on the preferences of the communities they interact with. It was shown that, if each meta-agent follows the proposed algorithm individually, its asymptotic regret can be made arbitrarily small after sufficient interaction with other communities. If all meta-agents follow the proposed algorithm, their global behavior is attracted to the correlated equilibria set of the formulated game. An important property of both aspects considered in this paper is their ordinal nature, which provides a useful approximation to human behavior. Finally, we illustrated the application of the proposed schemes in a real-world example using the energy market, and provided a numerical example to detect malicious agents in an online social network.

An alternative formulation for the interaction of local and global decision makers uses Bayesian social learning [61], [62]. It is worthwhile devising social learning models for homophilic communities.

Appendix A
Proof of Theorem 2.2

Consider a dataset $\mathcal{D}$ that satisfies homophilic behavior (6). Given $\mathcal{D}$, the inequalities (6) have a feasible solution. Denote the solutions by $\{\lambda_n^{ko} > 0, V_n^o\}$. Substituting $\mathbf{x}_n^k = \mathbf{y}_n^k - \mathbf{w}_n^k$ into the inequalities obtained from the solution of (6) yields

$$\begin{aligned} &V_m^o - V_n^o - \textstyle\sum_{k=1}^K \lambda_n^{ko} \mathbf{p}_n'(\mathbf{y}_m^k - \mathbf{y}_n^k) \\ &\quad \le \textstyle\sum_{k=1}^K \lambda_n^{ko} \mathbf{p}_n'(\mathbf{w}_n^k - \mathbf{w}_m^k). \end{aligned} \tag{37}$$

The goal is to compute an upper bound on the r.h.s. that is independent of λ_n^{ko}. We have

$$\begin{aligned} &\textstyle\sum_{k=1}^K \lambda_n^{ko} \mathbf{p}_n'(\mathbf{w}_n^k - \mathbf{w}_m^k) \\ &\quad \le \textstyle\sum_{k=1}^K \lambda_n^{ko} |\mathbf{p}_n'(\mathbf{w}_n^k - \mathbf{w}_m^k)| \\ &\quad \le \left[\textstyle\sum_{k=1}^K \lambda_n^{ko}\right] \left[\textstyle\sum_{k=1}^K |\mathbf{p}_n'(\mathbf{w}_n^k - \mathbf{w}_m^k)|\right] \\ &\quad \le \Lambda_n M, \end{aligned} \tag{38}$$

where $\Lambda_n = \sum_{k=1}^K \lambda_n^{ko}$, and M is defined by (13). Substituting (38) into (37) yields

$$\frac{1}{\Lambda_n}\left[V_m^o - V_n^o - \textstyle\sum_{k=1}^K \lambda_n^{ko} \mathbf{p}_n'(\mathbf{y}_m^k - \mathbf{y}_n^k)\right] \le M. \tag{39}$$

Let $\{\Phi^*\{\mathbf{y}\}, \lambda_n^{k^*}, V_n^*\}$ denote a solution of (12) given the noisy dataset $\mathcal{D}_{\text{obs}}$, defined in (9). Comparing (39) with the inequalities obtained from the solution of (12), it can be shown that $\{\Phi^*\{\mathbf{y}\} = M, \lambda_n^{k^*} = \lambda_n^{ko}, V_n^* = V_n^o\}$ is a feasible, but not necessarily optimal, solution of (12). Therefore, for $\mathcal{D}$ satisfying homophilic behavior (6), it must be the case that $\Phi^*\{\mathbf{y}\} \leq M$. This asserts, under the null hypothesis H_0, that $\Phi^*\{\mathbf{y}\}$ is upper bounded by M. For a given $\Phi^*\{\mathbf{y}\}$, the integral in (11) is the probability of $\Phi^*\{\mathbf{y}\} \leq M$; therefore, the conditional probability of rejecting H_0 when true is less then γ. ■

APPENDIX B
PROOF OF THEOREM 5.1

The proof is divided into three steps:

