QUICKEST TIME CHANGE DETECTION WITH SOCIAL LEARNING

Vikram Krishnamurthy

Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, V6T 1Z4, Canada. (email: vikramk@ece.ubc.ca).

ABSTRACT

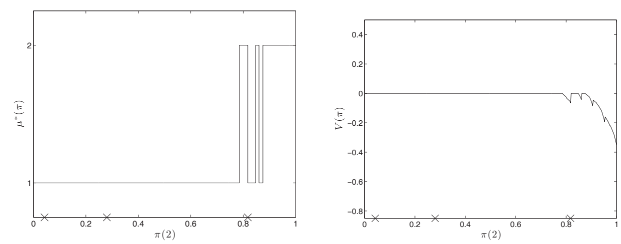
How does local and global decision making interact in detection theory? This paper considers multi-agent quickest time change detection with social learning. We show that the optimal decision exhibits a remarkable multi-threshold behavior within the space of Bayesian distributions. For small change probabilities, an explicit characterization of this behavior is obtained in terms of fixed points of the posterior update.

1. INTRODUCTION

Bayesian quickest time detection [1] involves detecting a geometrically distributed change time by optimizing the trade-off between false alarm frequency and delay penalty [2, 1]. This paper considers multi-agent quickest detection. *Given local decisions from agents performing social learning [3], how can a global decision maker achieve quickest change detection?* That is, each agent chooses its local decision by optimizing a local utility function (which depends on the public belief of the state and its local observation). Instead of revealing its posterior distribution of change, each agent reveals its local decision to subsequent agents. Subsequent agents update their public belief based on these local decisions (in a Bayesian setting), and the sequential procedure continues. How can such a multi-agent system detect a change in the underlying state and make a global decision to stop?

Classical quickest detection is a trivial case where agents reveal their local observation (instead of local decision) to subsequent agents. In quickest time change detection with social learning, 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 unusual behavior as outlined below.

Fig.1(a) gives a visual description of the optimal policy of quickest detection with social learning. It illustrates a *triple threshold policy* for geometric distributed change time. Complete details of this numerical example are given in Sec.4. The horizontal axis $\pi(2)$ is the posterior probability of no change. The vertical axis denotes the optimal decision: $u = 1$ denotes stop and declare change, while $u = 2$ denotes continue. The multi-threshold behavior of Fig.1(a) is unusual: *if it is optimal to declare a change for a particular posterior probability, it may not be optimal to declare a change when the posterior probability of a change is larger!* Fig.1(b) shows the associated non-concave value function obtained via stochastic dynamic programming. Fig.1 shows that social learning results in fundamentally different decision policies compared to classical quickest time detection (which has a single threshold).



(a) Optimal global decision policy $\mu^*(\pi)$

(b) Value function $V(\pi)$ for global decision policy

Fig. 1. Optimal decision policy for quickest time change detection based on social learning for geometric distributed change time, see Example 2 of Sec.4 for details. The optimal policy $\mu^*(\pi)$ is characterized by a triple threshold. The value function $V(\pi)$ is non-concave and discontinuous.

Related Works : In the last decade, social learning has been studied widely in economics to model the behavior of financial markets, crashes and booms, crowds and social networks, see [4, 3] and numerous references therein. The social learning setup is similar to Hellman's and Cover's 1970s paper [5]. Global decision making with social learning has recently been studied by several economists [3]. We address a related problem: if agents make (simple) decisions by optimizing a local utility, how can the global system achieve change detection?

Main Results: Sec.2 presents the multi-agent social learning protocol. The quickest time detection problem is formulated and the optimal stopping policy is characterized in terms of stochastic dynamic programming. The main result of Sec.3 is to characterize quickest time change detection policies when the probability of change, denoted ϵ, is small; see also [6]. Theorem 2 characterizes the multi-threshold structure of the optimal decision policy.

2. SOCIAL LEARNING PROTOCOL

Consider a countably infinite number of agents performing social learning to estimate an underlying state process x. (An identical setup holds if a finite number of agents are polled repeatedly in some pre-defined order, if agents pick local decisions based on the most recent public belief [7].) Each agent acts once in a sequential order indexed by $k = 1, 2, \ldots$.

