Stochastic approximation based methods for computing the optimal thresholds in remote-state estimation with packet drops

Jhelum Chakravorty, Jayakumar Subramanian and Aditya Mahajan

***Abstract*— A remote-state estimation system consisting of a sensor and an estimator is considered. The sensor observes a scalar Gauss-Markov process and at each time determines whether or not to transmit the state of the process. The transmission takes place over a packet drop channel. Previous results have established that the optimal transmission strategies are threshold based and optimal estimation strategies are Kalman-like. We propose stochastic approximation algorithms to compute the optimal thresholds for two setups: a Keifer-Wolfowitz based algorithm for the case when there is a cost associated with each transmission and a Robbins-Monro based algorithm for the case when there is a constraint on the expected number of transmissions. The results are verified by comparing against existing results for the no packet drop case.**

I. Introduction

In many *real-time* communication systems such as networked control systems, sensor surveillance networks, and transportation networks, etc., sequential data transmission takes place from node to node. In such applications the transmitter is often a battery powered device that transmits over a wireless packet-switched network, where the cost of turning the device on and transmitting a packet is much more significant compared to the size of the data packet. Therefore, the transmitter transmits intermittently but the transmitted packet is adequately big to communicate the current source realization. A remote estimator upon receiving the transmitted packet generates an estimate of the source realization in real-time. In such systems, there is a fundamental trade-off between communication cost and estimation accuracy.

When there are no packet-drops in the channel and the observed process is a first-order autoregressive process with unimodal and symmetric distribution, the structure of the suboptimal and optimal transmission and estimation strategies is known [1]–[5]. In particular, the optimal transmission scheme is threshold based and the optimal estimation scheme is Kalman-like.

The structural results mentioned above were generalized to channels with packet drops in [6]–[8]. Other variations of remote-state estimation over packet drop channels have been considered in [9]–[11].

In [12], algorithms for computing the optimal thresholds are proposed for channels with no packet drops. In particular, it is shown that for discrete-valued processes, the optimal thresholds can be computed by simple matrix calculations; for continuous-valued processes, the optimal thresholds can be computed by solving Fredholm integral equations of the second kind. In [6], the computational approach of [12] is generalized to channels with packet drops for discrete-valued processes. In principle, the results of [12] could work for continuous-valued processes. However, the resultant Fredholm integral equation would not be easy to solve because the kernel is discontinuous and the domain of the integral equation is $(-\infty, \infty)$.

In this paper, we propose stochastic approximation based methods to compute optimal thresholds for continuous-valued processes (and channels with packet drops). The proposed algorithms exploit the renewal property of the error process. For the case of no packet drops, the numerical results match the analytic results obtained in [12].

A. Model

Consider the following model of a discrete-time Markov process $\{X_t\}_{t=0}^{\infty}$ with the initial state $X_0 = 0$ and for $t \geq 0$, $X_{t+1} = aX_t + W_t$, where $a \in \mathbb{R}$ is the system parameter and $\{W_t\}_{t=0}^{\infty}$ is an i.i.d. Gaussian process with mean zero and variance σ^2. Let $\phi(\cdot)$ denote the probability distribution function of W_t.

A transmitter sequentially observes the process and at each time, chooses whether or not to transmit the current state. This decision is denoted by $U_t \in \{0, 1\}$, where $U_t = 0$ denotes no transmission and $U_t = 1$ denotes transmission. The decision to transmit is made using a *transmission strategy* $f = \{f_t\}_{t=0}^{\infty}$, where

$$U_t = f_t(X_{0:t}, U_{0:t-1}). \tag{1}$$

We use the short-hand notation $X_{0:t}$ to denote the sequence $(X_0, \dots, X_t)$. Similar interpretations hold for $U_{0:t-1}$.

If the transmitter decides to transmit (i.e., $U_t = 1$), X_t is transmitted over a wireless erasure channel and there is a probability $p_d \in (0, 1)$ that the transmitted packet is dropped. Let $S_t \in \{0, 1\}$ denote the state of the channel at time t. $S_t = 0$ denotes that the channel is in the OFF state and a transmitted packet will be dropped; $S_t = 1$ denotes that channel is in the ON state and a transmitted packet will be received. We assume that $\{S_t\}_{t\geq 0}$ is an i.i.d. process with $\mathbb{P}(S_t = 0) = p_d$. Moreover, $\{S_t\}_{t\geq 0}$ is independent of $\{X_t\}_{t\geq 0}$.

Transmission takes place using a TCP-like protocol, so there is an acknowledgement from the receiver to the transmitter when a packet is received successfully.[1] This means that the transmitter observes $H_t = U_t S_t$, which indicates whether the packet was successfully received by the receiver ($H_t = 1$) or not ($H_t = 0$)

When $H_t = 1$, the received symbol Y_t equals X_t; when $H_t = 0$, no symbol is received, which we denote by $Y_t = \mathfrak{E}$.

[1] The lack of an acknowledgement constitutes a negative acknowledgement (NACK), so a NACK does not need to be explicitly sent.

