Remote Estimation Over a Packet-Drop Channel With Markovian State

Jhelum Chakravorty, *Member, IEEE*, and Aditya Mahajan, *Senior Member, IEEE*

***Abstract*—We investigate a remote estimation problem in which a transmitter observes a Markov source and chooses the power level to transmit it over a time-varying packet-drop channel. The channel is modeled as a channel with Markovian state where the packet drop probability depends on the channel state and the transmit power. A receiver observes the channel output and the channel state and estimates the source realization. The receiver also feeds back the channel state and an acknowledgment for successful reception to the transmitter. We consider two models for the source—finite state Markov chains and first-order autoregressive processes. For the first model, using ideas from team theory, we establish the structure of optimal transmission and estimation strategies and identify a dynamic program to determine optimal strategies with that structure. For the second model, we assume that the noise process has unimodal and symmetric distribution. Using ideas from majorization theory, we show that the optimal transmission strategy is symmetric and monotonic and the optimal estimation strategy is like Kalman filter. Consequently, when there are a finite number of power levels, the optimal transmission strategy may be described using thresholds that depend on the channel state. Finally, we propose a simulation-based approach (renewal Monte Carlo) to compute the optimal thresholds and optimal performance and elucidate the algorithm with an example.**

***Index Terms*—Real-time communication, remote estimation, renewal theory, stochastic approximation, symmetric and quasi-convex value and optimal strategies.**

I. Introduction

A. Motivation and Literature Overview

NETWORK control systems are distributed systems where plants, sensors, controllers, and actuators are interconnected via a communication network. Such systems arise in a variety of applications such as Internet of Things, smart grids, vehicular networks, robotics, etc. One of the fundamental problem in network control system is *remote estimation*—how should a sensor (which observes a stochastic process) transmit its observations to a receiver (which estimates the state of the stochastic process) when there is a constraint on communication, either in terms of communication cost or communication rate.

Manuscript received July 23, 2018; revised February 24, 2019; accepted June 22, 2019. Date of publication July 1, 2019; date of current version April 23, 2020. This work was supported by the Natural Sciences and Engineering Research Council of Canada under Discovery Accelerator Grant 493011. This paper was presented in part at the 2016 IFAC Workshop on Distributed Estimation and Control in Networked Systems, Tokyo, Japan, Sep. 8-9, in part at the 2017 International Symposium of Information Theory, Aachen, Germany, Jun. 25–30, and in part at the 2017 American Control Conference, Seattle, WA, USA, May 24–26. Recommended by Associate Editor L. Zhang. *(Corresponding author: Jhelum Chakravorty.)*

The authors are with the Department of Electrical and Computer Engineering, the School of Computer Science, McGill University, Montreal, QC H3A 0G4, Canada (e-mail: jhelum.chakravorty@mail.mcgill.ca; aditya.mahajan@mcgill.ca).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors.

Digital Object Identifier 10.1109/TAC.2019.2926160

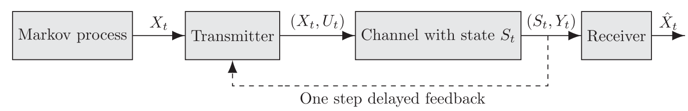

Fig. 1. Remote estimation over channel with state.

In this paper, we consider a remote estimation system as shown in Fig. 1. The system consists of a sensor and an estimator connected over a time-varying wireless fading channel. The sensor observes a Markov process and chooses the power level to transmit its observation to the remote estimator. Communication is noisy and the transmitted packet may get dropped according to a probability that depends on the channel state and the power level. When the packet is dropped the receiver generates an estimate of the state of the source according to previously received packets. The objective is to choose power control and estimation strategies to minimize a weighted sum of transmission power and estimation error.

Several variations of the abovementioned model have been considered in the literature. Models with noiseless communication channels have been considered in [1]–[6]. Since the channel is noiseless, these papers assume that there are only two power levels: power level 0, which corresponds to not transmitting; and power level 1, which corresponds to transmitting. Under slightly different modeling assumptions, these papers identify the structure of optimal transmission and estimation strategies for first-order autoregressive sources with unimodal noise and for higher order autoregressive sources with orthogonal dynamics and isotropic Gaussian noise. It is shown that the optimal transmission strategy is threshold based, i.e., the sensor transmits whenever the current *error* is greater than a threshold. It is also shown that the optimal estimation strategy is like Kalman filter: when the receiver receives a packet, the estimate is the received symbol; when it does not receive the packet, then the estimate is the one-step prediction based on the previous symbol. Quite surprisingly, these results show that there is no advantage in trying to extract information about the source realization from the choice of the power levels. The transmission strategy at the sensor is also called *event-triggered* communication because the sensor transmits when the event "error is greater than a threshold" is triggered. Models with independent and identically distributed packet-drop channels are considered in [7]–[9], where it is assumed that the transmitter has two power levels: ON or OFF. Remote estimation over additive noise channel is considered in [10].

