Abstract—We consider how local and global decision policies interact in quickest time change detection in multi-agent models of the order book. A monopolist market maker sets two-sided prices for an asset. The market evolves through the orders of trading agents. Agents observe local individual decisions of previous agents via an order book, combine these observed decisions with their noisy private signals about the asset, selfishly optimize their expected local utility, and then make their own individual decisions (whether to buy, sell or do nothing). Given this order book information, the goal is to achieve quickest change point detection when a shock occurs to the value of the asset. We provide a Bayesian formulation of the change point problem. Some structural results are given for the optimal policy.

I. INTRODUCTION

Market making has been studied extensively in [8], and more recently in a dynamical multi-period setting [6], [18]. [1]. Many markets employ a market maker to set prices. Exchanges such as NYSE use monopolist market makers to set prices and the practice has been shown to improve liquidity, particularly for less traded assets [8].

We are interested in an Agent Based Model (ABM) for the microstructure of asset prices. ABMs have been successfully used to capture the interaction of individual traders as well as their aggregate affect on the price discovery process. ABMs have been used to go beyond models such as random walks and capture empirical stylized facts observed in markets such as “fat tails”, correlation of returns and volatility clustering which implies long term memory [3].

In addition, we consider the role of Social Learning, or observational learning, as a key component in the dynamics of the market. The study of social learning in markets have lead to many interesting results such as herding and trend-following behavior [5], [15], [4], [17]. In models of social learning, the local decisions of agents are a combination of a noisy private signal and the history of all past actions. We assume that trading is performed by agents who adapt their behavior though social learning.

In this framework we are interested in the problem of quickest change point detection. Classical Bayesian quickest time detection involves detecting a geometrically distributed change time by optimizing the trade-off between false alarm frequency and delay penalty. Our setup is a generalization of this: rather than observations of the underlying state, given local decisions of trading agents that are performing social learning, how can a global decision maker achieve quickest time change detection to determine a shock in the asset value? The local decision determines the belief state which determines the global decision (stop or continue) which determines the local decision at the next time instant and so on. This interaction of local and global decision-making leads to discontinuous dynamics for the posterior probabilities (belief state) and unusual behavior as will be discussed. We will show that the optimal decision policy has multiple thresholds and the stopping regions are in general, non-convex.

Fig.1(a) gives a visual description of the optimal policy of social learning based quickest time change detection for geometric distributed change time, see [13] for details. The optimal policy \( \mu^*(\pi) \) is characterized by a triple threshold. The value function \( V(\pi) \) is non-concave and discontinuous.

This work was partially supported by NSERC.

V. Krishnamurthy and A. Aryan are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, V6T 1Z4, Canada. (email: vikramk@ece.ubc.ca).
from the trading pool and is allowed to place a buy, sell or no order with the market maker. We assume that no limit orders are placed. The trading pool of agents is assumed to be infinite. Agents are assumed to trade based on a noisy signal of the true value of the asset. We assume that the asset has a true underlying value \( Z \) which is known to all traders and the market maker at the start of the simulation. Once the simulation commences, the market maker does not receive direct information about \( Z \). It is only able to observe the public buy/sell local actions of agents. At a random time \( \tau^0 \), the asset experiences a jump change in its value to a value \( \lambda Z \) where \( \lambda > 0 \). We assume that agents receive noisy estimates of the new state through private observations and base their trading decision on this private information. The aim of the market maker is to optimally learn the value of the asset, and detect the change time (global decision) with minimal cost, through a social learning protocol based only on the observation agents’ past actions. 

A. The Multi-agent Social Learning Model

We consider a discrete time dealer market with a single asset. Consider a countable number of trading agents. Each agent acts once in a predetermined sequential order indexed by \( k = 1, 2, \ldots \). The index \( k \) can also be viewed as the discrete time instant when agent \( k \) acts. Let \( y_k \in \mathcal{Y} = \{1, 2, \ldots, Y\} \) denote the local (private) observation of agent \( k \) and \( a_k \in \mathcal{A} = \{1 \text{ (buy order)}, 2 \text{ (no trade)}, 3 \text{ (sell order)}\} \) denote the local decision agent \( k \) takes. Define the sigma algebras:

\[
\mathcal{H}_k \quad \sigma\text{-algebra generated by } (a_1, \ldots, a_{k-1}, y_k),
\]

\[
\mathcal{G}_k \quad \sigma\text{-algebra generated by } (a_1, \ldots, a_{k-1}, a_k). \tag{1}
\]