Step 1: The sequence $\{a_n^c, R_n^c\}$ is a Markov chain with state space $\mathcal{A}^c$, and transition probability matrix

$$\Pr\left(a_n^c = j \,\middle|\, a_{n-1}^c = i, R_n^c = R^c\right) = P_{ij}\left(R^c\right),$$

where the transition probability matrix $P(R^c) = [P_{ij}(R^c)]$ is defined by (27), and is continuous, irreducible and aperiodic for each R^c. Therefore, the Markov chain is ergodic in the sense that there is a unique stationary distribution $\boldsymbol{\psi}^c(R^c) = [\psi_1^c(R^c), \ldots, \psi_{A^c}^c(R^c)]$ such that $[P(R^c)]^n \to \mathbf{1}_{A^c}\boldsymbol{\psi}^c(R^c)$ as $n \to \infty$, which is a matrix with identical rows consisting of the stationary distribution. (The convergence in fact takes place geometrically fast.) In view of (27), it can be shown that $\boldsymbol{\psi}^c(R^c)$ satisfies

$$\textstyle\sum_{j\neq i} \psi_j^c(R^c)\, |r^c(j,i)|^+ = \psi_i^c(R^c) \sum_{j\neq i} |r^c(i,j)|^+ . \quad (40)$$

Before proceeding to the proof, a few definition and notations are in order.

Definition B.1: Let Y_n be an $\mathbb{R}^r$-valued random vector.

1) The sequence $\{Y_n\}$ is tight if for each $\varsigma > 0$, there exists a compact set $\mathcal{D}_\varsigma$ such that $\Pr(Y_n \in \mathcal{D}_\varsigma) \geq 1 - \varsigma$ for all n.

2) Y_n converges weakly to Y, denoted by $Y_n \Rightarrow Y$, if for any bounded and continuous function $f(\cdot)$, $\mathbb{E}f(Y_n) \to \mathbb{E}f(Y)$ as $n \to \infty$.

Definition B.1: A *differential inclusion* is a dynamical system of the form

$$\frac{dY}{dt} \in \mathcal{F}\left(Y\right),$$

where $Y \in \mathbb{R}^r$ and $\mathcal{F} : \mathbb{R}^r \to \mathbb{R}^r$ is a Marchaud map [58]. That is, i) the graph and domain of $\mathcal{F}$ are nonempty and closed; ii) the values $\mathcal{F}(X)$ are convex; and iii) the growth of $\mathcal{F}$ is linear: There exists $\kappa > 0$ such that, for every $Y \in \mathbb{R}^r$,

$$\textstyle\sup_{Z\in\mathcal{F}(Y)} \|Z\| \leq \kappa\left(1 + \|Y\|\right),$$

where $\|\cdot\|$ denotes any norm on $\mathbb{R}^r$.

By virtue of Prohorov's Theorem [59], we can extract convergent subsequences when tightness is verified. Therefore, we first prove tightness. The limit process is then characterized using a certain operator related to the limit martingale problem. We refer the reader to [47, Chapter 7] for further details on weak convergence and related matters.

To proceed, recall the continuous-time interpolated process $R^{c,\varepsilon}(t) = R_n^c$ for $t \in [n\varepsilon, (n+1)\varepsilon)$. We first prove the tightness. Consider (28) for the sequence of $(A^c)^2$-valued vectors resulted after rearranging the elements of R_n^c into a vector. Noting the boundedness of utility functions, and using Hölders and Gronwalls inequalities, for any $0 < T < \infty$, we obtain

$$\textstyle\sup_{n\leq T/\varepsilon} \mathbb{E}\left\|R_n^c\right\|^2 \leq \infty \quad (41)$$

where in the above and hereafter t/ε is understood to be the integer part of t/ε for each $t > 0$. By virtue of the tightness criteria [47], it suffices to verify

$$\lim_{\delta\to 0} \lim_{\varepsilon\to 0} \sup \mathbb{E}\{ \sup_{0\leq s\leq\delta} \mathbb{E}_t^\varepsilon \left\|R^{c,\varepsilon}(t+s) - R^{c,\varepsilon}(t)\right\|^2\} = 0,$$

where $\mathbb{E}_t^\varepsilon$ denotes conditional expectation given the σ-algebra generated by the ε-dependent past data up to time t. Noting the boundedness of $U^c(\cdot,\cdot,\boldsymbol{\mu}^c)$ for all values of $\boldsymbol{\mu}^c$, the above criteria can be easily verified.