The receiver sequentially observes $\{Y_t\}_{t=0}^{\infty}$ and generates an estimate $\{\hat{X}_t\}_{t=0}^{\infty}$, $\hat{X}_t \in \mathbb{R}$, using an *estimation strategy* $g = \{g_t\}_{t=0}^{\infty}$, i.e.,

$$\hat{X}_t = g_t(Y_{0:t}). \tag{2}$$

The fidelity of the estimation is measured by a per-step distortion $d(X_t - \hat{X}_t) = (X_t - \hat{X}_t)^2$.

B. Performance metrics

Given a transmission and estimation strategy (f, g) and a discount factor $\beta \in (0, 1]$, we define the expected distortion and the expected number of transmissions as follows. For $\beta \in (0, 1)$, the expected *discounted* distortion is given by

$$D_\beta(f, g) := (1-\beta)\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t d(X_t - \hat{X}_t) \;\Big|\; X_0 = 0\Big] \tag{3}$$

and for $\beta = 1$, the expected *long-term average* distortion is given by

$$D_1(f, g) := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{T-1} d(X_t - \hat{X}_t) \;\Big|\; X_0 = 0\Big]. \tag{4}$$

Similarly, for $\beta \in (0, 1)$, the expected *discounted* number of transmissions is given by

$$N_\beta(f, g) := (1-\beta)\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t U_t \;\Big|\; X_0 = 0\Big] \tag{5}$$

and for $\beta = 1$, the expected *long-term average* number of transmissions is defined similar to (4).

C. Problem formulations

Problem 1 (Costly communication) *Given a communication cost $\lambda \in \mathbb{R}_{>0}$, find a transmission and estimation strategy (f^*, g^*) such that*

$$C^*_\beta(\lambda) := C_\beta(f^*, g^*; \lambda) = \inf_{(f,g)} C_\beta(f, g; \lambda), \tag{6}$$

where $C_\beta(f, g; \lambda) := D_\beta(f, g) + \lambda N_\beta(f, g)$ is the total communication cost and the infimum in (6) *is taken over all history-dependent strategies.*

Problem 2 (Constrained communication) *Given a constraint $\alpha \in (0, 1)$, find a transmission and estimation strategy (f^*, g^*) such that*

$$D^*_\beta(\alpha) := D_\beta(f^*, g^*) = \inf_{(f,g): N_\beta(f,g)\le\alpha} D_\beta(f, g), \tag{7}$$

where the infimum is taken over all history-dependent strategies.

II. Preliminary results

A. Structure of optimal communication strategy

Theorem 1 (Structural results) *In Problem 1, we have:*

1) Structure of optimal estimation strategy*: The optimal estimation strategy for $\hat{X}_0 = 0$ and for $t > 0$ is as follows:*

$$\hat{X}_t = \begin{cases} Y_t, & \text{if } Y_t \neq \mathfrak{E} \\ a\hat{X}_{t-1}, & \text{if } Y_t = \mathfrak{E}. \end{cases}$$

We denote this strategy by g^*.

2) Structure of optimal transmission strategy*: Define $E_t := X_t - a\hat{X}_{t-1}$, which we call the* error process*. Then there exists a time-invariant threshold k such that the transmission strategy*

$$U_t = f^{(k)}(E_t) := \begin{cases} 1, & \text{if } |E_t| \ge k \\ 0, & \text{if } |E_t| < k \end{cases} \tag{8}$$

is optimal.

The proof of a finite horizon version of Theorem 1 follows from arguments similar to [3]–[5] (which considered channels with no packet drops). Generalization to infinite horizon follows from arguments similar to [12]. Similar results are proved in [6] for discrete sources and in [8] for channels with Markov packet drops.

Note that E_t is a regenerative process, whose time evolution can be written as

$$E_{t+1} = \begin{cases} aE_t + W_t, & \text{if } Y_t = \mathfrak{E} \\ W_t, & \text{if } Y_t \neq \mathfrak{E}. \end{cases}, \tag{9}$$

B. Performance of a threshold-based strategy

Let $\mathcal{F}^{(k)}$ denote the class of all time-homogeneous threshold-based strategies of the form (8). For $\beta \in (0, 1]$ and $e \in \mathbb{R}$, define the following for a system that starts in state e and follows strategy $f^{(k)}$:

- $L^{(k)}_\beta(e)$: the expected discounted distortion until the first successful reception;
- $M^{(k)}_\beta(e)$: the expected discounted time until the first successful reception;
- $K^{(k)}_\beta(e)$: the expected discounted number of transmissions until the first successful reception;
- $D^{(k)}_\beta(e)$: the expected discounted distortion;
- $N^{(k)}_\beta(e)$: the expected discounted number of transmissions;
- $C^{(k)}_\beta(e; \lambda)$: the expected discounted total cost, i.e.,

$$C^{(k)}_\beta(e; \lambda) = D^{(k)}_\beta(e) + \lambda N^{(k)}_\beta(e), \quad \lambda \ge 0.$$

$(L^{(k)}_\beta, M^{(k)}_\beta, K^{(k)}_\beta)$ and $(D^{(k)}_\beta, N^{(k)}_\beta)$ are related through renewal relationships. In particular, we have the following theorem:

Theorem 2 (Renewal relationships) *For any $\beta \in (0, 1]$ and $k \in \mathbb{R}_{>0}$, we have:*

$$D_\beta(f^{(k)}, g^*) = \frac{L^{(k)}_\beta(0)}{M^{(k)}_\beta(0)}, \quad N_\beta(f^{(k)}, g^*) = \frac{K^{(k)}_\beta(0)}{M^{(k)}_\beta(0)}.$$

The proof is similar to [12, Theorem 2].

