Adaptive Polling in Hierarchical Social Networks Using Blackwell Dominance

Sujay Bhatt and Vikram Krishnamurthy, *Fellow, IEEE*

***Abstract*—This paper presents adaptive polling algorithms and their analysis for social networks having a hierarchical influence structure. The adaptive polling problem on the social network is formulated as a partially observed Markov decision process (POMDP). Our main results exploit the structure of the polling problem to determine novel conditions for Blackwell dominance that arise in hierarchical social influence networks. The Blackwell dominance conditions enable the construction of myopic policies that provably upper bound the optimal policy of the POMDP for adaptive polling. Adaptive versions of intent polling and expectation polling are developed using Blackwell dominance, and they are inexpensive to implement. For polling problems not having a Blackwell dominance structure, the Le Cam deficiency is used to determine approximate Blackwell dominance; this is used to develop an adaptive version of the recently proposed Neighbourhood Expectation Polling algorithm. The proposed Blackwell dominance conditions also facilitate the comparison of Rényi divergence and Shannon capacity of more general channel structures that arise in polling hierarchical social influence networks. Numerical examples are provided to illustrate the adaptive polling policies with parameters estimated from YouTube data.**

***Index Terms*—Adaptive polling, POMDP, structural result, Blackwell dominance, myopic policy, intent polling, expectation polling, Neighborhood Expectation Polling, Shannon capacity, Le Cam deficiency.**

I. INTRODUCTION

BLACKWELL dominance and Le Cam deficiency are widely used in statistical analysis of estimators [1]–[3]; to characterize game-theoretic equilibria [4], and to construct myopic bounds in stochastic control [5], [6]. In this paper, we use Blackwell dominance to construct adaptive polling strategies for social networks exhibiting a hierarchical influence structure. The sufficient conditions we propose for the adaptive polling model also have useful information theoretic interpretations in terms of capacities and Rényi Divergence of more general channel structures.

Polling has applications in forecasting the outcome of an election [7], [8], estimating the fraction of individuals infected with a disease [9], and predicting the success of a particular product. Many social networks have a hierarchical influence structure; [10]–[14]. So, there is strong motivation to develop polling strategies that take into the account the inherent hierarchical social influence. This influence alters the opinions of the lower level individuals and hence affects the prediction or the poll estimate.

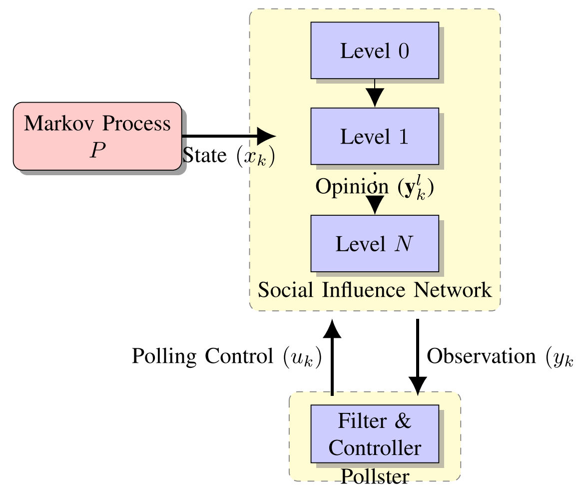

Fig. 1. The figure shows a simple hierarchical influence network where the individuals are grouped into $N+1$ levels Level 0, Level $1, \ldots,$ Level N in a hierarchical fashion. Each level influences the opinion of the level below it. The underlying state of nature x_k determines the opinion. A pollster samples observations y_k from the nodes having opinions $\mathbf{y}^l_k$, runs a local filter to compute the state estimate, and chooses a control to affect the (future) polling action. It is assumed that the pollster knows the number of hierarchical levels in the network and the corresponding node associations. The aim of the pollster is to estimate the underlying state by adapting its polling strategy to incur minimum polling cost.

This paper devises adaptive (feedback control based) polling strategies for the well studied polling algorithms (intent and expectation polling) that, in addition, take the hierarchical social influence and the time varying nature of the state into account; see Fig. 1. The adaptive polling problem is formulated as a partially observed Markov decision process (POMDP) to minimize the polling cost (measurement cost and uncertainty in the Bayesian state estimate).

A. Context. Blackwell Dominance

In general, POMDPs are computationally intractable[1] to solve [16]. The contribution of this paper is to exploit the

Manuscript received May 21, 2018; revised March 21, 2019; accepted May 6, 2019. Date of publication May 21, 2019; date of current version August 6, 2019. This research was supported in part by the Air Force Office of Scientific Research under grant FA9550-18-1-000, National Science Foundation under grant 1714180, and Army Research Office under grant W911NF-17-1-0335. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Waheed U. Bajwa. *(Corresponding author: Sujay Bhatt.)*

The authors are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: sh2376@cornell.edu; vikramk@cornell.edu).

Digital Object Identifier 10.1109/TSIPN.2019.2918442

[1]They are PSPACE hard requiring exponential computational cost (in sample path length) and memory [6], [15].

structure of the social influence network to construct computationally efficient *myopic policies* that provably upper bound the optimal POMDP polling policy. Construction of such myopic bounds involves using the concept of Blackwell dominance of the observation likelihoods.

Since the main results of the paper rely on Blackwell dominance, for convenience, we now define Blackwell dominance and some information theoretic consequences. Blackwell dominance formalizes the notion of which distribution (stochastic matrix) is *more informative* than the other.

Definition (Blackwell Dominance [6], [17]): A stochastic[2] matrix $B(1) \in \mathbb{P}(\mathcal{Y}^{(1)}|\mathcal{X})$ Blackwell dominates (more informative) another stochastic matrix $B(2) \in \mathbb{P}(\mathcal{Y}^{(2)}|\mathcal{X})$ written as $B(1) \succeq_B B(2)$, if

$$B(2) = B(1)R, \text{ for any stochastic matrix } R. \tag{1}$$

Blackwell dominance also has an information theoretic consequence: Consider the classic Discrete Memoryless Channel (DMC) [18] with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ denoted as $\mathbb{P}(\mathcal{Y}|\mathcal{X})$. Let $I(\mathcal{X};\mathcal{Y})$ denote the mutual information of the DMC. The post-processing of channel $B(1)$ in (1) is written as $\mathcal{X} \to \mathcal{Y}^{(1)} \to \mathcal{Y}^{(2)}$. Then from Data Processing Inequality [18], it follows that

$$B(1) \succeq_B B(2) \Rightarrow I(\mathcal{X};\mathcal{Y}^{(1)}) \geq I(\mathcal{X};\mathcal{Y}^{(2)}). \tag{2}$$

Theorem 1 below provides a relation between Blackwell Dominance and Shannon capacity.

Theorem 1 ([3], [19], [20]): For any two conditional distributions $B(1) \in \mathbb{P}(\mathcal{Y}^{(1)}|\mathcal{X})$ and $B(2) \in \mathbb{P}(\mathcal{Y}^{(2)}|\mathcal{X})$,

$$B(1) \succeq_B B(2) \Rightarrow \mathcal{C}^{(1)} \geq \mathcal{C}^{(2)}, \tag{3}$$

where the Shannon capacity $\mathcal{C}^{(i)}$ of a DMC is defined as

$$\mathcal{C}^{(i)} = \sup_{p_{\mathcal{X}}(x)} I(\mathcal{X};\mathcal{Y}^{(i)}),\ i = 1, 2. \tag{4}$$

Here $p_{\mathcal{X}}(x)$ is the marginal distribution over the input alphabet $\mathcal{X}$.

In this paper, we will characterize the capacity for more general channel structures that arise in polling hierarchical social influence networks. Also, Blackwell dominance is used to order the Rényi Divergence [18] of the observation likelihoods of these channels. These information theoretic consequences provide a ranking of these general channel structures in the order of their ability to distinguish the states.

B. Main Results and Organization

1) In Sec. II, the underlying state is modeled as a Markov chain and the adaptive polling problem is formulated as a POMDP. Open loop polling, where polling at a particular instant is not influenced by the information previously collected, is ineffective when the states evolve over time. In comparison, the proposed adaptive (feedback) polling algorithms utilize information previously collected to poll at the next instant. In a hierarchical social influence network, the nodes/ individuals at higher levels in the hierarchy are more influential and so provide more accurate information on the underlying state than the lower levels(see Fig. 1). The proposed adaptive polling algorithms for hierarchical social influence networks also takes this into account. We formulate adaptive generalizations of the *Intent Polling* and *Expectation Polling* methods[3] [21] in Sec. IV and Sec. V.
2) *Blackwell Dominance in Hierarchical Social Influence Networks:* As mentioned above, in general, solving a POMDP is computationally intractable (see Footnote 1). A key property of our POMDP model for adaptive polling is that it exhibits a Blackwell dominance structure. For such POMDPs, a myopic policy provably forms an upper bound to the optimal policy (Theorem 2). For the two adaptive polling algorithms considered, namely, intent polling and expectation polling, we present several novel sufficient conditions for Blackwell dominance involving matrix polynomial functions (Proposition IV.1) and ultrametric matrices (Proposition V.1).
3) *Information Theoretic Interpretations:* The Blackwell dominance relations in turn facilitates the comparison of Rényi Divergence and Shannon capacity of more general channels that arise naturally in hierarchical social influence networks. For example, Proposition IV.1 provides an interesting link between Hurwitz (stable) polynomials and Shannon capacity. While Blackwell Dominance helps compute computationally inexpensive policies that provably upper bound the computationally intractable optimal policy, the information theoretic consequences guide the choice of observation channels (likelihoods) for the pollster.
4) *Approximate Blackwell Dominance:* Blackwell dominance induces a partial order between two stochastic matrices; so not every pair of stochastic matrices is comparable. However, the upper bounds in Theorem 2 provide sufficient motivation to find a pair of matrices that are close to the given pair and are Blackwell comparable. Sec. VI defines the notion of closeness between stochastic matrices using *Le Cam deficiency*. Using this notion of approximate Blackwell dominance, we discuss how to design adaptive polling algorithms using Neighbourhood Expectation Polling [22].
5) *Performance Bounds and Ordinal Sensitivity:* The performance bounds of the mis-specified POMDP model and policy are provided in Sec. VII. This provides a way to obtain the sensitivity of costs for the misclassification of the nodes to different levels in the hierarchical social influence network. Finally, the ordinal sensitivity in polling hierarchical networks, namely, why some networks are inherently more expensive to poll than others.

[2]A $\mathcal{X} \times \mathcal{Y}$ matrix B is (row) stochastic if $\sum_j B_{ij} = 1$ for all $i \in \mathcal{X}, j \in \mathcal{Y}$, and $B_{ij} \in [0, 1]$.

[3][21]: *Intent Polling*– Who will you vote for? *Expectation Polling*– Who do you think will win?

C. Related Literature

[21] analyzes US presidential electoral college results from 1952–2008 where both intention and expectation polling were conducted and shows a remarkable result: In 77 cases where expectation and intent polling pointed to different winners, expectation polling was accurate 78% of the time! Unlike [21], we consider a Bayesian pollster that uses feedback control, and exploits the hierarchical influence structure along with the time evolution of the state (adaptive polling).

[23] analyzes a Bayesian approach to intent and expectation polling and illustrates how the posterior distribution of the leading candidate in the poll can be estimated based on incestuous estimates (each node summarizes the belief of its neighbours, which in turn are influenced by the nodes belief). Unlike [23], we consider hierarchical influence structure and feedback control, in the sense that current estimate dictates where and how to poll in the hierarchical network.

[24] investigates the role played by the network structure in polling by considering the trade-off between number of polled individuals and the bias introduced due to the network structure. [24] concludes that the estimators that consider the network structure into account are considerably more efficient than standard polling estimators. We take the (influence) structure of the network into account, but unlike [24], propose adaptive versions of the polling algorithms.

[25] presents a dynamic Bayesian forecasting method that systematically combines information from historical forecasting models in real time with results from the large number of state-level opinion surveys that are released publicly during the campaign. Similar to [25] we consider a dynamic polling method, but unlike [25] also take the influence structure of the social network into account.

II. Formulation of Adaptive Polling

Sec. II-A introduces the model for the adaptive polling problem, Sec. II-B provides a discussion of the model and illustrates the setup with a motivating example and Sec. II-C formulates the adaptive polling problem as a partially observed Markov decision process (POMDP).