Therefore, using Prohorov's Theorem [59], one can extract a convergent subsequence. For notational simplicity, we still denote the subsequence by $R^{c,\varepsilon}(\cdot)$ with limit $R^c(\cdot)$. Using the Skorohod representation theorem [47], with a slight abuse of notation, one can assume $(R^{c,\varepsilon}(\cdot), \boldsymbol{\mu}^{c,\varepsilon}(\cdot)) \to (R^c(\cdot), \boldsymbol{\mu}^c)$ in the sense of w.p.1 and the convergence is uniform on any finite interval. Note that the preferences $\boldsymbol{\mu}^c$ remains constant, which is due to the two timescale nature of the proposed scheme. It can be shown that, in two timescale systems, the slow component is quasi-static—remains almost constant—while analyzing the behavior of the fast timescale [47, Chapter 8]. The limit system $R^c(\cdot)$, as $\varepsilon \to 0$, can then be characterized using martingale averaging methods by

$$\tfrac{d}{dt} R^c \in \mathcal{G}^c(R^c, \boldsymbol{\mu}^c) - R^c, \quad (42)$$

where $\mathcal{G}^c = [g_{ij}^c]$ is an $A^c \times A^c$ matrix with elements: [see (43) at the bottom of the page]. In (43), $\boldsymbol{\psi}^{-c}$ denotes the joint strategy of all meta-agents excluding c, and $\Delta\mathcal{A}^{-c}$ represents the simplex of all such strategies. The details are tedious and are omitted for brevity. The interested reader is referred to [47, Chapter 8] and [60, Appendix B] for more details and similar proofs.

Step 2: Next, we examine stability of the limit system (42), and show its set of global attractors comprises the negative orthant. Define the Lyapunov function

$$D(R^c) = \tfrac{1}{2} \textstyle\sum_{i,j\in\mathcal{A}^c} \left(|r^c(i,j)|^+\right)^2 . \quad (44)$$

Taking the time-derivative, and using (42), yields: [see (45) at the bottom of the next page]. Separating the summation over the

$$g_{ij}^c\left(R^c, \boldsymbol{\mu}^c\right) = \{\left[U^c\left(j, \boldsymbol{\psi}^{-c}, \boldsymbol{\mu}^c\right) - U^c\left(i, \boldsymbol{\psi}^{-c}, \boldsymbol{\mu}^c\right)\right] \psi_i^c(R^c); \boldsymbol{\psi}^{-c} \in \Delta\mathcal{A}^{-c}\} \quad (43)$$

two terms, using (40), $\sum_{i,j} a_{ij} = \sum_{j,i} a_{ji}$, and $|r_n^c(i,i)|^+ = 0$, it can be shown that the first term on the r.h.s. of (45) is equal to zero. Therefore,

$$\tfrac{d}{dt} D(R^c) = -D(R^c). \tag{46}$$

Consequently, the negative orthant $\mathbb{R}_-$ is globally asymptotically stable for the limit system (42), and

$$\lim_{t\to\infty} \operatorname{dist}\left[R^c(t), \mathbb{R}_-\right] = 0. \tag{47}$$