Let $y_k \in \mathbb{Y} = \{1, 2, \ldots, Y\}$ denote the local (private) observation of agent k and $a_k \in \mathbb{A} = \{1, 2\}$ denote the local decision agent k takes. Define

$$\begin{aligned} \mathcal{H}_k &\quad \sigma\text{-algebra generated by } (a_1, \ldots, a_{k-1}, y_k), \\ \mathcal{G}_k &\quad \sigma\text{-algebra generated by } (a_1, \ldots, a_{k-1}, a_k). \end{aligned} \quad (1)$$

The social learning model [8, 3] comprises:

1. *Absorbing-state Markov chain*: We model the change point τ^0 by a geometric distribution by a two state Markov chain x_k on state space $\mathbb{X} = \{1, 2\}$. Here state '1' is an absorbing state and denotes the state after the jump change. The states 2 is state that x resides in before the jump. The transition probability matrix is $P = \begin{bmatrix} 1 & 0 \\ 1 - P_{22} & P_{22} \end{bmatrix}$. Let the "change time" τ^0 denote the time at which x_k enters the absorbing state 1, i.e., $\tau^0 = \inf\{k : x_k = 1\}$.

2. *Local Observation*: Agent's k local observation $y_k \in \mathbb{Y} = \{1, \ldots, Y\}$ is obtained from the observation likelihood distribution $B_{xy} = P(y_k = y | x_k = x)$.

3. Private belief: Using local observation y_k, agent k updates its private belief $\pi_k^P = (\pi_k^P(i),\ i \in \mathbb{X})$ where

$$\pi_k^P(i) = P(x_k = i | a_1, \ldots, a_{k-1}, y_k). \quad (2)$$

Thus the private belief is the posterior distribution of the underlying state given the past actions and current observation. It is computed by agent k as $\pi_k^P = T(\pi_{k-1}, y_k)$ where

$$T(\pi, y) = \frac{B_y P' \pi}{\sigma(\pi, y)},\ \sigma(\pi, y) = \mathbf{1}' B_y P' \pi. \quad (3)$$

Here $B_y = \text{diag}(B_{1y}, B_{2y})$ is a diagonal matrix for each $y \in \mathbb{Y}$. π_{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 $a_k \in \mathbb{A} = \{1, 2\}$ to minimize myopically its expected cost. Let $c(i, a)$ denote the cost incurred if the agent picks local decision a when the underlying state is $x = i$. Denote $c_a = [c(1, a),\ c(2, a)]'$. Then agent k chooses local decision a_k greedily to minimize its expected cost:

$$a_k = a(\pi_{k-1}, y_k) = \arg\min_{a \in \mathbb{A}} \mathbb{E}\{c(x, a) | \mathcal{H}_k\} = \arg\min_{a \in \mathbb{A}} \{c_a' \pi_k^P\}.$$

5. *Social learning Public Belief*: Finally agent k broadcasts its local decision a_k. Subsequent agents $\bar{k} > k$ use decision a_k to update their public belief of the underlying state x_k as follows: Define the public belief π_k as the posterior distribution of the state x given all actions taken up to time k.

$$\pi_k = \mathbb{E}\{x_k | \mathcal{G}_k\} = (\pi_k(i),\ i \in \mathbb{X}), \quad \pi_k(i) = P(x = i | a_1, \ldots a_k).$$

Then agents $\bar{k} > k$ update their public belief according to the following "social learning Bayesian filter": $\pi_k = T^{\pi_{k-1}}(\pi_{k-1}, a_k)$, where

$$T^\pi(\pi, a) = \frac{R_a^\pi P' \pi}{\sigma(\pi, a)},\ \sigma(\pi, a) = \mathbf{1}_X' R_a^\pi P' \pi \quad (4)$$

We use the notation $T^\pi(\cdot)$ to point out that the above Bayesian update map depends explicitly on the belief state π. This is a key difference compared to classical quickest detection (3) where the Bayesian update map $T(\cdot)$ does not change with belief state π. In (4), R_a^π denotes the diagonal matrix $R_a^\pi = \text{diag}(R_{i,a}^\pi,\ i \in \mathbb{X})$ where

$$R_{i,a}^\pi = P(a_k = a | x_k = i, \pi_{k-1} = \pi) \quad (5)$$

denotes the conditional probability that agent k chose action a given state i. We call $R_{i,a}^\pi$ as the *local decision likelihood probabilities* in analogy to observation likelihood probabilities B_{iy} in classical detection.

