Average cost optimal threshold strategies for remote estimation with communication cost

Jhelum Chakravorty and Aditya Mahajan

Abstract— **In this paper, we consider a remote sensing system that consists of a sensor and an estimator. A sensor observes a first order Markov source and must communicate it to a remote estimator. Communication is noiseless but expensive. At each time, based on the history of its observations and decisions, the sensor chooses whether to transmit or not. If the sensor does not transmit, then the estimator must estimate the Markov process using its past observations. We study the average cost problem in the light of vanishing discount approach. The problem was studied previously by Lipsa and Martins, IEEE TAC, 2011 and by Nayyar et al, IEEE TAC, 2013, where it was shown that the optimal estimation policy is Kalman-like and the optimal communication policy is to communicate when the estimation error is greater than a threshold. In the discounted set-up, we had earlier characterized the optimal policy and the optimal thresholds as a function of communication cost. The average cost problem is investigated as the limiting case of the discounted cost problem as the discount factor approaches one. The average cost and the optimal values of the thresholds are provided in terms of the communication cost. Lastly, we present an example of birth-death Markov chain to illustrate our results.**

I. Introduction

A. Motivation

In this paper, we consider a model that captures a fundamental trade-off between communication cost and estimation accuracy. This model is motivated by applications in smart grids and environmental monitoring.

In smart grids, it is envisioned that smart meters will measure the energy consumption in households and communicate these measurements to an aggregator which will use this information for demand response etc. In such a scenario, it is important not to flood the communication network with measurement information by communicating periodically. Instead, one can model the signaling overhead as a cost and optimally trade-off communication cost with estimation accuracy.

In environmental monitoring, a sensor network is used to measure an environmental variable such as rainfall, soil moisture, etc. Energy consumption at the sensor is an important consideration in such systems because it is expensive to replace the sensor battery. Thus, to conserve battery, it is important not to transmit periodically. Instead, one can model the energy consumed while communicating as a cost and optimally trade-off communication cost with estimation accuracy.

Jhelum Chakravorty and Aditya Mahajan are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada. {jhelum.chakravorty@mail.mcgill.ca, aditya.mahajan@mcgill.ca}

Similar scenarios also arise in other applications such as networked control systems. Consider the following model that captures the above scenarios. A sensor observes a first order Markov process and must communicate it to a remote estimator. Communication is noiseless but expensive. At each time, based on the history of its observations and decisions, the sensor chooses whether to transmit or not. If the sensor does not transmit, then the estimator must estimate the Markov process using its past observations. The objective is to minimize a weighted combination of communication cost and estimation error.

The remote estimation is conceptually difficult due to information decentralization. When the sensor decides not to communicate, its decision is based on the current value of the Markov source. So, even when the estimator does not receive an observation, the 'absence of observation' conveys some information about the Markov process. Such signaling problems are known to be notoriously hard, with the Witsenhausen's counterexample [1] being the most famous example.

The above model has received considerable attention in the literature. The most closely related papers are [2]–[5], which are briefly summarized below. Other related work includes models where the sensor is allowed to sleep for a pre-specified amount of time [6] and models where the estimator decides when the sensor should transmit [7]–[9]. The set-up of this paper is also related to the censoring sensors considered in hypothesis testing [10], [11] (where the sensor takes one measurement and decides whether to transmit it or not) and real-time source coding [12], [13] (where the sensor must transmit a quantized version of the observation).

In [2], the authors considered a remote estimation problem where the sensor could communicate a finite number of times. They assumed that the sensor used a threshold policy to decide when to communicate and determined the optimal estimation policy and the value of the thresholds. [14] have generalized the results of [2] to delays and transmission constraints imposed by real data networks. In [3], the authors considered remote estimation of Gaussian processes. They assumed a particular form of the estimator and show that the estimation error is a sufficient statistic for the sensor.

In [4] too, the authors considered remote estimation of Gaussian processes but do not impose any assumption on the communication or estimation policy. They use ideas from majorization theory to show that the optimal estimation policy is Kalman-like and the optimal communication policies are threshold based. In [5], the authors considered remote estimation of countable state Markov processes where the sensor harvests energy to communicate. They show that if the Markov process is symmetric in an appropriate sense, then the results of [4] continue to hold. Both [4] and [5] identified dynamic programs to find the optimal thresholds. The authors of [15] adopted a formulation that is similar to [4]. They consider a networked control problem with transmission costs, where they adopt a Kalman-like estimator and show, using dynamic programming, that, for such a pre-determined choice of estimator, the optimal pre-processor is a memoryless function of the state estimation error. In contrast to [4], the problem analyzed in [15] dealt also with the multidimensional case.

Threshold based transmission policies may be viewed as event-based transmission policies: transmission takes place when an event (estimation error greater than a threshold) takes place. Such event-based transmission has also received considerable attention in the literature, a detailed overview of which is given in [16]. In recent years, various event-triggering policies have been proposed and analyzed for different problem formulations, for both stochastic and deterministic set-ups. Given below are some of the contributions in stochastic event-based sampling.

[17] showed that the optimal event-triggering rules are given by the solutions to the optimal stopping-time problems and derived the jointly optimal stopping times and control signals for first-order systems. An extension to this work was carried out in [18] for a Markov state process with hard communication constraints, where it was shown that the optimal thresholds can be chosen one at a time using solutions to a nested sequence of optimization problems each with a single threshold as its decision variable, and in [19] for second-order systems, where the optimal threshold were computed by minimizing the aggregate performance for a given weight.