The social learning model [2], [4] comprises of the following ingredients:

1. **Dynamics of Asset Value:** The state \( x_k \) represents the underlying asset value that changes at time \( \tau^0 \). We model the change point \( \tau^0 \) by a phase type (PH) distribution. PH-distributions form a dense subset for the set of all distributions [16] and so can be used to approximate change times with arbitrary distribution. This is done by constructing a multi-state Markov chain as follows: Assume the underlying state \( x_k \) evolves as a Markov chain on the finite state space \( \mathcal{X} = \{1, \ldots, X\} \). Here state ‘1’ is an absorbing state and denotes the state after the jump change. The states 2, \ldots, \( X \) can be viewed as a single composite state that \( x \) resides in before the jump. These states can also be labelled with the unit indicator vectors \( e_1, e_2, \ldots, e_X \). The initial distribution is \( \pi_0 = (\pi_0(i), i \in \mathcal{X}) \), \( \pi_0(i) = P(x_0 = i) \). We are only interested in the case where the change occurs after a least one measurement, so assume \( \pi_0(1) = 0 \). So the transition probability matrix \( P \) is of the form

\[
P = \begin{bmatrix} 1 & 0 \\ \mathcal{P}_{X-1} & \mathcal{P}_{(X-1) \times (X-1)} \end{bmatrix} \tag{2}
\]

Let the “change time” \( \tau^0 \) denote the time at which \( x_k \) enters the absorbing state 1, i.e.,

\[
\tau^0 = \inf \{ k : x_k = 1 \}. \tag{3}
\]

At this random time \( \tau^0 \), the asset experiences a jump change in its value. The distribution of the change time \( \tau^0 \) is equivalent to the distribution of the absorption time to state 1 and is given by

\[
\nu_0 = \pi_0(1), \quad \nu_k = \pi_0^k \mathcal{P}^{k-1} \mathcal{P}, \quad k \geq 1 \tag{4}
\]

where \( \pi_0 = [\pi_0(2), \ldots, \pi_0(X)]' \). So by appropriately choosing the pair \((\pi_0, \mathcal{P})\) and state space dimension \( X \), one can approximate any given discrete distribution on \([0, \infty)\) by the distribution \(\{\nu_k, k \geq 0\}\); see [16, 240-243]. To ensure \( \tau^0 \) is finite, assume states 2, 3, \ldots, \( X \) are transient. For a 2-state Markov chain, the change time \( \tau^0 \) is geometric distributed.

Denote the 2-dimensional cost vector \( g \) associated with state 1 and states 2, \ldots, \( X \) as

\[
g = (\lambda Z, Z)' \quad \text{where } \lambda > 0
\]

Recall \( Z \) is the true underlying value of the asset. The constant \( \lambda \) models the relative size of the jump change in the asset value time \( \tau^0 \).

2. **Agent’s Private Observation:** Agent \( k \)'s local (private) observation \( y_k \in \mathcal{Y} = \{1, \ldots, Y\} \) is a noisy measurement of the true value of the asset. It is obtained from the observation likelihood distribution

\[
B_{xy} = P(y_k = y|x_k = x) \tag{5}
\]

The states 2, 3, \ldots, \( X \) are fictitious and are defined to generate the PH-distributed change time \( \tau^0 \). So states 2, 3, \ldots, \( X \) are indistinguishable in terms of the observation \( y \). That is, \( P(y|2) = P(y|3) = \cdots = P(y|X) \) for all \( y \in \mathcal{Y} \).