In this paper, we consider a remote estimation problem over packet-drop channel with Markovian state. We assume that the receiver observes the channel state and feeds it back to the transmitter with one step delay. Preliminary results for this model are presented in [11], where attention was restricted to a binary state channel with two input power values (ON or OFF). In this paper, we consider arbitrary number of channel states and power levels. A related paper is [12], in which a remote estimation over packet-drop channels with Markovian state is considered. It is assumed that the sensor and the receiver know the channel state. It is shown that optimal estimation strategies are like Kalman filter. A detailed comparison with [12] is presented in Section VI-A.

Several approaches for computing the optimal transmission strategies have been proposed in the literature. For noiseless channels, these include dynamic programming based approaches [4]–[6], approximate dynamic programming based approaches [13], and renewal theory based approaches [14]. It is shown in [15] that for event-triggered scheduling, the posterior density follows a generalized closed skew normal distribution. For Markovian channels (when the state is not observed), a change of measure technique to evaluate the performance of an event-triggered scheme is presented in [16]. In this paper, we present a renewal theory based Monte Carlo approach for computing the optimal thresholds. A preliminary version of the results was presented in [9] for a channel with independent and identically distributed packet drops.

If the transmitter always transmits at a constant power level, the remote estimation model reduces to that of Kalman filtering with intermittent observations. For channels with independent and identically distributed packet drops, necessary and sufficient conditions for mean-square stability are presented in [17]–[19]. For channels with Markov packet drops, sufficient conditions for peak-covariance stability and mean-square stability are presented in [20]–[23].

B. Contributions

In this paper, we investigate team optimal transmission and estimation strategies for remote estimation over time varying packet-drop channels. We consider two models for the source: finite state Markov source and first order autoregressive source (over either integers or reals). Our main contributions are as follows.

1) For finite sources, we identify sufficient statistics for both the transmitter and the receiver and obtain a dynamic programming decomposition to compute optimal transmission and estimation strategies.
2) For autoregressive sources, we identify qualitative properties of optimal transmission and estimation strategies. In particular, we show that the optimal estimation strategy is like Kalman filter and the optimal transmission strategy only depends on the current source realization and the previous channel state (and does not depend on the receiver's belief of the source). Furthermore, when the channel state is stochastically monotone (see Assumption 1 for definition), then for any value of the channel state, the optimal transmission strategy is symmetric and quasi-convex in the source realization. Consequently, when the power levels are finite, the optimal transmission strategy is threshold based, where the thresholds only depend on the previous channel state.
3) We show that the abovementioned qualitative properties extend naturally to infinite horizon models.
4) For infinite horizon models, we present a renewal theory based Monte Carlo algorithm to evaluate the performance of any threshold-based strategy. We then combine it with a simultaneous perturbation based stochastic approximation algorithm to compute the optimal thresholds. We illustrate our results with a numerical example of a remote estimation problem with a transmitter with two power levels and a Gilbert–Elliott erasure channel.
5) We show that the problem of transmitting over one of m available independent and identically distributed packet-drop channels (at a constant power level) can be considered as special case of our model. We show that there exist thresholds $\{k_t^{(i)}\}_{i=1}^m$, such that it is optimal to transmit over channel i if the *error state* $E_t \in [k_t^{(i)}, k_t^{(i+1)})$. See Section VI-C for details.

C. Notation

We use uppercase letters to denote random variables (e.g., X, Y, etc), lowercase letters to denote their realizations (e.g., x, y, etc.). $\mathbb{Z}$, $\mathbb{Z}_{\geq 0}$, and $\mathbb{Z}_{>0}$ denote, respectively, the sets of integers, of nonnegative integers, and of positive integers. Similarly, $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{R}_{>0}$ denote, respectively, the sets of reals, of nonnegative reals, and of positive reals. For any set $\mathbb{A}$, $\mathbb{1}_{\{x\in\mathbb{A}\}}$ denotes the indicator function of event $x \in \mathbb{A}$, i.e., $\mathbb{1}_{\{x\in\mathbb{A}\}}$ is 1 if $x \in \mathbb{A}$, else 0. $|\mathbb{A}|$ denotes the cardinality of set $\mathbb{A}$. $\mathbb{P}$ and $\mathbb{E}$ denote the probability and expectation of a random variable. $\Delta(\mathbb{X})$ denotes the space of probability distributions of $\mathbb{X}$. For any vector $v \in \mathbb{R}^n$, $v_i \in \mathbb{R}$ denotes the ith component of v. For any vector v and an interval $A = [a, b]$ of $\mathbb{R}$, $w = [v]_A$ means that w_i equals a if $v_i \leq a$; equals v_i if $v_i \in (a, b)$; and equals b if $v_i \geq b$. Given a Borel subset $A \subseteq \mathbb{R}$ and a density π, we use the notation $\pi(A) := \int_A \pi(e)de$. For any vector $\mathbf{v}$, $\nabla_{\mathbf{v}}$ denotes the derivative with respect to $\mathbf{v}$.

II. System Model and Problem Formulation

A. Communication System

We consider a remote estimation system shown in Fig. 1. The different components of the system are explained as follows.

1) Source Model: The source is a first-order time-homogeneous Markov chain $\{X_t\}_{t\geq 0}$, $X_t \in \mathbb{X}$. We consider two models for the source.