In a finite horizon LQG optimal control set-up, [20] considered optimal closed-loop design for event-based systems, with a joint selection of the event triggering policy and control law. They showed that the optimal scheduler sequence can be computed by solving a deterministic dynamic program. In a subsequent work, [21], the authors have shown that the optimal controller is a certainty equivalence controller consisting of linear gains and the optimal state estimator at the event-trigger is given by the Kalman filter. The event-trigger consists of a Kalman filter and a copy of the affine-linear predictor at the controller.

The rest of the paper is organized as follows. In Section II, to study the discounted cost problem, we revisit the model of [4] and [5] and look at it from a slightly different point of view. Consequently, we use the results of [22], where we use the idea of calibration from multi-armed bandits. We identify the value of the communication cost for which one is indifferent between two consecutive threshold policies. Using these values, we obtain the range of communication costs for which a particular policy is optimal. The main result is then discussed in Section III, where the average cost problem is studied as an extension to the discounted cost problem as the discount factor goes to 1. Using vanishing discount approach, we give the average cost and the optimal thresholds as functions of the communication cost. We show that the communication index provides the complete characterization of the optimal communication policy for all values of the communication cost.

B. Notation

$\mathbb{Z}$ denotes the set of integers and $\mathbb{N}$ denotes the set of natural numbers. $x_{1:t}$ is a short hand for the vector $(x_1, \dots, x_t)$. For a matrix A, A_{ij} denotes the (i,j)-th element of A and A_i denotes the i-th row of A. Note that unlike the standard notation, in our notation the indices to denote an element of a matrix take both positive and negative values. Furthermore, $A^\intercal$ denotes the transpose of A. I_k denotes the identity matrix of dimension $k \times k$, $k \in \mathbb{N}$. $\mathbf{1}_k$ denotes $k \times 1$ vector of ones. $\langle v, w \rangle$ denotes the inner product between vectors v and w, $\mathbb{P}(\cdot)$ denotes the probability of an event, $\mathbb{E}[\cdot]$ denotes the expectation of a random variable, and $\mathbb{1}\{\cdot\}$ denotes the indicator function of a statement.

C. Problem Formulation

Consider a remote sensing system, which consists of a sensor and an estimator. The sensor observes the state of a first-order Markov process $\{X_t\}_{t=0}^{\infty}$, $X_t \in \mathbb{Z}$[1], with transition matrix P. Assume that the Markov process starts in state x_0 that is known to the sensor and the estimator.

At time t, the sensor decides between two alternatives: either to transmit the current state X_t and incur a cost c or not transmit and incur no cost. The sensor's decision is denoted by $U_t \in \{0, 1\}$, where $U_t = 0$ denotes no transmission and $U_t = 1$ denotes transmitting the current state. The transmitted symbol Y_t is given by

$$Y_t = \begin{cases} X_t, & \text{if } U_t = 1 \\ \epsilon, & \text{if } U_t = 0 \end{cases}$$

where ϵ means no transmission.

The sensor's decision is generated as follows:

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

where f_t is called the *communication rule* at time t and the collection $\mathbf{f} = (f_1, f_2, \cdots)$ is called the *communication policy*.

The estimator observes the transmitted symbols and generates an estimate $\hat{X}_t \in \mathbb{Z}$ as follows:

$$\hat{X}_t = g_t(Y_{1:t}) \tag{2}$$

where g_t is called the *estimation rule* at time t and the collection $\mathbf{g} = (g_1, g_2, \cdots)$ is called the *estimation policy*.

We are interested in the following optimization problem:

Problem 1: In the above model, given the transition matrix P of the Markov process, the communication cost c and a

[1]The results generalize to $X_t \in \mathbb{R}^n$, but for ease of exposition we restrict our discussion to $X_t \in \mathbb{Z}$

distortion function $\ell : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}_+$, choose transmission and estimation policies $(\mathbf{f}, \mathbf{g})$ to minimize the average cost

$$J(\mathbf{f}, \mathbf{g}) = \lim_{T \to \infty} \mathbb{E}\Big[\frac{1}{T}\sum_{t=0}^{T}\big[cU_t + \ell(X_t, \hat{X}_t)\big] \;\Big|\; X_0 = x_0\Big] \tag{3}$$

To analyze the above average cost optimization problem, first we define the following discounted cost optimization problem as an intermediate step.

Problem 2: In the above model, given the transition matrix P of the Markov process, the communication cost c and a distortion function $\ell : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}_+$, choose transmission and estimation policies $(\mathbf{f}, \mathbf{g})$ to minimize the discounted cost

$$J(\mathbf{f}, \mathbf{g}) = \mathbb{E}\Big[\sum_{t=0}^{\infty}\beta^t\big[cU_t + \ell(X_t, \hat{X}_t)\big] \;\Big|\; X_0 = x_0\Big] \tag{4}$$

where $\beta \in (0, 1)$ is the discount factor.

In the next section, we state some relevant results for the discounted cost problem that were derived in [23]. The analysis of Problem 1 is then discussed as an extension to Problem 2.

D. Preliminary results on structure of optimal policies

We want to use vanishing discount approach to study Problem 1. So, we consider Problem 2 first. Slight variations of the above model have been considered in [4], [5] where the objective was to minimize the finite horizon cost

$$\mathbb{E}\Big[\sum_{t=0}^{T}\big[cU_t + \ell(X_t, \hat{X}_t)\big] \;\Big|\; X_0 = x_0\Big]. \tag{5}$$

The model in [4] assumed a Gauss-Markov source with square error distortion while the model in [5] assumed that sensor had limited battery that was replenished using energy harvesting. Both these papers established the structure of optimal communication and estimation policies under the following assumptions.