3. **Private belief:** Using local observation \( y_k \), trading agent \( k \) updates its private belief \( \pi_k^P \) defined as

\[
\pi_k^P = (\pi_k^P(i), i \in \mathcal{X}), \quad \pi_k^P(i) = P(x_k = i|a_1, \ldots, a_{k-1}, y_k) \tag{6}
\]

initialized by \( \pi_0 \). Thus the private belief is the posterior distribution of the underlying state given the past local decisions and current observation. It is computed by agent \( k \) according to the following Hidden Markov Model (HMM) filter: \( \pi_k^P = T(\pi_{k-1}, y_k) \).

\[
T(\pi, y) = \frac{B_{xy} P^\pi}{\sigma(\pi, y)} \quad \sigma(\pi, y) = 1' B_y P^\pi \tag{7}
\]

\[
B_y = \text{diag}(B_{1y}, \ldots, B_{Xy}) (X \times X \text{ diagonal matrix})
\]

Also \( \pi_{k-1} \) denotes the public belief available at time \( k-1 \) (defined in Step 5 below).

4. **Agent’s local decision:** Agent \( k \) then makes local decision \( a_k \in \mathcal{A} = \{1 \text{ (buy order)}, 2 \text{ (no trade)}, 3 \text{ (sell order)}\} \) to minimize myopically its expected cost of trading. To formulate this, let \( c(i, a) \) denote the non-negative cost incurred if the agent picks local decision \( a \) when the underlying state is \( x = i \). Denote the local decision \( X \)-dimensional cost

\[
c_a = \begin{bmatrix} c(1, a) & c(2, a) & \cdots & c(X, a) \end{bmatrix} \tag{8}
\]

Then agent \( k \) chooses local decision \( a_k \) greedily to minimize its expected cost:

\[
a_k = a(\pi_{k-1}, y_k) = \arg\min_{a \in \mathcal{A}} \mathbb{E}\{c(x, a)|\mathcal{H}_k\} \tag{9}
\]
In quickest change detection, since states 2, 3, . . . , X are indistinguishable in terms of observation y, we assume that \( c(2, a) = c(3, a) = \cdots = c(X, a) \) for each \( a \in \mathcal{A} \). Based on the posted bid/ask prices, the local cost vectors \( c_a \) for action \( a \in \{1, 2, 3\} \) are
\[
    c_1 = g - \bar{p}, \quad c_2 = I(p \leq g \leq \bar{p}), \quad c_3 = p - g \quad (10)
\]
Here \( \bar{p} \) and \( \bar{p} \) are the bid and ask price vectors set by the market maker as described below. Also \( I(\cdot) \) denotes the indicator vector elementwise.

At the beginning of the simulation we assume that the bid/ask prices bracket the intrinsic value, \( Z \in [\bar{p}, \bar{p}] \).

Furthermore, we assume that the jump size is strictly greater than the bid/ask spread, \( |1 - \lambda| > (\bar{p} - p)/Z \). If this condition is violated then action 2 dominates in all information states. Under these two assumptions, the costs of one action is dominated by the other two and can therefore be neglected. In the case where \( \lambda < 1 \), costs for action 1, \( c_1 \), are dominated by the cost vector of actions 2 and 3. When \( \lambda < 1 \), it is action 3 which is dominated. The associated cost matrix restricted to the two action reduction is sub-modular.

5. Market Making Mechanism: At each time period, the market maker sets bid and ask prices, \( \theta_k = (\bar{p}_k, \bar{p}_k) \), at which he will respectively buy and sell one unit of the asset. The bid/ask prices may be thought of as activation levels for the agent’s actions. The dealer sets prices based on its belief of the true value \( Z \). The knowledge of \( Z \) is driven only by the public information available, that is the past actions of agents. \( [9] \) suggests that the zero profit market maker should set prices according to \( \bar{p}_k = E[Z|a_{k-1} = \text{sell}] \) and \( \bar{p}_k = E[Z|a_{k-1} = \text{buy}] \). The key ingredient is the market maker’s belief of the asset value \( Z \) and the dynamics of this belief as agents act in the market, that is a mechanism to update \( \theta_{k+1} = M(\theta_k, a_k) \). In the case of change detection we assume that the time scale of the price update by the market maker is much slower than the time scale of arriving agents. This allows us to consider \( \theta \) as a constant over the time horizon of interest to change detection.