A. Adaptive Polling as a POMDP

We consider the typical framework for information diffusion and formation of opinions in a social network. The underlying state (true sentiment underlying social media message, popularity of a product/political party, quality of commercial product) evolves over time stochastically [26]–[30]. This underlying state is observed by the individuals in the social network through tweets, political commentary blogs and videos, or reviews on social media. Using the available information and interaction with neighbours, individuals form opinions about the underlying state. The population is classified into $N+1$ levels depending on the hierarchical influence as shown in Fig. 1. The population is sampled sequentially by a pollster to gather the information on the underlying state.

How should the pollster poll the hierarchical social network to estimate the state while minimizing the polling cost (measurement cost and uncertainty in the Bayesian state estimate)? We formulate this adaptive polling problem as a partially observed Markov decision process (POMDP). POMDPs provide a principled framework for sequential decision making problems with feedback control in partially observed domains. This formalism as a POMDP casts the adaptive polling problem as a stochastic control problem. We refer to [6, Chapter 7] for a detailed treatment of POMDPs - due to space restrictions we give a very terse description.

The POMDP for adaptive polling is specified by

$$\theta = (\mathcal{X}, \mathcal{Y}, \mathbb{Y}, \mathcal{U}, C, P, O(u), \rho), \tag{5}$$

where $\mathcal{X}$ denotes the state space, $\mathcal{Y}$ denotes the observation space, $\mathbb{Y}$ denotes the opinion space, $\mathcal{U}$ denotes the control/ action space, C denotes the state-action cost matrix, P is the state transition matrix, $O(u)$ is the control dependent observation distribution/ likelihood matrix and $\rho \in [0,1)$ is the economic discount factor.

We now describe the above 8 components of the model (5):

State: Let $x_k \in \mathcal{X} = \{1, 2, \ldots, X\}$ denote a Markov chain evolving at discrete time instants $k = 0, 1, \cdots$ on a finite state space. As mentioned previously, the state models the time evolving ground truth (sentiment, popularity, quality) quantized into a finite number of levels.

Transition matrix: Let P denote the transition probability matrix of the Markov chain x_k with elements

$$P_{ij} = \mathbb{P}(x_{k+1} = j | x_k = i), \; i, j \in \mathcal{X}, \forall\, k. \tag{6}$$

Pollster's control/actions: $\mathcal{U} = \{1, 2, \ldots, U\}$ denotes the set of possible controls (actions), with $u_k \in \mathcal{U}$ denoting the action chosen at time k. For example, the action can denote the choice of the hierarchical level the pollster seeks the opinion.

Polling cost: Let $C(x_k, u_k)$ denote the instantaneous cost incurred by the pollster for taking action u_k when in state x_k. This models the measurement cost and quality (accuracy) of the polling algorithm. For example, conducting surveys and opinion polls incurs a measurement cost and the type of poll conducted affects the quality of the information gathered.

Pollster's observation distribution: Let $\mathcal{Y}$ denote a finite set of observations with $y_k \in \mathcal{Y}$ representing the observations of the underlying state $x_k \in \mathcal{X}$. The observations $y_k \in \mathcal{Y}$ for the pollster model the information on the state gathered via views/ sentiments expressed by the nodes or individuals in the hierarchical social network (see Fig. 1). Let $O(u)$ denote the observation probability matrix with elements

$$O_{ij}(u) = \mathbb{P}(y_{k+1} = j | x_{k+1} = i, u_k = u), \; i \in \mathcal{X}, j \in \mathcal{Y}, \forall k. \tag{7}$$

At each time k, the pollster receives an observation y_k on the underlying state x_k after taking an action u_k. The observation matrix/ distribution $O(u)$ models the likelihood of the observations $y_k \in \mathcal{Y}$ given the state $x_k \in \mathcal{X}$, and is different for different polling algorithms.

The observations obtained by the pollster are the *opinions* about the state provided by the nodes. We now discuss the opinion dynamics, the corresponding opinion distribution, and how the observation distribution can be expressed in terms of the opinion distributions:

i) ***Opinion dynamics:*** Let $\mathbf{y}_k^l \in \mathbb{Y}$ denote the opinion of nodes at level l of the hierarchical network (see Fig. 1). Here $|\mathbb{Y}| = |\mathcal{X}|$. The opinion dynamics in Fig. 1 proceeds according to the following protocol for $k = 0, 1, \cdots$
   a) The state x_k evolves on time scale k.
   b) Opinions $\mathbf{y}_k^l$, for $l = 1, 2, \ldots, N$, are formed at the Level l at the fast time-scale $\bar{k} = k + l\delta$, where $0 < \delta \ll 1$.
   c) At time $k+1$, state transitions to x_{k+1}.
   We assume that $N\delta \ll 1$, where the number of levels in Fig. 1 is $N+1$. This implies that the state x_k is evolving over a slower time-scale than the time-scale over which the opinions are formed across the network given in Fig. 1.

ii) ***Opinion distribution:*** The opinions at different levels in the hierarchical social influence network are formed via information diffusion as follows [31], [32]: opinion at the highest level $\mathbf{y}_k^0$ is directly influenced by the state x_k. Opinion $\mathbf{y}_k^l$, $l \geq 0$ influences $\mathbf{y}_k^{l+1}$ (see Fig. 1). This is modeled probabilistically as $\mathbb{P}(\mathbf{y}_k^{l+1} = j | \mathbf{y}_k^l = i)$.
Let the opinion distribution at Level 0 be given by the stochastic matrix B having elements
$$B_{ij} = \mathbb{P}(\mathbf{y}_k^0 = j | x_k = i), \ i \in \mathcal{X}, j \in \mathbb{Y}, \forall k. \tag{8}$$
The opinions at levels $l \in \{1, \ldots, N\}$ in the hierarchical network are directly influenced by the preceding levels (see Fig. 1). The opinions at levels $l \in \{1, \ldots, N\}$ are given by $(B_l)_{ij} = \mathbb{P}(\mathbf{y}_k^l = j | x_k = i)$ for $i \in \mathcal{X}, j \in \mathbb{Y}$, and $\forall k$. The opinions at Level l are determined by the opinion distribution via the following decomposition
$$(B_l)_{ij} = \sum_{m \in \mathcal{Y}} \mathbb{P}(\mathbf{y}_k^l = j | \mathbf{y}_k^{l-1} = m) \times \mathbb{P}(\mathbf{y}_k^{l-1} = m | x_k = i). \tag{9}$$
For tractability, assume[4] that the confusion matrix between successive levels is modeled using the same time-homogeneous opinion distribution B in (8). So the opinions at levels $l \in \{0, 1, \ldots, N\}$ have an opinion distribution
$$B_l = B^{l+1}, \tag{10}$$
where B_l denotes the opinion distribution at level l.

iii) ***Observation distribution via Opinion distribution:*** Since the observations for the pollster, to update the estimate of the state, are the opinions from the nodes, the observation distribution (7) is directly related to the opinion distribution (8). For example, in case of adaptive intent polling (Sec. IV below), $O(u) = Bf_u(B)$, where f_u is any matrix polynomial, where the probabilities with which the nodes at different levels in the hierarchical social network are polled are proportional to the co-efficients of the polynomial f_u; and in case of adaptive expectation polling (Sec. V below), $O(u) = B_l^{l_u/l}$, where the nodes at level l are polled to provide information on the nodes at level l_u.

[4]This is a modeling assumption, and Example 1 in Sec. VIII shows how to estimate such a structure using a modified EM algorithm. For the case where it is known a priori that the distribution (confusion matrix) between the hierarchical levels are different, results in Sec. VI on Approximate Blackwell Dominance can be used to obtain the policies and performance bounds on the proposed polling algorithms in Sec. IV and Sec. V.

Polling Objective: The actions taken by the pollster influences the noisy state-observations via the selection of the observation distribution. The goal of the pollster is to choose an action, based on the history of past actions and observations, that minimizes the expected costs incurred over time. We consider the following infinite horizon discounted cost for specifying the objective [6, Chapter 7]:
$$J_\mu(\pi_0; \theta) = \mathbb{E}_\mu \left\{ \sum_{k=0}^{\infty} \rho^k C(x_k, u_k = \mu(\mathcal{I}_k)) | \pi_0 \right\}. \tag{11}$$
Here $J_\mu(\pi_0; \theta)$ denotes the expected cumulative cost with respect to the stationary (time-independent) policy μ, $\mathcal{I}_k = \{\pi_0, u_0, y_1, \ldots, u_{k-1}, y_{k-1}\}$ denotes the history of past actions and observations and $\pi_0 = (\pi_0(i), i \in \mathcal{X})$, where $\pi_0(i) = \mathbb{P}(x_0 = i)$ is the initial probability distribution over the state space. The objective of the pollster is to find the optimal stationary policy μ^* such that
$$J_{\mu^*}(\pi_0; \theta) = \inf_{\mu \in \boldsymbol{\mu}} J_\mu(\pi_0; \theta) \tag{12}$$
where $\boldsymbol{\mu}$ denotes the class of stationary policies.

Discounting in Polling: The parameter $\rho \in [0, 1)$ is an economic discount factor that determines the way the polling cost is counted towards the *polling value* defined in (12), and affects the optimal policy μ^*. Choosing $\rho = 0$ implies that the pollster is myopic, and only minimizes the instantaneous polling cost $C(x, u)$ without considering future polling costs. Choosing $\rho > 0$ implies that the pollster geometrically weighs the polling costs incurred in the future.

Summary: We have formulated adaptive polling as a POMDP parametrized by θ, defined in (5). The infinite horizon objective (11) is for notational convenience. Our main result will exploit Blackwell dominance to construct a myopic upper bound to the optimal policy μ^* in (12). For a finite horizon formulation - the optimal policy is non-stationary - but all subsequent results in this paper continue to hold.

B. Discussion of Model

1) We assume that the POMDP model θ in (5) is known. This implies that the number of hierarchical levels and nodes-level associations are known to the pollster. Otherwise the problem becomes an adaptive stochastic control problem which is intractable to solve. Note that Blackwell dominance is a class type result - even if the observation probabilities (7) are not known exactly, as long as the Blackwell dominance condition is satisfied, the main result (Theorem 2 below) holds.

2) The opinion dynamics are such that the entire network holds the view on the state x_k at time k, between times k and $k+1$. This modeling assumption has two implications: (i) It enables the decomposition of the opinion distributions (9), (ii) It implies that during the sampling instant, the information gathered by the pollster, from anywhere in the hierarchical network, pertains to the same underlying state.
3) Opinions from higher levels or more reputable sources (see Fig. 1), are more informative (Blackwell sense) and hence the information acquisition is costlier compared to lower levels. This is motivated by study in [33], which shows that information acquisition from more informative distributions (in the sense of Blackwell) is more costly. The intuition is that nodes at a higher level pay more attention to acquire information to form an opinion, and hence require commensurate compensation to divulge that information to the pollster. The assumption on the cost $C(x,u)$ captures this intuition.
4) *Example:* Consider estimating the severity of a natural disaster like an earthquake, modeled as as $x(\text{Severity}) = \{\text{Low Damage, High Damage}\}$ via Twitter. The pollster's observations and opinions formed by the nodes at level l could be modeled as $y = \mathbf{y}^l = \{\text{Opinion.1, Opinion.2}\}$. The pollster incurs a measurement cost for obtaining information from the people who tweet, and a cost for the accuracy or the uncertainty reduction in the state estimate. In this example application, polling opinions from the influencers (knowledgeable participants) incurs a higher measurement (processing) cost but their opinions will result in a larger reduction in uncertainty in the state estimate. The pollster's objective is to estimate the disaster intensity by polling observations from the participant pool while incurring the least polling cost on average.

C. Stochastic Dynamic Programming for Adaptive Polling

In this section, the solution of (12) is formulated as a stochastic dynamic programming problem over the $X-$ dimensional unit simplex $\Pi(X) = \{\pi : \pi(i) \in [0,1], \sum_{i=1}^{X} \pi(i) = 1\}$ of posterior probabilities (beliefs); see [6] for details.

Belief State Formulation: Let π_k denote the belief at time k and the i^{th} element $\pi_k(i)$ is:

$$\pi_k(i) = \mathbb{P}(x_k = i|\mathcal{I}_k) \tag{13}$$

where $x = \{1,2,3,\ldots,X\}$ denotes the state space and $\mathcal{I}_k = \{\pi_0, u_0, y_1, \ldots, u_{k-1}, y_{k-1}\}$ denotes the history of past actions and observations. The belief (13) is computed from the opinions gathered by the pollster, and is a sufficient statistic [6] for the history of actions and opinions $\{u_1, y_1, \ldots, u_{k-1}, y_{k-1}\}$. The dynamics of the POMDP is given by the Bayesian filtering update

$$\pi_k = T(\pi_{k-1}, y_k, u_k), \text{ where } T(\pi, y, u) = \frac{O_y(u)P'\pi}{\mathbf{1}'O_y(u)P'\pi} \tag{14}$$

and $O_y(u) = \text{diag}(O_{1y}(u), \ldots, O_{Xy}(u))$. Here $\mathbf{1}$ is the column vector of 1s and P' denotes the matrix transpose.