C. Characterization of the optimal solutions

Let $\partial_k D^{(k)}_\beta$, $\partial_k N^{(k)}_\beta$ and $\partial_k C^{(k)}_\beta$ denote the derivative[2] of $D^{(k)}_\beta$, $N^{(k)}_\beta$ and $C^{(k)}_\beta$ with respect to k.

[2]Following [12], one can show that $D^{(k)}_\beta$, $N^{(k)}_\beta$ and $C^{(k)}_\beta$ are differentiable in k.

The following two theorems characterize the optimal performance for Problems 1 and 2. The proof is similar to [12].

Theorem 3 *For $\beta \in (0,1]$, we have the following.*

1) *If the pair (λ, k) satisfies the following*

$$\lambda \partial_k N_\beta^{(k)}(0) + \partial_k D_\beta^{(k)}(0) = 0, \tag{10}$$

then, the strategy $(f^{(k)}, g^*)$ is optimal for Problem 1 with communication cost λ. Furthermore, for any $k > 0$, there exists a $\lambda \geq 0$ that satisfies (10).

2) *The optimal performance $C_\beta^*(\lambda)$ is continuous, concave and increasing function of λ.*

Theorem 4 *For any $\beta \in (0,1]$ and $\alpha \in (0,1)$, let $k_\beta^*(\alpha) \in \mathbb{R}_{\geq 0}$ be such that*

$$N_\beta^{(k_\beta^*(\alpha))}(0) = \alpha. \tag{11}$$

Such a $k_\beta^*(\alpha)$ exists and we have the following:

1) *The strategy $(f^{(k_\beta^*(\alpha))}, g^*)$ is optimal for Problem 2 with constraint α.*
2) *The distortion-transmission function $D_\beta^*(\alpha)$ is continuous, convex and decreasing in α and is given by*

$$D_\beta^*(\alpha) = D_\beta^{(k_\beta^*(\alpha))}(0). \tag{12}$$

III. Motivation for the current work

According to Theorem 2, computing $L_\beta^{(k)}(0)$, $K_\beta^{(k)}(0)$ and $M_\beta^{(k)}(0)$ is sufficient to compute $D_\beta^{(k)}(0)$ and $N_\beta^{(k)}(0)$ (and therefore, compute the performance of strategy $f^{(k)}$). In [12], which considers the case of no packet drops (i.e., $p_d = 0$), $L_\beta^{(k)}(0)$ and $M_\beta^{(k)}(0)$ were computed by solving the balance equations for the truncated Markov chain. These balance equations corresponded to Fredholm integral equations of the second kind. Using this exact policy evaluation, the optimal thresholds were identified by a binary search over k.

When $p_d \neq 0$, the balance equations for the truncated Markov chain still correspond to Fredholm integral equations of the second kind, but it is not straightforward to solve them numerically because the integration kernel is discontinuous and the integration domain is $(-\infty, \infty)$. For this reason, we investigate an alternative computational approach. The main idea behind our proposed solution is to replace the exact policy evaluation by a Monte Carlo based approximate policy evaluation and to replace the binary search for the optimal threshold by a stochastic approximation iteration. In particular, we use Kiefer-Wolfowitz algorithm [13] to solve (10) and Robbins-Monro algorithm [14] to solve (11). The details are presented in the next section.

IV. Stochastic approximation algorithms

A. Noisy policy evaluation

The first step to develop a stochastic approximation algorithm to identify the optimal thresholds is to replace the exact policy evaluation by an approximate policy evaluation. The simplest way to do so is to use sample path average. In particular, let $\{(E_t^{(k)}, U_t^{(k)})\}_{t\geq 0}$ denote the sample paths of the error process and the transmission process under

Algorithm 1: Algorithm for noisy policy evaluation

function *MonteCarloEvaluation*(k, K)
  input : Threshold $k \in \mathbb{R}_{>0}$
    Number of episodes $\mathsf{K} \in \mathbb{Z}_{>0}$
  output : Estimate $\hat{L}_\beta^{(k,\mathsf{K})}$ of $L_\beta^{(k)}(0)$
    Estimate $\hat{M}_\beta^{(k,\mathsf{K})}$ of $M_\beta^{(k)}(0)$
    Estimate $\hat{K}_\beta^{(k,\mathsf{K})}$ of $K_\beta^{(k)}(0)$
  initialize: $\hat{L} = 0$, $\hat{M} = 0$, $\hat{K} = 0$
  for *iteration* $i = 1$ *upto* K **do**
    Set $t = 0$, $\ell = 0$, $m = 0$, $\kappa = 0$, $E_0 = 0$
    while *true* **do**
      $S_{t+1} \sim \text{Bernoulli}(p_d)$
      if $|E_t| < k$ *or* $S_{t+1} = 0$ **then**
        $\ell \leftarrow \ell + \beta^t E_t^2$
        $m \leftarrow m + \beta^t$
        $\kappa \leftarrow \kappa + \beta^t \mathbb{1}_{\{|E_t| \geq k\}}$
      else
        $\kappa \leftarrow \kappa + \beta^t$
        break
      $E_{t+1} = aE_t + W_t$, where $W_t \sim \mathcal{N}(0, \sigma^2)$
      $t \leftarrow t + 1$
    $\hat{L} \leftarrow \hat{L} + \ell$, $\hat{M} \leftarrow \hat{M} + m$, $\hat{K} \leftarrow \hat{K} + \kappa$
  return $(\hat{L}/\mathsf{K}, \hat{M}/\mathsf{K}, \hat{K}/\mathsf{K})$