- **(A1)** The transition matrix P of the Markov process is a symmetric and banded Toeplitz matrix with decaying terms, i.e.

$$P_{ij} = \begin{cases} p_{|i-j|}, & \text{if } |i-j| \le b \\ 0, & \text{if } |i-j| > b \end{cases}$$

  where b is fixed and $p_0 > p_1 > \cdots > p_b$.
- **(A2)** The distortion function is given by $\ell(X_t, \hat{X}_t)$, where ℓ is monotone increasing and even function given by $\ell(X_t, \hat{X}_t) = \mathbb{1}\{X_t \neq \hat{X}_t\}$ or $\ell(X_t, \hat{X}_t) = |X_t - \hat{X}_t|^d$ for some $d > 0$. With a slight abuse of notation, we use $\ell(X_t - \hat{X}_t)$ to denote the distortion function.

Remark 1: Instead of assumption (A1), it was assumed in [5] that the Markov process evolves according to

$$X_{t+1} = X_t + N_t$$

where N_t has an even and a.s.u. (almost symmetric and unimodal) distribution. That assumption was motivated by the model in [4] where the (real-valued) Markov process evolves according to

$$X_{t+1} = aX_t + N_t$$

where N_t is a zero-mean Gaussian process. Our assumption (A1) is equivalent the assumption of [5]. In addition, [5] had assumed that the distribution of X_0 is even and a.s.u. We do not need that assumption because we assume that x_0 is fixed.

Definition 1: Let Z_t denote the most recent transmitted value of the Markov process. The process $\{Z_t\}_{t=0}^{\infty}$ evolves in a controlled Markov manner as follows:

$$Z_0 = x_0$$

and

$$Z_t = \begin{cases} X_t, & \text{if } U_t = 1 \\ Z_{t-1}, & \text{if } U_t = 0. \end{cases}$$

Note that since U_t can be inferred from the transmitted symbol Y_t, the estimator can also keep track of Z_t as follows:

$$Z_0 = x_0$$

and

$$Z_t = \begin{cases} Y_t, & \text{if } Y_t \neq \epsilon \\ Z_{t-1}, & \text{if } Y_t = \epsilon. \end{cases}$$

Theorem 1: *[5, Theorem 2], [4, Proposition 1] Consider the finite horizon version of Problem 2 under assumptions (A1) and (A2). The process $\{Z_t\}_{t=0}^{T}$ is a sufficient statistic at the estimator and an optimal estimation policy is given by:*

$$\hat{X}_t = g_t(Z_t) = Z_t. \tag{6}$$

In general, the optimal estimation policy depends on the choice of the communication policy and vice-versa. Theorem 1 shows that when the Markov process and the distortion function satisfy appropriate symmetry assumptions, the optimal estimation policy can be specified in closed form. Consequently, we can fix the estimator to be of the above form, and consider the centralized problem of identifying the best communication policy.

Definition 2: Let $E_t = X_t - Z_{t-1}$. The process $\{E_t\}_{t=0}^{\infty}$ evolves in a controlled Markov manner as follows:

$$\mathbb{P}(E_{t+1} = n \mid E_t = e, U_t = u) = \begin{cases} P_{0n} & \text{if } u = 1 \\ P_{en} & \text{if } u = 0. \end{cases}$$

Theorem 2: *[5, Theorem 3], [4, Theorem 1] Consider the finite horizon version of Problem 2 under assumptions (A1) and (A2). The process $\{E_t\}_{t=1}^{T}$ is a sufficient statistic at the sensor and an optimal communication policy is characterized by a sequence of thresholds $\{k_t\}_{t=1}^{T}$ i.e.,*

$$U_t = f_t(E_t) = \begin{cases} 1, & \text{if } |E_t| \ge k_t \\ 0, & \text{if } |E_t| < k_t. \end{cases}$$

The optimal policy mentioned above is given by the solution of the following dynamic programming

$$V_{T+1}(\cdot;\cdot) = 0 \tag{7}$$

and for $t = T, \dots, 0$

$$\begin{aligned} V_t(e;c) = \min\Big\{ & c + \sum_{n=-\infty}^{\infty} P_{0n} V_{t+1}(n;c), \\ & \ell(e) + \sum_{n=-\infty}^{\infty} P_{en} V_{t+1}(n;c) \Big\}. \end{aligned} \tag{8}$$

We are interested in the sensitivity of the optimal policy to a change in the communication cost c. For that reason, we parametrize value function with the communication cost c.

Lemma 3: *[23, Lemma 5] For all $t = 1, \cdots, T$, the value function $V_t(\cdot;\cdot)$ is even and increasing on $\mathbb{Z}_{\geq 0}$, i.e.*

$$V_t(-e;c) = V_t(e;c) \leq V_t(e+1;c),\ \forall e \in \mathbb{Z}_{\geq 0}.$$

E. Contributions of this paper

Theorem 2 shows that the optimal communication policy is threshold based and, hence, easy to implement. However, we still need to solve an appropriate dynamic program to identify the thresholds. In this paper, we first consider the infinite horizon discounted cost problem. From Theorem 1, the optimal estimation policy is time-invariant. As an intermediate step to study the average cost problem, we first mention the relevant results of [22] for discounted cost problem, which showed that under appropriate conditions the optimal communication policy is time-invariant threshold policy and is given by the fixed point of a dynamic program.