As is well known in POMDPs, instantaneous cost $C(\pi_k, u_k)$ in terms of the belief π_k given by

$$C(\pi_k, u_k) = \sum_i C(x_k = i, u_k)\pi_k(i), \tag{15}$$

where π_k is the belief at time k. The costs $C(\pi, u)$ in (15) capture the cost of measurement and the uncertainty or error in the state estimate. Any non-linear cost can be used in the formulation of the polling problems. In this paper, we consider the following non-linear costs [6, Chapter 8] – entropy and state-estimation error – to illustrate the different formulations.

Associated with a stationary polling policy μ and initial belief $\pi_0 \in \Pi(X)$, the objective (11) can be re-expressed as:

$$J_\mu(\pi_0; \theta) = \mathbb{E}_\mu\left\{\sum_{k=0}^{\infty} \rho^k C(\pi_k, u_k = \mu(\pi_{k-1}))|\pi_0\right\}. \tag{16}$$

Our aim is to find the optimal stationary polling policy $\mu^* : \Pi(X) \to \mathcal{U}$ defined in (12).

Stochastic Dynamic Programming: Obtaining the optimal policy μ^* in (12) is equivalent to solving Bellman's stochastic dynamic programming equation [6, Chapter 7]:

$$\begin{aligned} \mu^*(\pi) &= \arg\min_{u\in\mathcal{U}} \mathcal{Q}(\pi, u) \\ J_{\mu^*}(\pi; \theta) &= V(\pi) = \min_{u\in\mathcal{U}} \mathcal{Q}(\pi, u), \text{ where} \\ \mathcal{Q}(\pi, u) &= C(\pi, u) + \rho \sum_{y\in\mathcal{Y}} V(T(\pi, y, u))\sigma(\pi, y, u). \end{aligned} \tag{17}$$

Here $\sigma(\pi, y, u) = \mathbf{1}'O_y(u)P'\pi$ is the measure on the observation alphabet $\mathcal{Y}$, $V(\pi)$ is the value function denoting the minimum expected cost, $T(\pi, y, u)$ is the Bayesian filtering update (14), $\mathcal{Q}(\pi, u)$ is the state-action value function ($\mathcal{Q}-$ function), $C(\pi, u)$ is the state-action cost in terms of the belief state (15) and ρ is the discount factor.

Since the belief space $\Pi(X)$ is a continuum, Bellman's equation (17) does not translate into practical solution methodologies as $V(\pi)$ needs to be evaluated at each $\pi \in \Pi(X)$. The computation of optimal policy of the POMDP is P-SPACE hard [16]. Also, the costs $C(\pi, u)$ capture the cost of measurement and the uncertainty or error in the state estimate, and hence are non-linear in the belief. In order to capture the uncertainty in the estimate it is necessary to use a non-linear cost on the belief state. The non-linearity is required so that the costs are zero at the vertices of the belief space $\Pi(X)$ (reflecting perfect state estimation) and largest at the centroid of the belief space (most uncertain estimate). This results in a non-standard[5] POMDP. This motivates the construction of *optimal upper bound policy* $\bar{\mu}(\pi)$ to the optimal policy $\mu^*(\pi)$ that is inexpensive to compute. In the remainder of the section, we construct such upper bound policies in terms of easily computable myopic policies using *Blackwell dominance*, for adaptive polling problems formulated as POMDPs.

[5]POMDP solvers can only handle POMDPs with linear costs, see [6, Chapter 7].

III. Meta-Theorems for Adaptive Polling

Given the POMDP model for adaptive polling (5) and the polling objective (12), the aim of this section is to describe the key meta theorems that will be used in subsequent sections to develop efficient adaptive polling algorithms. In this section, we provide two main theorems using Blackwell dominance: (i) A structural result using Blackwell dominance for the adaptive polling POMDP, i.e., characterization of achievable performance without brute force computations but using mathematical analysis, (ii) An information theoretic consequence of Blackwell dominance, namely, Rényi divergence, and why this is useful for the pollster.

A. Main Result. Optimality of Myopic Polling Policies

The aim of the pollster is to estimate the (underlying) evolving state x_k by incurring minimum cumulative cost (16). The pollster employs the control $u_k = \mu^*(\pi_{k-1})$ to obtain opinions ($y_k \in \mathcal{Y}$) from the nodes, and then updates the belief $\pi_{k-1} \to \pi_k$ about the underlying state $x_k \in \mathcal{X}$ using (14).

Define the myopic policy $\bar{\mu}(\pi)$ as

$$\bar{\mu}(\pi) = \underset{u\in\mathcal{U}}{\arg\min}\, C(\pi, u) \tag{18}$$

Theorem 2 below provides sufficient conditions on the observation distribution of the pollster $O(u)$ so that a myopic polling policy (18) upper bounds the optimal polling policy in (12).

Theorem 2 (Optimality of Myopic Policies via Blackwell Dominance): Consider the adaptive polling problem formulated in Sec. II-A as a POMDP. Assume that the cost $C(\pi, u)$ is concave in π. Suppose $O(u) \succeq_B O(u+1)\, \forall u \in \mathcal{U}$. Then $\bar{\mu}(\pi)$ defined in (18) is an upper bound to the optimal polling policy $\mu^*(\pi)$ defined in (17), i.e., $\mu^*(\pi) \leq \bar{\mu}(\pi)$ for all $\pi \in \Pi$. In particular, for belief states where $\bar{\mu}(\pi) = 1$, the myopic policy coincides with the optimal policy $\mu^*(\pi)$.

Discussion: The concavity of the costs $C(\pi, u)$ for each polling action u implies that the value function $V(\pi)$ is concave [6, Theorem 8.4.1] on $\Pi(X)$. This together with Jensen's inequality is used to establish Theorem 2. Theorem 2 is a well known structural result for POMDPs [6, Chapter 14, Sec.14.7], and it says the following:

i) The instantaneous costs satisfying $C(\pi, 1) < C(\pi, u)$ for $u = 2, \ldots, U$ does not trivially imply that the myopic policy $\bar{\mu}(\pi)$ in (18) coincides with the optimal policy $\mu(\pi)$, since the optimal policy applies to the cumulative cost function involving an infinite horizon trajectory of the dynamical system. But when $O(u) \succeq_B O(u+1)\, \forall\, u \in \mathcal{U}$ and the costs are concave, the myopic policy coincides and forms a provably optimal upper bound to the computationally intractable optimal policy. The concavity of the costs imply that extremes are better than averages, reflecting perfect state estimation; and captures the effect of increasing marginal utility.

ii) The trivial sub-optimal policy $\hat{\mu}(\pi) = 1\, \forall \pi \in \Pi(X)$ is also an upper bound to the optimal policy - but a useless bound because $\mu^*(\pi) = 1 \Rightarrow \hat{\mu}(\pi) = 1$. In comparison, the upper bound constructed via Theorem 2 (Blackwell dominance) says that $\bar{\mu}(\pi) = 1 \Rightarrow \mu^*(\pi) = 1$, which is a much more useful construction. Thus, the myopic polling policy forms a provably optimal upper bound to the computationally intractable optimal policy.

B. Blackwell Dominance and Rényi Divergence Interpretation

In this section, we will discuss the information theoretic consequences of Blackwell dominance. While Blackwell Dominance helps compute inexpensive policies that provably upper bound the computationally intractable optimal policy, the information theoretic consequences guide the choice of observation channels (likelihoods) for the pollster.

Rényi Divergence is a generalization of the Kullback-Leibler divergence [18], and it measures the dissimilarity between two distributions. Theorem 3 below shows the relation between Rényi Divergence and Blackwell dominance.

With a slight abuse of notation in (7), let $O_i(u)$ denote the i^{th} row of the observation likelihood matrix $O(u)$. In words, $O_i(u)$ is the distribution over the observation alphabet $\mathcal{Y}$ conditional on the state $x = i$.

Definition: (Rényi Divergence) For an observation likelihood $O(u)$, the Rényi Divergence of order $\alpha \in [0, 1)$ for $i, j \in \mathcal{X}$ is defined as

$$D_\alpha(O_i(u)||O_j(u)) = \frac{1}{\alpha - 1} \log \sum_{y\in\mathcal{Y}} O_{iy}^{\alpha}(u) O_{jy}^{\alpha-1}(u). \tag{19}$$

Theorem 3 (Ordering of Rényi Divergence): If the observation distribution for the pollster satisfy $O(u) \succeq_B O(u+1)\, \forall u \in \mathcal{U}$, then for any $i, j \in \mathcal{X}$:

$$D_\alpha(O_i(u)||O_j(u)) \geq D_\alpha(O_i(u+1)||O_j(u+1))\, \forall u \in \mathcal{U}. \tag{20}$$

Discussion: Theorem 3 says that when $O(u) \succeq_B O(u+1)$, conditional on the state, the observation distributions are more dissimilar in case of $O(u)$. Here, more the dissimilarity, better the pollster is able to distinguish the states. In other words, Theorem 3 provides a ranking of channel structures in terms of their ability to distinguish the states. In Sec. IV and Sec. V, we discuss the ordering of Rényi Divergence for more general channels that arise in hierarchical social influence networks, where the information theoretic consequences guide the choice of the observation distributions for the pollster.

IV. Adaptive Intent Polling Algorithm

In intent polling [21], to decide between two states, the sampled individuals are asked "who will you vote for?". In this section, we develop an adaptive version of intent polling [21]: the resulting algorithm (Algorithm 1 below) is designed for hierarchical social networks with time-varying state of nature.

We present novel sufficient conditions for Blackwell dominance in the context of adaptive intent polling. These conditions enable the application of Theorem 2 to determine myopic policies that upper bound the optimal adaptive intent polling policy. These myopic polices are used for polling in Algorithm 1, which is inexpensive to implement.

Algorithm 1: Adaptive Intent Polling for Pollster.

1 **Polling Policy:** Compute the myopic adaptive intent polling policy $\bar{\mu}_I : \Pi(X) \rightarrow \mathcal{U}$ that maps beliefs to polling actions.
2 For an initial belief π_0, **Loop** $k = 1, 2, \cdots$:
3 **Polling Action:** Polling action $u_k = \bar{\mu}_I(\pi_{k-1})$ is a choice of a distribution $\beta^{(u_k)}$, where $\beta^{(u_k)} = (\beta_0^{(u_k)}, \beta_1^{(u_k)}, \cdots \beta_N^{(u_k)})$ and $\sum_i \beta_i^{(u_k)} = 1$ and $N+1$ is the number of levels in the network. Poll a node at level l with a probability $\beta_l^{(u_k)}$ and ask the following question to obtain the observation y_k:

“What does it (a node at level l) think the state is?”

4 **State-estimation:** Estimate the state π_k using the Bayesian filtering update (14) with observation distribution $O(u_k) = B f_{u_k}(B)$, where B is the opinion distribution (8) and $f_{u_k}(z)$ is a Hurwitz polynomial.
5 **Polling Cost:** Incur an intent polling cost $C(\pi_k, u_k) = S(\beta^{(u_k)}) + \eta_e(\pi_k, u_k)$ that is composed of measurement and entropy costs respectively.
6 **End**

In the adaptive intent polling (Algorithm 1), the pollster adapts the intent polling policies, namely, the probabilities with which the nodes at different levels in the hierarchical social influence network are polled. This affects the observation distribution $O(u)$, and hence the state estimate (see Fig. 1).