policy $f^{(k)}$ and T be a large number. Then, $D_\beta^{(k)}(0) \approx (1-\beta)\sum_{t=0}^{T} \beta^t d(E_t^{(k)})$, $N_\beta^{(k)}(0) \approx (1-\beta)\sum_{t=0}^{T} \beta^t U_t^{(k)}$, and $C_\beta^{(k)}(0;\lambda) = D_\beta^{(k)}(0) + \lambda N_\beta^{(k)}(0)$.

For the discounted case, using naive approach leads to numerical difficulties as one needs to compute β^t for large t, which makes the term very small. To circumvent this, we estimate $L_\beta^{(k)}(0)$, $M_\beta^{(k)}(0)$ and $K_\beta^{(k)}(0)$ by Monte Carlo evaluations and then use the renewal relationship of Theorem 2 to approximate $D_\beta^{(k)}(0)$ and $N_\beta^{(k)}(0)$.

The Monte Carlo evaluations are done by averaging over K episodes. In each episode, the error process starts at $E_0 = 0$ and evolves under strategy $f_t^{(k)}$. The episode ends at the stopping time $\tau^{(k)}$ of the first successful reception. Let $\{(E_{n,t}^{(k)}, U_{n,t}^{(k)})\}_{t\geq 0}$ denote the sample path of the error process and the transmission process in episode n. Then,

$$L_\beta^{(k)}(0) \approx \frac{1}{\mathsf{K}} \sum_{n=1}^{\mathsf{K}} \sum_{t=0}^{\tau^{(k)}-1} \beta^t d(E_{n,t}^{(k)}), \tag{13}$$

$$M_\beta^{(k)}(0) \approx \frac{1}{\mathsf{K}} \sum_{n=1}^{\mathsf{K}} \sum_{t=0}^{\tau^{(k)}-1} \beta^t, \tag{14}$$

$$K_\beta^{(k)}(0) \approx \frac{1}{\mathsf{K}} \sum_{n=1}^{\mathsf{K}} \sum_{t=0}^{\tau^{(k)}} \beta^t U_{n,t}^{(k)}. \tag{15}$$

Then, $D_\beta^{(k)}(0)$, $N_\beta^{(k)}(0)$, and $C_\beta^{(k)}(0;\lambda)$ can be computed using the expressions in Theorem 2. The complete details for this evaluation are shown in Algorithm 1.

Algorithm 2: Algorithm for costly communication

input : Initial guess $\mathsf{k}_{\mathsf{init}} \in \mathbb{R}_{>0}$;
Number of episodes $\mathsf{K} \in \mathbb{Z}_{>0}$
Number of iterations $\mathsf{N}_{\mathsf{iterations}} \in \mathbb{Z}_{>0}$
output : Optimal threshold k°
initialize: $k^\circ = \mathsf{k}_{\mathsf{init}}$
for *iteration* $i = 1$ *upto* $\mathsf{N}_{\mathsf{iterations}}$ **do**
  Pick δ as a small non-negative real
  $k^\circ_+ = k^\circ + \delta$
  $k^\circ_- = k^\circ - \delta$
  $[\hat{L}_+, \hat{M}_+, \hat{K}_+]$ = MonteCarloEvaluation$\big(k^\circ_+, \mathsf{K}\big)$
  $[\hat{L}_-, \hat{M}_-, \hat{K}_-]$ = MonteCarloEvaluation$\big(k^\circ_-, \mathsf{K}\big)$
  Compute C_+, C_- using Theorem 2
  $\partial_k \mathsf{C} = (\mathsf{C}_+ - \mathsf{C}_-)/2\delta$
  Compute γ_i using ADAM [15]
  $k^\circ \leftarrow k^\circ - \gamma_i \partial_k \mathsf{C}$
return k°

Let $\hat{L}^{(k,\mathsf{K})}_\beta$, $\hat{M}^{(k,\mathsf{K})}_\beta$ and $\hat{K}^{(k,\mathsf{K})}_\beta$ denote the right hand sides of (13), (14) and (15). Then, (13)–(15) can be written as

$$L^{(k)}_\beta(0) = \hat{L}^{(k,\mathsf{K})}_\beta + \xi^L_{\mathsf{K}}, \quad M^{(k)}_\beta(0) = \hat{M}^{(k,\mathsf{K})}_\beta + \xi^M_{\mathsf{K}},$$
$$K^{(k)}_\beta(0) = \hat{K}^{(k,\mathsf{K})}_\beta + \xi^K_{\mathsf{K}},$$

where ξ^L_{K}, ξ^M_{K} and ξ^K_{K} are approximation errors that go to zero as K goes to infinity. Define estimates $\hat{D}^{(k,\mathsf{K})}_\beta$, $\hat{N}^{(k,\mathsf{K})}_\beta$, and $\hat{C}^{(k,\mathsf{K})}_\beta(\lambda)$ for $D^{(k)}_\beta(0)$, $N^{(k)}_\beta(0)$, and $C^{(k)}_\beta(0;\lambda)$ in terms of $\hat{L}^{(k)}_\beta$, $\hat{M}^{(k)}_\beta$ and $\hat{K}^{(k)}_\beta$ using renewal expressions given in Theorem 2.