6. Social learning Public Belief: Agent \( k \) local decision \( a_k \) is recorded in the order book. Subsequent agents \( \tilde{k} > k \) use decision \( a_k \) to update their public belief of the underlying state \( x_k \) as follows: Define the public belief \( \pi_k \) as the posterior distribution of the state \( x \) given all local decisions taken up to time \( k \).

\[
    \pi_k = \mathbb{E}\{x_k|G_k\} = (\pi_k(i), i \in \mathcal{X}), \quad \pi_k(i) = P(x = i|a_1, \ldots, a_k) \quad (11)
\]
initialized by \( \pi_0 \). Then agents \( \tilde{k} > k \) update their public belief according to the following “social learning Bayesian filter”:
\[
    \pi_k = T^{\pi_k-1}(\pi_{k-1}, a_k), \quad \text{where } T^{\pi}(\pi, a) = \frac{R^\pi_a P^{\pi_a}}{\sigma(\pi, a)}, \quad \sigma(\pi, a) = 1_\mathcal{X} R^\pi_a P^{\pi_a} \quad (12)
\]
We use the notation \( T^{\pi}(\cdot) \) to point out that the above Bayesian update map depends explicitly on the belief state \( \pi \). This is a key difference compared to the HMM filter (7) where the Bayesian update map \( T(\cdot) \) does not depend explicitly on belief state \( \pi \). In (12), \( R^\pi_a \) denotes the diagonal matrix \( R^\pi_a = \text{diag}(R^\pi_{i,a}, i \in \mathcal{X}) \) where

\[
    R^\pi_{i,a} = P(a_k = a|x_k = i, \pi_{k-1} = \pi) \quad (13)
\]
denotes the conditional probability that agent \( k \) chose local decision \( a \) given state \( i \). We call \( R^\pi_{i,a} \) as the local decision likelihood probabilities in analogy to observation likelihood probabilities \( B_{iy} \) (5) in classical filtering.

Clearly observing the local decision \( a_k \) taken by agent \( k \) yields information about its local observation \( y_k \). That is, \( a_k \) serves as a surrogate observation of the underlying state \( x_k \).

The following lemma summarizes how subsequent agents use \( a_k \) to compute the local decision likelihood probabilities \( R^\pi_{i,a} \) in the social learning filter. The proof is omitted.

Lemma 1: The local decision likelihood probability matrix \( R^\pi \) in the social learning Bayesian filter (12) is computed as

\[
    R^\pi = BM^\pi \quad \text{where } \quad M^\pi = \begin{cases} P(a|y, \pi) = \prod_{a \in \mathcal{A} - \{a\}} I(c'_a B_y P'\pi < c'_a B_y P\pi), & (14) \end{cases}
\]

Here \( R^\pi \) is a \( \mathcal{X} \times \mathcal{A} \) matrix, \( B, B_y \) are the private observation probabilities defined in (5), (7), \( c_a, c_a \) are the local cost vectors defined in (8), and \( I(\cdot) \) denotes the indicator function.

The main implication of Lemma 1 is that the social learning Bayesian filter (12) is discontinuous in the belief state \( \pi \), due to the presence of indicator functions in (14). The likelihood probabilities \( R^\pi \) in (13) are an explicit function of the belief state \( \pi \) – this is stark contrast to the standard quickest detection problems where the observation distribution is not an explicit function of the posterior distribution.

Remark: The public belief belongs to the unit \( X - 1 \) dimensional simplex denoted as \( \Pi(X) \triangleq \{ \pi \in \mathbb{R}^X : 1^\pi_x \pi = 1, 0 \leq \pi(i) \leq 1 \text{ for all } i \in \mathcal{X} \} \).

B. Quickest Time Detection: Global Costs

With the above social learning based local decision framework, we now formulate the quickest time detection problem faced by the global decision maker. At each time \( k \), given the public belief \( \pi_k \), let \( u_k \) denote the global decision taken:

\[
    u_k = \mu(\pi_k) \in \{1 \text{ (announce change)}, 2 \text{ (continue)} \} \quad (15)
\]
Thus the global decision \( u_k \) is \( G_k \) measurable, where \( G_k \) is defined in (1). In (15), the policy \( \mu \) belongs to the class of stationary decision policies denoted \( \mu \). Below we formulate the costs incurred when taking these global decisions \( u_k \).