The average cost problem is then studied as an extension to the discounted cost problem as the discount factor goes to 1. Using vanishing discount approach, we give the average cost and the optimal thresholds as functions of the communication cost. We show that the communication index provides the complete characterization of the optimal communication policy for all values of the communication cost.

II. Results for the discounted cost set-up

A. Dynamic program for infinite horizon set-up

The structural result of Theorems 1 and 2 extend to the infinite horizon set-up. To show that the corresponding optimal policy is time-homogeneous, we assume the following:

(A3) There exists positive and finite constant μ_1 and μ_2 and a function $w\colon \mathbb{Z} \to \mathbb{R}$ such that for all $e \in \mathbb{Z}$

$$\max\{c, \ell(e)\} \leq \mu_1 w(e);$$

and

$$\max\Big\{ \sum_{n=-\infty}^{\infty} P_{en} w(n), \sum_{n=-\infty}^{\infty} P_{0n} w(n) \Big\} \leq \mu_2 w(e).$$

Theorem 4: *[22] Consider Problem 2 under assumptions (A1), (A2) and (A3). The process $\{E_t\}_{t=1}^{\infty}$ is a sufficient statistic at the sensor and an optimal communication policy is characterized by a time-invariant threshold k, i.e.,*

$$U_t = f(E_t) = \begin{cases} 1, & \text{if } |E_t| \geq k \\ 0, & \text{if } |E_t| < k. \end{cases} \tag{9}$$

The optimal policy mentioned above is given by the unique fixed point of the following dynamic programming

$$\begin{aligned} V(e;c) = \min\Big\{ & c + \beta \sum_{n=-\infty}^{\infty} P_{0n} V(n;c), \\ & \ell(e) + \beta \sum_{n=-\infty}^{\infty} P_{en} V(n;c) \Big\}. \end{aligned} \tag{10}$$

We are interested in the sensitivity of the optimal policy to a change in the communication cost c. For that reason, we parametrize value function with the communication cost c.

B. Performance of a threshold policy

Let $\mathcal{F}$ denote the class of all time-homogeneous threshold-based policies of the type (9). Let $f_k \in \mathcal{F}$ denote the policy with threshold k, $k \in \mathbb{N}$, i.e.

$$f_k(e) = \begin{cases} 1, & \text{if } |e| \geq k \\ 0, & \text{if } |e| < k. \end{cases}$$

Let $W_k(e;c)$ denote the performance of policy f_k when the system starts in state e and has a communication cost c. From standard results in Markov decision processes, W_k is the unique fixed point of the following fixed point equation:

$$W_k(e;c) = \begin{cases} c + \beta \sum_{n=-\infty}^{\infty} P_{0n} W_k(n;c), & \text{if } |e| \geq k, \\ \ell(e) + \beta \sum_{n=-\infty}^{\infty} P_{en} W_k(n;c), & \text{if } |e| < k \end{cases} \tag{11}$$

Define square matrices $P^{(k)}$ and $Q^{(k)}$ and a column vector $\ell^{(k)}$ indexed by $I^{(k)} = \{-k+1, \cdots, k-1\}$ as follows:

$$P_{ij}^{(k)} = P_{ij}, \quad i,j \in I^{(k)}; \tag{12}$$

$$Q^{(k)} = [I_{2k-1} - \beta P^{(k)}]^{-1}; \tag{13}$$

$$\ell^{(k)} = [\ell(-k+1), \ell(-k+2), \cdots, \ell(k-2), \ell(k-1)]^{\intercal}. \tag{14}$$

Lemma 5: *[23] For all $n \in \mathbb{N}$,*

$$\Big\langle (P^{(k)})_0^n, \mathbf{1}_{2k-1} \Big\rangle < \Big\langle (P^{(k+1)})_0^n, \mathbf{1}_{2k+1} \Big\rangle, \tag{15}$$

$$\Big\langle (P^{(k)})_0^n, \ell^{(k)} \Big\rangle < \Big\langle (P^{(k+1)})_0^n, \ell^{(k+1)} \Big\rangle. \tag{16}$$

Let τ_k denote the stopping time when the Markov process starting at state 0 at time $t = 0$ enters the set $\{e \in \mathbb{Z} : |e| \geq k\}$. Define L_k and T_k as follows

$$L_k = \mathbb{E}\Big[\sum_{t=0}^{\tau_k - 1} \beta^t \ell(E_t) \,\Big|\, E_0 = 0 \Big]$$

and

$$T_k = \frac{1 - \mathbb{E}[\beta^{\tau_k}]}{1 - \beta}.$$

Then, the closed form expression for L_k and T_k is given by the following lemma.

Lemma 6: *[22]*

1) *$W_k(0;c)$ can be written as*

$$W_k(0;c) = \frac{L_k + c}{(1-\beta)T_k} - c \tag{17}$$

2) *L_k and T_k defined above can be expressed in a closed form as follows:*

$$L_k = \left\langle Q_0^{(k)}, \ell^{(k)} \right\rangle; \tag{18}$$

$$T_k = \left\langle Q_0^{(k)}, \mathbf{1}_{2k-1} \right\rangle, \tag{19}$$

where $Q_0^{(k)}$ denotes the row with index 0 in $Q^{(k)}$. Substituting these in (17), we get a closed form expression of $W_k(0;c)$.

3) *$T_k < T_{k+1}$ and $L_k < L_{k+1}$.*

Lemma 7: *[22] The value function $V(et;\cdot)$, as given in (10, is a piece-wise linear and concave function of c for all $e \in \mathbb{Z}$.*

C. Characterization of optimal policy

(A4) L_k/T_k is increasing in k.