A. Formulation of Intent Polling Costs

The instantaneous cost in adaptive intent polling consists of two components– the measurement cost and the entropy cost (uncertainty in the state estimate):

a) *Measurement Cost:* Let $u \in \{1, 2, \ldots, U\}$ denote the choice of distributions (polling actions) $\beta^{(u)}$, where $\beta^{(u)} = (\beta_0^{(u)}, \beta_1^{(u)}, \cdots \beta_N^{(u)})$ and $\sum_i \beta_i^{(u)} = 1$. Here $\beta_l^{(u)}$ for $l = 0, 1, 2 \cdots, N$ denotes the probability of selecting a node from level l, having an opinion distribution B^{l+1}. Let $s(l)$ denote the measurement cost from level l. Since nodes at higher levels in the hierarchy (small l) provide more informative (in the Blackwell sense) observations, higher costs are associated with obtaining observations from these levels [33], i.e., $s(l) \geq s(l+1)$, and the average measurement cost for employing the polling algorithm $\beta^{(u)}$ is $S(\beta^{(u)}) = \sum_{l=0}^{N} \beta_l^{(u)} s(l)$.

b) *Entropy Cost:* The entropy cost models the uncertainty in the state estimate π in (13), and is given as

$$\eta_e(\pi, u) = -\gamma_1(u) \sum_{i=1}^{2} \pi(i) \log_2 \pi(i) + \gamma_2(u)$$

for $\pi_k(i) \in (0, 1)$ and $\eta_e \stackrel{\Delta}{=} 0$ for $\pi_k(i) = \{0, 1\}$. Here $\gamma_1, \gamma_2 > 0$ are user defined scalar weights.

Since more informative opinions lead to larger reduction in uncertainty, $\gamma_1(u) > \gamma_1(u+1)$ and $\gamma_2(u+1) > \gamma_2(u)$.

The net instantaneous cost $C(\pi, u)$ in (15) incurred by the pollster in case of adaptive intent polling is thus given as:

$$C(\pi, u) = S(\beta^{(u)}) + \eta_e(\pi, u). \tag{21}$$

The cost (21) expressed in terms of the belief state π captures the fact that a control with higher measurement cost should result in a smaller entropy (more reduction in uncertainty) cost and vice versa.

B. Matrix polynomials and Blackwell Dominance

The aim of this section is to provide a rationale for choosing the intent polling actions (distributions) $\beta^{(u)}$, $u \in \mathcal{U}$, with $f_u(z) = \sum_{l=0}^{N} \beta_l^{(u)} z^l$ denoting the polynomial associated with action distributions $\beta^{(u)}$. It is shown that when $f_u(z)$ is Hurwitz stable, there exists a Blackwell dominance relation between the observation distributions (7) for the pollster. Let $\mathcal{P}_N = \{h | h(z) = \sum_{i=0}^{N} \beta_i z^i, \sum_{i=0}^{N} \beta_i = 1, \beta_i \geq 0\}$ denote the set of all polynomials with co-efficients that constitute a convex combination.

Proposition IV.1: Let Q be a stochastic matrix. For $n > m$, let $p(z) \in \mathcal{P}_n$ and $q(z) \in \mathcal{P}_m$ be two polynomials such that all the roots of $q(z)$ are roots of $p(z)$. If $q(z)$ and $p(z)$ are Hurwitz, then $q(Q) \succeq_B p(Q)$.

Discussion: The above result is interesting since it shows that Hurwitz stability is a sufficient condition for Blackwell dominance. Consider two polynomial channels $p(Q)$ and $q(Q)$ where Q is any stochastic matrix. Proposition IV.1 says that polling with observation channel $q(Q)$ is always better than polling with channel $p(Q)$. In adaptive intent polling (Theorem 4), the degree of the polynomial is the same as the number of levels in the hierarchy (Fig. 1). A polling action in adaptive intent polling corresponds to choosing the (normalized) co-efficients of a polynomial, and these coefficients are the probabilities of polling from the various levels of the social network.

According to Proposition IV.1, if the polynomials are Hurwitz and have common factors, a Blackwell dominance relation exists between their corresponding matrix polynomials. If, however, $p(z) \in \mathcal{P}_n$ is not a Hurwitz polynomial, then $q(Q) \succeq_B p(Q)$ only if the polynomial $q(z) \in \mathcal{P}_m$ is the single quadratic factor ($m = 2$) corresponding to any conjugate pair of zeros of $p(z)$ having smallest argument in magnitude; see [34].

C. Main Result. Myopic Policies for Adaptive Intent Polling

Our main result on adaptive intent polling is Theorem 4 below. It shows that when it cheaper for the pollster to (myopically) listen to the polynomial channel that provides largest reduction in uncertainty on the state, it is indeed optimal to do that. Polynomial channels are parallel cascaded channels that model the communication medium between the pollster and the nodes of a social network having a hierarchical influence structure as in Fig. 1, when the pollster polls all levels of the hierarchical network as in intent polling.

Let $f_u(z) = \sum_{l=0}^{N} \beta_l^{(u)} z^l$ denote the polynomial corresponding to the polling policy $\beta^{(u)}$. For an opinion distribution B (defined in (8)), let the matrix polynomials be $f_u(B)\, \forall u \in \mathcal{U}$.

Theorem 4 (Adaptive Intent Polling): Consider the adaptive intent polling problem with costs specified in (21). Let the observation distribution for the pollster be $O(u) = Bf_u(B)\, \forall u \in \mathcal{U}$. Assume that the polynomial $f_U(z) \in \mathcal{P}_N$ is Hurwitz.[6]

a) Then, $O(u) \succeq_B O(u+1)\, \forall u \in \mathcal{U}$.

b) By Theorem 2, the myopic intent polling policy $\bar{\mu}_I(\pi)$ forms an upper bound to the optimal intent polling policy $\mu_I^*(\pi)$, i.e., $\mu_I^*(\pi) \leq \bar{\mu}_I(\pi)$ for all $\pi \in \Pi$. In particular, for belief states where $\bar{\mu}_I(\pi) = 1$, the myopic policy coincides with the optimal policy $\mu_I^*(\pi)$.

Discussion: The instantaneous cost for adaptive intent polling (21) is concave in π by definition. The proof of Theorem 4 follows from Proposition IV.1 and Theorem 2. The adaptive intent polling algorithm employed by the pollster determines how the opinions are gathered, and the opinions are distributed as $O(u)$ for the pollster. For an opinion distribution B, the observation distribution of the pollster in case of adaptive intent polling is given as $O(u) = Bf_u(B)$, where $f_u(B) = \sum_{l=0}^{N} \beta_l^{(u)} B^l$ and nodes at level l are sampled with probability $\beta_l^{(u)}$. The matrix polynomial $f_u(B)$ has an identity observation likelihood for the co-efficient $\beta_0^{(u)}$. This motivates the choice $O(u) = Bf_u(B)\, \forall u \in \mathcal{U}$.

Proposition IV.1 provides a justification for the polynomial $f_U(z)$ to be Hurwitz. If $f_U(z)$ is Hurwitz, then a way to compute $f_g(z)$ for $g \in \{U-1, \ldots, 2, 1\}$ is by successive long-division of $f_U(z)$ by linear or quadratic factors of $f_U(z)$. If we know that the polynomial $f_U(z)$ is Hurwitz, then the polynomial $f_{U-1}(z)$ obtained by removing any linear or quadratic factor from $f_U(z)$ is also Hurwitz. From Proposition IV.1, we know that if two polynomials are Hurwitz, there is a Blackwell dominance relation between them. From Theorem 4, the observation distribution for the pollster is ordered as

$$O(U-1) = B \cdot f_{U-1}(B) \succeq_B B \cdot f_U(B) = O(U).$$

A similar procedure can be carried out to obtain observation distributions for $u = U-2, \ldots, 1$ as long as the number of levels $N+1$ is greater than U in the hierarchical social network, as there are $N+1$ roots for a polynomial of degree $N+1$.

D. Information Theoretic Interpretation

The aim of this section is to provide an interesting link between Hurwitz stability and channel capacity in terms of Blackwell dominance. Let $I(\mathcal{X}; \mathcal{Y}^{(u)})$ denote the mutual information of channel $f_u(B)$ and $\mathcal{C}^{(u)}$ denote the capacity defined in (4). Let $f_u^i(B)$ denote the i^{th} row of the matrix polynomial $f_u(B)$.

Proposition IV.2: If the channel error probabilities (likelihoods) for the pollster satisfy $f_u(B) \succeq_B f_{u+1}(B)\, \forall u \in \mathcal{U}$, then

a) Shannon Capacity Ordering: $\mathcal{C}^{(u)} \geq \mathcal{C}^{(u+1)}\, \forall u \in \mathcal{U}$.

b) Rényi Divergence Ordering:

$$D_\alpha(f_u^i(B)||f_u^j(B)) \geq D_\alpha(f_{u+1}^i(B)||f_{u+1}^j(B))$$

for all $u \in \mathcal{U}$ and for all $i, j \in \mathcal{X}$.

Discussion: The proof of Proposition IV.2 follows from Theorem 1 and Theorem 3. From Proposition IV.2, the Hurwitz polynomial channels are ordered such that the channel that is a sub channel of the other results in a larger reduction in uncertainty on the state.

Together with Proposition IV.1, Proposition IV.2 provides an interesting link between Hurwitz (stable) polynomials and channel capacity. From Proposition IV.1, those polling actions that result in Hurwitz (stable) polynomials allow decomposition of channels into sub channels that have higher capacity from Proposition IV.2.

[6]A polynomial f is Hurwitz if all its zeroes lie in the open left half-plane of the complex plane, and all its co-efficients have the same sign.

Algorithm 2: Adaptive Expectation Polling for Pollster.

1 **Polling Policy:** Compute the myopic adaptive expectation polling policy $\bar{\mu}_E : \Pi(X) \to \mathcal{U}$ that maps beliefs to polling actions.
2 For an initial belief π_0, **Loop** $k = 1, 2, \cdots$:
3 **Polling Action:** Polling action $u_k = \bar{\mu}_E(\pi_{k-1})$ is a choice of level in the network. Poll a node at level l and ask the following question to obtain the observation y_k:

"what does a node at level l think

the nodes at level $j(< l)$ would report the state as?"

4 **State-estimation:** Estimate the state using the Bayesian filtering update (14) with observation distribution $O(u_k) = B_l^{l_{u_k}/l}$, where B_l is the opinion distribution (10) and B in (8) is an ultrametric matrix. Here nodes at level l are polled to provide information of the nodes at level l_{u_k}.
5 **Polling Cost:** Incur an expectation polling cost $C(\pi_k, u_k) = S(\beta^{(u_k)}) + \eta_2(\pi_k, u_k)$ that is composed of measurement and state estimation error respectively.
6 **End**

V. Adaptive Expectation Polling

In expectation polling [21], to decide between two states, the sampled individuals are asked "who will your friends vote for?". In a hierarchical network, this can be seen as asking "who will your more influential friends vote for?". In this section, we develop an adaptive version of expectation polling [21]: the resulting algorithm (Algorithm 2 below) is designed for hierarchical social influence networks with time-varying state of nature.

We present novel sufficient conditions for Blackwell dominance in the context of adaptive expectation polling. These conditions involving ultrametric matrices enable the application of Theorem 2 to determine myopic policies that upper bound the optimal adaptive expectation polling policy. The myopic policies are used for polling in Algorithm 2, which is inexpensive to implement.

Algorithm 2 below is a more sophisticated version of standard expectation polling, for multiple states and hierarchical social networks.

In adaptive expectation polling (Algorithm 2), the pollster controls the observation distribution $O(u)$ by choosing different levels to gather the opinion, and this in turn affects the estimate of the state (see Fig. 1).

A. Formulation of Expectation Polling Costs

The instantaneous cost in adaptive expectation polling consists of two components– the measurement cost and the uncertainty in the state estimate:

a) *Measurement Cost:* Let $u \in \{1, 2, \ldots, U\}$ denote the choice of levels. In adaptive expectation polling, unlike adaptive intent polling, not all levels are polled. The pollster selects a level l and asks the nodes at level l to provide information about the other levels. Let $S(u)$ denote the measurement cost for action u. Since more informative opinions are costlier to obtain [33], from Theorem 5(i) below, $S(u) \geq S(u+1)\, \forall u \in \mathcal{U}$.

b) *State-Estimation error:* The state-estimation error incurred in choosing action u is modelled as

$$\eta_2(\bar{x}, u) = w_u \|\bar{x} - \pi\|_2. \tag{22}$$

The scalar $w_u > 0$ allows the costs associated with different controls/ or the levels to be weighed differently. In (22), π denotes the posterior distribution updated according to (14) and $\bar{x} \in \{e_1, e_2, \ldots, e_X\}$, where e_i is the unit indicator vector. Note that this is an alternate representation of the state space $\mathcal{X}$. In (22) using the law of iterated expectation [6, Chapter 8, Sec.8.4.2], [35, Lemma 3.2], $\eta_2(\pi, u)$ can be expressed in terms of the belief π as follows:[7]

$$\eta_2(\pi, u) = w_u\,(1 - \pi'\pi) \tag{23}$$

Since more informative opinions lead to smaller state-estimation error, from Theorem 5(i) below, $w_{u+1} > w_u$.