Note that the stochastic approximation algorithms that we describe next work under mild assumptions on ξ^L_{K}, ξ^M_{K} and ξ^K_{K}. Therefore, the number K of episodes need not be large. In our experiments that we report later, we choose K as 1000.

B. Computing thresholds for costly communication using stochastic approximation

In our subsequent discussion, we assume the following:

(A1) $C^{(k)}_\beta(0;\lambda)$ is convex in k.

(A2) $\mathbb{E}[\mathsf{C}^{(k,\mathsf{K})}_\beta(\lambda)] = C^{(k)}_\beta(0;\lambda)$.

We verified through simulation that (A1) holds. (A2) holds if $\mathsf{C}^{(k,\mathsf{K})}_\beta(\lambda)$ is an unbiased estimator of $C^{(k)}_\beta(0;\lambda)$, which we verified through simulations.

According to Theorem 3, a threshold k is optimal if $\partial_k C^{(k)}_\beta(0;\lambda) = 0$. Using Algorithm 1, we can obtain a noisy "measurement" $\hat{C}^{(k,\mathsf{K})}_\beta(\lambda)$ of $C^{(k)}_\beta(0;\lambda)$. Using this noisy measurement, it is possible to search for the optimal threshold using the Kiefer-Wolfowitz algorithm [13], which is a first-order stochastic gradient descent algorithm that works as follows.

We start with an initial guess k°_0 of the optimal threshold. Let k°_i denote our guess at the beginning of iteration i. During iteration i, we obtain a noisy measurement of the gradient $\partial_k C^{(k)}_\beta(0;\lambda)$ using the finite difference $\Delta^{(k^\circ_i,\mathsf{K})}_i = \hat{C}^{(k^\circ_i+\delta,\mathsf{K})}_\beta(\lambda) - \hat{C}^{(k^\circ_i-\delta,\mathsf{K})}_\beta(\lambda)$ and update our guess as follows:

$$k^\circ_{i+1} = k^\circ_i - \gamma_i \Delta^{(k^\circ_i,\mathsf{K})}_i, \tag{16}$$

where γ_i are learning rates that satisfy $\sum_{i=1}^\infty \gamma_i = \infty$ and $\sum_{i=1}^\infty \gamma_i^2 < \infty$. See Algorithm 2 for details.

Theorem 5 *Under assumptions (A1)–(A2), the threshold iterates k°_i of Algorithm 2 converge almost surely to the optimal threshold, i.e., $\lim_{i\to\infty} k^\circ_i = k^*(\lambda)$, a.s., where $k^*(\lambda)$ is optimal threshold for Problem 1.*

The proof follows immediately from [13].

The rate of convergence of the Kiefer-Wolfowitz algorithm is sensitive to the choice of learning rates. We use ADAM (Adaptive Moments) [15] to adaptively tune the learning rate based on the "measurements" $\Delta^{(k^\circ_i,\mathsf{K})}_i$.

C. Computing thresholds for constrained communication using stochastic approximation

First, we note the following facts:

(F1) $\mathsf{M}^{(k,\mathsf{K})}_\beta$ is increasing with k and $\mathsf{K}^{(k,\mathsf{K})}_\beta$ is decreasing with k.

(F2) $\mathbb{E}[\mathsf{M}^{(k,\mathsf{K})}_\beta] = M^{(k)}_\beta(0)$ and $\mathbb{E}[\mathsf{K}^{(k,\mathsf{K})}_\beta] = K^{(k)}_\beta(0)$.

(F1) can be proved using an argument similar to the one used in [12]. (F2) holds by definition.

According to Theorem 4, a threshold k is optimal if $\alpha M^{(k)}_\beta(0) = K^{(k)}_\beta(0)$. Using Algorithm 1, we can obtain noisy "measurements" of $M^{(k)}_\beta(0)$ and $K^{(k)}_\beta(0)$. Using these noisy measurements, it is possible to search for the optimal threshold using the Robbins-Monro algorithm [14], which is a first-order stochastic root-finding algorithm that works as follows.

We start with an initial guess k°_0 of the optimal threshold. Let k°_i denote our guess at the beginning of iteration i. During iteration i, we obtain a noisy measurement $\hat{M}^{(k,\mathsf{K})}_\beta$ of $M^{(k)}_\beta(0)$ and $\hat{K}^{(k,\mathsf{K})}_\beta$ of $K^{(k)}_\beta(0)$ and update our guess as follows:

$$k^\circ_{i+1} = k^\circ_i - \gamma_i \left(\alpha \hat{M}^{(k^\circ_i,\mathsf{K})}_i - \hat{K}^{(k^\circ_i,\mathsf{K})}_i \right), \tag{17}$$

where γ_i are learning rates that satisfy $\sum_{i=1}^\infty \gamma_i = \infty$ and $\sum_{i=1}^\infty \gamma_i^2 < \infty$. See Algorithm 3 for details.