Theorem 8: *[22] Define c_k as*

$$c_k = \left(\frac{L_{k+1}}{T_{k+1}} - \frac{L_k}{T_k}\right) \Bigg/ \left(\frac{1}{T_k} - \frac{1}{T_{k+1}}\right). \tag{20}$$

Suppose (A4) holds and $\{c_k\}_{k=1}^{\infty}$ is an increasing sequence. Then, for all $c \in (c_k, c_{k+1}]$ such that $c_k \neq c_{k+1}$, the policy f_{k+1} is discounted cost optimal.

III. Main results for the average cost set-up

We follow the vanishing discount approach and show that the optimal policy for the average cost set-up is given as a limit of the discounted cost set-up as the discount factor $\beta \to 1$. This approach relies on the following theorem. Note that to capture the dependence of the discount factor, in the sequel we use the subscript or superscript β with $W_{\beta,k}$, $Q^{(\beta,k)}$, $L_{\beta,k}$, $T_{\beta,k}$ and $c_{\beta,k}$.

Theorem 9: *[24, Theorem 7.2.3] Let V_β and g_β be the value function and optimal policy for a countable state MDP for the discounted cost set-up with discount factor β. Suppose the value function V_β satisfies the following SEN conditions.*

(S1) *There exists a refernece state $e_0 \in \mathbb{Z}$, such that $(1-\beta)V_\beta(e_0) < \infty$ for all $\beta \in (0,1)$.*

(S2) *Define $h_\beta(e) = V_\beta(e) - V_\beta(e_0)$. There exists a function $M : \mathbb{Z} \to \mathbb{R}$ such that $h_\beta(e) \leq M(e)$ for $e \in \mathbb{Z}$ and $\beta \in (0,1)$.*

(S3) *There exists a nonegative (finite) constant L such that $-L \leq h_\beta(e)$ for $e \in \mathbb{Z}$ and $\beta \in (0,1)$.*

Then,

i The optimal average cost is given by

$$J \equiv \lim_{\beta \to 1^-} (1-\beta)V_\beta(e). \tag{21}$$

The limit exists and does not depend on e.

ii Let $\bar{g}$ be any limit point of g_β. Then $\bar{g}$ is average cost optimal.

We will show that the value functions in the discounted cost set-up derived in Section II satisfy the SEN conditions. To show that, we introduce the notion of a z-standard policy.

Definition 3: Consider a Markov chain with state space $\mathcal{X}$. For $i, j \in \mathcal{X}$. Let m_{ij} denote the expected time to go from i to j for the first time and c_{ij} denotes the expected cost of a first passage from i to j. A Markov chain is called *z-standard*, $z \in \mathcal{X}$, if $m_{ij} < \infty$ and $c_{ij} < \infty$ for all $i \in \mathcal{X}$.

Definition 4: Let d be a (possibly randomized) stationary policy for a Markov decision process. Then d is a *z-standard policy* if the Markov chain induced by d is z-standard.

Lemma 10: *The policy f_0 is 0-standard.*

Proof: The policy f_0 is essentially an 'always transmit' policy. Then, for any starting state e, the induced Markov chain goes to the state 0 instantaneously and then goes to some next state n with probability P_{0n}. Hence we have $m_{e0} < \infty$. Also, the expected cost is given by $c_{e0} = \frac{c}{1-\beta} < \infty$. Hence, f_0 is 0-standard. ∎

Proposition 11: *The value function V_β (defined by (10)) satisfies SEN conditions.*

Proof: Proposition 7.5.3 of [24] states that if there exists a z-standard policy, then (S1)-(S2) hold for reference state z. Hence in our model (S1)-(S2) hold for reference state 0. Furthermore, an immediate consequence of Lemma 3 that $V_\beta(e) \geq V_\beta(0)$. Hence (S3) holds for $L = 0$. ∎

Let $J(c)$ denote the optimal average cost when the communication cost is c. Since our model satisfies the SEN conditions (Proposition 11), the optimal average cost $J(c)$ and an optimal policy are given by Theorem 9. In particular

$$J(c) = \lim_{\beta \to 1^-} (1-\beta)V_\beta(0). \tag{22}$$

Theorem 8 gives the value of communication costs for which policy f_{k+1} is optimal. Proposition 11 gives the corresponding value function (which equals $W_{\beta,k}$). In the next section, we combine these two results to identify a closed form expression for $J(c)$ and the optimal policy.

A. Structure of optimal policies

Recall the definition of $P^{(k)}$ and $\ell^{(k)}$ given by (12) and (14) and define

$$\bar{Q}^{(k)} = [I_{2k-1} - P^{(k)}]^{-1}.$$

Note that $\bar{Q}^{(k)}$ exists since $P^{(k)}$ is substochastic [25].

Define

$$\bar{L}_k = \left\langle \bar{Q}_0^{(k)}, \ell^{(k)} \right\rangle \tag{23}$$

$$\bar{T}_k = \left\langle \bar{Q}_0^{(k)}, \mathbf{1}_{2k-1} \right\rangle \tag{24}$$

where $\bar{Q}_0^{(k)}$ denotes the row with index 0 in $\bar{Q}^{(k)}$ and

$$\bar{c}_k = \left(\frac{\bar{L}_{k+1}}{\bar{T}_{k+1}} - \frac{\bar{L}_k}{\bar{T}_k}\right) \Bigg/ \left(\frac{1}{\bar{T}_k} - \frac{1}{\bar{T}_{k+1}}\right). \tag{25}$$

Note that $\bar{c}_k$ is well-defined as the following lemma holds.