Subsequently, we study asymptotic stability by looking at the case where $\varepsilon \to 0$, $n \to \infty$, and $\varepsilon n \to \infty$. Nevertheless, instead of considering a two-stage limit by first letting $\varepsilon \to 0$ and then $t \to \infty$, we study $R^{c,\varepsilon}(t + t_\varepsilon)$ and require $t_\varepsilon \to \infty$ as $\varepsilon \to 0$. Define $\widehat{R}^{c,\varepsilon}(\cdot) = R^{c,\varepsilon}(\cdot + t_\varepsilon)$. It can be shown that $\widehat{R}^{c,\varepsilon}(\cdot)$ is also tight. For any $T \le \infty$, take a weakly convergent subsequence of $\{\widehat{R}^{c,\varepsilon}(\cdot), \widehat{R}^{c,\varepsilon}(\cdot - T)\}$, and denote its limit by $(\widehat{R}^c(\cdot), \widehat{R}^c_T(\cdot))$. Note that $\widehat{R}^c(0) = \widehat{R}^c_T(T)$. The value of $\widehat{R}^c_T(0)$ may be unknown, but the set of all possible values (over all T and convergent subsequences) belong to a tight set. Using this, the stability condition and the results of Step 1, for any $\vartheta > 0$ there is a $T_\vartheta < \infty$ such that for all $T > T_\vartheta$, $\operatorname{dist}[\widehat{R}^c_T(T), \mathbb{R}_-] \ge 1 - \vartheta$. This implies that $\operatorname{dist}[\widehat{R}^c(0), \mathbb{R}_-] \ge 1 - \vartheta$. Therefore, $R^{c,\varepsilon}(\cdot + t_\varepsilon) \to \mathbb{R}_-$ in probability. This completes the proof of the first part in Theorem 5.1.

Step 3: The final step shows that the convergence of the regrets of individual meta-agents to negative orthant provides the necessary and sufficient condition for convergence of their global behavior to the correlated equilibria set. Recall $R^{c,\varepsilon}(\cdot)$ and $\mathbf{z}^{\varepsilon}(\cdot)$, defined in (32). Using (25) and (30), and for any fixed global preference vector $\boldsymbol{\mu}^c$, we obtain: [see (48) at the bottom of the page], where $z^{c,\varepsilon}(i, \mathbf{a}^{-c})(t)$ denotes the interpolated empirical distribution of meta-agent c taking action i and the rest choosing $\mathbf{a}^{-c}$. On any convergent subsequence $\{\mathbf{z}_{n'}\}_{n'>0} \to \boldsymbol{\pi}$, with slight abuse of notation, let $\mathbf{z}^{\varepsilon}(t) = \mathbf{z}_{n'}$ and $R^{c,\varepsilon}(t) = R^c_{n'}$ for $t \in [n'\varepsilon, (n'+1)\varepsilon)$. Consequently, for any $\boldsymbol{\mu}^c$, and for all $i, j \in \mathcal{A}^c$ and $c \in \mathcal{C}$,

$$\lim_{t\to\infty} r^{c,\varepsilon}(i,j)(t) = \sum_{\mathbf{a}^{-c}} \pi^c\left(i, \mathbf{a}^{-c}\right)\left[U^c\left(j, \mathbf{a}^{-c}, \boldsymbol{\mu}^c\right) - U^c\left(i, \mathbf{a}^{-c}, \boldsymbol{\mu}^c\right)\right]. \tag{49}$$

Here, $\pi^c\left(i, \mathbf{a}^{-c}\right)$ denotes the probability of meta-agent c choosing action i and the rest playing $\mathbf{a}^{-c}$. Finally, combining (49) with the final result of Step 2 and comparing with (24) completes the proof of the second part in Theorem 5.1. ■

References

[1] C. Shalizi and A. Thomas, "Homophily and contagion are generically confounded in observational social network studies," *Sociol. Methods Res.*, vol. 40, no. 2, pp. 211–239, 2011.

[2] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, pp. 415–444, 2001.

[3] D. Fudenberg and D. K. Levine, "Learning and equilibrium," *Annu. Rev. Econ.*, vol. 1, pp. 385–420, 2009.

[4] W. Diewert, "Afriat and revealed preference theory," *Rev. Econ. Stud.*, vol. 40, no. 3, pp. 419–425, 1973.

[5] R. Blundell, "How revealing is revealed preference?" *J. Eur. Econ. Assoc.*, vol. 3, nos. 2–3, pp. 211–235, 2005.

[6] W. Diewert, "Afriat's theorem and some extensions to choice under uncertainty," *Econ. J.*, vol. 122, no. 560, pp. 305–331, 2012.

[7] R. J. Aumann, "Correlated equilibrium as an expression of Bayesian rationality," *Econometrica*, vol. 55, no. 1, pp. 1–18, 1987.