The net instantaneous cost $C(\pi, u)$ in (15) incurred by the pollster in adaptive expectation polling is thus given as:

$$C(\pi, u) = S(u) + \eta_2(\pi, u) \tag{24}$$

The cost (24) expressed in terms of the belief state π models the fact that asking the nodes at level i to provide information on the opinions of nodes at levels $j (< i)$ is costly, but more informative – smaller state estimation error.

[7]
$$J_\mu(\pi_0) = \mathbb{E}_\mu\left\{\sum_{k=0}^{\infty} \rho^k \mathbb{E}\{C(\bar{x}_k, u_k)\,|\mathcal{I}_k\}|\pi_0\right\}.$$
$$J_\mu(\pi_0) = \mathbb{E}_\mu\left\{\sum_{k=0}^{\infty} \rho^k \sum_{i=1}^{X} C(e_i, u_k)\pi_k(i)\,|\pi_0\right\} \text{ from (13).}$$

The instantaneous cost is thus given as $\sum_{i=1}^{X} C(e_i, u_k)\pi_k(i)$, where $C(e_i, u_k) = S(u_k) + w_{u_k}\|e_i - \pi_k\|_2$. By noting that $\|e_i - \pi_k\|_2 = \sum_{m=1}^{X} |e_i(m) - \pi_k(m)|^2$, the equation (23) follows from (22) by simple algebraic manipulation.

B. Fractional Exponents of Stochastic Matrices and Blackwell Dominance

The aim of this section is to provide a rationale for choosing the expectation polling actions, which correspond to sampling nodes at a particular level and soliciting information from other levels (see Fig. 1). It is shown fractional matrix powers can model requesting information from hidden levels in a hierarchical social influence network. When the opinion distribution is an ultrametric matrix, there is a relation between the fractional matrix powers and Blackwell dominance.

Definition: (Ultrametric Matrix [36]) A square stochastic matrix Q is *ultrametric* if

1) Q is symmetric.
2) $Q_{ij} \geq \min\{Q_{ik}, B_{kj}\}$ for all i, j, k.
3) $Q_{ii} > \min Q_{ik}$ for all $k \neq i$.

For any ultrametric matrix Q, the K^{th} primary root, $Q^{1/K}$, is also stochastic for any positive integer K; see [36].

Proposition V.1: For any ultrametric matrix Q, the following hold for any positive integer j:

a) $Q^{j/K} \succeq_B Q^j$.
b) $Q^{j/K} \succeq_B Q^{(j+1)/K} \ldots \succeq_B Q^{(j+K-1)/K}$
c) $Q^{j/(K+1)} \succeq_B Q^{j/(K)}$.
d) $Q \succeq_B Q^{j/K}$, for all $j > K$.

Discussion: Clearly, any integer power of a stochastic matrix is a stochastic matrix. Proposition V.1 says that fractional power of certain stochastic matrices, namely ultrametric, are also stochastic. In adaptive expectation polling (Theorem 5), polling actions correspond to choosing different levels in the hierarchy (Fig. 1) and soliciting opinions of nodes at other levels. In Proposition V.1, $Q^{j+1/K+1}$ can be used to interpret the notion of node at level K providing information on nodes' opinions at level j, and hence provides a way to order the likelihoods corresponding to different polling actions. According to Proposition V.1, when the opinion distribution B in (8) is ultrametric, there exists a Blackwell dominance relation between the observation distributions of the pollster.

C. Main Result. Myopic policies for Adaptive Expectation Polling

Our main result in adaptive expectation polling is Theorem 5 below. It shows that when it is cheaper for the pollster to (myopically) listen to the ultrametric channel that provides the most information on the state, it is optimal to do so. Ultrametric channels are (hidden) cascaded channels that model the communication medium between the pollster and the nodes of a social network having a hierarchical influence structure as in Fig. 1, when the pollster seeks opinions formed at the hidden levels from the levels that are easily accessible.

Theorem 5 (Adaptive Expectation Polling): Consider the adaptive expectation polling problem with costs specified in (24). Assume that the opinion distribution B (defined in (8)) is ultrametric. Let the observation distributions for the pollster be $O(u) = B_l^{l_u/l}\, \forall u \in \mathcal{U}$.

a) For the choice of levels $l_u > l_{u+1}$, we have

$$O(u) \succeq_B O(u+1)\, \forall u \in \mathcal{U}.$$

b) By Theorem 2, the myopic expectation polling policy $\bar{\mu}_E(\pi)$ forms an upper bound to the optimal expectation polling policy $\mu^*_E(\pi)$, i.e., $\mu^*_E(\pi) \leq \bar{\mu}_E(\pi)$ for all $\pi \in \Pi$. In particular, for belief states where $\bar{\mu}_E(\pi) = 1$, the myopic policy coincides with the optimal policy $\mu^*_E(\pi)$.

Discussion: The instantaneous cost for adaptive expectation polling (24) is concave in π by definition. The proof of Theorem 5 follows from Proposition V.1 below and Theorem 2. The expectation polling algorithm employed by the pollster determines how the opinions are gathered, and the opinions are distributed as $O(u)$ for the pollster. Proposition V.1 below provides a justification for the opinion distribution B to be ultrametric. Note that B_l denotes the opinion distribution at level l, i.e., $B_l = B^{l+1}$ from Fig. 1. For any $K > 0$, clearly $B_K^{j+1/K+1} = B_j$. This motivates the choice of the observation distribution of the pollster in case of adaptive expectation polling as $O(u) = B_l^{l_u/l}$, where nodes at level l are polled to provide information of the nodes at level l_u. It is easiest (see Sec. VIII) to poll nodes at level N, so a convenient choice is $O(u) = B_{N+1}^{l_u/N+1}$.

D. Information Theoretic Interpretation

The aim of this section is to provide a link between ultrametric channels (hidden channels) and Shannon capacity in terms of Blackwell dominance. Let $I(\mathcal{X};\mathcal{Y}^{(l_u)})$ denote the mutual information of the ultrametric channel $Q^{l_u/K}$ and $\mathcal{C}^{(l_u)}$ denote the capacity defined in (4). Let $Q_i^{l_u/K}$ denotes the i^{th} row of the channel $Q^{l_u/K}$.

Proposition V.2: If the channel error probabilities (likelihoods) for the pollster satisfy $Q^{l_u/K} \succeq_B Q^{l_v/K}$ for any $K > 0$, we have

i) Shannon Capacity Ordering: $\mathcal{C}^{(l_u)} \geq \mathcal{C}^{(l_v)}$ for $l_u > l_v$.

ii) Rényi Divergence Ordering:

$$D_\alpha(Q_i^{l_u/K} || Q_j^{l_u/K}) \geq D_\alpha(Q_i^{l_v/K} || Q_j^{l_v/K})$$

for all $u \in \mathcal{U}$ and for all $i, j \in \mathcal{X}$.

Discussion: The proof of Proposition V.2 follows from Theorem 1 and Theorem 3. Proposition V.2 provides an ordering of Rényi Divergence and Shannon capacity between ultrametric channels $Q^{l_u/K}$, $K > 0$, $\forall u \in \mathcal{U}$. From Proposition V.2, the ultrametric channels are ordered such that the information of nodes at Level 0, for example, revealed by the nodes at Level $N(\neq 0)$ result in a larger reduction in uncertainty on the state, than opinions from nodes at Level $N+1(\neq 0)$.

VI. Approximate Blackwell Dominance

So far we have discussed sufficient conditions for Blackwell dominance; when these conditions hold, the optimal adaptive polling policy is provably upper bounded by a myopic policy. *A natural question is: Can efficient polling methods be developed when Blackwell dominance does not hold exactly?*

Algorithm 3: Approximate Blackwell Dominance.

Let $\mathcal{M}$ denotes the set of all stochastic matrices.
Initialize: $O(1) = \hat{O}(1)$
For $u \in \{1, 2, \ldots, U-1\}$**, do**:
$R^*_{u+1} = \arg\min_{R \in \mathcal{M}} \|O(u+1) - \hat{O}(u)R\|_\infty$
$\hat{O}(u+1) = \hat{O}(u)R^*_{u+1}$
end
Output: $\hat{O}(u)$ for $u \in \mathcal{U}$.

This section discusses approximate Blackwell dominance and its applications in a novel polling method called adaptive neighborhood expectation polling. The main idea involves Le Cam deficiency.

A. Le Cam Deficiency

Given a collection of matrices, it is important to check whether there exists a Blackwell dominance relation, as Theorem 2 can used to compute inexpensive policies. In this section, an approximation procedure using *Le Cam deficiency* is provided. *Le Cam deficiency* enables to calculate the closest matrix that is Blackwell comparable.

Definition: (Le Cam deficiency) For any two stochastic matrices W and H, the *Le Cam deficiency* is

$$\delta(W, H) \triangleq \inf_{R \in \mathcal{M}} \|W - HR\|_\infty, \tag{25}$$

where $\mathcal{M}$ denotes the set of all stochastic matrices and $\|\cdot\|_\infty$ denotes the induced norm.

The $\inf$ in (25) is achieved – this can be shown using *Le Cam randomization* criterion [37]. The Le Cam deficiency is an approximation measure that quantifies the loss when using one observation distribution instead of the other. There is no loss if there exists a mechanism able to convert the observations from one distribution to the other.

(25) can be solved as a convex optimization problem using *CVXOPT* toolbox in Python or *CVX* in Matlab. Solving (25) yields observation distributions that are Blackwell comparable.

Consider a POMDP model $\theta = (\mathcal{X}, \mathcal{Y}, P, O(u), C, \rho)$, where $O(u)$ for $u = \{1, 2, \ldots, U\}$ are observation matrices that are not Blackwell comparable. Consider an approximation $\gamma = (\mathcal{X}, \mathcal{Y}, P, O(1), \hat{O}(\hat{u}), C, \rho)$, where $\hat{u} = \mathcal{U}/\{1\}$ and the observations distributions are such that

$$O(1) \succeq_B \hat{O}(2) \cdots \succeq_B \hat{O}(U). \tag{26}$$

Algorithm 3 details a procedure to compute observation distributions that share a Blackwell dominance relation (26).

B. Applications of Approximate Blackwell Dominance

Algorithm 3 can be used to design POMDPs for adaptive polling that have observation distributions that are not Blackwell comparable – for example, when the polling distributions in case of adaptive intent polling are not Hurwitz, when the opinion distributions are not ultrametric in case of adaptive expectation polling, when the pollster has a choice between different polling algorithms over the polling horizon, etc.

1) Adaptive Neighborhood Expectation Polling: Here each polled node gathers the opinion from other nodes at the same level on each state and reports the opinion fraction to the pollster. The question asked by the pollster in case of adaptive NEP polling is

> "*what does a node at level l think the fraction in favor of different states is, at level l?*"

This polling algorithm is a more sophisticated version of Neighborhood Expectation Polling (NEP) [22]. NEP is a polling algorithm to decide between two states where the pollster asks the following question [22]: "what is a nodes' estimate of the fraction of votes for a particular candidate?".

In the case of adaptive NEP polling, the pollster controls the observation distribution $O(u)$ by choosing different levels to gather the information in the form of fractions, and this in turn affects the estimate of the state (see Fig. 1).

Remark: In case of adaptive NEP polling, the nodes report opinion fractions to the pollster. If instead, the nodes report probabilities with $\mathcal{Y} = [0,1]^{|\mathcal{X}|}$, there is a possibility that the pollster receives biased information. There is a disjunction effect – the beliefs about the state change when aggregated differently. This is the well known *Simpson's Paradox*; see [38].