(i) Cost of announcing change and stopping: If global decision \( u_k = 1 \) is chosen, then the social learning protocol of Sec.II-A terminates. If \( u_k = 1 \) is chosen before the change point \( \tau^0 \), then a false alarm penalty is incurred. The false alarm event \( \cup_{k \geq 2} \{x_k = 1\} \cap \{u_k = 1\} = \{x_k \neq 1\} \cap \{u_k = 1\} \) represents the event that a change is announced before the change happens at time \( \tau^0 \). To evaluate the false alarm penalty, let \( f_i I(x_k = i, u_k = 1) \) denote the cost of a false
alarm in state $i$, $i \in X$, where $f_i \geq 0$. Of course, $f_1 = 0$ since a false alarm is only incurred if the stop action is picked in states $2, \ldots, X$. The expected false alarm penalty is

$$C(\pi_k, u_k = 1) = \sum_{i \in X} f_i \mathbb{E}\{I(x_k = i, u_k = 1) | \mathcal{G}_k\} = f^\prime \pi_k,$$

where $f^\prime = (f_1, \ldots, f_X)'$, $f_1 = 0$. \hfill (16)

The false alarm vector $f$ is chosen with increasing elements so that states further from state 1 incur larger penalties. (Obviously $f_i \geq 0$ since $f_1 = 0$).

(ii) Delay cost of continuing: If global decision $u_k = 2$ is taken then the social learning protocol of Sec.II-A continues to time $k + 1$. A delay cost is incurred when the event $\{x_k = 1, u_k = 2\}$ occurs, i.e., no change is declared at time $k$, even though the state has changed at time $k$. The expected delay cost is

$$C(\pi_k, u_k = 2) = d \mathbb{E}\{I(x_k = 1, u_k = 2) | \mathcal{G}_k\} = de^\prime \pi_k,$$

where $d > 0$ denotes the delay cost. \hfill (17)

C. Quickest Time Detection Objective

Let $(\Omega, \mathcal{F})$ be the underlying measurable space where $\Omega = (X \times \mathbb{U} \times \mathbb{Y})^\infty$ is the product space, which is endowed with the product topology and $\mathcal{F}$ is the corresponding product sigma-algebra. For any $\tau_0 \in \Pi(X)$, and policy $\mu \in \mu$, there exists a (unique) probability measure $\mathbb{P}_\tau_0^\mu$ on $(\Omega, \mathcal{F})$. Let $\mathbb{E}_\tau_0^\mu$ denote the expectation with respect to the measure $\mathbb{P}_\tau_0^\mu$.

Let $\tau$ denote a stopping time adapted to the sequence of $\sigma$-algebras $\mathcal{G}_k, k \geq 1$, see (1). That is, with $u_k$ determined by decision policy (15),

$$\tau = \{\inf k : u_k = 1\}. \hfill (18)$$

For each initial distribution $\tau_0 \in \Pi(X)$, and policy $\mu$, the following cost is associated:

$$J_\mu(\tau_0) = \mathbb{E}_\tau_0^\mu\{\sum_{k=1}^{\tau-1} \rho^{k-1} C(\pi_k, u_k = 2) + \rho^{\tau-1} C(\pi_\tau, u_\tau = 1)\}$$

where $\tau \geq 0$. \hfill (19)

Here $\rho \in [0, 1]$ denotes an economic discount factor. Since $C(\pi_1, \pi), C(\pi_2, \pi)$ are non-negative and bounded for all $\pi \in \Pi(X)$, stopping is guaranteed in finite time, i.e., $\tau$ is finite with probability 1 for any $\rho \in [0, 1]$ (including $\rho = 1$).

**Kolmogorov–Shiryaev criterion:** Suppose $X = \{1, 2\}$ implying that the change time $\tau^0$ is geometrically distributed. Choose the false alarm vector $f = (f_2, f_{2,2}) = (0, 1)'$ where $f_2$ is a positive constant, delay cost (17), and discount factor $\rho = 1$. Then the quickest time objective (19) assumes the classical Kolmogorov–Shiryaev criterion for detection of disorder:

$$J_\mu(\tau_0) = d \mathbb{E}_\tau_0^\mu\{(\tau - \tau_0)^+\} + f_2 \mathbb{P}_{\tau_0}^\mu(\tau < \tau_0). \hfill (20)$$

However, unlike classical quickest detection, the posterior (public belief) $\pi$ has discontinuous dynamics given by the social learning Bayesian filter (12). (Recall from (12), (14) that the dynamics of public belief $\pi$ depend on the local decision costs $c_a$).