[8] O. N. Gharehshiran, V. Krishnamurthy, and G. Yin, "Distributed tracking of correlated equilibria in regime switching noncooperative games," *IEEE Trans. Autom. Control*, vol. 58, no. 10, pp. 2435–2450, Oct. 2013.

[9] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, 2000.

[10] S. Hart and A. Mas-Colell, "A general class of adaptive strategies," *J. Econ. Theory*, vol. 98, no. 1, pp. 26–54, 2001.

[11] S. Hart, A. Mas-Colell, and Y. Babichenko, *Simple Adaptive Strategies: From Regret-Matching to Uncoupled Dynamics*. Singapore: World Scientific, 2013.

[12] R. Blundell, M. Browning, I. Crawford, B. Rock, F. Vermeulen, and L. Cherchye, "Sharp for SARP: Nonparametric bounds on the behavioural and welfare effects of price changes," *Amer. Econ. J. Microecon.*, vol. 7, no. 1, pp. 43–60, 2015.

[13] L. Cherchye, B. Rock, and F. Vermeulen, "Opening the black box of intrahousehold decision making: Theory and nonparametric empirical tests of general collective consumption models," *J. Political Econ.*, vol. 117, no. 6, pp. 1074–1104, 2009.

[14] G. Coricelli, R. J. Dolan, and A. Sirigu, "Brain, emotion and decision making: The paradigmatic example of regret," *Trends Cognit. Sci.*, vol. 11, no. 6, pp. 258–265, 2007.

[15] A. P. Steiner and A. D. Redish, "Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task," *Nat. Neurosci.*, vol. 17, no. 7, pp. 995–1002, 2014.

[16] S. Afriat, "The construction of utility functions from expenditure data," *Int. Econ. Rev.*, vol. 8, no. 1, pp. 67–77, 1967.

[17] H. Varian, "Non-parametric tests of consumer behaviour," *Rev. Econ. Stud.*, vol. 50, no. 1, pp. 99–110, 1983.

[18] V. Krishnamurthy and W. Hoiles, "Afriat's test for detecting malicious agents," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 801–804, Dec. 2012.

[19] R. Deb, "Interdependent preferences, potential games and household consumption," University Library of Munich, Munich, Germany, MPRA Paper 6818, Jan. 2008.

[20] M. O. Jackson and A. Wolinsky, "A strategic model of social and economic networks," *J. Econ. Theory*, vol. 71, no. 1, pp. 44–74, 1996.

[21] M. Kearns, "Graphical games," in *Algorithmic Game Theory*, N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2007, vol. 3, pp. 159–180.

[22] C. Griffin and A. Squicciarini, "Toward a game theoretic model of information release in social media with experimental results," in *Proc. IEEE Symp. Sec. Privacy Workshops*, 2012, pp. 113–116.

[23] S. Hart and A. Mas-Colell, "A reinforcement procedure leading to correlated equilibrium," in *Economics Essays: A Festschrift for Werner Hildenbrand*. Heidelberg, Germany: Springer, 2001, pp. 181–200.

[24] E. Cartwright and M. Wooders, "Correlated equilibrium, conformity and stereotyping in social groups," *J. Public Econ. Theory*, vol. 16, no. 5, pp. 743–766, 2014.

[25] J. Xu and M. V. der Schaar, "Social norm design for information exchange systems with limited observations," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 11, pp. 2126–2135, Dec. 2012.

[26] P. Maillé and B. Tuffin, *Telecommunication Network Economics: From Theory to Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

$$\tfrac{d}{dt} D(R^c) = \sum_{i,j\in\mathcal{A}^c} \left|r_n^c(i,j)\right|^+ \left[U^c\left(j, \boldsymbol{\psi}^{-c}, \boldsymbol{\mu}^c\right) - U^c\left(i, \boldsymbol{\psi}^{-c}, \boldsymbol{\mu}^c\right)\right] \psi_i^c(R^c) - \sum_{i,j\in\mathcal{A}^c} \left|r_n^c(i,j)\right|^+ r_n^c(i,j) \tag{45}$$

$$r^{c,\varepsilon}(i,j)(t) = \sum_{\mathbf{a}^{-c}} z^{c,\varepsilon}(i, \mathbf{a}^{-c})(t)\left[U^c\left(j, \mathbf{a}^{-c}, \boldsymbol{\mu}^c\right) - U^c\left(i, \mathbf{a}^{-c}, \boldsymbol{\mu}^c\right)\right] \tag{48}$$