The local decision likelihood probability matrix R^π in the social learning filter (4) is computed as $R^\pi = BM^\pi$ where $M_{y,a}^\pi \triangleq P(a|y, \pi) = \prod_{\tilde{a} \in \mathbb{A} - \{a\}} I(c_a' B_y P' \pi < c_{\tilde{a}}' B_y P' \pi)$. Here R^π is a $\mathbb{Y} \times \mathbb{A}$ matrix, B_y is defined in (3) and $I(\cdot)$ denotes the indicator function. The likelihood probabilities R^π in (5) are an explicit and discontinuous function of the belief state π – this is stark contrast to the standard quickest detection problems where the observation distribution is not an explicit function of the posterior distribution.

Global Costs: At each time k, given the public belief π_k, let $u_k = \mu(\pi_k) \in \{1 \text{ (announce change and stop)}, 2 \text{ (continue)}\}$ denote the global decision. 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 terminates. If $u_k = 1$ is chosen before the change point τ^0, then a false alarm penalty is incurred. The false alarm event $\{x_k = 2\} \cap \{u_k = 1\} = \{x_k \neq 1\} \cap \{u_k = 1\}$ represents the event that a change is announced before the change happens at time τ^0. With $f > 0$, the expected false alarm penalty is

$$\bar{C}(\pi_k, u_k = 1) = f \mathbb{E}\{I(x_k = 2, u_k = 1) | \mathcal{G}_k\} = [0\ f]\, \pi_k.$$

(ii) *Delay cost of continuing*: If global decision $u_k = 2$ is taken then the social learning protocol continues to time $k+1$. A delay cost is incurred when the event $\{x_k = e_1, u_k = 2\}$ occurs, i.e., no change is declared at time k, even though the state has changed at time k. The expected delay cost is

$$\bar{C}(\pi_k, u_k = 2) = d\, \mathbb{E}\{I(x_k = e_1, u_k = 2) | \mathcal{G}_k\} = d e_1' \pi_k$$

where $d > 0$ denotes the delay cost.

Quickest Time Detection Objective: Define $\tau = \{\inf k : u_k = 1\}$. For each initial distribution $\pi_0 \in \Pi(X)$, and policy μ, the following cost is associated:

$$J_\mu(\pi_0) = \mathbb{E}_{\pi_0}^\mu \{\sum_{k=1}^{\tau-1} \rho^{k-1} \bar{C}(\pi_k, u_k = 2) + \rho^{\tau-1} \bar{C}(\pi_\tau, u_\tau = 1)\}. \quad (6)$$

Here $\rho \in [0, 1]$ denotes an economic discount factor. If $\rho = 1$, we obtain the classical Kolmogorov–Shiryaev criterion for detection of disorder [1] is

$$J_\mu(\pi_0) = d \mathbb{E}_{\pi_0}^\mu \{(\tau - \tau^0)^+\} + f\, \mathbb{P}_{\pi_0}^\mu (\tau < \tau^0). \quad (7)$$

Unlike classical quickest detection, the posterior π has discontinuous dynamics given by the social learning filter (4).

The goal is to determine the change time τ^0 with minimal cost, that is, compute the optimal policy $\mu^* \in \boldsymbol{\mu}$ to minimize (6), where $J_{\mu^*}(\pi_0) = \inf_{\mu \in \boldsymbol{\mu}} J_\mu(\pi_0)$. Considering the above cost (6), the optimal stationary policy $\mu^* : \Pi(X) \to \{1,2\}$ and associated value function $\bar{V}(\pi)$ are the solution of the following "Bellman's dynamic programming equation"

$$C(\pi,1) = 0, \quad C(\pi,2) = \bar{C}(\pi,2) - \mathbf{f}'\pi + \rho \mathbf{f}' P' \pi \tag{8}$$

$$\mu^*(\pi) = \arg\min_{u \in \mathbb{U}} Q(\pi,u), \; J_{\mu^*}(\pi_0) = V(\pi_0) \tag{9}$$

$$V(\pi) = \min_{u \in \{1,2\}} Q(\pi,u), \text{ where } Q(\pi,1) = C(\pi,1) = 0$$

$$Q(\pi,2) = C(\pi,2) + \rho \sum_{a \in \mathbb{A}} V\left(T^\pi(\pi,a)\right) \sigma(\pi,a).$$

3. QUICKEST DETECTION WITH SMALL CHANGE PROBABILITIES

Here we consider quickest time change detection when the underlying state has a small probability of change. Theorem 2 and Corollary 1 show that the quickest-time detection policy for change probability ϵ, yields a cost that is within $O(\epsilon)$ of the optimal cost for sequential detection of a constant state.

3.1. Polytope Structure and Main Assumptions

Although in general there are 2^Y possible R^π matrices, we now show that by introducing assumptions (A1), (A2) and (S) below, there are only $Y+1$ distinct R^π matrices.

