

Sequential Detection of Market Shocks With Risk-Averse CVaR Social Sensors

Vikram Krishnamurthy, *Fellow, IEEE*, and Sujay Bhatt

Abstract—This paper considers a statistical signal processing problem involving agent-based models of financial markets, which at a microlevel are driven by socially aware and risk-averse agents. These agents trade (buy or sell) stocks at each trading instant by using the decisions of all previous agents (social learning) in addition to a private (noisy) signal they receive on the value of the stock. We are interested in the following: 1) modelling the dynamics of these risk averse agents and 2) sequential detection of a market shock based on the behaviour of these agents. Structural results that characterize social learning under a risk measure, conditional value-at-risk (CVaR), are presented and formulation of the Bayesian change point detection problem is provided. The structural results exhibit two interesting properties: 1) risk averse agents herd more often than risk neutral agents and 2) the stopping set in the sequential detection problem is nonconvex. The framework is validated on data from the Yahoo! Tech Buzz game dataset and it is revealed that 1) the model identifies the value changes based on agent’s trading decisions. 2) Reasonable quickest detection performance is achieved when the agents are risk-averse.

Index Terms—Conditional value at risk (CVaR), social learning filter, market shock, quickest detection, agent based models, monotone Bayesian update, coherent risk measure, POMDP.

I. INTRODUCTION

FINANCIAL markets evolve based on the behaviour of a large number of interacting entities. Understanding the interaction of these agents is therefore essential in statistical inference from financial data. This motivates the study of “agent based models” for financial markets. Agent based models are useful for capturing the global behaviour of highly interconnected financial systems by simulating the behaviour of the local interacting systems [1]–[4]. Unlike standard economic models which emphasize the equilibrium properties of financial markets, agent based models stress local interactions and out-of-equilibrium dynamics that may not reach equilibrium in the long run [5]. Agent based models are commonly used to determine the conditions that lead a group of interacting agents to form an aggregate behaviour [6]–[9] and to model stylized facts like correlation of returns and volatility clustering [10], [11]. Agent based models have also been used model anomalies that the standard approaches fail to explain like “fat tails”, absence of simple arbitrage, gain/loss asymmetry and leverage effects [12], [13].

Manuscript received October 14, 2015; revised February 16, 2016; accepted March 16, 2016. Date of publication April 14, 2016; date of current version August 12, 2016. The guest editor coordinating the review of this manuscript and approving it for publication was Dr. Emmanuelle Jay.

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

Digital Object Identifier 10.1109/JSTSP.2016.2548995

In this paper, we are interested in developing agent based models for studying global events in financial markets where the underlying value of the stock experiences a jump change (shock). Market shocks are known to affect stock market returns [14], cause fluctuations in the economy [15] and necessitate market making [16]. Therefore detecting shocks is essential and when the interacting agents are acting based on private signals and complete history of other agents’ trading decisions, it is non-trivial [17].

The problem of market shock detection in the presence of social learning considered in this paper is different from a standard signal processing (SP) problem in the following ways:

- 1) Agents (or social sensors) influence the behaviour of other agents, whereas in standard SP sensors typically do not affect other sensors.
- 2) Agents reveal quantized information (decisions) and have dynamics, whereas in standard SP sensors are static with the dynamics modelled in the state equation.
- 3) Standard SP is expectation centric. In this paper we use *coherent risk measures* which generalizes the concept of expected value and is much more relevant in financial applications. Such coherent risk measures [18] are now widely used in finance to model risk averse behaviour.

Properties 1 and 2 above are captured by social learning models. Such social learning models, where agents face *fixed* prices, are considered in [9], [19]–[21]. They show that after a finite amount of time, an informational cascade takes place and all subsequent agents choose the same action regardless of their private signal. Models where agents act sequentially to optimize local costs (to choose an action) and are socially aware were considered in [7], [22]. This paper considers a similar model, but, in order to incorporate property 3 above (risk averse behaviour), we will replace the classical social learning model of expected cost minimizers with that of risk averse minimizers. The resulting risk-averse social learning filter has several interesting (and unusual) properties that will be discussed in the paper.

A. Main Results and Organization

Section II presents the social learning agent based model and the market observer’s objective for detecting shocks. The formulation involves the interaction of local and global decision makers. Individual agents perform social learning and the market observer seeks to determine if the underlying asset value has changed based on the agent behaviour. The shock in the asset value changes at a phase distributed time (which generalizes

geometric distributed change times). The problem of market shock detection considered in this paper is different from the classical Bayesian quickest detection [23]–[25] where, local observations are used to detect the change. Quickest detection in the presence of social learning was considered in [17] where it was shown that making global decisions (stop or continue) based on local decisions (buy or sell) leads to discontinuous value function and the optimal policy has multiple thresholds. However, unlike [17] which deals with expected cost, we consider a more general measure to account for the local agents' attitude towards risk.

It is well documented in various fields like economics [26], behavioural economics, psychology [27] that people prefer a certain but possibly less desirable outcome over an uncertain but potentially larger outcome. To model this risk averse behaviour, commonly used risk measures¹ are Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR), Entropic risk measure and Tail value at risk; see [28]. We consider social learning under CVaR risk measure. CVaR [29] is an extension of VaR that gives the total loss given a loss event and is a coherent risk measure [18]. In this paper, we choose CVaR risk measure as it exhibits the following properties [18], [29]: (i) It associates higher risk with higher cost. (ii) It ensures that risk is not a function of the quantity purchased, but arises from the stock. (iii) It is convex. CVaR as a risk measure has been used in solving portfolio optimization problems [30], [31] credit risk optimization [32] and also order execution [33]. For an overview of risk measures and their application in finance, see [28].

Section III provides structural results which characterize the social learning under CVaR risk measure and its properties. We show that, under reasonable assumptions on the costs, the trading decisions taken by socially aware and risk-averse agents are ordinal functions of their private observations and monotone in the prior information. This implies that the Bayesian social learning follows simple intuitive rules. The change point detection problem is formulated as a Market Observer² seeking to detect a shock in the stock value (modelled as a Markov chain) by balancing the natural trade-off between detection delay and false alarm.

Section IV discusses the unusual properties exhibited by the CVaR social learning filter and explores the link between local and global behaviour in agent based models for detection of market shocks. We show that the stopping region for the sequential detection problem is non-convex; this is in contrast to standard signal processing quickest detection problems where the stopping set is convex.

Finally, Section V discusses an application of the agent based model and change detection framework in a stock market data set. We use a data set from Tech Buzz Game which is a stock

¹A risk measure $\varrho : \mathcal{L} \rightarrow \mathbb{R}$ is a mapping from the space of measurable functions to the real line which satisfies the following properties: (i) $\varrho(0) = 0$. (ii) If $S_1, S_2 \in \mathcal{L}$ and $S_1 \leq S_2$ a.s then $\varrho(S_1) \leq \varrho(S_2)$. (iii) if $a \in \mathbb{R}$ and $S \in \mathcal{L}$, then $\varrho(S + a) = \varrho(S) + a$. The risk measure is coherent if in addition ϱ satisfies: (iv) If $S_1, S_2 \in \mathcal{L}$, then $\varrho(S_1 + S_2) \leq \varrho(S_1) + \varrho(S_2)$. (v) If $a \geq 0$ and $S \in \mathcal{L}$, then $\varrho(aS) = a\varrho(S)$. The expectation operator is a special case where subadditivity is replaced by additivity.

²Market observer could be, for example, a securities dealer (investment bank or syndicate) that underwrites the stock which is later traded in a secondary market.

market simulation launched by Yahoo! Research and O'Reilly Media on March 15, 2005 to gain insights into forecasting high-tech events and trades. The game uses Dynamic pari-mutuel markets (DPM) as its trading mechanism. DPMs are known to provide accurate predictions in field studies on – price formation in election stock markets [34], mechanism design for sales forecasting [35] and betting in sports markets [36], [37].

II. CVAR SOCIAL LEARNING MODEL AND MARKET OBSERVER'S OBJECTIVE

This section presents the Bayesian social learning model and defines the objective of the market observer. As will be shown later in Section III, the model results in ordinal decision making thereby mimicking human behavior and the risk measure captures a trader's attitude towards risk.

A. CVaR Social Learning Model

The market micro-structure is modelled as a discrete time dealer market motivated by algorithmic and high-frequency tick-by-tick trading [38]. There is a single traded stock or asset, a market observer and a countable number of trading agents. The asset has an initial true underlying value $x_0 \in \mathcal{X} = \{1, 2, \dots, X\}$. The market observer does not receive direct information about $x \in \mathcal{X}$ but only observes the public buy/sell actions of agents, $a_k \in \mathcal{A} = \{1(\text{buy}), 2(\text{sell})\}$. The agents themselves receive noisy private observations of the underlying value x and consider this in addition to the trading decisions of the other agents visible in the order book [39], [40], [41]. At a random time, τ^0 determined by the transition matrix P , the asset experiences a jump change in its value to a new value. The aim of the market observer is to detect the change time (global decision) with minimal cost, having access to only the actions of these socially aware agents. Let $y_k \in \mathcal{Y} = \{1, 2, \dots, Y\}$ denote agent k 's private observation. The initial distribution is $\pi_0 = (\pi_0(i), i \in \mathcal{X})$ where $\pi_0(i) = \mathbb{P}(x_0 = i)$.