[27] N. Alon, M. Feldman, A. Procaccia, and M. Tennenholtz, “A note on competitive diffusion through social networks,” *Inf. Process. Lett.*, vol. 110, no. 6, pp. 221–225, 2010.
[28] R. Deb, “A testable model of consumption with externalities,” *J. Econ. Theory*, vol. 144, no. 4, pp. 1804–1816, 2009.
[29] W. Hoiles and V. Krishnamurthy, “Nonparametric demand forecasting and detection of demand-responsive consumers,” *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 695–704, Mar. 2014.
[30] D. Monderer and L. Shapley, “Potential games,” *Games Econ. Behav.*, vol. 14, no. 1, pp. 124–143, 1996.
[31] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[32] A. Izenman, “Review papers: Recent developments in nonparametric density estimation,” *J. Amer. Stat. Assoc.*, vol. 86, no. 413, pp. 205–224, 1991.
[33] P. Peretti, “Testing the significance of the departures from utility maximization,” *Macroecon. Dyn.*, vol. 9, no. 3, pp. 372–397, 2005.
[34] R. Rosenthal, “A class of games possessing pure-strategy Nash equilibria,” *Int. J. Game Theory*, vol. 2, no. 1, pp. 65–67, 1973.
[35] J. C. Spall, *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*. Hoboken, NJ, USA: Wiley, 2003.
[36] M. Ester, H. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise,” in *Proc. ACM/SIGKDD Conf. Knowl. Discovery Data Min.*, 1996, vol. 96, pp. 226–231.
[37] J. Lee and S. Leyffer, *Mixed Integer Nonlinear Programming*. New York, NY, USA: Springer, 2011, vol. 154.
[38] C. DAmbrosio and A. Lodi, “Mixed integer nonlinear programming tools: A practical overview,” *Quart. J. Oper. Res*, vol. 9, no. 4, pp. 329–349, 2011.
[39] P. Bonami, M. Kilinç, and J. Linderoth, “Algorithms and software for convex mixed integer nonlinear programs,” in *Mixed Integer Nonlinear Programming*. New York, NY, USA: Springer, 2012, pp. 1–39.
[40] R. Blundell, M. Browning, and I. Crawford, “Best nonparametric bounds on demand responses,” *Econometrica*, vol. 76, no. 6, pp. 1227–1262, 2008.
[41] P. Boggs and J. Tolle, “Sequential quadratic programming,” *Acta Numer.*, vol. 4, pp. 1–51, 1995.
[42] E. Gudes, N. Gal-Oz, and A. Grubshtein, “Methods for computing trust and reputation while preserving privacy,” in *Data and Applications Security XXIII*, E. Gudes and J. Vaidya, Eds. Heidelberg, Germany: Springer, 2009, vol. 5645, pp. 291–298.
[43] L. Mui, “Computational models of trust and reputation: Agents, evolutionary games, and social networks,” Ph.D. dissertation, Elect. Eng. Comput. Eng., MIT, Cambridge, MA, USA, 2002.
[44] V. Bala and S. Goyal, “A noncooperative model of network formation,” *Econometrica*, vol. 68, no. 5, pp. 1181–1229, 2000.
[45] Y. Zhang and M. V. der Schaar, “Strategic networks: Information dissemination and link formation among self-interested agents,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 6, pp. 1115–1123, Jun. 2013.
[46] P. Golle, K. Leyton-Brown, I. Mironov, and M. Lillibridge, “Incentives for sharing in peer-to-peer networks,” in *Electronic Commerce*, L. Fiege, G. Mühl and U. Wilhelm, Eds. Heidelberg, Germany: Springer, 2001, vol. 2232, pp. 75–87.
[47] H. J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed. Berlin, Germany: Springer-Verlag, 2003.
[48] M. Granovetter, “The impact of social structure on economic outcomes,” *J. Econ. Perspect.*, vol. 19, no. 1, pp. 33–50, 2005.
[49] C. Ibars, M. Navarro, and L. Giupponi, “Distributed demand management in smart grid with a congestion game,” in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, 2010, pp. 495–500.
[50] A. Ramachandran, N. Feamster, and S. Vempala, “Filtering spam with behavioral blacklisting,” in *Proc. 14th ACM Conf. Comput. Commun. Sec.*, 2007, pp. 342–351.
[51] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, “The rise of social bots,” arXiv preprint, 2014.
[52] S. Khattak, N. Ramay, K. Khan, A. Syed, and S. Khayam, “A taxonomy of botnet behavior, detection, and defense,” *IEEE Commun. Surveys Tuts.*, vol. 16, no. 2, pp. 898–924, Oct. 2014.
[53] S. Silva, R. Silva, R. Pinto, and R. Salles, “Botnets: A survey,” *Comput. Netw.*, vol. 57, no. 2, pp. 378–403, 2013.
[54] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, “Design and analysis of a social botnet,” *Comput. Netw.*, vol. 57, no. 2, pp. 556–578, 2013.
[55] A. Wang, “Detecting spam bots in online social networking sites: A machine learning approach,” in *Data and Applications Security and Privacy XXIV*. New York, NY, USA: Springer, 2010, pp. 335–342.
[56] J. Dickerson, V. Kagan, and V. Subrahmanian, “Using sentiment to detect bots on Twitter: Are humans more opinionated than bots?” in *Proc. IEEE/ACM Int. Conf. Adv. Soc. Netw. Anal. Min.*, 2014, pp. 620–627.
[57] D. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. Hoboken, NJ, USA: Wiley, 2015.
[58] J. P. Aubin and A. Cellina, *Differential Inclusions: Set-Valued Maps and Viability Theory*. Berlin, Germany: Springer-Verlag, 1984.
[59] P. Billingsley, *Convergence of Probability Measures*. Hoboken, NJ, USA: Wiley, 1968.
[60] V. Krishnamurthy, O. N. Gharehshiran, and M. Hamdi, “Interactive sensing and decision making in social networks,” *Found. Trends Signal Process.*, vol. 7, nos. 1–2, pp. 1–196, 2014.
[61] V. Krishnamurthy, “Quickest detection POMDPs with social learning: Interaction of local and global decision makers,” *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5563–5587, Aug. 2012.
[62] V. Krishnamurthy and H. V. Poor, “A tutorial on interactive sensing in social networks,” *IEEE Trans. Comput. Soc. Syst.*, vol. 1, no. 1, pp. 3–21, Mar. 2014.