Algorithm 3: Algorithm for constrained communication

input : Initial guess $\mathsf{k}_{\mathsf{init}} \in \mathbb{R}_{>0}$;
Number of episodes $\mathsf{K} \in \mathbb{Z}_{>0}$
Number of iterations $\mathsf{N}_{\mathsf{iterations}} \in \mathbb{Z}_{>0}$
output : Optimal threshold k°
initialize: $k^\circ = \mathsf{k}_{\mathsf{init}}$
for *iteration* $i = 1$ *upto* $\mathsf{N}_{\mathsf{iterations}}$ **do**
  $\gamma_i = 1/i$
  $[\hat{L}, \hat{M}, \hat{K}]$ = MonteCarloEvaluation$\big(k^\circ, \mathsf{K}\big)$
  $k^\circ \leftarrow k^\circ - \gamma_i\big(\alpha\hat{M} - \hat{K}\big)$
return k°

TABLE I: Comparative results for costly communication using Stochastic Approximation (SA) and Fredholm Integral Equations of second kind (FIE) for $a = 1$ and $p_d = 0$.

(a) $\beta = 0.9$

λ	Threshold k^* SA	FIE	Error (Absolute)	Performance $C^*_\beta(\lambda)$ SA	FIE	Error (Absolute)
100	4.9355	4.9298	5.7×10^{-3}	5.2511	5.2511	9.1×10^{-6}
200	6.3221	6.3086	1.4×10^{-2}	6.5221	6.5221	3.5×10^{-5}
300	7.3421	7.3289	1.3×10^{-2}	7.2208	7.2208	2.4×10^{-5}
400	8.2118	8.1764	3.5×10^{-2}	7.6654	7.6652	1.4×10^{-4}
500	8.9469	8.9177	2.9×10^{-2}	7.9700	7.9700	7.2×10^{-5}
600	9.5830	9.5854	2.5×10^{-3}	8.1886	8.1886	4.7×10^{-7}
700	10.0803	10.1984	1.2×10^{-1}	8.3515	8.3507	8.0×10^{-4}

(b) $\beta = 1.0$

λ	Threshold k^* SA	FIE	Error (Absolute)	Performance $C^*_\beta(\lambda)$ SA	FIE	Error (Absolute)
100	4.3438	4.3446	7.9×10^{-4}	7.8540	7.8540	4.2×10^{-7}
200	5.283	5.2841	8.3×10^{-4}	11.2327	11.2327	4.2×10^{-7}
300	5.9340	5.9136	2.0×10^{-2}	13.8265	13.8262	2.7×10^{-4}
400	6.4079	6.4004	7.5×10^{-3}	16.0131	16.0131	3.7×10^{-5}
500	6.8028	6.8028	4.4×10^{-5}	17.9399	17.9399	4.4×10^{-9}
600	7.1487	7.1485	1.1×10^{-4}	19.6810	19.6810	1.9×10^{-9}
700	7.4569	7.4534	3.5×10^{-3}	21.2829	21.2829	8.0×10^{-6}

TABLE II: Comparative results for constrained communication using Stochastic Approximation (SA) and Fredholm Integral Equations of second kind (FIE) for $a = 1$ and $p_d = 0$.

(a) $\beta = 0.9$

α	Threshold k^* SA	FIE	Error (Absolute)	Performance $D^*_\beta(\alpha)$ SA	FIE	Error (Absolute)
0.1	2.2230	2.2217	1.3×10^{-3}	0.9293	0.9283	9.9×10^{-4}
0.2	1.4416	1.4404	1.2×10^{-3}	0.3954	0.3947	7.0×10^{-4}
0.3	1.0586	1.0620	3.4×10^{-3}	0.1974	0.1989	1.5×10^{-3}
0.4	0.8014	0.8057	4.3×10^{-3}	0.0989	0.1003	1.4×10^{-3}
0.5	0.6017	0.5981	3.5×10^{-3}	0.0460	0.0453	7.4×10^{-4}
0.6	0.4357	0.4395	3.7×10^{-3}	0.0186	0.0190	4.6×10^{-4}
0.7	0.2823	0.2808	1.5×10^{-3}	0.0052	0.0052	8.1×10^{-5}
0.8	0.1396	0.1465	6.8×10^{-3}	0.0006	0.0007	9.9×10^{-5}

(b) $\beta = 1.0$

α	Threshold k^* SA	FIE	Error (Absolute)	Performance $D^*_\beta(\alpha)$ SA	FIE	Error (Absolute)
0.1	2.5396	2.5391	5.7×10^{-4}	1.3677	1.3671	5.8×10^{-4}
0.2	1.6020	1.5991	2.9×10^{-3}	0.5485	0.5464	2.0×10^{-3}
0.3	1.1713	1.1719	6.2×10^{-4}	0.2767	0.2770	3.4×10^{-4}
0.4	0.9014	0.9033	1.9×10^{-3}	0.1477	0.1485	7.7×10^{-4}
0.5	0.6994	0.6958	3.6×10^{-3}	0.0767	0.0756	1.0×10^{-3}
0.6	0.5334	0.5371	3.7×10^{-3}	0.0365	0.0373	7.2×10^{-4}
0.7	0.3884	0.3906	2.2×10^{-3}	0.0148	0.1500	2.4×10^{-4}
0.8	0.2540	0.2563	2.3×10^{-3}	0.0043	0.0044	1.2×10^{-4}