The goal of the global decision maker is to determine the change time $\tau^0$ with minimal cost, that is, compute the optimal global decision policy $\mu^* \in \mu$ to minimize (19), where

$$J_{\mu^*}(\tau_0) = \inf_{\mu \in \mu} J_\mu(\tau_0).$$

D. Stochastic Dynamic Programming Formulation

The optimal stationary policy $\mu^* : \Pi(X) \rightarrow \{1, 2\}$ and associated value function $V(\pi)$ of the stopping time problem (19) are the solution of “Bellman’s dynamic programming equation”:

$$V(\pi) = \min\{\bar{C}(\pi, 1), \bar{C}(\pi, 2) + \rho \sum_{a \in A} V(T^\pi(\pi, a))\sigma(\pi, a)\}.$$

However, unlike classical quickest detection, the posterior (public belief) $\pi$ has discontinuous dynamics given by the social learning Bayesian filter (12). (Recall from (12), (14) that the dynamics of public belief $\pi$ depend on the local decision costs $c_a$).

Recalling the notation in Sec.II-A, we list the following assumptions.
The observation distribution \( B_{xy} = p(y|x) \) is TP2, i.e., all second order minors of matrix \( B \) are non-negative. (A2) The transition probability matrix \( P \) is TP2. (All second order minors of \( P \) are non-negative).

(A3) The elements of vector \( C \) in (21) are decreasing. A sufficient condition is that for \( j \geq i \) and \( i \geq 2 \) the false alarm vector \( f \) and delay penalty \( d \) satisfy \( f_i \geq \max \{ 1, \rho P^t e_i - d \} \) and \( f_j - f_i \geq \rho P^t (e_j - e_i) \).

(S) The local decision cost vector \( c_{ij} \) in (8) is submodular. That is, the elements \( c(i, a) \) satisfy \( c(1, 2) > c(1, 1) \) and \( c(2, 2) < c(2, 1) \). (Recall from Sec.II-A that \( c(2, a) = c(3, a) = \ldots = c(X, a) \) in quickest detection problems with \( P, A \) in quickest detection problems with \( P, \text{d}-\text{distributed change} \text{times} \).

Discussion of Assumptions:

Assumption (A1): The requirement that \( P(y|x) \) is TP2 with respect to states \( \{1, 2\} \) and \( y \in \mathcal{Y} \) holds for numerous examples [10]. Examples include quantized Gaussians, quantized exponential distributions, Binomial, Poisson, etc. For example consider quantized Gaussians. Suppose \( B_{iy} = P(y|x = i) = \frac{b_y}{\sum_{y \in Y} b_y} \), where \( b_y = \exp \left( -\frac{1}{2} \frac{(y - g_i)^2}{\Sigma} \right), \Sigma > 0 \), and \( g_1 < g_2 \). Then (A1) holds.

Assumption (A2) always holds trivially for \( X = 2 \). For \( X > 2 \), see [11] for numerous examples. Consider the tridiagonal transition probability matrix \( P \) with \( p_{ij} = 0 \) for \( j \geq i + 2 \) and \( j \leq i - 2 \). A necessary and sufficient condition for tridiagonal \( P \) to be TP2 is that \( p_{i+1,i+1} p_{i+1,i+1} \geq p_{i,i+1} p_{i+1,i} \). Such a diagonally dominant tridiagonal matrix satisfies Assumption (A2).

Assumption (A3) is a sufficient condition for \( C(\pi, 2) \) to be decreasing in \( \pi \) with respect to the monotone likelihood ratio order. We will use (A3) in Sec.IV to obtain sufficient conditions for a threshold policy. Assumption (A3) always holds for the geometric distributed change times (\( X = 2 \)). For PH-distributed change times (\( X > 2 \), Assumption (A3) can be viewed as design constraints the decision maker needs to take into account so that quickest detection with PH-distributed change times has a threshold policy [13]. Feasible values for the elements of \( f \) are straightforwardly obtained using a LP solver such as linprog in Matlab.