The agent based model has the following dynamics:

1. *Shock in the asset value:* At time $\tau^0 > 0$, the asset experiences a jump change (shock) in its value due to exogenous factors. The change point τ^0 is modelled by a *phase type (PH) distribution*. The family of all PH-distributions forms a dense subset for the set of all distributions [42] i.e., for any given distribution function F such that $F(0) = 0$, one can find a sequence of PH-distributions $\{F_n, n \geq 1\}$ to approximate F uniformly over $[0, \infty)$. The PH-distributed time τ^0 can be constructed via a multi-state Markov chain x_k with state space $\mathcal{X} = \{1, \dots, X\}$ as follows: Assume state '1' is an absorbing state and denotes the state after the jump change. The states $2, \dots, X$ (corresponding to beliefs e_2, \dots, e_X) can be viewed as a single composite state that x resides in before the jump. So $\tau^0 = \inf \{k : x_k = 1\}$ and the transition probability matrix P is of the form

$$P = \begin{bmatrix} 1 & 0 \\ \underline{P}_{(X-1) \times 1} & \bar{P}_{(X-1) \times (X-1)} \end{bmatrix} \quad (1)$$

The distribution of the absorption time to state 1 is

$$\nu_0 = \pi_0(1), \quad \nu_k = \bar{\pi}'_0 \bar{P}^{k-1} \underline{P}, \quad k \geq 1, \quad (2)$$

where $\bar{\pi}_0 = [\pi_0(2), \dots, \pi_0(X)]'$. The key idea is that by appropriately choosing the pair (π_0, P) and the associated state space dimension X , one can approximate any given discrete distribution on $[0, \infty)$ by the distribution $\{\nu_k, k \geq 0\}$; see [42, pp. 240–243]. The event $\{x_k = 1\}$ means the change point has occurred before time k according to PH-distribution (2). In the special case when x is a 2-state Markov chain, the change time τ^0 is geometrically distributed.

2. *Agent's Private Observation:* Agent k 's private (local) observation denoted by y_k is a noisy measurement of the true value of the asset. It is obtained from the observation likelihood distribution as,

$$B_{xy} = \mathbb{P}(y_k = y | x_k = x) \quad (3)$$

3. *Private Belief update:* Agent k updates its private belief using the observation y_k and the prior public belief $\pi_{k-1}(i) = \mathbb{P}(X = i | a_1, \dots, a_{k-1})$ as the following Hidden Markov Model update

$$\eta_k = \frac{B_{y_k} P' \pi_{k-1}}{\mathbf{1}' B_{y_k} P' \pi_{k-1}} \quad (4)$$

where $\mathbf{1}$ denotes the X -dimensional vector of ones.

4. *Agent's trading decision:* Agent k executes an action $a_k \in \mathcal{A} = \{1(\text{buy}), 2(\text{sell})\}$ to myopically minimize its cost. Let $c(i, a)$ denote the cost incurred if the agent takes action a when the underlying state is i . Let the local cost vector be

$$c_a = [c(1, a)c(2, a) \dots c(X, a)] \quad (5)$$

The costs for different actions are taken as

$$c(i, j) = p_j - \beta_{ij} \text{ for } i \in \mathcal{X}, j \in \mathcal{A} \quad (6)$$

where β_{ij} corresponds to the agent's demand. Here demand is the agent's desire and willingness to trade at a price p_j for the stock. Here p_1 is the quoted price for purchase and p_2 is the price demanded in exchange for the stock. We assume that the price is the same during the period in which the value changes. As a result, the willingness of each agent only depends on the degree of uncertainty on the value of the stock.

Remark 1: The analysis provided in this paper straightforwardly extends to the case when different agents are facing different prices like in an order book [39]–[41]. For notational simplicity we assume the cost are time invariant.

The agent considers measures of risk in the presence of uncertainty in order to overcome the losses incurred in trading. To illustrate this, let $c(x, a)$ denote the loss incurred with action a while at unknown and random state $x \in \mathcal{X}$. When an agent solves an optimization problem involving $c(x, a)$ for selecting the best trading decision,

it will take into account not just the expected loss, but also the ‘‘riskiness’’ associated with the trading decision a . The agent therefore chooses an action a_k to minimize the CVaR measure³ of trading as

$$a_k = \operatorname{argmin}_{a \in \mathcal{A}} \{ \text{CVaR}_\alpha(c(x_k, a)) \} \quad (7)$$

$$= \operatorname{argmin}_{a \in \mathcal{A}} \{ \min_{z \in \mathbb{R}} \{ z + \frac{1}{\alpha} \mathbb{E}_{y_k} [\max\{(c(x_k, a) - z), 0\}] \} \}$$

Here $\alpha \in (0, 1]$ reflects the degree of risk-aversion for the agent (the smaller α is, the more risk-averse the agent is). Define

$$\mathcal{H}_k := \sigma\text{-algebra generated by } (a_1, a_2, \dots, a_{k-1}, y_k) \quad (8)$$

\mathbb{E}_{y_k} denotes the expectation with respect to private belief, i.e., $\mathbb{E}_{y_k} = \mathbb{E}[\cdot | \mathcal{H}_k]$ when the private belief is updated after observation y_k .

5. *Social Learning and Public belief update:* Agent k 's action is recorded in the order book and hence broadcast publicly. Subsequent agents and the market observer update the public belief on the value of the stock according to the social learning Bayesian filter as follows

$$\pi_k = T^{\pi_{k-1}}(\pi_{k-1}, a_k) = \frac{R_{a_k}^{\pi_{k-1}} P' \pi_{k-1}}{\mathbf{1}' R_{a_k}^{\pi_{k-1}} P' \pi_{k-1}} \quad (9)$$

Here, $R_{a_k}^{\pi_{k-1}} = \operatorname{diag}(\mathbb{P}(a_k | x = i, \pi_{k-1}), i \in \mathcal{X})$, where $\mathbb{P}(a_k | x = i, \pi_{k-1}) = \sum_{y \in \mathcal{Y}} \mathbb{P}(a_k | y, \pi_{k-1}) \mathbb{P}(y | x_k = i)$ and

$$\mathbb{P}(a_k | y, \pi_{k-1}) = \begin{cases} 1 & \text{if } a_k = \operatorname{argmin}_{a \in \mathcal{A}} \text{CVaR}_\zeta(c(x_k, a)); \\ 0 & \text{otherwise.} \end{cases}$$

Note that π_k belongs to the unit simplex $\Pi(X) \triangleq \{\pi \in \mathbb{R}^X : \mathbf{1}'_X \pi = 1, 0 \leq \pi \leq 1 \text{ for all } i \in \mathcal{X}\}$.

6. *Market Observer's Action:* The market observer (securities dealer) seeks to achieve quickest detection by balancing delay with false alarm. At each time k , the market observer chooses action⁴ u_k as

$$u_k \in \mathcal{U} = \{1(\text{stop}), 2(\text{continue})\} \quad (10)$$

Here ‘Stop’ indicates that the value has changed and the dealer incorporates this information before selling new issues to investors. The formulation presented considers a general parametrization of the costs associated with detection delay and false alarm costs. Define

$$\mathcal{G}_k := \sigma\text{-algebra generated by } (a_1, a_2, \dots, a_{k-1}, a_k). \quad (11)$$

³For the reader unfamiliar with risk measures, it should be noted that CVaR is one of the ‘big’ developments in risk modelling in finance in the last 15 years. In comparison, the value at risk (VaR) is the percentile loss namely, $\text{VaR}_\alpha(x) = \min\{z : F_x(z) \geq \alpha\}$ for cdf F_x . While CVaR is a coherent risk measure, VaR is not convex and so not coherent. CVaR has other remarkable properties [29]: it is continuous in α and jointly convex in (x, α) . For continuous cdf F_x , $\text{CVaR}_\alpha(x) = \mathbb{E}\{X | X > \text{VaR}_\alpha(x)\}$. Note that the variance is not a coherent risk measure.

⁴It is important to distinguish between the ‘local’ decisions a_k of the agents and ‘global’ decisions u_k of the market observer. Clearly the decisions a_k affect the choice of u_k as will be made precise below.

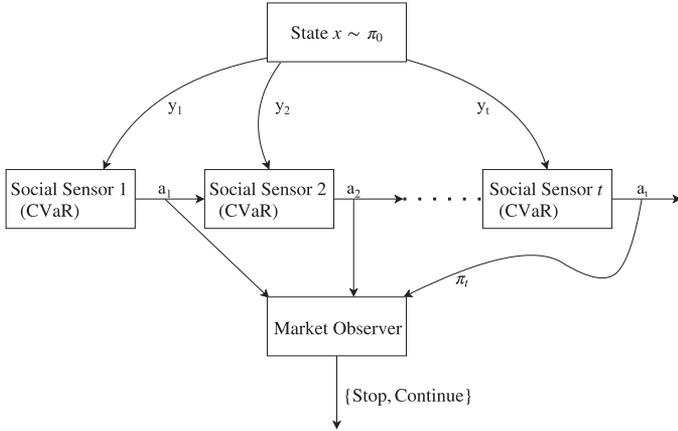


Fig. 1. Sequential detection with risk-averse social sensors. Each social sensor receives a noisy observation on the state and chooses an action to minimize its CVaR measure of trading. The social sensors communicate their actions to subsequent sensors. The market observer seeks to determine if there is a change in the value of the underlying asset from the actions of the sensors.

- i) *Cost of Stopping*: The asset experiences a jump change (shock) in its value at time τ^0 . If the action $u_k = 1$ is chosen before the change point, a false alarm penalty is incurred. This corresponds to the event $\bigcup_{i \geq 2} x_k = i \cap u_k = 1$. Let \mathcal{I} denote the indicator function. The cost of false alarm in state $i, i \in \mathcal{X}$ with $f_i \geq 0$ is thus given by $f_i \mathcal{I}(x_k = i, u_k = 1)$. The expected false alarm penalty is

$$\begin{aligned} C(\pi_k, u_k = 1) &= \sum_{i \in \mathcal{X}} f_i \mathbb{E}\{\mathcal{I}(x_k = i, u_k = 1) | \mathcal{G}_k\} \\ &= \mathbf{f}' \pi_k \end{aligned} \quad (12)$$

where $\mathbf{f} = (f_1, \dots, f_X)$ and it is chosen with increasing elements, so that states further from '1' incur higher false alarm penalties. Clearly, $f_1 = 0$.

- ii) *Cost of delay*: A delay cost is incurred when the event $\{x_k = 1, u_k = 2\}$ occurs, i.e., even though the state changed at k , the market observer fails to identify the change. The expected delay cost is

$$\begin{aligned} C(\pi_k, u_k = 2) &= d \mathbb{E}\{\mathcal{I}(x_k = 1, u_k = 2) | \mathcal{G}_k\} \\ &= d e_1' \pi_k \end{aligned} \quad (13)$$

where $d > 0$ is the delay cost and e_1 denotes the unit vector with 1 in the first position.