The adaptive NEP polling algorithm deployed by the pollster determines how the opinions are gathered, and the observations for the pollster are tuples reported by the nodes that indicate the fraction in favor of each state. Channels specified by multinomial distributions model the likelihood of opinion counts in favor of different states from different nodes at the same level. Let $\mathcal{N} \in \{1, 2, \ldots, \mathbb{N}\}$ denote the number of nodes accessible (friends with) to nodes at each level in the hierarchical social influence network. This models the possibility of different individuals or nodes having different friends with $\mathbb{N}$ denoting a finite maximum number. Let the observation alphabet for the pollster be $\mathcal{Y} = \{(\frac{\boldsymbol{n}_1}{\mathcal{N}}, \frac{\boldsymbol{n}_2}{\mathcal{N}}, \ldots, \frac{\boldsymbol{n}_X}{\mathcal{N}}) \,\forall \mathcal{N} : \boldsymbol{n}_i \in \mathbb{Z}_+, \sum_i \boldsymbol{n}_i = \mathcal{N}\}$, where $\mathbb{Z}_+$ denotes the set of non-negative integers. Let $O(l)$ denote the opinion fraction that the pollster receives from level l, and has elements

$$(O(l))_{ij} = \mathbb{P}(y^l_{k+1} = j | x_{k+1} = i, \mathcal{N}_j), i \in \mathcal{X}, j \in \mathcal{Y}.$$

Here,

$$j = \left(\frac{\boldsymbol{n}_1^{(j)}}{\mathcal{N}_j}, \frac{\boldsymbol{n}_2^{(j)}}{\mathcal{N}_j}, \ldots, \frac{\boldsymbol{n}_X^{(j)}}{\mathcal{N}_j}\right), \mathcal{N}_j \in \{1, 2, \ldots, \mathbb{N}\},$$

$$\sum_h \boldsymbol{n}_h^{(j)} = \mathcal{N}_j.$$

$$\mathbb{P}\left(y^l_{k+1} = j | x_{k+1} = i, \mathcal{N}_j\right) = \frac{\mathcal{N}_j!}{\boldsymbol{n}_1^{(j)}! \times \cdots \times \boldsymbol{n}_X^{(j)}!} \times \prod_{h=1}^{X} (B_l)_{ih}^{\boldsymbol{n}_h^{(j)}}. \tag{27}$$

Here $\mathcal{N}_j$ and $\boldsymbol{n}_i^{(j)}$ indicate the total and the number in favor of $x = i$ reported and B_l denotes the opinion distribution (10) at level l. The likelihood in (27) is the well known *multinomial distribution*.

The observation distributions (27) are not necessarily Blackwell ordered, but it is intuitive that the opinion fractions in (27) from nodes at level i are more informative than opinion fractions from nodes at level $j(>i)$ in Fig. 1 owing to obvious Blackwell dominance relation of opinion distributions B_l for $l = i, j$ in (10). However, Algorithm 3 can be used to obtain approximate Blackwell dominance of observation distributions (27).

2) Adaptive Polling With Choice: In this section, we establish that expectation polling from the lowest level (least informative) and seeking opinions about the highest level is better (more informative) than intent polling (here, the pollster seeks information from all levels). Depending on the availability of access to different levels for the pollster, it can switch between polling algorithms.

For example, when using intent polling on an organizational network (implicitly hierarchical in nature), the executive levels might become inaccessible during IPOs or financial crisis. Then, the pollster can switch to listening the inside information from the lower levels (expectation polling), to estimate the underlying state of nature.

Let the opinion distribution B (defined in (8)) be ultrametric and $f_2(z) \in \mathcal{P}_N$ be *any* polynomial. Let the true POMDP model be $\theta = (\mathcal{X}, \mathcal{Y}, \mathbb{Y}, P, O(1), O(2), C)$ and the approximation be $\gamma = (\mathcal{X}, \mathcal{Y}, \mathbb{Y}, P, O(1), \hat{O}(2), C)$. Let $\mu(\cdot;\gamma)$ denote the policy parameterized by the approximate model γ.

Proposition (Adaptive Expectation v/s Intent): Let $O(1) = B^{l_1/N+1}_{N+1}$, and $O(2) = Bf_2(B)$ for some l_1 and f_2, denote the observation distributions in case of adaptive expectation polling and adaptive intent polling respectively.

i) The approximate Blackwell ordering using Algorithm 3 is
$$O(1) \succeq_B \hat{O}(2).$$
ii) The myopic polling policy $\bar{\mu}(\pi;\gamma)$ is an upper bound to the optimal polling policy $\mu^*(\pi;\gamma)$, i.e., $\mu^*(\pi;\gamma) \leq \bar{\mu}(\pi;\gamma)$ for all $\pi \in \Pi$.

Discussion: For $u = 1$, the pollster chooses expectation polling and hence listens to an ultrametric channel, and for $u = 2$, the pollster chooses intent polling and hence listens to a polynomial channel. As $O(2) = Bf_2(B)$, we have $B \succeq_B O(2)$. Note that since $O(1) = B^{l_1/N+1}_{N+1}$, when $l_1 = 1$ (nodes at Level N are polled to provide opinion of nodes at Level 0), $O(1) = B^{1/N+1}_{N+1} = B \succeq_B O(2)$. This implies that expectation polling is more informative than intent polling.

For $l_1 > 1$, there is no apparent comparison of ultrametric and polynomial channels. However, Algorithm 3 can be used to design POMDPs for adaptive polling for arbitrary l_1 and f_2.

VII. Performance Bounds and Ordinal Sensitivity

In Sec. VI, we discussed an approximation procedure to compute a POMDP model for an adaptive polling problem that has a Blackwell dominance structure and is close (Le Cam sense) to the true POMDP. Sec. VII-A provides performance bounds on the comparison of POMDPs for adaptive polling.

Sec. VII-B provides the ordinal sensitivity in polling, i.e., an ordering of the cumulative costs with respect to the variation in opinion distributions B (defined in (8)).

A. Performance Bounds on Adaptive Polling

Let $\theta = (\mathcal{X}, \mathcal{Y}, \mathbb{Y}, P, O(u), C, \rho)$ denote the given POMDP model for adaptive polling and $\gamma = (\mathcal{X}, \mathcal{Y}, \mathbb{Y}, P, \hat{O}(u), C, \rho)$ denote the POMDP model for adaptive polling having a Blackwell dominance relation between the observation distributions. Let $J_{\mu^*(\gamma)}(\pi;\theta)$ and $J_{\mu^*(\gamma)}(\pi;\gamma)$ be defined as in (16), and denote the cumulative costs incurred by the two models θ and γ respectively, when using the polling policy $\mu^*(\gamma)$. Let $J_{\mu^*(\theta)}(\pi;\theta)$ and $J_{\mu^*(\theta)}(\pi;\gamma)$ be defined as in (16), and denote the cumulative costs incurred by the two models θ and γ respectively, when using the polling policy $\mu^*(\theta)$. Theorem 6 below provides a bound on the deviations from the optimal cost and policy performance of the POMDP models for adaptive polling.

Theorem 6: Consider two POMDP models $\theta = (\mathcal{X}, \mathcal{Y}, \mathbb{Y}, P, O(u), C, \rho)$ and $\gamma = (\mathcal{X}, \mathcal{Y}, \mathbb{Y}, P, \hat{O}(u), C, \rho)$ for adaptive polling. Then for the mis-specified model and mis-specified policy, the following sensitivity bounds hold:

$$\text{Mis-specified Model: } \sup_{\pi\in\Pi} |J_{\mu^*(\gamma)}(\pi;\gamma) - J_{\mu^*(\gamma)}(\pi;\theta)| \leq G\|\gamma-\theta\|. \tag{28}$$

$$\text{Mis-specified Policy: } J_{\mu^*(\gamma)}(\pi;\theta) \leq J_{\mu^*(\theta)}(\pi;\theta) + 2G\|\gamma-\theta\|. \tag{29}$$

Here $G = \max_{i\in\mathcal{X},u} \frac{C(e_i,u)}{1-\rho}$ and e_i denotes the indicator vector with a '1' in the i^{th} position, and

$$\|\gamma-\theta\| = \max_u \max_i \sum_y \sum_j P_{ij} |O_{jy}(u) - \hat{O}_{jy}(u)|.$$

Discussion: Theorem 6 provides uniform bounds on the additional cost incurred for using parameters that are Blackwell comparable in place of the given parameters of the POMDP for adaptive polling. The proof follows from arguments similar to [6, Theorem 14.9.1], and is omitted.

So far it was assumed that the pollster has complete knowledge of the node-level associations. However, if a set of nodes are misclassified to a different level by the pollster, then the pollster is essentially updating the belief using different observation distributions. Theorem 6 can be used to compute the performance bounds for this misclassification as well.

B. Ordering of Hierarchical Social Influence Networks

So far we have discussed two types of polling algorithms on a single hierarchical social influence network. In this section, we briefly discuss how to order hierarchical influence networks that differ in the opinion distributions B, according to the expected polling cost. Theorem 7 below shows that some networks are inherently more expensive to poll than others; it defines a partial order over networks that results in an ordering of the cost of polling.

Let the POMDP model of the hierarchical influence network $\mathbb{H}_i$ for $i = 1, 2, \cdots$ be θ_i, where the tuple $\theta_i = (\mathcal{X}, \mathcal{Y}, \mathbb{Y}, P, O^{(i)}, C)$. Let $\mu_i^*(\pi;\theta_i)$ denote the optimal polling policy on each of the network, and let $J_{\mu_i^*(\theta_i)}(\pi;\theta_i)$ denote the corresponding optimal cumulative cost.

Theorem 7 (Ordinal sensitivity in Polling): Consider two hierarchical networks $\mathbb{H}_1$ and $\mathbb{H}_2$. Let the POMDPs for adaptive polling of each hierarchical network have the observation distributions that satisfy $O^{(1)} \succeq_B O^{(2)}$. Then

$$J_{\mu_1^*(\theta_1)}(\pi;\theta_1) \leq J_{\mu_2^*(\theta_2)}(\pi;\theta_2). \tag{30}$$

Here $O^{(1)} \succeq_B O^{(2)}$ denotes $O^{(1)}(u) \succeq_B O^{(2)}(u) \, \forall\, u \in \mathcal{U}$.

Discussion: The proof of Theorem 7 follows from arguments similar to Theorem 14.8.1 in [6], and is omitted. Since the observation likelihood for the pollster ($O^{(i)} \forall i$) depends on the opinion distribution (10), Theorem 7 provides a way to compare the cumulative costs of hierarchical influence networks with different opinion distributions. The result is useful, in that, a hierarchical influence network that has more informative opinion distribution at every level compared to another hierarchical influence network is cheaper to poll on average as the nodes provide more informative opinions.

VIII. Numerical Examples

The main results of this paper involve using Blackwell dominance to construct myopic policies that provably upper bound the optimal adaptive polling policy. In this section, the performance of this myopic upper bound is illustrated using numerical examples for adaptive polling. As discussed in Sec. II-A, the discount factor ρ determines the way the polling cost is counted towards the polling value $J_{\mu^*}(\pi_0)$ defined in (12) when using the optimal policy $\mu^*(\pi)$. Since the computationally inexpensive myopic policy is used for polling in Algorithm 1 and Algorithm 2, instead of the optimal policy $\mu^*(\pi)$, the performance loss and sensitivity (both defined below) in terms of the polling value $J_{\mu^*}(\pi_0)$ is evaluated for different values of the discount factor.

Let $J_{\bar{\mu}}(\pi_0)$ denote the discounted costs associated with the myopic policy $\bar{\mu}(\pi)$. We consider the following two measures for measuring the effectiveness of the myopic polling policy:

i) The percentage loss in optimality due to using the myopic policy $\bar{\mu}$ instead of optimal policy μ^* is

$$\mathcal{L}_1 = \frac{J_{\bar{\mu}}(\pi_0) - J_{\mu^*}(\pi_0)}{J_{\mu^*}(\pi_0)}. \tag{31}$$

In (31), the total average cost is evaluated using 1000 Monte carlo simulations over a horizon of 100 time units. The optimal cost $J_{\mu^*}(\pi_0)$ is calculated as in (12).

ii) Let Π_1^s represent the set of belief states for which $C(\pi,1) < C(\pi,u) \, \forall u = 2,\ldots,U$. So on the set Π_1^s, the myopic policy coincides with the optimal policy $\mu^*(\pi)$. What is the performance loss outside the set Π_1^s? Define the following discounted cost

$$\tilde{J}_{\mu^*}(\pi_0) = \mathbb{E}\left\{\sum_{k=1}^{\infty} \rho^{k-1} \tilde{C}\left(\pi_k, \mu^*(\pi_k)\right)\right\}, \text{ where } \rho \in [0,1),$$

$$\tilde{C}\left(\pi, \mu^*(\pi)\right) = \begin{cases} C(\pi,1) & \pi \in \Pi_1^s \\ C(\pi,1) + w_2\eta_2(\pi,2) & \pi \notin \Pi_1^s \end{cases}$$

Clearly a lower bound for the percentage loss in optimality due to using the myopic policy $\bar{\mu}$ instead of optimal policy μ^* is

$$\mathcal{L}_2 = \frac{J_{\bar{\mu}}(\pi_0) - \tilde{J}_{\mu^*}(\pi_0)}{\tilde{J}_{\mu^*}(\pi_0)}. \quad (32)$$

In (32), the cumulative discounted cost is evaluated using 1000 Monte carlo simulations over a horizon of 100 time units.

Here μ^* is the optimal policy of the non-standard (non-linear cost) POMDP, and is solved using POMDP algorithms in [6, Chapter 8, Sec.8.4.4].