Lemma 12: *$\bar{L}_k < \bar{L}_{k+1}$ and $\bar{T}_k < \bar{T}_{k+1}$.*

Proof: Since $P^{(k)}$ is a sub-stochastic matrix, $\bar{Q}^{(k)}$ exists and is given by

$$\bar{Q}^{(k)} = I + P^{(k)} + (P^{(k)})^2 + \cdots . \tag{26}$$

Thus,

$$\begin{aligned}\bar{T}_k &= \left\langle \bar{Q}_0^{(k)}, \mathbf{1}_{2k-1} \right\rangle \\ &= \left\langle I_0, \mathbf{1}_{2k-1} \right\rangle + \left\langle P_0^{(k)}, \mathbf{1}_{2k-1} \right\rangle \\ &\qquad + \left\langle (P^{(k)})_0^2, \mathbf{1}_{2k-1} \right\rangle + \cdots \end{aligned} \tag{27}$$

$$\begin{aligned}\bar{L}_k &= \left\langle \bar{Q}_0^{(k)}, \ell^{(k)} \right\rangle \\ &= \left\langle I_0, \ell^{(k)} \right\rangle + \left\langle P_0^{(k)}, \ell^{(k)} \right\rangle \\ &\qquad + \left\langle (P^{(k)})_0^2, \ell^{(k)} \right\rangle + \cdots \end{aligned} \tag{28}$$

The result follows from (27), (28) and Lemma 5. ■

Lemma 13: *We have the following*

(i) $\lim_{\beta\to 1^-} Q^{(\beta,k)} = \bar{Q}^{(k)}$
(ii) $\lim_{\beta\to 1^-} L_{\beta,k} = \bar{L}_k$
(iii) $\lim_{\beta\to 1^-} T_{\beta,k} = \bar{T}_k$
(iv) $\lim_{\beta\to 1^-} c_{\beta,k} = \bar{c}_k$.

Proof: (i) follows from the continuity of $Q^{(\beta,k)}$ in β; (ii) and (iii) follow from (i) and the continuity of inner product; since $c_{\beta,k}$ is continuous in β, (iv) follows from (ii) and (iii). ■

Theorem 14: *Suppose that (A4) holds and $\{\bar{c}_k\}$ is increasing. Then for all $c \in (\bar{c}_k, \bar{c}_{k+1}]$ the policy f_{k+1} is optimal and corresponding average cost is given by*

$$J(c) = \frac{\bar{L}_{k+1} + c}{\bar{T}_{k+1}}. \tag{29}$$

Proof: Consider a $c \in (\bar{c}_k, \bar{c}_{k+1}]$. From Lemma 13, it follows that $\exists \beta^* \in (0,1)$ such that $\forall \beta \in (\beta^*, 1)$, $c \in (c_{\beta,k}, c_{\beta,k+1}]$. By Theorem 8, the policy f_{k+1} is optimal for $\beta \in (\beta^*, 1)$. Hence, by Theorem 9 (ii), the policy f_{k+1} is also optimal for the average cost set-up.

Since policy f_{k+1} is optimal for $\beta \in (\beta^*, 1)$, the value function is given by $V_{\beta,k+1}(e;c)$. By Theorem 9 (i) and (17), we get that

$$J(c) = \lim_{\beta\to 1^-} (1-\beta) V_{\beta,k+1}(0;c) = \frac{\bar{L}_{k+1} + c}{\bar{T}_{k+1}}.$$

■

Proposition 15: *$J(c)$ is piece-wise linear, increasing and concave in c.*

Proof: From the expression of $J(c)$ as given in (29), we see that $J(c)$ is a linear and increasing in c, for all $c \in (\bar{c}_k, \bar{c}_{k+1})$. Also, from Theorem 14 we have that

$$\begin{aligned} J(\bar{c}_k^-) &= \frac{\bar{L}_k + \bar{c}_k}{\bar{T}_k} \\ J(\bar{c}_k^+) &= \frac{\bar{L}_{k+1} + \bar{c}_k}{\bar{T}_{k+1}}. \end{aligned}$$

Following (25), we have that $J(\bar{c}_k^-) = J(\bar{c}_k^+)$ and hence $J(\cdot)$ is a continuous function of c.

Now, let us consider two consecutive intervals $(\bar{c}_k, \bar{c}_{k+1}]$ and $(\bar{c}_{k+1}, \bar{c}_{k+2}]$. We know that

$$\begin{aligned} J(c_k^+) &= \frac{\bar{L}_{k+1} + \bar{c}_k}{\bar{T}_{k+1}} \\ J(c_{k+1}) = J(c_{k+1}^-) &= \frac{\bar{L}_{k+1} + \bar{c}_{k+1}}{\bar{T}_{k+1}} \\ = J(c_{k+1}^+) &= \frac{\bar{L}_{k+2} + \bar{c}_{k+1}}{\bar{T}_{k+2}} \\ J(c_{k+2}) = J(c_{k+2}^-) &= \frac{\bar{L}_{k+2} + \bar{c}_{k+2}}{\bar{T}_{k+2}}. \end{aligned}$$