Fig. 1 illustrates the above social learning model in which the information exchange between the risk-averse social sensors is sequential.

B. Market Observer's Quickest Detection Objective

The market observer chooses its action at each time k as

$$u_k = \mu(\pi_k) \in \{1(\text{stop}), 2(\text{continue})\} \quad (14)$$

where μ denotes a stationary policy. For each initial distribution $\pi_0 \in \Pi(X)$ and policy μ , the following cost is associated

$$J_\mu(\pi_0) = \mathbb{E}_{\pi_0}^\mu \left\{ \sum_{k=1}^{\tau-1} \rho^{k-1} C(\pi_k, u_k = 2) + \rho^{\tau-1} C(\pi_k, u_k = 1) \right\} \quad (15)$$

Here $\rho \in [0, 1]$ is the discount factor which is a measure of the degree of impatience of the market observer. (As long as ρ is non-zero, stopping is guaranteed in finite time and so $\rho = 1$ is allowed.)

Given the cost, the market observer's objective is to determine τ^0 with minimum cost by computing an optimal policy μ^* such that

$$J_{\mu^*}(\pi_0) = \inf_{\mu \in \mathcal{M}} J_\mu(\pi_0) \quad (16)$$

The sequential detection problem (16) can be viewed as a partially observed Markov decision process (POMDP) where the belief update is given by the social learning filter.

C. Stochastic Dynamic Programming Formulation

The optimal policy of the market observer $\mu^* : \Pi(X) \rightarrow \{1, 2\}$ is the solution of (15) and is given by Bellman's dynamic programming equation as follows:

$$\begin{aligned} V(\pi) &= \min \left\{ C(\pi, 1), C(\pi, 2) + \rho \sum_{a \in \mathcal{A}} V(T^\pi(\pi, a)) \sigma(\pi, a) \right\} \\ \mu^*(\pi) &= \operatorname{argmin} \left\{ C(\pi, 1), C(\pi, 2) + \rho \sum_{a \in \mathcal{A}} V(T^\pi(\pi, a)) \sigma(\pi, a) \right\} \end{aligned} \quad (17)$$

where $T^\pi(\pi, a) = \frac{R_a^\pi P' \pi}{\mathbf{1}' R_a^\pi P' \pi}$ is the CVaR-social learning filter and $\sigma(\pi, a) = \frac{\mathbf{1}' R_a^\pi P' \pi}{\mathbf{1}' R_a^\pi P' \pi}$ is the normalization factor of the Bayesian update. $C(\pi, 1)$ and $C(\pi, 2)$ from (12) and (13) are the market observer's costs. As $C(\pi, 1)$ and $C(\pi, 2)$ are non-negative and bounded for $\pi \in \Pi(X)$, the stopping time τ is finite for all $\rho \in [0, 1]$.

The aim of the market observer is then to determine the stopping set $\mathcal{S} = \{\pi \in \Pi(X) : \mu^*(\pi) = 1\}$ given by:

$$\mathcal{S} = \left\{ \pi : C(\pi, 1) < C(\pi, 2) + \rho \sum_{a \in \mathcal{A}} V(T^\pi(\pi, a)) \sigma(\pi, a) \right\}$$

The dynamic programming equation (17) is similar to that for stopping time POMDP except that the belief update is given by a CVaR social learning filter. As will be shown below, because of the social learning dynamics, quite remarkably, \mathcal{S} is not necessarily a convex set. This is in stark contrast to classical quickest detection where the stopping region is always convex irrespective of the change time distribution [43].

III. PROPERTIES OF CVAR SOCIAL LEARNING FILTER

This section discusses the main results regarding the structural properties of the CVaR social learning filter and highlights

the significant role it plays in charactering the properties of market observer's value function and optimal policy. According to Theorem 1, risk-averse agents take decisions that are monotone and ordinal in the observations and monotone in the prior; and its monotone ordinal behaviour implies that a Bayesian model chosen in this paper is a useful idealization.

A. Assumptions

The following assumptions will be used throughout the paper:

(A1) Observation matrix B and transition matrix P are TP2 (all second order minors are non-negative)

(A2) Agents' local cost vector c_a is sub-modular. That is $c(x, 2) - c(x, 1) \leq c(x + 1, 2) - c(x + 1, 1)$.

The matrices being TP2 [44] ensures that the public belief Bayesian updates can be compared [45] and sub-modular [46] costs ensure that if it is less risky to choose $a = 2$ when in x , it is also less risky to choose it when in $x + 1$.

B. Properties of CVaR Social Learning Filter

The $\mathcal{Y} \times \mathcal{A}$ local decision likelihood probability matrix R^π (analogous to observation likelihood) can be computed as

$$R^\pi = BM^\pi, \text{ where } M_{y,a}^\pi \triangleq \mathbb{P}(a|y, \pi) \quad (18)$$

$$\mathbb{P}(a|y, \pi) = \mathcal{I}(\text{CVaR}_\alpha(c(x_k, a)) < \text{CVaR}_\alpha(c(x_k, a')))$$

where $a' = A - \{a\}$. Here \mathcal{I} denotes the indicator function.

Let $H^\alpha(y, a) = \text{CVaR}_\alpha(c(x_k, a))$ denote the cost with CVaR measure, associated with action a and observation y for convenience i.e.,

$$H^\alpha(y, a) = \min_{z \in \mathbb{R}} \left\{ z + \frac{1}{\alpha} \mathbb{E}_y[\max\{(c(x, a) - z), 0\}] \right\} \quad (19)$$

Here $\mathbb{E}_y = \mathbb{E}[\cdot | \mathcal{H}_k]$. y indicates the dependence of \mathbb{E} and hence H^α on the observation. Let $a^*(\pi, y) = \text{argmin } H^\alpha(y, a)$ denote the optimal action of the agent with explicit dependence on the distribution and observation.

The following result says that agents choose a trading decision that is monotone and ordinal in their private observation. Humans typically convert numerical attributes to ordinal scales before making decisions. For example, it does not matter if the cost of a meal at a restaurant is \$200 or \$205; an individual would classify this cost as "high". Also credit rating agencies use ordinal symbols such as AAA, AA, A.

Theorem 1: Under (A1) and (A2), the action $a^*(\pi, y)$ made by each agent is increasing and hence ordinal in y for any prior belief π . Under (A2), $a^*(\pi, y)$ is increasing in π with respect to the monotone likelihood ratio order (Definition 1 in the appendix).

The proof is given in the appendix. Theorem 1 says that agents exhibit monotone ordinal behaviour. The condition that $a^*(\pi, y)$ is monotone in the observation y is required to characterize the local decision matrices on different regions in the belief space which is stated next.

Theorem 2: Under (A1) and (A2), there are at most $Y + 1$ distinct local decision likelihood matrices R^π and the belief

space $\Pi(X)$ can be partitioned into the following $Y + 1$ polytopes:

$$\begin{aligned} \mathcal{P}_1^\alpha &= \{\pi \in \Pi(X) : H(1, 1) - H(1, 2) \geq 0\} \\ \mathcal{P}_l^\alpha &= \{\pi \in \Pi(X) : H(l-1, 1) - H(l-1, 2) < 0 \\ &\quad \cap H(l, 1) - H(l, 2) \geq 0\}, l = 2, \dots, Y \\ \mathcal{P}_{Y+1}^\alpha &= \{\pi \in \Pi(X) : H(Y, 1) - H(Y, 2) < 0\} \end{aligned} \quad (20)$$

Also, the matrices R^π are constant (with respect to π) on each of these polytopes.

The proof is given in the appendix. Theorem 2 is required to specify the policy for the market observer. Indeed it leads to unusual behavior (non-convex) stopping regions in quickest detection as described in Section IV-B.

IV. SOCIAL LEARNING AND CHANGE DETECTION FOR RISK-AVERSE AGENTS

This section illustrates the properties of the risk-averse social learning filter which leads to a non-convex value function and therefore non-convex stopping set of quickest detection.

A. Social Learning Behavior of Risk Averse Agents

The following discussion highlights the relation between risk-aversion factor α and the regions \mathcal{P}_l^α . For a given risk-aversion factor α , Theorem 2 shows that there are at most $Y + 1$ polytopes on the belief space. It was shown in [17] that for the risk neutral case with $X = 2$, and $P = I$ (the value is a random variable) the intervals \mathcal{P}_1^α and \mathcal{P}_3^α correspond to the herding region and the interval \mathcal{P}_2^α corresponds to the social learning region. In the herding region, the agents take the same action as the belief is frozen. In the social learning region there is observational learning. However, when the agents are optimizing a more general risk measure (CVaR), the social learning region is different for different risk-aversion factors. The social learning region for the CVaR risk measure is shown in Fig. 2. It can be observed from Fig. 2 that \mathcal{P}_1^α becomes smaller, \mathcal{P}_2^α becomes smaller and \mathcal{P}_3^α becomes larger as α decreases. The following parameters were chosen:

$$B = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}, P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, c = \begin{bmatrix} 1 & 2 \\ 3 & 0.5 \end{bmatrix}$$

This can be interpreted as risk-averse agents showing a larger tendency to go with the crowd rather than "risk" choosing the other action. With the same B and c parameters, but with transition matrix

$$P = \begin{bmatrix} 1 & 0 \\ 0.1 & 0.9 \end{bmatrix}$$

the social learning region is shown in Fig. 3. From Fig. 3, it is observed that when the state is evolving and when the agents are sufficiently risk-averse, social learning region is very small. It can be interpreted as: agents having a strong risk-averse attitude don't prefer to "learn" from the crowd; but rather face the same consequences, when $P \neq I$.