Omid Namvar Gharehshiran received the Ph.D. degree from the University of British Columbia, Vancouver, BC, Canada, in 2015. He is currently holding the NSERC Postdoctoral Fellowship with the Actuarial Science and Mathematical Finance Group, Department of Statistical Sciences, University of Toronto, Toronto, ON, Canada. His research interests include stochastic optimization and control, games, and learning theory. He is a former member of the Statistical Signal Processing Group, University of British Columbia.

William Hoiles (S’15) received the M.A.Sc. degree from the Department of Engineering Science, Simon Fraser University, Burnaby, BC, Canada, in 2012, and the Ph.D. degree in electrical and computer engineering from the University of British Columbia, Vancouver, BC, Canada, in 2015. He is currently a Postdoctoral Researcher in electrical and computer engineering with the University of British Columbia. His research interests include social sensors and the bioelectronic interface.

Vikram Krishnamurthy (S’90–M’91–SM’99–F’05) received the Ph.D. degree from the Australian National University, Canberra, A.C.T., Australia, in 1992. He is currently a Professor and holds the Canada Research Chair at the Department of Electrical Engineering, University of British Columbia, Vancouver, BC, Canada. His research interests include statistical signal processing, computational game theory, and stochastic control in social networks. He served as a Distinguished Lecturer for the IEEE Signal Processing Society and the Editor-in-Chief of the IEEE JOURNAL SELECTED TOPICS IN SIGNAL PROCESSING. He was the recipient of an Honorary Doctorate from KTH (Royal Institute of Technology), Sweden, in 2013.