Theorem 6 *The threshold iterates k°_i of Algorithm 3 converge almost surely to the optimal thresholds, i.e., $\lim_{i\to\infty} k^\circ_i = k^*(\alpha)$, a.s., where $k^*(\alpha)$ is optimal threshold for Problem 2.*

The proof follows immediately from [14].

Here we found that using the learning rates $\gamma_i = 1/i$ yields fast convergence and hence we did not use ADAM to adapt the learning rates.

V. Numerical results

In all the results reported below, $a = 1$ and $\sigma^2 = 1$. The code for the experiments is available at [16].

A. Channels with no packet drops (for validation)

We start by comparing the proposed stochastic approximation algorithms with the exact algorithm of [12].

For costly communication, we consider $\beta \in \{0.9, 1.0\}$ and $\lambda \in \{100, 200, \dots, 700\}$. We set the number of episodes in Algorithm 1 to 1000 and number of iterations in Algorithm 2 to 10,000. The corresponding thresholds are shown in Table I.

The optimal thresholds obtained by Fredholm integral equations (as proposed in [12]) are also shown in Table I. The thresholds obtained by stochastic approximation are within 10^{-2} of the optimal for most cases. We also compute the total cost $C^{(k)}_\beta(0; \lambda)$ (by solving Fredholm integral equation) for both cases. The cost of the thresholds obtained by stochastic approximation is less than 10^{-3} from the optimal cost.

For constrained communication, we consider $\beta \in \{0.9, 1.0\}$ and $\alpha \in \{0.1, 0.2, \dots, 0.8\}$. The number of episodes in Algorithm 1 is set to 1. The corresponding thresholds are shown in Table II.

As in the case of costly communication, we compare the thresholds and the performance $D^{(k)}_\beta(0)$ obtained by stochastic approximation with those obtained by Fredholm integral equations. The thresholds obtained by stochastic approximation are within 10^{-3} or less of the optimal.

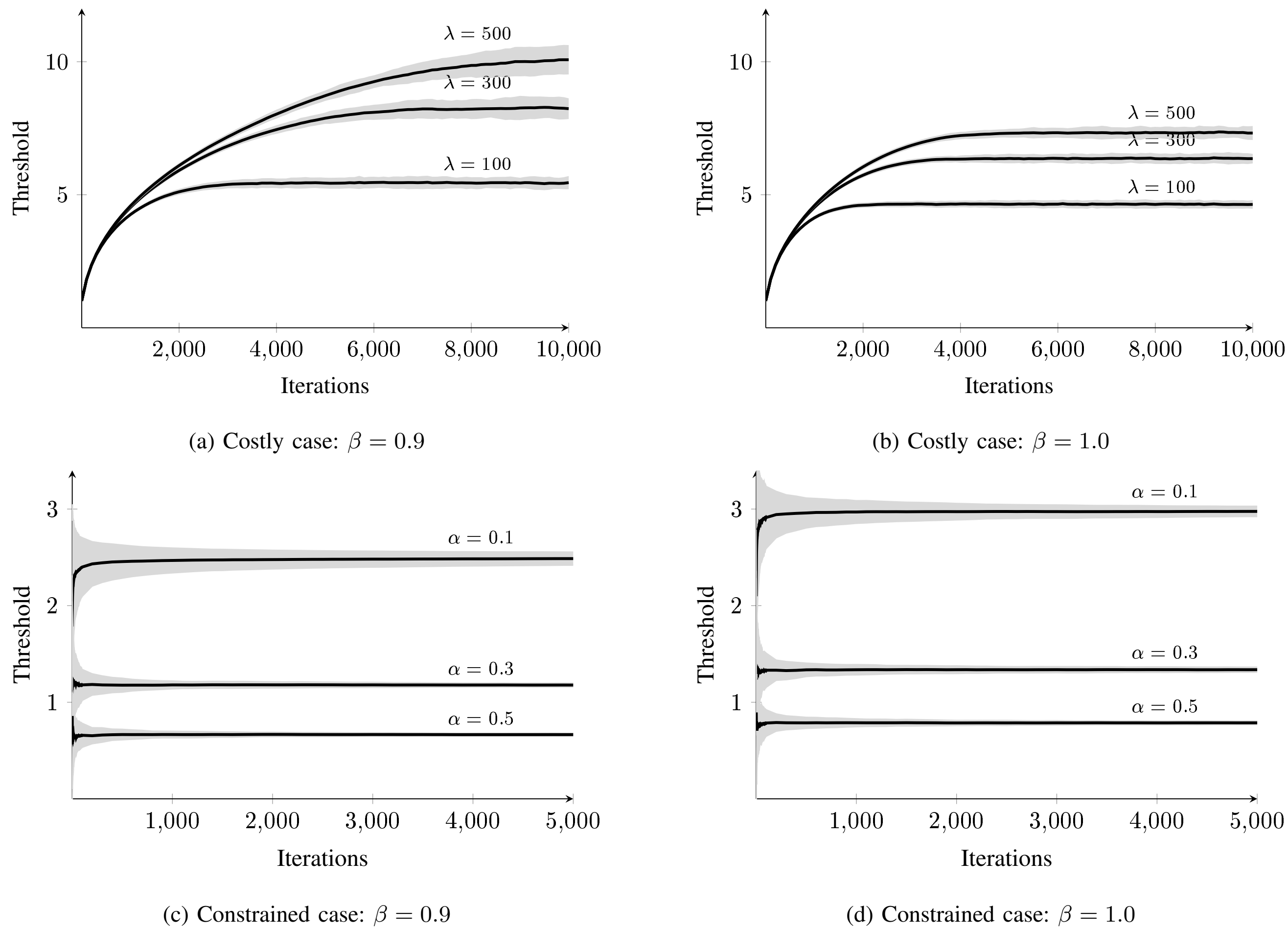
These results show that the results obtained by stochastic approximation algorithms are accurate.