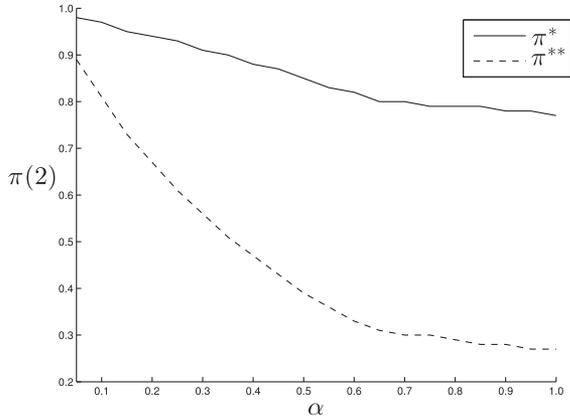


Fig. 2. The social learning region for the risk-aversion parameter $\alpha \in (0, 1]$. It can be seen that the curves corresponding to π^{**} and π^* do not intersect and their separation (social learning region) varies with α . Here $P = I$, i.e. the value is a random variable.

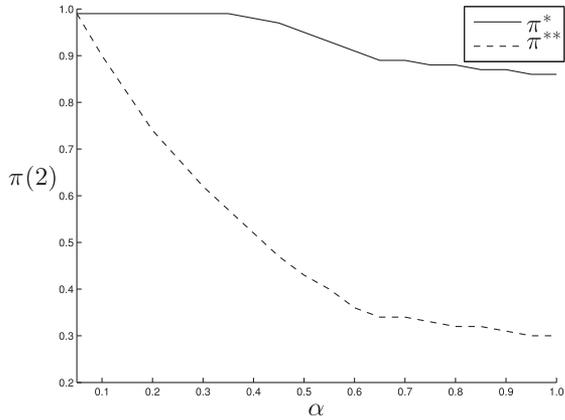


Fig. 3. The social learning region for the risk-aversion parameter $\alpha \in (0, 1]$. It can be seen that the social learning region is absent when agents are sufficiently risk-averse and is larger when the stock value is known to change, i.e. $P \neq I$.

B. Nonconvex Stopping Set for Market Shock Detection

We now illustrate the solution to the Bellman's stochastic dynamic programming equation (17), which determines the optimal policy for quickest market shock detection, by considering an agent based model with two states. Clearly the agents (local decision makers) and market observer interact – the local decisions a_k taken by the agents determines the public belief π_k and hence determines decision u_k of the market observer via (14).

From Theorem 2, the polytopes $\mathcal{P}_1^\alpha, \mathcal{P}_2^\alpha$ and \mathcal{P}_3^α are subsets of $[0, 1]$. Under (A1) and (A2), $\mathcal{P}_1^\alpha = [0, \pi^{**}(2)]$, $\mathcal{P}_2^\alpha = [\pi^{**}(2), \pi^*(2)]$, $\mathcal{P}_3^\alpha = [\pi^*(2), 1]$, where π^{**} and π^* are the belief states at which $H^\alpha(2, 1) = H^\alpha(2, 2)$ and $H^\alpha(1, 1) = H^\alpha(1, 2)$ respectively. From Theorem 2 and (17), the value function can be written as,

$$\begin{aligned} V(\pi) = & \min\{C(\pi, 1), C(\pi, 2) + \rho V(\pi) \mathcal{I}(\pi \in \mathcal{P}_1^\alpha) \\ & + \rho \sum_{a \in \mathcal{A}} V(T^\pi(\pi, a)) \sigma(\pi, a) \mathcal{I}(\pi \in \mathcal{P}_2^\alpha) \\ & + \rho V(\pi) \mathcal{I}(\pi \in \mathcal{P}_3^\alpha)\} \end{aligned}$$

The explicit dependence of the filter on the belief π results in discontinuous value function. The optimal policy in general has

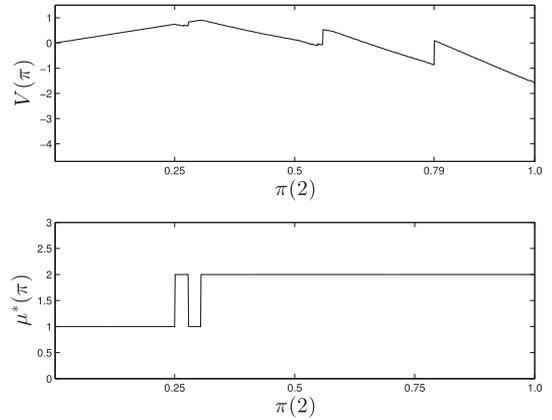


Fig. 4. The value function $V(\pi)$ and the double threshold optimal policy $\mu^*(\pi)$ are plotted over $\pi(2)$. The significance of the double threshold policy is that the stopping regions are non-convex. The implication of non-convex stopping set for the market observer is that - if he believes that it is optimal to stop, it need not be optimal to stop when his belief is larger.

multiple thresholds and the stopping region in general is non-convex.

Example 1: Fig. 4 displays the value function and optimal policy for a toy example having the following parameters:

$$B = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}, P = \begin{bmatrix} 1 & 0 \\ 0.06 & 0.94 \end{bmatrix}, c = \begin{bmatrix} 1 & 2 \\ 2.5 & 0.5 \end{bmatrix}$$

The parameters for the market observer are chosen as: $d = 1.25$, $\mathbf{f} = [0 \ 3]$, $\alpha = 0.8$ and $\rho = 0.9$.

From Fig. 4 it is clear that the market observer has a double threshold policy and the value function is discontinuous. The double threshold policy is unusual from a signal processing point of view. Recall that $\pi(2)$ depicts the posterior probability of no change. The market observer “changes its mind” - it switches from no change to change as the posterior probability of change decreases! Thus the global decision (stop or continue) is a non-monotone function of the posterior probability obtained from local decisions in the agent based model. The example illustrates the unusual behaviour of the social learning filter.

C. Multi-state Markov chains

The structural results for the risk averse social learning filter, namely Theorem 1 and Theorem 2, apply to multi-state Markov chains. However, in numerical examples, to illustrate the optimal policy, we have used 2-state Markov chains to model the stock value. Multi-state Markov chain examples can also be considered, but the numerical solution is substantially more expensive and one has to resort to suboptimal methods such as open loop feedback control (OLFC) [47] to compute a policy. In [17], structural results for the optimal policy in the risk neutral case are considered. It is of interest to generalize these results to the risk averse case considered in this paper.

V. DATASET EXAMPLE

Here, we illustrate multi-agent quickest change detection by considering a dataset from the Tech Buzz Game, which is a stock market simulation launched by Yahoo! Research

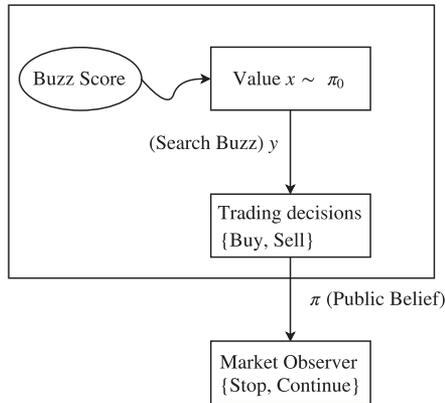


Fig. 5. Model of Tech-Buzz game. Daily Buzz Scores and the individual trading decisions of all the agents is available in the dataset. Value of the stock is calculated using the method in [49]. In the simulated model, the public belief is updated after every trading decision and the Market Observer announces a change based on the optimal policy.

and O'Reilly Media on March 15, 2005 to gain insights into forecasting high-tech events and trades. The overall setup is described in Fig. 5.

A. Tech Buzz Game Pricing Mechanism

The Tech Buzz game uses Dynamic Pari-Mutuel Market (DPM) as its trading mechanism. DPM was developed in [48] as a mechanism for risk allocation and information speculation. A DPM is a hybrid between pari-mutuel (i.e., redistributive - guaranteed to pay out exactly the money taken in) and a continuous double auction (CDA) market. DPM is designed to have infinite liquidity, like pari-mutuel markets, wherein the traders can always purchase shares of any stock at any time at a price automatically set by the mechanism. Like in CDA, DPM incorporates information arriving over time. A market using DPM as its trading mechanism changes the price for the stock based on the demand for the stock: price increases when the demand increases. The price is set by the market using a *price function* [48], which has the flexibility to accommodate the properties desired. Tech Buzz game consisted of multiple sub-markets trading stocks of contemporary rival technologies. DPM played the role of a market maker that accepts orders at its current price and adjusts the price after each order. In the Tech Buzz game, DPM set the price by equating the ratio of prices of any two stocks (within the sub-market) by the ratio of number of shares outstanding for the two stocks at any time of the market. So, if $\mathbf{q} = [q_1, q_2, \dots, q_n]$ is the vector of outstanding shares for the n stocks in a sub-market, the price for stock i is given by the price function defined by [49]:

$$p_i(\mathbf{q}) = \frac{\nu q_i}{\sqrt{\sum_{j=1}^n q_j^2}} \forall i \quad (21)$$

where $\nu > 0$ is a free parameter. The traders buy or sell the stocks depending on the price to maximize their utility. Tech Buzz dataset was chosen to demonstrate the framework as the individual actions for the duration of the game was made available by Yahoo.

B. Simulation Model

The stock market simulation is modelled as shown in Fig. 5. A stock's "buzz score" is an indicator of the number number of buzz searches over the past seven days, as a percentage of all searches in the same market [49]. Thus, if searches for the stock named "SKYPE" make up 80 percent of all Yahoo! searches in the telecommunication application software market, then SKYPE's buzz score is 80. The buzz scores of all technologies within a market always add up to 100. The scores reflect the "ground truth", based on which the value of the stock is calculated. The payout and dividend are directly proportional to the buzz score. The value of a stock is a function of the payout of the stock and its dividend [49]. The state x_k is chosen to represent value of the stock, with $x_k = 1$ indicating a high valued stock and $x_k = 2$ indicating a low valued stock. At each trading instant, the traders (or players) have access to the current *search* buzz associated with each of the stocks measured by the number of users searching information on it at Yahoo Search. The noisy observations, y_k , are chosen as the search buzz which is a proxy for the popularity (sentiment) of the stock [49]. The choice of probabilities for the observation matrix B was motivated by the experimental evidence provided in [50], that when there is social learning "alone", the trading rate was 71% based on peer effects. Since the local decision likelihood matrix $R^\pi = B$ in the social learning region (in our model), the parameters were chosen as

$$B = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

The probabilities in the transition matrix P were chosen to reflect the time window considered. For tractability, it is assumed that all agents have the same attitude towards risk, i.e., same risk-aversion parameter α in (19).