We list the following assumptions.

(A1) The observation distribution $B_{xy} = p(y|x)$ is TP2. (All second order minors of matrix B are non-negative).

(S) The local costs $c(i,a)$ incurred by individual agents satisfy $c(1,2) > c(1,1)$ and $c(2,2) < c(2,1)$.

Assumption (A1) holds for numerous examples, see [9]. Examples include quantized Gaussians, quantized exponential distributions, Binomial, Poisson, etc. *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 $a = 2$.

Theorem 1. *Under (A1), (A2), (S), the belief space can be partitioned into $Y+1$ intervals denoted $\mathcal{P}_1, \ldots, \mathcal{P}_{Y+1}$ where*

$$\begin{aligned} \mathcal{P}_1 &= \{\pi \in \Pi(X) : (c_1 - c_2)' B_1 P' \pi > 0\} \\ \mathcal{P}_l &= \{\pi \in \Pi(X) : (c_1 - c_2)' B_{l-1} P' \pi < 0 \\ &\quad \cap \; (c_1 - c_2)' B_l P' \pi > 0\}, l = 2, \ldots, Y \\ \mathcal{P}_{Y+1} &= \{\pi \in \Pi(X) : (c_1 - c_2)' B_Y P' \pi < 0\} \end{aligned} \tag{10}$$

On each such interval $\mathcal{P}_l$, R^π, $\pi \in \mathcal{P}_l$ is a constant. ■

As an immediate consequence of Theorem 1, on each interval we will denote

$$R^l = R^\pi = BM^l = BM^\pi, \pi \in \mathcal{P}_l, l = 1, \ldots, Y+1 \tag{11}$$

$$\eta_y = \{\pi \in \Pi(X) : (c_1 - c_2)' B_y P' \pi = 0\}, y = 1, \ldots, Y.$$

3.2. Quickest Detection with Small Probability of Change

Here we consider quickest detection with social learning for the following special case: $\mathbb{X} = \mathbb{Y} = \mathbb{A} = \{1,2\}$, $P = \begin{bmatrix} 1 & 0 \\ \epsilon & 1-\epsilon \end{bmatrix}$. Here the change probability $\epsilon \ll 1$ is a small non-negative scalar. The analysis in this subsection proceeds as follows:

Step 1: For $\epsilon = 0$, the problem is a simple sequential detection problem for state 1. We characterize the multi-threshold behavior of the optimal decision policy in Theorem 2 below.

Step 2: It is then shown that for small ϵ, the optimal value function is within $O(\epsilon)$ of the value function for the case of zero change probability (Corollary 1). So, the optimal policy computed for zero change probability yields performance close to that of the optimal quickest detection policy.

Step 1: Sequential Detection of Static State: In line with above plan, consider the sequential detection problem for state 1 with social learning with $X = \mathbb{Y} = \mathbb{A} = \{1,2\}$, $P = I$. The state x is a random variable chosen at $k = 0$ with distribution π_0 and remains constant for $k > 0$. The goal is to detect and announce state 1 if $x_0 = 1$ based on noisy observations. The global decision $u_k = \mu(\pi_k) \in \{1 \text{ (stop)}, 2 \text{ (continue)}\}$ is a function of the public belief π_k.

The 2-dimensional belief state $\pi = [1 - \pi(2), \pi(2)]$ is parametrized by the scalar $\pi(2) \in [0,1]$, i.e., $\Pi(X)$ is the interval $[0,1]$. Let $[1 - \eta_y(2), \eta_y(2)]$ denote the belief state corresponding to η_y. If (A1) and (S) hold, then $\mathcal{P}_3 = (0, \eta_2(2))$, $\mathcal{P}_2 = (\eta_2(2), \eta_1(2))$, $\mathcal{P}_1 = (\eta_1(2), 1)$.

Define $q = T^{\eta_1}(\eta_1, 1)$. The following lemma that characterizes useful structural properties of the social learning filter.

Lemma 1. *Consider the social learning filter (4). Assume (A1), (S) hold. If B is symmetric, then η_1 and η_2 are fixed points of the composite Bayesian map: $\eta_1 = T^q(T^{\eta_1}(\eta_1, 1), 2)$, $\quad \eta_2 = T^q((T^{\eta_2}(\eta_2, 2), 1)$*

The implication of the above lemma is that $\Pi(X)$ can be partitioned into 4 intervals, namely $[e_1, \eta_2)$, $[\eta_2, q)$, $[q, \eta_1)$ and $[\eta_1, e_2]$. The main result below characterizes the multi-threshold global decision policy $\mu^*(\pi)$ on these 4 intervals.