Assumption (S) is only required for the problem to be non-trivial. If (S) does not hold and \( c(i, 1) < c(i, 2) \) for \( i = 1, 2 \), then local decision \( a = 1 \) will always dominate decision \( a = 2 \) and the problem reduces to a standard quickest detection problem where the observed local decision \( a = 1 \) yields no information about the state. Assumption (S) implies \( c(x, 2) < c(x, 1) \) is decreasing in \( x \in \{1, 2\} \), i.e., the local cost \( c(x, a) \) is submodular which implies the zero crossing condition that is important in the proof of Theorem 1.

Theorem 1: Under (A1), (A2), (S), the belief state space \( \Pi(X) \) defined in (7?) can be partitioned into at most \( Y + 1 \) non-empty polytopes denoted \( \mathcal{P}_1, \ldots, \mathcal{P}_{Y+1} \) where

\[
\mathcal{P}_1 = \{ \pi \in \Pi(X) : (c_1 - c_2)'B_1 P^t \pi \geq 0 \} \tag{24}
\]

\[
\mathcal{P}_i = \{ \pi \in \Pi(X) : (c_1 - c_2)'B_{i-1} P^t \pi < 0 \} \cap \{ \pi \in \Pi(X) : (c_1 - c_2)'B_i P^t \pi \geq 0 \}, \quad i = 2, \ldots, Y \tag{25}
\]

\[
\mathcal{P}_{Y+1} = \{ \pi \in \Pi(X) : (c_1 - c_2)'B_Y P^t \pi < 0 \}
\]

On each such polytope, the local decision likelihood matrix \( R^\pi \) defined in (14) is a constant with respect to belief state \( \pi \).

The proof of the above theorem is in [13].

Define the following \( Y \) hyperplanes that are subsets of \( \Pi(X) \): For \( 1, 2 \in \mathcal{Y} \)

\[\eta_{1,2,y} = \{ \pi \in \Pi(X) : (c_1 - c_2)'B_y P^t \pi = 0 \}, \quad y = 1, \ldots, Y. \tag{26}\]

As a consequence of Theorem 1 and (14), there are only \( X + 1 \) possible decision likelihood matrices \( R^\pi \). one per polytope \( \mathcal{P}_i, i = 1, \ldots, Y + 1 \). We will denote these decision likelihood matrices as

\[R^i = R^\pi = BM^i = BM^\pi, \pi \in \mathcal{P}_i, i = 1, \ldots, Y + 1. \tag{27}\]

IV. Quickest Time Detection for Geometric and PH-distributed Change Time

Assumption (PH): Recall fictitious states \( 2, \ldots, X \) (corresponding to belief states \( e_2, \ldots, e_X \)) are used to model the PH-distribution in (4). It therefore makes sense to constrain the model parameters so that the global decision policy \( \mu^*(\pi) \) at the belief states \( e_2, \ldots, e_X \) are identical (and similarly for the local decisions taken in social learning). Throughout this section, when considering PH-distributed change times, we make the following assumption.

(PH)(i) \( C'e_i < 0 \) for \( i = 2, \ldots, Y \). (ii) \( e_2, \ldots, e_X \) lie in polytope \( \mathcal{P}_1 \).

Assumption (PH)(i) says that the optimal policy \( \mu^*(\pi) \) treats each of the fictitious states \( 2, \ldots, X \) identically – they all lie outside the stopping set \( \mathcal{S} \). In similar vein, (PH)(ii) requires that individual agents making local decisions treat the fictitious states \( i = 2, \ldots, X \) identically, i.e., they lie to the left of each hyperplane \( \eta_{ij}, y = 1, \ldots, Y \). Obviously, (PH) holds trivially for \( X = 2 \) (geometric case) - otherwise the quickest change problem would be degenerate.

A. Existence of a single threshold switching curve

Theorem 2 below shows that the stopping set is characterized by a single threshold curve on the belief space. The threshold coincides with the classical quickest time detection problem with non-informative observations.