Tech Buzz game has an order book indicating: trader id, date and time of the transaction, the stock traded, number of shares bought or sold, cost of the transaction, and price of the stock before the transaction. The order book information is available from the dataset. Individual agents choose to buy ($a = 1$) or sell ($a = 2$) depending on the past history of actions, price and search buzz to minimize their CVaR measure of trading. On each day the stock is traded, we consider only the agent that buys or sells maximum shares and record its trading decision (positive or negative values in the dataset). This is reasonable assumption since the agent trading maximum shares ("big players" in finance) will significantly influence the public belief.

The cost parameters in (5), were chosen to reflect the intuition that purchasing a high valued product at a high price will maximize the utility⁵ (minimize the cost). The costs of all the traders are assumed to be the same for simplicity. The costs for the market observer, (12) and (13), were chosen to accommodate the desired trade-off between false alarm and delay penalty. With these parameters, the market observer seeks to determine if there is a change in the value of the underlying asset from the actions of these agents using the public belief.

⁵With additional data regarding budget constraints, tools from expected utility theory [51], [52] and revealed preferences [53], [54], [55], may be used to estimate the parameters in the agents' utility/cost function.

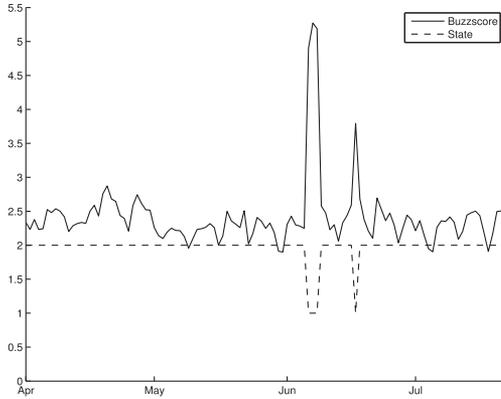


Fig. 6. Quantized values of the stock *SKYPE* and the scaled buzz scores during April - July is shown. It is seen that the value changed in the month of June, 2005. The market observer's aim is to detect the change in the value, i.e., when $x_k = 1$, using only the trading decisions.

C. Dataset and Quickest Detection

From the Yahoo! Dataset⁶, the buzz score for the stocks *SKYPE* and *IPOD* trading in markets *VOIP* and *PORTMEDIA* respectively, was obtained for the period from April 1, 2005 to July 27, 2005. Quantized values of the stocks *SKYPE* and *IPOD* during the period is shown in the Fig. 6 and Fig. 9 along with the scaled buzz score.

1. *Quickest Detection for SKYPE*: From Fig. 6, it is seen that the stock value changed during the month of June, 2005. To apply the quickest detection protocol, we consider a window from May 17, 2005 to June 8, 2005. It is observed that the price was (almost) constant during this period with a value close to 13 per stock. The trading decisions (along with the value) and the public belief during this period are shown in Fig. 7.

The cost parameters chosen according to the rationale described in Section V-B are:

$$c = \begin{bmatrix} 0.5 & 1 \\ 1 & 0.5 \end{bmatrix}, \mathbf{f} = [0 \ 2], d = 0.8$$

The probabilities in the transition matrix P were chosen to reflect the time window considered. As $\mathbb{E}\{\tau_0\} = 25$, the transition matrix assuming a geometrically distributed change time is:

$$P = \begin{bmatrix} 1 & 0 \\ 0.04 & 0.96 \end{bmatrix}$$

It was observed that the state changed on June 6, 2005 and for a risk-aversion factor of $\alpha = 0.45$, it was detected on June 7, 2005. The value function and the optimal policy for the market observer are shown in Fig. 8. As seen from Fig. 8, the optimal policy is a threshold. In general, however, the optimal policy in a social learning model can be non-convex as shown in Fig. 4. The stopping set corresponds to $\pi(2) \in [0, 0.354]$. The regions $\pi(2) \in [0, 0.34]$ and $\pi(2) \in [0.76, 1]$ correspond to the regions where social learning is absent. It can be observed that the value function is discontinuous. This implies that small changes

⁶“Yahoo! Webscope”, http://research.yahoo.com/Academic_Relations_ydata-yrbuzzgame-transactions-period1-v1.0_ydata-yrbuzzgame-buzzscores-period1-v1.0

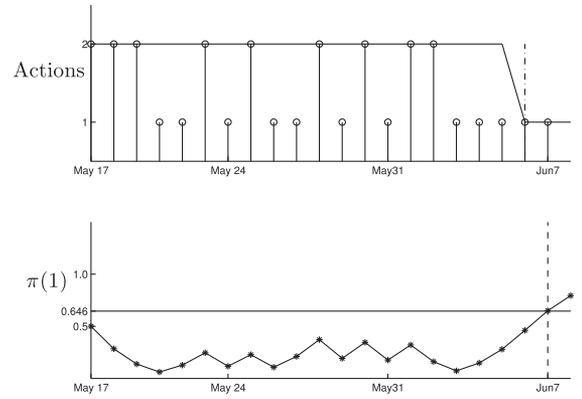


Fig. 7. The daily trading decisions of the agents is shown with the corresponding belief update. Here $a = 1$ corresponds to buying and $a = 2$ corresponds to selling the stock. Since $\pi(2) \in [0, 0.354]$ is the stopping region, it corresponds to $\pi(1) \geq 0.646$. The change was detected one day after it occurred.

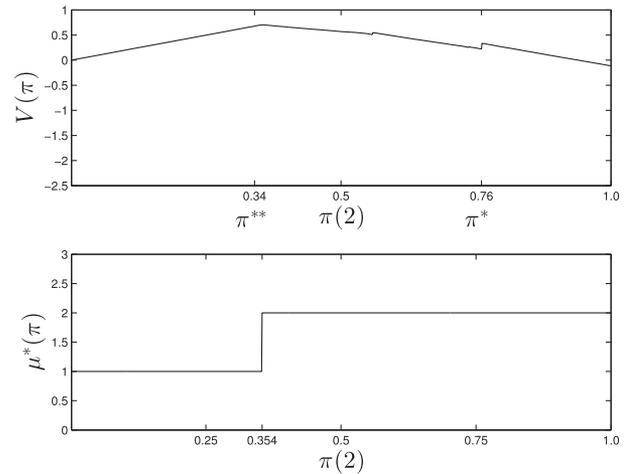


Fig. 8. Value function $V(\pi)$ and the optimal policy $\mu^*(\pi)$ are plotted over $\pi(2)$ for $\alpha = 0.45$. Here π^{**} and π^* correspond to the boundary points of the social learning region. $\mu^*(\pi) = 1$ corresponds to stop and $\mu^*(\pi) = 2$ corresponds to continue. $\pi(2) \in [0, 0.354]$ corresponds to the stopping region.

in the public belief can result in large changes in the optimal cost accrued by the market observer. This unusual feature is again due to the social learning dynamics; in classical quickest detection the optimal cost is a continuous function of the belief.

2. *Quickest Detection for IPOD*: From Fig. 9, it is seen that the stock value changed during April and July. To apply the quickest detection protocol, we consider a window from July 2, 2005 to July 10, 2005. It is observed that the price was (almost) constant during this period with a value close to 17 per stock. The trading decisions (along with the value) and the public belief during this period are shown in Fig. 10.

The cost parameters chosen according to the rationale described in Section V-B are:

$$c = \begin{bmatrix} 0.5 & 1 \\ 1 & 0.5 \end{bmatrix}, \mathbf{f} = [0 \ 1.8], d = 0.95$$

The probabilities in the transition matrix P were chosen to reflect the time window considered. As $\mathbb{E}\{\tau_0\} = 9$, the

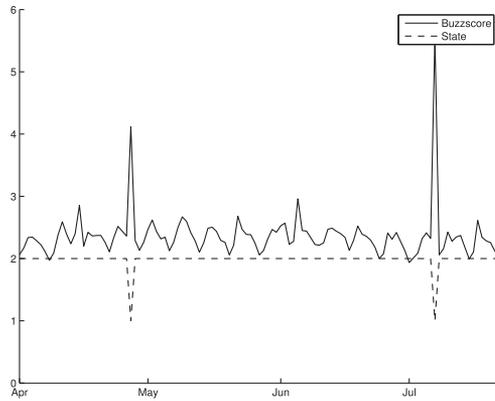


Fig. 9. Quantized values of the stock *IPOD* and the scaled buzz scores during April - July is shown. It is seen that the stock value changed during April, 2005 and July, 2005. The market observer’s aim is to detect the change in the value, i.e. when $x_k = 1$, using only the trading decisions.

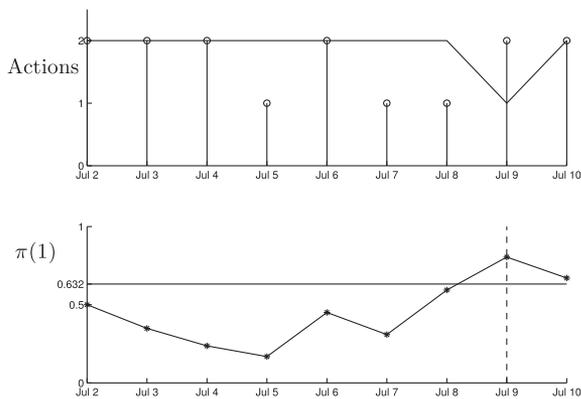


Fig. 10. The daily trading decisions of the agents is shown with the corresponding belief update. Here $a = 1$ corresponds to buying and $a = 2$ corresponds to selling the stock. Since $\pi(2) \in [0, 0.368]$ is the stopping region, it corresponds to $\pi(1) \geq 0.632$. The change was detected on the day it occurred with a higher penalty on the delay.

transition matrix assuming a geometrically distributed change time is:

$$P = \begin{bmatrix} 1 & 0 \\ 0.11 & 0.89 \end{bmatrix}$$

It was observed that the state changed on July 9, 2005 and for a risk-aversion factor of $\alpha = 0.45$, it was detected on July 9, 2005. It is seen that when the delay penalty is increased, the change is detected on the same day. The value function and the optimal policy for the market observer are shown in Fig. 11. $\pi(2) \in [0, 0.368]$ corresponds to the stopping set.