Theorem 2. *Consider the sequential detection problem for $\epsilon = 0$. Suppose agents make local decisions via social learning. Assume (A1), (S) hold. The optimal global decision policy $\mu^*(\pi)$ has the following properties:*

(i) For $\pi \in \mathcal{P}_1 \cup \mathcal{P}_3$, $\mu*$ has a threshold structure:

$$\mu^*(\pi) = \begin{cases} 2 & \text{if } \pi(2) > \pi^*(2) \\ 1 & \text{otherwise} \end{cases} \quad \text{where } \pi^*(2) = \frac{d}{f(1-\rho) + d}$$

Also for $\pi \in \mathcal{P}_1 \cup \mathcal{P}_3$, the value function is $V(\pi) = \min\{0, C(\pi,2)/(1-\rho)\}$.

(ii) If B is symmetric, then for $\pi \in \mathcal{P}_2$, the global decision policy has the following structure:

(a) For $\pi \in [\eta_2(2), q(2)]$, $V(\pi)$ is concave and there is at most one interval where $\mu^*(\pi) = 1$.

(b) For $\pi \in [q(2), \eta_1(2)]$, $V(\pi)$ is concave and there is at most one interval where $\mu^*(\pi) = 1$. ■

The implication of the above theorem is that the optimal policy has up to three thresholds.

Step 2: Quickest Time Detection bound for small ϵ: Given the characterization in Theorem 2 for $\epsilon = 0$, we now consider the quickest change detection problem for small ϵ. Let $V_{\mu_\epsilon^*}(\pi)$ denote the cost incurred by the optimal policy μ_ϵ^* with transition probabilities $P_{22}^\epsilon = 1 - \epsilon$. Denote the three intervals $\mathcal{P}_l^\epsilon$, $l = 1, 2, 3$, defined in (10). The following result bounds $|V_{\mu_0^*}(\pi) - V_{\mu_\epsilon^*}(\pi)|$. Note that $\mu_0^*(\pi)$ is characterized in Theorem 2 and $P^0 = I$.

Corollary 1. *Consider the social learning quickest detection problem with small probability of change. Then, for initial belief $\pi \in \mathcal{P}_l^\epsilon \cap \mathcal{P}_l^0$, $l = 1, 2, 3$, the optimal policy μ_0^* (characterized in Theorem (2)) has a total cost $V_{\mu_0^*}(\pi)$ that constitutes an $O(\epsilon)$ upper-bound to the optimal cost $\bar{V}_{\mu_\epsilon^*}(\pi)$ incurred in the quickest detection problem. More specifically,*

$$\bar{V}_{\mu_0^*}(\pi) - \bar{V}_{\mu_\epsilon^*}(\pi) \leq \frac{4\rho\epsilon}{(1-\rho)^2} \max(d, f). \blacksquare \qquad (12)$$

Discussion: The implication of (12) is that the simple policy $\mu_0^*(\pi)$ of Theorem 2 is near optimal for quickest time detection with social learning when ϵ is small. The above bound is tight in the sense that for $\epsilon = 0$, the optimal costs $\bar{V}_{\mu_0^*}(\pi)$ and $\bar{V}_{\mu_\epsilon^*}(\pi)$ coincide. The proof of Corollary 1 follows from Theorem 2 of [10]. Theorem 2 of [10] shows that

$$\bar{V}_{\mu_0^*}(\pi) \leq \bar{V}_{\mu_\epsilon^*}(\pi) + \frac{2\rho}{(1-\rho)^2} \|\bar{C}(\pi, u)\|_\infty \sup_i \|[P^\epsilon - P^0]_{ij}\, R_a^l\|_1 \qquad (13)$$

where the $\|\cdot\|_1$ induced matrix norm is with respect to the (j, a) elements. Since from Theorem 2, the value function is piecewise linear, (13) applies. Clearly

$$\sup_i \|[P^\epsilon - P^0]_{ij}\, R_a^l\|_1 = \epsilon \max(B_{11} + B_{21},\, B_{12} + B_{22}) \leq 2\epsilon.$$

Substituting $\|\bar{C}(\pi, u)\|_\infty = \max(d, f)$ in (13) yields (12).