Hence, the slope of J in $(\bar{c}_k, \bar{c}_{k+1}]$ is given by

$$\begin{aligned} &\frac{J(c_{k+1}) - J(c_k^+)}{c_{k+1} - c_k} \\ &= \left(\frac{\bar{L}_{k+1} + \bar{c}_{k+1}}{\bar{T}_{k+1}} - \frac{\bar{L}_{k+1} + \bar{c}_k}{\bar{T}_{k+1}}\right) \Big/ (\bar{c}_{k+1} - \bar{c}_k) \\ &= \frac{1}{\bar{T}_{k+1}}. \end{aligned}$$

Similarly, we have that the slope of J in $(\bar{c}_{k+1}, \bar{c}_{k+2}]$ is $1/\bar{T}_{k+2}$. Since by Lemma 12 $\bar{T}_{k+1} < \bar{T}_{k+2}$, we have that the slope of J is decreasing in c. This completes the proof. ■

IV. An Example: Aperiodic, Symmetric Birth-Death Markov Chain

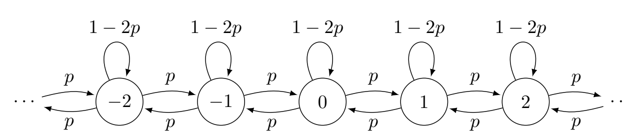

Fig. 1. A birth-death Markov chain

Consider a birth-death Markov chain, shown in Fig. 1, with transition probabilities as follows:

$$P_{ij} = \begin{cases} p, & \text{if } j = i+1, i-1 \\ 1-2p, & \text{if } j = i \\ 0, & \text{otherwise.} \end{cases}$$

where $p \in (0, \frac{1}{2})$. Let the distortion function $\ell(\cdot)$ to be $\ell(e) = |e|$. Note that P and ℓ satisfy (A1) and (A2).

Lemma 16: *(A3) is always satisfied for the above model. The values of the function $w(\cdot)$ and the parameters μ_1 and μ_2 are given by:*

$$w(e) = \max\{c, \ell(e)\}, \mu_1 = 1,$$
$$\mu_2 = \max\{1 - 2p + 2p/c, \ell(2)\}$$

Proof: The result may be verified separately for $\ell(e) = \mathbb{1}\{e \neq 0\}$ and $\ell(e) = |e|^d$ by substitution. ∎

Since (A3) is satisfied, the optimal communication policy is a time-homogeneous threshold policy. Hence, the framework of sections II-B–II-C is applicable. The communication indices $\{c_{\beta,k}\}_{k=1}^{\infty}$ depend on the values of $L_{\beta,k}$ and $T_{\beta,k}$, which in turn depend on the matrix $Q^{(\beta,k)}$. Since, $Q^{(\beta,k)}$ is the inverse of a tridiagonal symmetric Toeplitz matrix, an explicit formula for its elements is available [26].

Lemma 17: *Define $D = -2 - (1-\beta)/(\beta p)$ and $\lambda = \cosh^{-1}(-D/2)$. As $\beta \to 1^-$, we have*

$$\bar{D} = \lim_{\beta \to 1^-} D = -2.$$

Then $\bar{Q}^{(k)}$ is given by

$$\bar{Q}_{ij}^{(k)} = \frac{(2k + i + j - |j - i|)(2k - i - j - |j - i|)}{8kp}, \quad (30)$$

where $i, j \in I^{(k-1)}$. In particular $\bar{Q}_{0j}^{(k)}$ is given by:

$$\bar{Q}_{0j}^{(k)} = \frac{(2k + j - |j|)(2k - j - |j|)}{8kp}. \quad (31)$$

Proof: The matrix $I_{2k-1} - P^{(k)} \in \mathbb{R}^{2k-1 \times 2k-1}$ is a symmetric tridiagonal matrix given by

$$I_{2k-1} - P^{(k)} = -p \begin{bmatrix} \bar{D} & 1 & 0 & 0 & \cdots & \cdots & 0 \\ 1 & \bar{D} & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \bar{D} & 1 & \cdots & \cdots & 0 \\ 0 & \cdots & 1 & \bar{D} & 1 & \cdots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 & \bar{D} & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & \bar{D} \end{bmatrix}$$

$\bar{Q}^{(k)}$ is the inverse of the above matrix. The inverse of tridiagonal Toeplitz matrix in the above form are computed in closed form in [26].

Following [26], we get the result of the lemma as a limiting case of the $Q^{(\beta,k)}$ by solving finite difference equation associated with the one-dimensional Poisson equation with Dirichlet boundary conditions for the case $\bar{D} = -2$, i.e. for $\lambda = 0$. ∎

Lemma 18: *For the above birth-death Markov chain with $\ell(e) = |e|$,*

$$\bar{L}_k = \frac{k(k^2 - 1)}{6p} \quad (32a)$$

$$\bar{T}_k = \frac{k^2}{2p} \quad (32b)$$

$$\bar{c}_k = \frac{k(k+1)(k^2 + k + 1)}{6p(2k+1)}. \quad (32c)$$

Proof: Combining the results of Lemmas 6 and 17, we get that

$$\bar{L}_k = \sum_{e=-k+1}^{k-1} \bar{Q}_{0e}^{(k)} \ell(e),$$
$$\text{and } \bar{T}_k = \sum_{e=-k+1}^{k-1} \bar{Q}_{0e}^{(k)},$$

where $Q_{0e}^{(\beta,k)}$ is given by (31). Simplifying the above expressions we get (32a) and (32b). (32c) is obtained by plugging these expressions for $\bar{L}_k$ and $\bar{T}_k$ in (25). ∎

Lemma 19: *For the above birth-death Markov chain with $\ell(e) = |e|$, $\bar{c}_k$ is increasing in k.*

Proof: Using (32c) we have

$$\bar{c}_{k+1} - \bar{c}_k = \frac{(k+1)(6k^3 + 20k^2 + 16k + 6)}{6p(2k+3)(2k+1)}.$$

Hence $\bar{c}_{k+1} \geq \bar{c}_k$ since $k \in \mathbb{Z}_{\geq 0}$. ∎

For the above birth-death Markov chain, the average cost $J(\cdot)$ is piece-wise linear, increasing and concave function of the communication cost c which is shown in Fig. 2. The values of the corresponding $\bar{c}_k$ are shown in the abscissa ($\bar{c}_1 = 1.11$ is not shown due to dearth of space).