VI. CONCLUSION

The paper provided a Bayesian formulation of the problem of quickest detection of change in the value of a stock using the decisions of socially aware risk averse agents. From a signal processing point of view, the formulation and solutions presented here are non-standard due to the three properties described in Section I. The quickest detection problem was shown to be non-trivial - the stopping region is in general

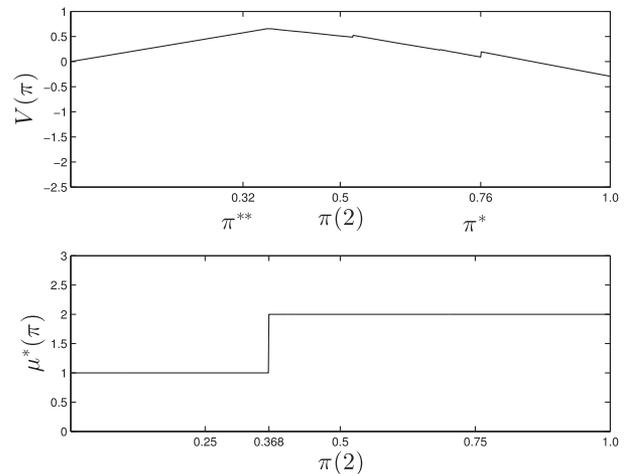


Fig. 11. Value function $V(\pi)$ and the optimal policy $\mu^*(\pi)$ are plotted over $\pi(2)$ for $\alpha = 0.45$. Here π^{**} and π^* correspond to the boundary points of the social learning region. $\mu^*(\pi) = 1$ corresponds to stop and $\mu^*(\pi) = 2$ corresponds to continue. $\pi(2) \in [0, 0.368]$ corresponds to the stopping region.

non-convex when the agents’ risk attitude was accounted for by considering a coherent risk measure, CVaR. Results which characterize the structural properties of social learning under the CVaR risk measure were provided and the importance of these results in understanding the global behaviour was discussed. It was observed that the behaviour of these risk-averse agents is, as expected, different from risk neutral agents. Risk averse agents herd sooner and don’t prefer to “learn” from the crowd, i.e. social learning region is smaller the more risk-averse the agents are. There is an opportunity to apply the framework to study the behaviour of interacting agents in online prediction markets such as Iowa Electronic Markets⁷, Trade Sports and Foresight Exchange.

APPENDIX A
PRELIMINARIES AND DEFINITIONS

Definition 1 MLR Ordering [56] (\geq_r): Let $\pi_1, \pi_2 \in \Pi(X)$ be any two belief state vectors. Then $\pi_1 \geq_r \pi_2$ if

$$\pi_1(i)\pi_2(j) \leq \pi_2(i)\pi_1(j), \quad i < j, i, j \in \{1, \dots, X\}.$$

Definition 2 First-Order Stochastic Dominance (\geq_s): Let $\pi_1, \pi_2 \in \Pi(X)$ be any two belief state vectors. Then $\pi_1 \geq_s \pi_2$ if

$$\sum_{i=j}^X \pi_1(i) \geq \sum_{i=j}^X \pi_2(i) \text{ for } j \in \{1, \dots, X\}.$$

Lemma 3: [56] $\pi_2 \geq_s \pi_1$ iff for all $v \in \mathcal{V}$, $v'\pi_2 \leq v'\pi_1$, where \mathcal{V} denotes the space of X - dimensional vectors v , with non-increasing components, i.e. $v_1 \geq v_2 \geq \dots v_X$.

Lemma 4: [56] $\pi_2 \geq_s \pi_1$ iff for all $v \in \mathcal{V}$, $v'\pi_2 \geq v'\pi_1$, where \mathcal{V} denotes the space of X - dimensional vectors v , with non-decreasing components, i.e. $v_1 \leq v_2 \leq \dots v_X$.

⁷The authors thank an anonymous reviewer for this suggestion.

Let $\Pi(X) \triangleq \{\pi \in \mathbb{R}^X : \mathbf{1}'_X \pi = 1, 0 \leq \pi(i) \leq 1 \text{ for all } i \in X\}$

Definition 3 Submodular function [46]: A function $f : \Pi(X) \times \{1, 2\} \rightarrow \mathbb{R}$ is submodular if $f(\pi, u) - f(\pi, \bar{u}) \leq f(\bar{\pi}, u) - f(\bar{\pi}, \bar{u})$, for $\bar{u} \leq u, \bar{\pi} \geq_r \pi$.

Definition 4 Single Crossing Condition [46]: A function $g : \mathcal{Y} \times \mathcal{A} \rightarrow \mathbb{R}$ satisfies a single crossing condition in (y, a) if

$$g(y, a) - g(y, \bar{a}) \geq 0 \Rightarrow g(\bar{y}, a) - g(\bar{y}, \bar{a}) \geq 0$$

for $\bar{a} > a$ and $\bar{y} > y$. For any such function g ,

$$a^*(y) = \underset{a}{\operatorname{argmin}} g(y, a) \text{ is increasing in } y. \quad (22)$$

Theorem 5 [46] If $f : \Pi(X) \times \{1, 2\} \rightarrow \mathbb{R}$ is sub-modular, then there exists a $u^*(\pi) = \underset{u \in \{1, 2\}}{\operatorname{argmin}} f(\pi, u)$ satisfying,

$$\bar{\pi} \geq_r \pi \Rightarrow u^*(\pi) \leq u^*(\bar{\pi})$$

APPENDIX B PROOFS

The following lemmas are required to prove Theorem 1 and Theorem 2. The results will be proved for general state and observation spaces having two actions.

Lemma 6: For a finite state and observation alphabet, $\underset{z \in \mathbb{R}}{\operatorname{argmin}} \{z + \frac{1}{\alpha} \mathbb{E}_y[\max\{(c(x, a) - z), 0\}]\}$ is equal to $c(i, a)$ for some $i \in \{1, 2, \dots, X\}$.

Proof: Let η_y be the belief update (p.m.f) with observation y , i.e, $\eta_y(i) = \mathbb{P}_y(x = i)$. Let $F_y(x)$ denote the cumulative distribution function. For simplicity of notation, let $h_y(z) = z + \frac{1}{\alpha} \mathbb{E}_y[\max\{(c(x, a) - z), 0\}]$. The extremum of $h_y(z)$ is attained where the derivative is zero. It is obtained as follows.

$$\begin{aligned} h_y(z) &= z + \frac{1}{\alpha} \mathbb{E}_y[\max\{(c(x, a) - z), 0\}] \\ h'_y(z) &= 1 + \frac{1}{\alpha} \lim_{\Delta z \rightarrow 0} \frac{\mathbb{E}_y[\max\{c(x, a) - z - \Delta z, 0\}] - \mathbb{E}_y[\max\{c(x, a) - z, 0\}]}{\Delta z} \\ &= 1 + \frac{1}{\alpha} \mathbb{E}_y \\ &\quad \times \left(\lim_{\Delta z \rightarrow 0} \frac{\max\{c(x, a) - z - \Delta z, 0\} - \max\{c(x, a) - z, 0\}}{\Delta z} \right) \\ &= 1 + \frac{1}{\alpha} \mathbb{E}_y (0 \times \mathcal{I}_{0 > (c(x, a) - z)} - 1 \times \mathcal{I}_{(c(x, a) - z) > 0}) \\ &= 1 - \frac{1}{\alpha} \mathbb{P}_y(c(x, a) > z). \end{aligned}$$

Also, $h''_y(z) = \frac{1}{\alpha} \frac{d}{dz} (F_y(z))$ and therefore $h''_y(z) \geq 0$. We have, $\underset{z \in \mathbb{R}}{\operatorname{argmin}} \{h_y(z)\} = \{z : \mathbb{P}_y(c(x, a) > z) = \alpha\}$. Since X is a random variable, $c(x, a)$ is a random variable with realizations $c(i, a)$ for $i \in \{1, \dots, X\}$. Hence $z = c(i, a)$ for some $i \in \{1, 2, \dots, X\}$. ■

The result of Lemma 6 is similar to Proposition 8 in [57]. It was shown in [45] that $\eta_{y+1} \geq_r \eta_y$. Also, MLR dominance implies first order dominance, i.e, $\eta_{y+1} \geq_s \eta_y$.

Lemma 7: Let l and k be the indices such that

$$\underset{z \in \mathbb{R}}{\operatorname{argmin}} \{h_y(z)\} = c(l, a)$$

$$\underset{z \in \mathbb{R}}{\operatorname{argmin}} \{h_{y+1}(z)\} = c(k, a)$$

For all $y \in \{1, 2, \dots, Y\}, k \geq l$.

Proof: Proof is by contradiction. From Lemma 6, we have $F_y(c(l, a)) = 1 - \alpha$ and $F_{y+1}(c(k, a)) = 1 - \alpha$. Suppose $l > k$. We know that $F_{y+1}(z)$ is a monotone function in z . Since $l > k$, $F_{y+1}(c(l, a)) > 1 - \alpha$. But, by definition of first order stochastic dominance, $F_y(z) \geq F_{y+1}(z)$ for all z . Therefore, $F_y(c(l, a)) \geq F_{y+1}(c(l, a)) > 1 - \alpha$, a contradiction. ■

From Lemma 6 and equation (19), we have

$$H^\alpha(y, 2) = c(l, 2) + \frac{1}{\alpha} \sum_{i=1}^{l-1} \eta_y(i) (c(i, 2) - c(l, 2)),$$

$$H^\alpha(y+1, 2) = c(k, 2) + \frac{1}{\alpha} \sum_{i=1}^{k-1} \eta_{y+1}(i) (c(i, 2) - c(k, 2))$$

Lemma 8: $H^\alpha(y, 2) \geq H^\alpha(y+1, 2)$ if $\alpha \geq 1 - \mathbb{P}_y(x=X)$.