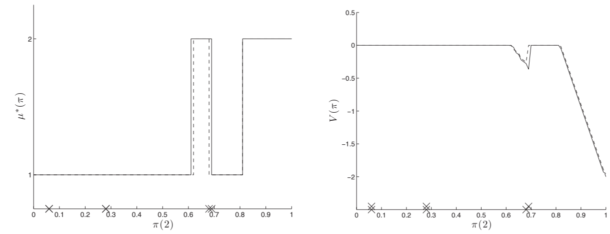
4. NUMERICAL RESULTS

Example 1: Consider $\mathbb{X} = \mathbb{Y} = \mathbb{A} = \{1, 2\}$, $\rho = 0.8$, $d = 1.8$, $f = 2$, $B = \begin{bmatrix} 0.85 & 0.15 \\ 0.15 & 0.85 \end{bmatrix}$, $c = \begin{bmatrix} 1 & 2 \\ -1 & -3.57 \end{bmatrix}$.

Fig.2 shows the optimal policies μ_0^* (Theorem 2) and μ_ϵ^* (optimal quickest detection policy) together with optimal costs $V_{\mu_0^*}(\pi)$ and $V_{\mu_\epsilon^*}(\pi)$ for change probability $\epsilon = 0.005$. As can be seen the quickest detection optimal policy and costs are very close to that specified by Theorem 2.

Example 2: Here we illustrate the multiple threshold policies inherent in social learning (this example was mentioned in Sec.1). We chose the social learning model with parameters $\mathbb{X} = \{1, 2\}$, $\mathbb{Y} = \{1, 2, 3\}$, $\mathbb{A} = \{1, 2\}$, $B = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$, $P = \begin{bmatrix} 1 & 0 \\ 0.05 & 0.95 \end{bmatrix}$, $c = \begin{bmatrix} 1 & 2 \\ -1 & -3.57 \end{bmatrix}$. For the global quickest time detection parameters we chose $\rho = 0.99$, delay $d = 1.25$, false alarm vector $\mathbf{f} = 3e_2$ (i.e., $f = 3$). It is easily checked that (A1) and (S) hold.

The optimal policy $\mu^*(\pi)$ is shown in Fig.1(a) and comprises of a triple threshold policy. The 'x' in Fig.1(a) and (b) are the values of $\eta_2(2)$, $q(2)$ and $\eta_1(2)$, respectively.



(a) Optimal global decision policies $\mu_0^*(\pi)$ and $\mu_\epsilon^*(\pi)$

(b) Value functions for global decision policy

Fig. 2. Optimal decision policy for quickest time change detection with social learning and small probability of change. The policies and optimal costs for $\epsilon = 0.005$ (solid line) are very close to $\epsilon = 0$ (broken line).

5. CONCLUSIONS

Motivated by understanding how local and global decision makers interact, this paper analyses quickest time detection when agents perform social learning. It was shown that the optimal global decision policy has a remarkable multi-threshold behavior. The optimal threshold policy was characterized explicitly in Theorem 2 and Corollary 1. Numerical examples of the multi-threshold behavior were presented.

6. REFERENCES

[1] A.N. Shiryayev, *Optimal stopping rules*, Springer-Verlag, 1978.

[2] H.V. Poor and O. Hadjiliadis, *Quickest Detection*, Cambridge, 2008.

[3] C. Chamley, *Rational herds: Economic models of social learning*, Cambridge, 2004.

[4] D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi, "Cascades in networks and aggregate volatility," Tech. Rep., National Bureau of Economic Research, 2010.

[5] T. Cover and M. Hellman, "The two-armed-bandit problem with time-invariant finite memory," *Information Theory, IEEE Transactions on*, vol. 16, no. 2, pp. 185–195, 1970.

[6] A.G. Tartakovsky and V.V. Veeravalli, "General asymptotic Bayesian theory of quickest change detection," *Theory of Probability and its Applications*, vol. 49, no. 3, pp. 458–497, 2005.

[7] V. Krishnamurthy, "Bayesian sequential detection with phase-distributed change time and nonlinear penalty – a lattice programming pomdp approach," *IEEE Trans. Inform. Theory*, vol. 57, no. 3, Oct. 2011, http://arxiv.org/abs/1011.5298.

[8] S. Bikchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as information cascades," *Journal of Political Economy*, vol. 100, pp. 992–1026, 1992.

[9] S. Karlin and Y. Rinott, "Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions," *Journal of Multivariate Analysis*, vol. 10, pp. 467–498, 1980.

[10] S. Ross, M. Izadi, M. Mercer, and D. Buckeridge, "Sensitivity analysis of POMDP value functions," in *IEEE International Conference on Machine Learning and Applications ICMLA'09*, 2009.