1) Structural Result: We make the following assumptions. Recall the global decision maker’s cost vector \( C \) is defined in (21). Let \( \tilde{\eta}_j, j = 1, \ldots, X - 1 \) denote the \( X - 1 \) vertices of the intersection of hyperplane \( \eta_{ij} \) (defined in (26)) with \( \Pi(X) \).

\[
(c_1 - c_2)'B_Y P^t \tilde{\eta}_j \leq 0 \quad \text{for} \quad j = 1, \ldots, X - 1. \tag{C2}\]

\[
(c_1 - c_2)'B_Y P^t \pi < 0 \quad \text{for} \quad \pi \in \Pi(X) \backslash \{ \tilde{\eta}_j \}. \tag{C3}\]

The linear hyperplane \( \{ \pi : C'e_0 = 0 \} \) lies in polytope \( \mathcal{P}_{Y+1} \).

The following is the main result. The proof is in [13].

Theorem 2: Consider the social learning based quickest detection model \( (P, B, C, \rho) \). Assume (A1), (A2), (S) and (PH) hold. The optimal policy \( \mu^*(\pi) \) has the following structure

(i) Under (C3), \( \mu^*(\pi) = 2 \) for \( \pi \notin \mathcal{P}_{Y+1} \).

(ii) Under (C2) and (C3), the stopping set \( \mathcal{S} \) is as convex
subset of polytope $P_{Y+1}$. Therefore the boundary of $S$ is differentiable almost everywhere.

(iii) For geometric-distributed change time ($X = 2$), under (C3), the optimal policy is identical to that of the Kolmogorov–Shiryaev criterion (20) with uniformly distributed observation probabilities.

(iv) Under (A3), (C2), (C3), on the polytope $P_{Y+1}$, $\mu^*(\pi)$ has the following structure:

$$\pi_1, \pi_2 \in P_{Y+1}$$

and $\pi_1 \succeq \pi_2 \Rightarrow \mu^*(\pi_1) \geq \mu^*(\pi_2)$

(28)

(Here $\succeq$ denotes the monotone likelihood ratio order). Hence the boundary of the stopping set $S$ intersects any line segment $L(e_1, \bar{\pi})$ or $L(e_X, \bar{\pi})$ at most once (see geometric interpretation below).

2) \textbf{Discussion of Theorem 2 and assumptions:} Assumption (C3) localizes the decision threshold to polytope $P_{Y+1}$. As a consequence of (C3), $C^*\pi < 0$ on all polytopes except $P_{Y+1}$. Therefore on these polytopes, $\mu^*(\pi) = 2$. Thus statement (i) is obvious.

Assumption (C2) together with (A1), (A2), (S) and (PH) ensures that the polytope $P_{Y+1}$ is closed under the belief state mapping $T^\pi(\pi, a)$. That is, $\pi \in P_{Y+1}$ implies $T^\pi(\pi, a) \in P_{Y+1}$ for all $a$. Note that Assumption (C2) holds trivially for $X = 2$.

Assumptions (C2) and (C3) allow us to show that the value function $V(\pi)$ is concave on $P_{Y+1}$. Then Statement (ii), namely convexity of the stopping set $S$, follows from arguments in [14].

Statement (iii) is straightforward to show. The local decision likelihood probabilities on $P_{Y+1}$ are uniform since the local decision yields no information about the state. Thus under (C3) the threshold is identical to the classical quickest detection threshold for the Kolmogorov–Shiryaev criterion (20) with uniformly distributed observation probabilities.

The proof of Statement (iv) is given in [13].

V. Numerical Results

In this section we provide simulation results of the algorithm. Figure 2 shows the sample path of a simulated realization of the model. We can observe the change in the underlying state, the sequence of observations, agent actions and the filtered belief state. In this example, the underlying price starts at $Z = 2$ and at a random time the value drops to $Z = 0$. In the example below, the change time is $\tau^0 = 25$. The following parameters were used in the simulation: $p = .99$, $p^b = 1.5$, $p^a = 2.1$, $\Sigma = 1$, $f = 0.25$, $d = 0.1$ and $p = 0.75$.

Remark: For discussion, complete proofs and numerical studies of social learning based quickest detection see [12], [13].

REFERENCES