Proof: From the definitions of $H^\alpha(y, 2)$ and $H^\alpha(y+1, 2)$ we have,

$$\begin{aligned} H^\alpha(y, 2) - H^\alpha(y+1, 2) &= c(l, 2) - c(k, 2) \\ &+ \frac{1}{\alpha} \sum_{i=1}^{l-1} \eta_y(i) (c(i, 2) - c(l, 2)) + \frac{1}{\alpha} \sum_{i=1}^{k-1} \eta_{y+1}(i) (c(k, 2) - c(i, 2)) \\ &\geq c(l, 2) - c(k, 2) + \frac{1}{\alpha} \sum_{i=1}^{l-1} \eta_y(i) (c(i, 2) - c(l, 2)) \\ &+ \frac{1}{\alpha} \sum_{i=1}^{k-1} \eta_y(i) (c(k, 2) - c(i, 2)) \end{aligned} \quad (23)$$

Equation (23) follows from Lemma 3 and can be simplified as

$$\begin{aligned} H^\alpha(y, 2) - H^\alpha(y+1, 2) &\geq c(l, 2) - c(k, 2) + \\ &\frac{1}{\alpha} \sum_{i=1}^{l-1} \eta_y(i) (c(k, 2) - c(l, 2)) + \frac{1}{\alpha} \sum_{i=l}^{k-1} \eta_y(i) (c(k, 2) - c(i, 2)) \\ &\geq c(l, 2) - c(k, 2) - \frac{1}{\alpha} \Gamma' \eta_y \end{aligned}$$

where Γ is such that $\Gamma_i = c(l, 2) - c(k, 2)$ for $i = 1, \dots, l-1$ and $\Gamma_i = c(i, 2) - c(k, 2)$ for $i = l, \dots, k-1$. Clearly, $\Gamma_i \geq 0$ and decreasing. Right hand side of inequality attains its maximum when $k = X$ and $l = 1$ and $\Gamma_i = c(l, 2) - c(k, 2)$ for all i . Therefore, we have

$$\begin{aligned} H^\alpha(y, 2) - H^\alpha(y+1, 2) &\geq c(l, 2) - c(k, 2) - \frac{1}{\alpha} \Gamma' \eta_y \\ &\geq (c(l, 2) - c(k, 2)) - \frac{1}{\alpha} (c(l, 2) - c(k, 2)) (1 - \mathbb{P}_y(x = X)) \end{aligned}$$

After rearrangement we have,

$$\begin{aligned} H^\alpha(y, 2) - H^\alpha(y+1, 2) &\geq \\ &\frac{\alpha - (1 - \mathbb{P}_y(x = X))}{\alpha} (c(l, 2) - c(k, 2)) \end{aligned}$$

Since $\alpha \geq 1 - \mathbb{P}_y(x = X)$ and $(c(l, 2) - c(k, 2)) \geq 0$ (follows from Lemma 7 and assumption (A2)), we have $H^\alpha(y, 2) \geq H^\alpha(y + 1, 2)$. ■

From Lemma 6 and (19), we have

$$H^\alpha(y, 1) = c(l, 1) + \frac{1}{\alpha} \sum_{i=l+1}^X \eta_y(i)(c(i, 1) - c(l, 1)),$$

$$H^\alpha(y + 1, 1) = c(k, 1) + \frac{1}{\alpha} \sum_{i=k+1}^X \eta_{y+1}(i)(c(i, 1) - c(k, 1))$$

Lemma 9: $H^\alpha(y + 1, 1) \geq H^\alpha(y, 1)$ if $\alpha \geq 1 - \mathbb{P}_{y+1}(x = X)$.

Proof: From the definitions of $H^\alpha(y + 1, 1)$ and $H^\alpha(y, 1)$ we have,

$$\begin{aligned} H^\alpha(y + 1, 1) - H^\alpha(y, 1) &= c(k, 1) - c(l, 1) + \\ &\frac{1}{\alpha} \sum_{i=k+1}^X \eta_{y+1}(i)(c(i, 1) - c(k, 1)) \\ &- \frac{1}{\alpha} \sum_{i=l+1}^X \eta_y(i)(c(i, 1) - c(l, 1)) \\ &\geq c(k, 1) - c(l, 1) + \\ &\frac{1}{\alpha} \sum_{i=k+1}^X \eta_{y+1}(i)(c(i, 1) - c(k, 1)) \\ &- \frac{1}{\alpha} \sum_{i=l+1}^X \eta_{y+1}(i)(c(i, 1) - c(l, 1)) \end{aligned} \quad (24)$$

Equation (24) follows from Lemma 4 and can be simplified as

$$\begin{aligned} H^\alpha(y + 1, 1) - H^\alpha(y, 1) &\geq c(k, 1) - c(l, 1) + \\ &\frac{1}{\alpha} \sum_{i=k+1}^X \eta_{y+1}(i)(c(i, 1) - c(k, 1)) \\ &- \frac{1}{\alpha} \sum_{i=l+1}^X \eta_{y+1}(i)(c(i, 1) - c(l, 1)) \\ &\geq c(k, 1) - c(l, 1) - \frac{1}{\alpha} \Delta' \eta_{y+1} \end{aligned}$$

where Δ is such that $\Delta_i = c(i, 1) - c(l, 1)$ for $i = l, \dots, k$ and $\Delta_i = c(k, 1) - c(l, 1)$ for $i = k + 1, \dots, X$. Clearly, $\Delta_i \geq 0$ and decreasing. Right hand side of inequality attains its maximum when $k = X$ and $l = 1$ and $\Delta_i = c(k, 1) - c(l, 1)$ for all i . Therefore, we have

$$\begin{aligned} H^\alpha(y + 1, 1) - H^\alpha(y, 1) &\geq (c(k, 1) - c(l, 1)) \\ &- \frac{1}{\alpha} (c(k, 1) - c(l, 1))(1 - \mathbb{P}_{y+1}(x = X)) \end{aligned}$$

After rearrangement we have,

$$\begin{aligned} H^\alpha(y + 1, 1) - H^\alpha(y, 1) &\geq \frac{\alpha - (1 - \mathbb{P}_{y+1}(x = X))}{\alpha} \\ &\times (c(k, 1) - c(l, 1)) \end{aligned}$$

Since $\alpha \geq 1 - \mathbb{P}_{y+1}(x = X)$ and $c(k, 1) - c(l, 1) \geq 0$ (follows from Lemma 7 and assumption (A2)), we have $H^\alpha(y + 1, 1) \geq H^\alpha(y, 1)$. ■

Lemma 10: Let $\alpha \geq (1 - \mathbb{P}_y(x = X))$. The function $H^\alpha(y, a)$ satisfies the single crossing condition i.e.,

$$(H^\alpha(y, 1) - H^\alpha(y, 2)) \geq 0 \Rightarrow (H^\alpha(y + 1, 1) - H^\alpha(y + 1, 2)) \geq 0$$

Proof: Assume $(H^\alpha(y, 1) - H^\alpha(y, 2)) \geq 0$. We have,

$$\begin{aligned} H^\alpha(y, 1) - H^\alpha(y, 2) &\geq 0 \\ \Rightarrow H^\alpha(y, 1) - H^\alpha(y + 1, 2) &\geq 0 \end{aligned} \quad (25)$$

Equation (25) follows from Lemma 8. And,

$$\begin{aligned} H^\alpha(y, 1) - H^\alpha(y + 1, 2) &\geq 0 \\ \Rightarrow H^\alpha(y + 1, 1) - H^\alpha(y + 1, 2) &\geq 0 \end{aligned} \quad (26)$$

Equation (26) follows from Lemma 9. ■

Lemma 10 is a crucial result which helps us to prove Theorem 1 and Theorem 2.

Proof of Theorem 1: From Lemma 10, $H^\alpha(y, a)$ satisfies the single crossing condition and hence is sub-modular in (y, a) . Using Theorem 5, we get $a^*(\pi, y) = \operatorname{argmin} H^\alpha(y, a)$ is increasing in y .

Proof of Theorem 2: From Lemma 10, $H^\alpha(y, a)$ satisfies the single crossing condition. It is easily verified that the belief states satisfy the following property

$$\begin{aligned} \{\pi : H^\alpha(y, 1) - H^\alpha(y, 2) \geq 0\} &\subseteq \\ \{\pi : H^\alpha(y + 1, 1) - H^\alpha(y + 1, 2) \geq 0\} \end{aligned} \quad (27)$$

Equation (27) says that the curves $\{\pi : H^\alpha(y, 1) - H^\alpha(y, 2) = 0\}$ for all $y \in \mathcal{Y}$ do not intersect. Also from (18) and (27), it is easily verified that there are at most $Y + 1$ local decision likelihood matrices R^π (can be less than $Y + 1$ when $H^\alpha(\bar{y}, 1) - H^\alpha(\bar{y}, 2) > 0$ for some $\bar{y} \in \mathcal{Y}$, for all π). The matrices R^π from (18) and Theorem 1 are constant on each of the $Y + 1$ polytopes.

REFERENCES

- [1] B. LeBaron, "Agent-based computational finance," in *Handbook of Computational Economics*, 2006, vol. 2, pp. 1187–1233.
- [2] B. LeBaron, "Agent-based computational finance: Suggested readings and early research," *J. Econ. Dyn. Control*, vol. 24, no. 5, pp. 679–702, 2000.
- [3] E. Samanidou, E. Zschischang, D. Stauffer, and T. Lux, "Agent-based models of financial markets," *Rep. Prog. Phys.*, vol. 70, no. 3, p. 409, 2007.
- [4] V. Alfi, M. Cristelli, L. Pietronero, and A. Zaccaria, "Minimal agent based model for financial markets I," *Eur. Phys. J. B*, vol. 67, no. 3, pp. 385–397, 2009.
- [5] L. Tesfatsion and K. L. Judd, *Handbook of Computational Economics: Agent-Based Computational Economics*. Amsterdam, The Netherlands: Elsevier, 2006, vol. 2.
- [6] R. Cont and J.-P. Bouchaud, "Herd behavior and aggregate fluctuations in financial markets," *Macroeconomic Dyn.*, vol. 4, no. 2, pp. 170–196, 2000.
- [7] C. Avery and P. Zemsky, "Multidimensional uncertainty and herd behavior in financial markets," *Amer. Econ. Rev.*, vol. 88, no. 4, pp. 724–748, Sep., 1998.
- [8] A. Park and H. Sabourian, "Herding and contrarian behavior in financial markets," *Econometrica*, vol. 79, no. 4, pp. 973–1026, 2011.