A. Example 1: Market Research. Adaptive Expectation Polling via YouTube

We describe how to estimate the revenue level that a movie generates based on the response received on the social media platform YouTube.

YouTube Dataset: A sample of 30 comedy movies from 2016–2018 were selected. For each of these movies, YouTube comments on their trailers that expressed personal opinions were collected using the Python YouTube API.[8] The sentiment associated with each of the comments was identified using sentiment analysis tool - textblob (http://textblob.readthedocs.org/en/dev/).

Hierarchical network modeling: The critics and those who see the movie before its release will influence the future movie goers by sharing opinions on social media platforms. So the critics are in Level 0 and the common movie goer is Level 1. So the number of levels in the hierarchical social network (Fig. 1) for this example is thus $N = 1$.

Polling algorithm – Expectation Polling: In adaptive expectation polling (Sec. V), to poll the common movie goers who provide their opinion on YouTube, the pollster asks the following question to estimate the performance of a movie:

> "*what does a node at Level 1 think the nodes at Level* 0 (*u=1*) *and Level* 1 (*u=2*) *would report the state as?*"

In other words, the pollster asks "what do you think?" and "what do they think?". So the polling action $u \in \{1, 2\}$ selects the opinion distributions $B_2^{u/2}$.

State – Popularity: The popularity of each of 30 movies is modeled as a 3 state Markov chain x, and depending on their box-office revenues (https://www.boxofficemojo.com/), each of these movies were assigned a state from the state-space $\mathcal{X} = \{\text{High}, \text{Medium}, \text{Low}\}$.

State transition matrix – Popularity changes: The popularity of a movie evolves over time due to a number of factors including release of a better advertised movie, release of a more anticipated movie, the gradual decline of the hype surrounding the movie, or increase in popularity after celebrity endorsement etc. This is modelled using a state transition matrix and the maximum likelihood estimate was computed using an Expectation Maximization algorithm with ultrametric constraints (see Appendix B).

Observation matrix – Sentiments: Prior to a movie's release, the production and the media house (proprietor) associated with the movie release a variety of promotional material, in the form of trailer videos, digital billboards, blogs, pre-screenings etc., to advertise the movie. A matrix consisting of number of positive, neutral and negative comments for state of each movie $\{\text{Good}, \text{Neutral}, \text{Bad}\}$ was formed. Using this matrix, the opinion matrix B_2, given in (33) was then obtained by using maximum likelihood estimation algorithm with (See Appendix B) ultrametric constraints. This can be used to obtain the opinion distribution B_1 of Level 0.

Parameters: The computed parameters (see Appendix B) for P, $O(1) = O^{1/2}(2)$, and $O(2)$ are as follows:

$$\begin{pmatrix} 0.9089 & 0.0281 & 0.0630 \\ 0.0346 & 0.9433 & 0.0221 \\ 0.0065 & 0.0138 & 0.9797 \end{pmatrix}, \begin{pmatrix} 0.6382 & 0.1809 & 0.1809 \\ 0.1809 & 0.6382 & 0.1809 \\ 0.1809 & 0.1809 & 0.6382 \end{pmatrix},$$
$$\begin{pmatrix} 0.4728 & 0.2636 & 0.2636 \\ 0.2636 & 0.4728 & 0.2636 \\ 0.2636 & 0.2636 & 0.4728 \end{pmatrix}. \quad (33)$$

The costs associated with actions $u = 1, 2$ are chosen as:

$$S(1) = 0.5, S(2) = 0.25, w_1 = 0.5, w_2 = 1. \quad (34)$$

The numerical values in (34) are real numbers that obey the ordinal relations below, and are chosen using the empirical evidence in [33]. The ordinal relations capture the fact that higher levels are more informative (Blackwell sense) and hence more costly. Note that the costs associated with the actions $u = 1$ and $u = 2$ in (34) assume the following structure: $w_1 \leq w_2$ model the accuracy of the observations and $S(1) \geq S(2)$ model the additional cost in expectation polling – nodes need to be compensated for exhausting their resources gathering information from different levels.

For a new (test) movie, depending on which level the observation is obtained from, the pollster updates the probability distribution over the states using the state transition matrix P and the corresponding estimated observation distribution matrix $O(1)$ or $O(2)$.

Performance evaluation: The probabilities in (33) and the costs in (34) constitute the POMDP parameters. Fig. 2 provides percentage loss in optimality $\mathcal{L}_1$ and $\mathcal{L}_2$, and the change in optimality $\frac{\partial \mathcal{L}_1}{\partial \rho}$ for different values of the discount factor $\rho \in [0, 1)]$. In the adaptive expectation polling algorithm (Algorithm 2), a myopic policy is used to poll the YouTube users. Of course, when $\rho = 0$, the myopic policy is the optimal policy. When $\rho > 0$, the pollster still adopts the computationally inexpensive myopic policy (which provably upper bounds the optimal policy via Theorem 2), while clearly compromising on the polling value $J_{\mu^*}(\pi_0)$. It is intuitive that the performance loss measured as the difference of the expected cost when using a myopic policy $J_{\bar{\mu}}(\pi_0)$ and the polling value $J_{\mu^*}(\pi_0)$ will increase with the discount factor as $\rho \to 1$; as evident in Fig. 2. Also, the performance loss sensitivity $\frac{\partial \mathcal{L}_1}{\partial \rho}$ is observed in Fig. 2 to be higher for changes in large values of the discount factor. These two observations imply that a more forward looking pollster, i.e., larger value of ρ, will incur higher losses than its

[8]https://gdata-python-client.googlecode.com/hg/pydocs/gdata.youtube.html

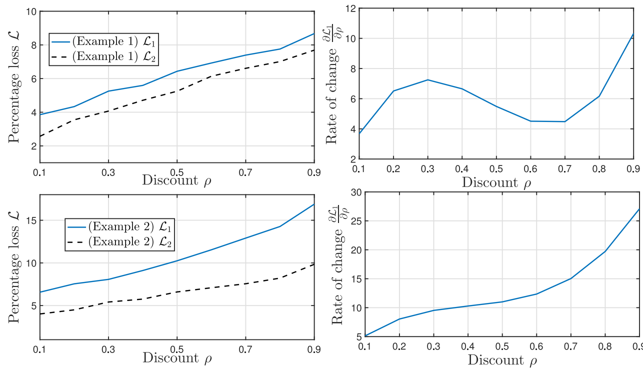

Fig. 2. The percentage loss in optimality $\mathcal{L}_1$ and $\mathcal{L}_2$, and the sensitivity of the performance loss to discount factor $\frac{\partial \mathcal{L}_1}{\partial \rho}$, is evaluated by simulation for different values of the discount factor ρ, when using a myopic policy for Example 1 and Example 2 respectively. $\mathcal{L}_2$ is a lower bound for $\mathcal{L}_1$. The performance loss is observed to be most sensitive to changes in large values of the discount factor.

short-sighted counterpart; for an inexpensive implementation of the polling algorithms.

B. Example 2: Large Dimensional Example. Adaptive Intent Polling With X = 20, Y = 20, U = 5 and N = 9

The Blackwell dominance structural result is particularly useful for large number of states and observation symbols since solving the POMDP (for the optimal policy) is intractable. Random stochastic matrices of size 20×20 were generated for the transition probability matrix P and the observation probability matrix B. The matrices are generated by stochastic simulation as follows: twenty (1×20) probability vectors were simulated from the Dirichlet distribution on a 19 dimensional unit simplex and stacked as rows. We know that B^l for $l = 2, \ldots, 10$ constitute the opinion distribution of level l. The observation distribution of the pollster $O(u) = \sum_{l=0}^{N} \beta_l^{(u)} B^{l+1}$, $\gamma_1 = [5, 4, 3, 2, 1]$ and $\gamma_2 = [1, 2, 3, 4, 5]$. The cost parameters conform to the ordinal relations that capture the fact that higher levels are more informative and hence more costly. Here the probability distributions are chosen as follows: $\beta^{(5)}$ is chosen as[9] and $\beta^{(u)}$ for $u = \{4, 3, 2, 1\}$ are obtained by successively removing the smallest root.

Fig. 2 provides (average) percentage loss in optimality $\mathcal{L}_1$ and $\mathcal{L}_2$, and the change in optimality $\frac{\partial \mathcal{L}_1}{\partial \rho}$ for different values of the discount factor ρ. In the adaptive intent polling algorithm (Algorithm 1), a myopic policy is used to poll the users. Of course, when $\rho = 0$, the myopic policy is the optimal policy. When $\rho > 0$, the pollster still adopts the computationally inexpensive myopic policy, while clearly compromising on the polling value $J_{\mu^*}(\pi_0)$.

IX. Conclusions

This paper considered the problem of adaptive (stochastic feedback control based) polling in hierarchical social networks, formulated as a partially observed Markov decision process (POMDP). POMDPs are intractable to solve. The key idea of the paper was to exploit Blackwell dominance to construct myopic bounds that provably upper bound the optimal polling policy. We presented two main results. First, the notion of Blackwell dominance was extended to the case of polynomial observation likelihoods (channels) described by matrix polynomials. This was used to develop an adaptive intent polling algorithm that is inexpensive to implement. Second, the notion of Blackwell dominance was extended to the case of ultrametric observation likelihoods (channels) described by fractional matrix powers. This was used to develop an adaptive expectation polling algorithm that is inexpensive to implement.

This extension of Blackwell dominance to more general channels that arise in hierarchical social influence networks was used to provide a natural ordering of Rényi Divergence and Shannon capacity. These information theoretic consequences provide a ranking of these general channel structures in the order of their ability to distinguish the states, and hence guide the choice of observation distributions for the pollster.

We discussed approximate Blackwell dominance based on Le Cam deficiency to facilitate the comparison of the different polling algorithms, and situations where a Blackwell dominance relation is absent. This was used to provide an adaptive generalization of neighborhood expectation polling to hierarchical social influence networks, where the notion of Blackwell dominance was extended to the case of multinomial distributions of observation likelihoods. We also provided performance bounds on the cumulative cost and polling policy, when the model parameters are mis-specified. Finally, we illustrated the results and the performance of the myopic polling policy using a YouTube social media dataset.

Appendix A
Proofs

Proof of Theorem 2: Denote by $y^{(u)}$ as the observations recorded when using action u. Then $O(u+1) = O(u)R$ implies the following

$$\mathbb{P}\left(y^{(u+1)}|x\right) = \sum_{y^{(u)}} \mathbb{P}\left(y^{(u+1)}|y^{(u)}\right) \mathbb{P}\left(y^{(u)}|x\right) \tag{35}$$

For notational convenience, let $T(\pi, y, u)$ be written as $T(\pi, y^{(u)} = y)$. Observe that,

$$\begin{aligned} T\left(\pi, y^{(u+1)} = y\right) &= \frac{O_{u+1}(y)P'\pi}{\sigma\left(\pi, y^{(u+1)} = y\right)} \\ &= \sum_r \Lambda(r) T(\pi, y^{(u)} = r) \end{aligned} \tag{36}$$

where $\Lambda(r)$ is a probability mass function w.r.t r and defined as

$$\Lambda(r) = \mathbb{P}\left(y^{(u+1)} = y|y^{(u)} = r\right) \frac{\sigma\left(\pi, y^{(u)} = r\right)}{\sigma\left(\pi, y^{(u+1)} = y\right)} \tag{37}$$

[9] $\beta^{(5)}$ = [25/1296, 1555/15552, 3461/15552, 86925/311040, 13627/ 62208, 11617/103680, 437/11520, 2671/311040, 73/62208, 29/ 311040, 1/311040].