B. Channel with packet drops

We repeat the experiments of the previous section with $p_d = 0.3$. To understand the variability of stochastic approximation across different runs, we run each experiment 100 times and plot the mean and standard deviation of the thresholds versus the number of iterations in Fig. 2. For ease of visualization, we only show the results for a subset of values of λ and α. For both costly and constrained communication, there is very little variation across multiple runs. It takes about 9000 iterations to converge for costly communication and 3000 iterations for constrained communication.

VI. Conclusion and discussion

We present stochastic approximation algorithms to compute optimal thresholds for remote state estimation over communication channels with packet drops. The inner loops



(a) Costly case: $\beta = 0.9$

(b) Costly case: $\beta = 1.0$

(c) Constrained case: $\beta = 0.9$

(d) Constrained case: $\beta = 1.0$

Fig. 2: The sample paths for costly and constrained cases for $p_d = 0.3$. Here the bold lines represent the sample means for 100 runs and the shaded regions correspond to mean $\pm$ twice the standard deviation across the runs (i.e., the 95% confidence interval).

of these algorithms use Monte Carlo evaluation to get a noisy estimate of the performance of a threshold-based strategy.

Stochastic approximation algorithms scale well to multi-dimensional setup, where the Kiefer-Wolfowitz algorithm can be replaced by Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm [17] which requires only two random samples to estimate the gradient.

References

[1] O. C. Imer and T. Basar, "Optimal estimation with limited measurements," in *Joint 44th IEEE Conference on Decision and Control and European Control Conference*, vol. 29, Dec 2005, pp. 1029–1034.

[2] Y. Xu and J. P. Hespanha, "Optimal communication logics in networked control systems," in *43rd IEEE Conference on Decision and Control*, vol. 4, Dec 2004, pp. 3527–3532.

[3] G. M. Lipsa and N. C. Martins, "Remote state estimation with communication costs for first-order LTI systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 9, pp. 2013–2025, Sep 2011.

[4] A. Nayyar, T. Basar, D. Teneketzis, and V. V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2246–2260, Sep 2013.

[5] A. Molin and S. Hirche, "An iterative algorithm for optimal event-triggered estimation," in *4th IFAC Conference on Analysis and Design of Hybrid Systems*, June 2012, pp. 64–69.

[6] J. Chakravorty and A. Mahajan, "Remote state estimation with packet drop," in *6th IFAC Workshop on Distributed Estimation and Control in Networked Systems*, Sep 2016.

[7] X. Ren, J. Wu, K. H. Johansson, G. Shi, and L. Shi, "Infinite horizon optimal transmission power control for remote state estimation over fading channels," *arxiv:1604.08680*, Apr 2016.

[8] J. Chakravorty and A. Mahajan, "Structure of optimal strategies for remote estimation over Gilbert-Elliott channel with feedback," *arxiv:1701.05943*, Jan 20 2017.

[9] L. Shi and L. Xie, "Optimal sensor power scheduling for state estimation of Gauss-Markov systems over a packet-dropping network," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2701–2705, May 2012.

[10] S. Dey, A. Chiuso, and L. Schenato, "Remote estimation subject to packet loss and quantization noise," in *52nd IEEE Conference on Decision and Control*, Dec 2013, pp. 6097–6104.

[11] Y. Li, D. E. Quevedo, V. Lau, and L. Shi, "Online sensor transmission power schedule for remote state estimation," in *52nd IEEE Conference on Decision and Control*, Dec 2013, pp. 4000–4005.

[12] J. Chakravorty and A. Mahajan, "Fundamental limits of remote estimation of Markov processes under communication constraints," *Accepted in IEEE Transactions on Automatic Control*, 2017.

[13] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Statist.*, vol. 23, no. 3, pp. 462–466, Sep 1952.

[14] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol. 22, no. 3, pp. 400–407, Sep 1951.

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arxiv:1412.6980*, Jan 2017.

[16] https://github.com/adityam/remote-estimation-packet-drops.

[17] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Trans. Autom. Control*, vol. 37, no. 3, pp. 332–341, Mar 1992.