- [9] C. Chamley, *Rational Herds: Economic Models of Social Learning*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [10] R. Cont, "Volatility clustering in financial markets: Empirical facts and agent-based models," in *Long Memory in Economics*. New York, NY, USA: Springer, 2007, pp. 289–309.
- [11] T. Lux and M. Marchesi, "Volatility clustering in financial markets: A microsimulation of interacting agents," *Int. J. Theor. Appl. Finance*, vol. 3, no. 4, pp. 675–702, 2000.
- [12] D. Challet et al., *Minority Games: Interacting Agents in Financial Markets*. London, U.K.: Oxford Univ. Press, 2013.
- [13] M. Cristelli, L. Pietronero, and A. Zaccaria, "Critical overview of agent-based models for economics," arXiv preprint arXiv:1101.1847, 2011.
- [14] N. Apergis and S. M. Miller, "Do structural oil-market shocks affect stock prices?" *Energy Econ.*, vol. 31, no. 4, pp. 569–575, 2009.
- [15] S. Gilchrist, V. Yankov, and E. Zakrajšek, "Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets," *J. Monetary Econ.*, vol. 56, no. 4, pp. 471–493, 2009.
- [16] S. Das and M. Magdon-Ismail, "Adapting to a market shock: Optimal sequential market-making," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 361–368.
- [17] V. Krishnamurthy, "Quickest detection POMDPs with social learning: Interaction of local and global decision makers," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5563–5587, Aug. 2012.
- [18] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Math. Finance*, vol. 9, no. 3, pp. 203–228, 1999.
- [19] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *J. Political Economy*, vol. 100, no. 5, pp. 992–1026, Oct. 1992.
- [20] I. Welch, "Sequential sales, learning, and cascades," *J. Finance*, vol. 47, no. 2, pp. 695–732, Jun. 1992.
- [21] A. V. Banerjee, "A simple model of herd behavior," *Quart. J. Econ.*, vol. 107, no. 3, pp. 797–817, Aug. 1992.
- [22] L. R. Glosten, "Insider trading, liquidity, and the role of the monopolist specialist," *J. Business*, vol. 62, no. 2, pp. 211–235, Apr. 1989.
- [23] A. N. Shiryaev and A. Aries, *Optimal Stopping Rules*. New York, NY, USA: Springer, 2007, vol. 8.
- [24] H. V. Poor and O. Hadjiladis, *Quickest Detection*. Cambridge, U.K.: Cambridge Univ. Press, 2009, vol. 40.
- [25] M. Frisén, "Optimal sequential surveillance for finance, public health, and other areas," *Sequential Anal.*, vol. 28, no. 3, pp. 310–337, 2009.
- [26] R. A. Cohn, W. G. Lewellen, R. C. Lease, and G. G. Schlarbaum, "Individual investor risk aversion and investment portfolio composition," *J. Finance*, vol. 30, no. 2, pp. 605–620, May 1975.
- [27] B. Donkers and A. V. Soest, "Subjective measures of household preferences and financial decisions," *J. Econ. Psychol.*, vol. 20, no. 6, pp. 613–642, 1999.
- [28] S. Mitra and T. Ji, "Risk measures in quantitative finance," *Int. J. Bus. Continuity Risk Manage.*, vol. 1, no. 2, pp. 125–135, 2010.
- [29] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *J. Risk*, vol. 2, pp. 21–41, 2000.
- [30] P. Krokmal, J. Palmquist, and S. Uryasev, "Portfolio optimization with conditional value-at-risk objective and constraints," *J. Risk*, vol. 4, no. 2, pp. 43–68, 2002.
- [31] C. Lim, H. D. Sherali, and S. Uryasev, "Portfolio optimization by minimizing conditional value-at-risk via nondifferentiable optimization," *Comput. Optim. Appl.*, vol. 46, no. 3, pp. 391–415, 2010.
- [32] F. Andersson, H. Mausser, D. Rosen, and S. Uryasev, "Credit risk optimization with conditional value-at-risk criterion," *Math. Program.*, vol. 89, no. 2, pp. 273–291, 2001.
- [33] Y. Feng, F. Rubio, and D. Palomar, "Optimal order execution for algorithmic trading: A CVaR approach," in *Proc. IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2012, pp. 480–484.
- [34] R. Forsythe, T. A. Rietz, and T. W. Ross, "Wishes, expectations and actions: a survey on price formation in election stock markets," *Journal of Economic Behavior & Organization*, vol. 39, no. 1, pp. 83–110, 1999.
- [35] C. R. Plott and K. Y. Chen, "Information aggregation mechanisms: Concept, design and implementation for a sales forecasting problem," *California Inst. Technol., Division of the Humanities and Social Sciences*, Working Papers 1131, 2002.
- [36] R. H. Thaler and W. T. Ziemba, "Parimutuel betting markets: Racetracks and lotteries," *J. Econ. Perspect.*, vol. 2, no. 2, pp. 161–174, 1988.
- [37] J. M. Gandar, W. H. Dare, C. R. Brown, and R. A. Zuber, "Informed traders and price variations in the betting market for professional basketball games," *J. Finance*, vol. 53, no. 1, pp. 385–401, Feb. 1998.
- [38] A. Cartea and S. Jaimungal, "Modelling asset prices for algorithmic and high-frequency trading," *Applied Mathematical Finance*, vol. 20, no. 6, pp. 512–547, 2013.
- [39] A. N. Akansu and M. U. Torun, *A Primer for Financial Engineering: Financial Signal Processing and Electronic Trading*. New York, NY, USA: Academic, 2015.
- [40] M. Avellaneda and S. Stoikov, "High-frequency trading in a limit order book," *Quant. Finance*, vol. 8, no. 3, pp. 217–224, 2008.
- [41] V. Krishnamurthy and A. Aryan, "Quickest detection of market shocks in agent based models of the order book," in *Proc. IEEE 51st Conf. Decision Control (CDC)*, Dec. 2012, pp. 1480–1485.
- [42] M. F. Neuts, *Structured Stochastic Matrices of MG-1 Type and Their Applications*. New York, NY, USA: Marcel Dekker, 1989.
- [43] V. Krishnamurthy, "Bayesian sequential detection with phase-distributed change time and nonlinear penalty—A POMDP lattice programming approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 7096–7124, Oct. 2011.
- [44] S. Karlin and Y. Rinott, "Classes of orderings of measures and related correlation inequalities: I. Multivariate totally positive distributions," *J. Multivariate Anal.*, vol. 10, no. 4, pp. 467–498, 1980.
- [45] W. S. Lovejoy, "Some monotonicity results for partially observed Markov decision processes," *Oper. Res.*, vol. 35, no. 5, pp. 736–743, 1987.
- [46] D. M. Topkis, *Supermodularity and Complementarity*. Princeton, NJ, USA: Princeton Univ. Press, 1998.
- [47] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA, USA: Athena Scientific, 1995, vol. 1.
- [48] D. M. Pennock, "A dynamic pari-mutuel market for hedging, wagering, and information aggregation," in *Proc. 5th ACM Conf. Electron. Commerce*, 2004, pp. 170–179.
- [49] Y. Chen, D. M. Pennock, and T. Kasturi, "An empirical study of dynamic pari-mutuel markets: Evidence from the tech buzz game," in *Proc. Web Mining Web Usage Anal. Workshop (WebDKK)*, Las Vegas, NV, USA, 2008.
- [50] L. Bursztyjn, F. Ederer, B. Ferman, and N. Yuchtman, "Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions," *Econometrica*, vol. 82, pp. 1273–1301, 2014.
- [51] D. A. Graham, "Estimating the state dependent utility function," *Nat. Resources J.*, vol. 23, p. 649, 1983.
- [52] W. N. Evans and W. K. Viscusi, "Estimation of state-dependent utility functions using survey data," *Rev. Econ. Statist.*, vol. 73, no. 1, pp. 94–104, 1991.
- [53] S. N. Afriat, "The construction of utility functions from expenditure data," *Int. Econ. Rev.*, vol. 8, no. 1, pp. 67–77, 1967.
- [54] H. R. Varian, *Intermediate Microeconomics: A Modern Approach*. New York, NY, USA: Norton, 2014.
- [55] G. Charness, U. Gneezy, and A. Imas, "Experimental methods: Eliciting risk preferences," *J. Econ. Beh. Organ.*, vol. 87, pp. 43–51, 2013.
- [56] A. Müller and D. Stoyan, *Comparison Methods for Stochastic Models and Risks*. Hoboken, NJ, USA: Wiley, 2002, vol. 389.
- [57] R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *J. Banking Finance*, vol. 26, no. 7, pp. 1443–1471, 2002.



Vikram Krishnamurthy (F'05) received the Ph.D. degree from the Australian National University, Canberra, ACT, Australia, in 1992. He is currently a Professor and holds the Canada Research Chair in statistical signal processing with the Department of Electrical Engineering, University of British Columbia, Vancouver, BC, Canada. His research interests include statistical signal processing and stochastic control with applications in social networks, and dynamical models for protein molecules in biosensing devices. He has served as a Distinguished Lecturer for the IEEE Signal Processing Society and the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING. He received a Honorary Doctorate from KTH (Royal Institute of Technology), Stockholm, Sweden, in 2013. He is author of the book: *Partially Observed Markov Decision Processes* (Cambridge University Press, 2016).



Sujay Bhatt received the M.Tech. degree from Indian Institute of Technology Bombay, Mumbai, India. He is currently working with Vikram Krishnamurthy at the University of British Columbia, Vancouver, BC, Canada. His research interests include social learning, Bayesian learning over networks, and game theory.