The following inequality follows from the concavity of $V(\pi)$ and (37)

$$V\left(T\left(\pi, y^{(u+1)} = y\right)\right) = V\left(\sum_r \Lambda(r) T(\pi, y^{(u)} = r)\right)$$
$$V\left(T\left(\pi, y^{(u+1)} = y\right)\right) \geq \sum_r \Lambda(r) V\left(T(\pi, y^{(u)} = r)\right) \tag{38}$$

Following completes the proof of Theorem 2 using (38).

$$\begin{aligned}&\sum_y \sigma(\pi, y^{(u+1)} = y) V\left(T\left(\pi, y^{(u+1)} = y\right)\right)\\ &\geq \sum_y \sum_r \Lambda(r) V\left(T(\pi, y^{(u)} = r)\right) \sigma(\pi, y^{(u+1)} = y)\\ &= \sum_r V\left(T\left(\pi, y^{(u)} = r\right)\right) \sigma\left(\pi, y^{(u)} = r\right)\end{aligned} \tag{39}$$

$\therefore C(\pi, 1) \leq C(\pi, u)\, \forall u \Rightarrow \mu^*(\pi) = 1 \Rightarrow \mu^*(\pi) \leq \bar{\mu}(\pi)$. ■

Proof of Theorem 3: Let $O(u) \succeq_B O(u+1)$ for $u \in \mathcal{U}$. From the definition of Rényi Divergence (19) we have [39]:

$$\begin{aligned}&D_\alpha(O_i(u+1)||O_j(u+1))\\ &\quad \leq \min\Big\{(1-\alpha) D(O_i(u+1)||O_j(u+1)),\\ &\qquad \alpha D(O_j(u+1)||O_i(u+1))\Big\}.\end{aligned} \tag{40}$$

We know that [40]:

$$\begin{aligned}&O(u) \succeq_B O(u+1)\\ &\quad \Rightarrow D(O_i(u)||O_j(u)) \geq D(O_i(u+1)||O_j(u+1)),\end{aligned} \tag{41}$$

for all $i, j \in \mathcal{X}$. From (40) and (41), the result follows. ■

Proof of Proposition IV.1: It is given that $p(z) \in \mathcal{P}_n$ and $q(z) \in \mathcal{P}_m$, with $n > m$. Clearly, $f(Q)$ and $g(Q)$ are stochastic matrices. Further, if the quotient polynomial $h(z) = \frac{f(z)}{g(z)} \in \mathcal{P}_{(n-m)}$, then it is easily seen that $g(Q) \succeq_B f(Q)$.

Since the polynomials $p(z)$ and $q(z)$ are Hurwitz, the quotient polynomial $h(z) = \frac{p(z)}{q(z)} = \sum_{i=0}^{(n-m)} \alpha_i z^i$ has positive coefficients; i.e., $\alpha_i > 0$. It suffices to prove that $h(z) \in \mathcal{P}_{(n-m)}$. It is clear that $p(1) = q(1) = 1$, which implies that $h(1) = 1$; i.e., $\sum_{i=0}^{(n-m)} \alpha_i = 1$. ■

Proof of Proposition V.1: We will only prove Theorem V.1 b and Theorem V.1 c.

For Theorem V.1 b, we have $Q^{(j+J)/K} = Q^{j/K} \times Q^{J/K}$. Therefore $Q^{j/K} \succeq_B Q^{(j+J)/K}$.

For Theorem V.1 c, we have $Q^{j/K} = Q^{j/K+1} \times Q^{j/K(K+1)}$. Therefore $Q^{j/K} \succeq_B Q^{j/K+1}$. ■

Appendix B
EM Algorithm With Ultrametric Constraints

The parameters of the POMDP are computed using a sequence of observations obtained from level N in Fig. 1. Specifically, we describe a modified version of the EM algorithm [41] is used to compute the maximum likelihood estimate of the tuple (P, B_{N+1}), where B_{N+1} is restricted to the space of ultrametric stochastic matrices. The opinion probability matrices at all other levels are computed by taking fractional exponents of B_{N+1}. In this modified EM algorithm, computing B_{N+1} requires maximizing an auxiliary likelihood function (of observation sequences) subject to ultrametric constraints (see Footnote 13) on B_{N+1}. However, the space of ultrametric stochastic matrices is non-convex because of constraint $B_{N+1}(i,j) \geq \min\{B_{N+1}(i,k), B_{N+1}(k,j)\}$.

The following reformulation based on the Big-M method in linear programming [42] is used to deal with the non-convex constraint. For all $i, j, k \in \mathcal{X}, i \neq j \neq k$:

$$B_{N+1}(i,j) \geq B_{N+1}(i,k) + M(1-\kappa), \tag{42}$$
$$B_{N+1}(i,j) \geq B_{N+1}(k,j) + M\kappa, \tag{43}$$
$$B_{N+1}(k,j) \geq B_{N+1}(i,k) + M(1-\kappa), \tag{44}$$
$$B_{N+1}(i,k) \geq B_{N+1}(k,j) + M\kappa, \tag{45}$$
$$\kappa \geq 0, \tag{46}$$
$$-\kappa \geq -1, \tag{47}$$

for some large positive value M. The resulting observation likelihood B_{N+1} is a stochastic and ultrametric matrix.

References

[1] L. Le Cam, *Asymptotic Methods in Statistical Decision Theory*. Berlin, Germany: Springer Science & Business Media, 2012.

[2] L. Le Cam and G. L. Yang, *Asymptotics in Statistics: Some Basic Concepts*. Berlin, Germany: Springer Science & Business Media, 2012.

[3] J. Cohen, J. Kempermann, and G. Zbaganu, *Comparisons of Stochastic Matrices With Applications in Information Theory, Statistics, Economics and Population*. Berlin, Germany: Springer Science & Business Media, 1998.

[4] D. Bergemann and S. Morris, "Bayes correlated equilibrium and the comparison of information structures in games," *Theor. Econ.*, vol. 11, no. 2, pp. 487–522, 2016.

[5] U. Rieder, "Structural results for partially observed control models," *Methods Models Oper. Res.*, vol. 35, no. 6, pp. 473–490, 1991.

[6] V. Krishnamurthy, *Partially Observed Markov Decision Processes*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[7] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with twitter: What 140 characters reveal about political sentiment," *Int. AAAI Conf. Web Social Media*, 2010, vol. 10, no. 1, pp. 178–185.

[8] N. Silver, *The Signal and the Noise: Why so Many Predictions Fail–But Some Don't*. Westminster, U.K.: Penguin, 2012.

[9] K. J. Gile, "Improved inference for respondent-driven sampling data with application to HIV prevalence estimation," *J. Am. Statist. Assoc.*, vol. 106, no. 493, pp. 135–146, 2011.

[10] L. López, J. F. Mendes, and M. A. Sanjuán, "Hierarchical social networks and information flow," *Physica A: Statist. Mech. Appl.*, vol. 316, no. 1, pp. 695–708, 2002.

[11] J. A. Almendral, L. López, and M. A. Sanjuán, "Information flow in generalized hierarchical networks," *Physica A: Statist. Mech. Appl.*, vol. 324, no. 1, pp. 424–429, 2003.

[12] C. Shirky, *Cognitive Surplus: Creativity and Generosity in a Connected Age*. Westminster, U.K.: Penguin, 2010.

[13] J. Borge-Holthoefer *et al.*, "Structural and dynamical patterns on online social networks: The Spanish May 15th movement as a case study," *PloS One*, vol. 6, no. 8, 2011, Art. no. e23883.

[14] C. L. Boyd, "The hierarchical influence of courts of appeals on district courts," *J. Legal Studies*, vol. 44, no. 1, pp. 113–141, 2015.

[15] W. S. Lovejoy, "A survey of algorithmic methods for partially observed Markov decision processes," *Ann. Oper. Res.*, vol. 28, no. 1, pp. 47–65, 1991.

[16] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of Markov decision processes," *Math. Oper. Res.*, vol. 12, no. 3, pp. 441–450, 1987.
[17] D. Blackwell, "Equivalent comparisons of experiments," *Ann. Math. Statist.*, vol. 24, no. 2, pp. 265–272, 1953.
[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.
[19] M. Raginsky, "Shannon meets Blackwell and Le Cam: Channels, codes, and statistical experiments," in *Proc. IEEE Int. Symp. Inf. Theory*, 2011, pp. 1220–1224.
[20] J. Rauh, P. K. Banerjee, E. Olbrich, J. Jost, N. Bertschinger, and D. Wolpert, "Coarse-graining and the Blackwell order," *Entropy*, vol. 19, no. 10, 2017, Art. no. 527.
[21] D. M. Rothschild and J. Wolfers, "Forecasting elections: Voter intentions versus expectations," 2011.
[22] B. Nettasinghe and V. Krishnamurthy, "What do your friends think? Efficient polling methods for networks using friendship paradox," 2018, arXiv:1802.06505.
[23] V. Krishnamurthy and W. Hoiles, "Online reputation and polling systems: Data incest, social learning, and revealed preferences," *IEEE Trans. Comput. Soc. Syst.*, vol. 1, no. 3, pp. 164–179, Sep. 2014.
[24] A. Dasgupta, R. Kumar, and D. Sivakumar, "Social sampling," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2012, pp. 235–243.
[25] D. A. Linzer, "Dynamic Bayesian forecasting of presidential elections in the states," *J. Am. Statist. Assoc.*, vol. 108, no. 501, pp. 124–134, 2013.
[26] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: Real-time event detection by social sensors," in *Proc. 19th Int. Conf. World Wide Web*, ACM, 2010, pp. 851–860.
[27] S. Zhao, L. Zhong, J. Wickramasuriya, and V. Vasudevan, "Human as real-time sensors of social and physical events: A case study of twitter and sports games," 2011, arXiv:1106.4300.
[28] K. T. Chang, W. Chen, and B. C. Tan, "Advertising effectiveness in social networking sites: Social ties, expertise, and product type," *IEEE Trans. Eng. Manage.*, vol. 59, no. 4, pp. 634–643, 2012.
[29] D. Wang *et al.*, "Using humans as sensors: An estimation-theoretic perspective," in *Proc. IPSN-14 13th Int. Symp. Inf. Process. Sensor Netw.*, 2014, pp. 35–46.
[30] P. Burnap, R. Gibson, L. Sloan, R. Southern, and M. Williams, "140 characters to victory?: Using twitter to predict the UK 2015 general election," *Electoral Studies*, vol. 41, pp. 230–233, 2016.
[31] D. J. Watts and P. S. Dodds, "Influentials, networks, and public opinion formation," *J. Consum. Res.*, vol. 34, no. 4, pp. 441–458, 2007.
[32] D. López-Pintado, "Influence networks," *Games Econ. Behav.*, vol. 75, no. 2, pp. 776–787, 2012.
[33] A. Caplin and M. Dean, "Revealed preference, rational inattention, and costly information acquisition," *Am. Econ. Rev.*, vol. 105, no. 7, pp. 2183–2203, 2015.
[34] R. Barnard, W. Dayawansa, K. Pearce, and D. Weinberg, "Polynomials with nonnegative coefficients," *Proc. Am. Math. Soc.*, vol. 113, no. 1, pp. 77–85, 1991.
[35] V. Krishnamurthy, "Algorithms for optimal scheduling and management of hidden Markov model sensors," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1382–1397, Jun. 2002.
[36] N. J. Higham and L. Lin, "On pth roots of stochastic matrices," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 448–463, 2011.
[37] V. Spokoiny and A. Shiryayev, *Statistical Experiments And Decision, Asymptotic Theory*, vol. 8. Singapore: World Scientific, 2000.
[38] C. R. Blyth, "On Simpson's paradox and the sure-thing principle," *J. Am. Statist. Assoc.*, vol. 67, no. 338, pp. 364–366, 1972.
[39] M. Naghshvar and T. Javidi, "Active sequential hypothesis testing," *Ann. Statist.*, vol. 41, no. 6, pp. 2703–2738, 2013.
[40] M. Sakaguchi, *Information Theory and Decision Making*. Washington, D.C., USA: Statist. Dept., George Washington University, 1964.
[41] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. Series B (Methodol.)*, pp. 1–38, 1977.
[42] I. Griva, S. G. Nash, and A. Sofer, *Linear and Nonlinear Optimization*, vol. 108. Philadelphia, PA, USA: SIAM, 2009.

Sujay Bhatt received the M.Tech degree in electrical engineering from the Indian Institute of Technology Bombay, Mumbai, India, in 2014 and the Ph.D. degree in electrical and computer engineering from Cornell University, Ithaca, NY, USA in 2019. His research interests are in network science, computational economics, statistical inference, stochastic control, and reinforcement learning.

Vikram Krishnamurthy (F'05) received the Ph.D. degree from Australian National University, Canberra ACT, Australia, in 1992. He is currently a Professor with Cornell Tech, New York, NY, USA, and the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. From 2002 to 2016, he was a Professor and Canada Research Chair with the University of British Columbia, Canada. His research interests include statistical signal processing and stochastic control in social networks and adaptive sensing. He was a Distinguished Lecturer for the IEEE Signal Processing Society and the Editor-in-Chief of the IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING. In 2013, he was awarded an Honorary Doctorate from KTH, Royal Institute of Technology, Sweden. He is author of the books *Partially Observed Markov Decision Processes* (Cambridge University Press, 2016) and *Dynamics of Engineered Artificial Membranes and Biosensors* (Cambridge University Press, 2018).