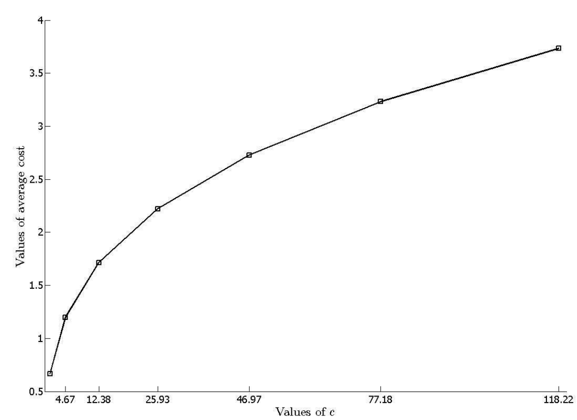

Fig. 2. The plot of average cost as a function of communication cost for the example with $p = 0.3$.

V. Conclusion

In this paper, we study a remote sensing problem with communication cost and investigate the average cost optimization problem. We follow vanishing discount approach and extend the results of infinite horizon discounted cost problem to obtain the desired results. We obtain an explicit characterization of the communication indices that represent the value of communication cost for which one is indifferent between two consecutive threshold policies. We provide closed form expressions of these communication thresholds and use them to completely characterize the optimal communication policy for all values of the communication cost.

REFERENCES

[1] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

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

[3] Y. Xu and J. P. Hespanha, "Optimal communication logics in networked control systems," *43rd IEEE Conference on Decision and Control*, vol. 4, no. 7, pp. 3527 – 3532, 1996.

[4] 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, 2011.

[5] 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, 2013.

[6] D. Shuman and M. Liu, "Optimal sleep scheduling for a wireless sensor network node," in *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, October 2006, pp. 1337–1341.

[7] M. Athans, "On the determination of optimal costly measurement strategies for linear stochastic systems," *Automatica*, vol. 8, no. 4, pp. 397–412, 1972.

[8] J. Geromel, "Global optimization of measurement startegies for linear stochastic systems," *Automatica*, vol. 25, no. 2, pp. 293–300, 1989.

[9] W. Wu, A. Araposthathis, and V. V. Veeravalli, "Optimal sensor querying: General markovian and LQG models with controlled observations," *IEEE Transactions on Automatic Control*, vol. 53, no. 6, pp. 1392–1405, 2008.

[10] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication rate scheme for distributed detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 554–568, April 1996.

[11] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Decentralized detection with censoring sensors," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1362–1373, April 2008.

[12] H. S. Witsenhausen, "On the structure of real-time source coders," *The Bell System Technical Journal*, vol. 58, no. 6, pp. 1437–1451, 1979.

[13] Y. Kaspi and N. Mehrav, "Structure theorems for real-time variable rate coding with and without side information," *IEEE Transactions on Information Theory*, vol. 58, no. 12, pp. 7135–7153, December 2012.

[14] T. Henningsson, E. Johannesson, and A. Cervin, "Sporadic event-based control of first-order linear stochastic systems," *Automatica*, vol. 44, no. 11, pp. 2890–2895, 2008.

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

[16] K. J. Åström, *Analysis and Design of Nonlinear Control Systems*. Berlin, Heidelberg: Springer, 2008, ch. Event based control.

[17] M. Rabi, "Packet based inference and control," *PhD thesis, Institute for Systems Research, University of Maryland*, 2006.

[18] M. Rabi, G. Moustakides, and J. Baras, "Adaptive sampling for linear state estimation," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 672–702, 2012.

[19] X. Meng and T. Chen, "Optimal sampling and performance comparison of periodic and event based impulse control," *IEEE Transactions of Automatic Control*, vol. 57, no. 12, pp. 3252–3259, 2012.

[20] A. Molin and S. Hirche, "On LQG joint optimal scheduling and control under communication constratints," in *Proceedings of 48th IEEE Conference on Decision and Control*, 2009, pp. 5832–5838.

[21] ——, "Structural characterization of optimal event-based controllers for linear stochastic systems," in *Proceedings of 49th IEEE Conference on Decision and Control*, 2010, pp. 3227–3233.

[22] J. Chakravorty and A. Mahajan, "On the optimal thresholds in remote state estimation with communication costs," *accepted in the Proceedings of 53rd IEEE Conference on Decision and Control*, 2014.

[23] ——, "Optimal real-time transmission of a markov source under constraints on the number of transmissions," *in preparation*, 2014.

[24] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*. NY, USA: John Wiley and Sons Inc., Publication, 1999.

[25] P. J. Weeda, *Finite generalized Markov programming*, ser. Mathematical Centre tracts. Amsterdam: Mathematisch Centrum, 1979, vol. 92.

[26] G. Y. Hu and R. F. O'Connell, "Analytical inversion of symmetric tridiagonal matrices," *Journal of Physics A: Mathematical and General*, vol. 29, no. 7, 1996.