1) *Finite state Markov source:* In this model, we assume that $\mathbb{X}$ is a finite set and denote the state transition matrix by P, i.e., for any $x, y \in \mathbb{X}$, $P_{xy} = \mathbb{P}(X_{t+1} = y \,|\, X_t = x)$.
2) *First-order autoregressive source:* In this model, we assume that $\mathbb{X}$ is either $\mathbb{Z}$ or $\mathbb{R}$. The initial state $X_0 = 0$ and for $t \geq 0$, the source evolves as

$$X_{t+1} = aX_t + W_t \tag{1}$$

where $a, W_t \in \mathbb{X}$ and $\{W_t\}_{t\geq 0}$ is an independent and identically distributed sequence where W_t is distributed according to a symmetric and unimodal distribution[1] μ.

2) Channel Model: The channel is a packet-drop channel with state. The state process $\{S_t\}_{t\geq 0}$, $S_t \in \mathbb{S}$, is a first-order time-homogeneous Markov chain with transition probability matrix Q. We assume that $\mathbb{S}$ is finite. This is a standard model for time-varying wireless channels [24], [25].

The input alphabet of the channel is $\mathbb{X}$ and the output alphabet is $\mathbb{Y} := \mathbb{X} \cup \{\mathfrak{E}\}$ where the symbols $\mathfrak{E}$ denotes that no packet was received. At time t, the channel output is denoted by Y_t.

The packet drop probability depends on the input power $U_t \in \mathbb{U}$, where $\mathbb{U}$ is the set of allowed power levels. We assume that $\mathbb{U}$ is a subset of $\mathbb{R}_{\geq 0}$ and $\mathbb{U}$ is either a finite set of the form $\{0, u_{(1)}, \dots, u_{\max}\}$ or an interval of the form $[0, u_{\max}]$, i.e., $\mathbb{U}$ is uncountable. When $U_t = 0$, it means that the transmitter does not send a packet. In particular, for any realization $(x_{0:T}, s_{0:T}, u_{0:T}, y_{0:T})$ of $(X_{0:T}, S_{0:T}, U_{0:T}, Y_{0:T})$, we have

$$\begin{aligned}\mathbb{P}(S_t = s_t \mid X_{0:t} = x_{0:t}, S_{0:t-1} = s_{0:t-1}, U_{0:t} = u_{0:t})\\ = \mathbb{P}(S_t = s_t \mid S_{t-1} = s_{t-1}) = Q_{s_{t-1}s_t}\end{aligned} \tag{2}$$

and

$$\begin{aligned}&\mathbb{P}(Y_t = y_t \mid X_{0:t} = x_{0:t}, S_{0:t} = s_{0:t}, U_{0:t} = u_{0:t})\\ &= \begin{cases} 1 - p(s_t, u_t), & \text{if } y_t = x_t\\ p(s_t, u_t), & \text{if } y_t = \mathfrak{E}\\ 0, & \text{otherwise}\end{cases}\end{aligned} \tag{3}$$

where $p(s_t, u_t)$ is the probability that a packet transmitted with power level u_t when the channel is in state s_t is dropped. We assume that the set $\mathbb{S}$ of the channel states is an ordered set where a larger state means a better channel quality. Then, for all $s \in \mathbb{S}$, $p(s, u)$ is (weakly) decreasing in u with $p(s, 0) = 1$ and $p(s, u_{\max}) \geq 0$. Furthermore, we assume that for all $u \in \mathbb{U}$, $p(s, u)$ is decreasing in s.

B. Decision Makers and the Information Structure

There are two decision makers in the system—the transmitter and the receiver. At time t, the transmitter chooses the transmit power U_t while the receiver chooses an estimate $\hat{X}_t \in \mathbb{X}$. Let I_t^1 and I_t^2 denote the information sets at the transmitter and the receiver, respectively.

The transmitter observes the source realization X_t. In addition, there is one-step delayed feedback from the receiver to the transmitter.[2] Thus, the information available at the transmitter is

$$I_t^1 = \{X_{0:t}, U_{0:t-1}, S_{0:t-1}, Y_{0:t-1}\}.$$

The transmitter chooses the transmit power U_t according to

$$U_t = f_t(I_t^1) = f_t(X_{0:t}, U_{0:t-1}, S_{0:t-1}, Y_{0:t-1}) \tag{4}$$

where f_t is called the *transmission rule* at time t. The collection $\boldsymbol{f} := (f_0, f_1, \dots)$ for all time is called the *transmission strategy*.

The receiver observes Y_t and, in addition, observes the channel state S_t. Thus, the information available at the receiver is

$$I_t^2 = \{S_{0:t}, Y_{0:t}\}.$$

[1]With a slight abuse of notation, when $\mathbb{X} = \mathbb{R}$, we consider μ to the probability density function and when $\mathbb{X} = \mathbb{Z}$, we consider μ to be the probability mass function.

[2]Note that feedback of Y_t requires 1 bit to indicate whether the packet was received or not and feedback of S_t requires $\lceil \log_2 |\mathbb{S}| \rceil$ bits.

The receiver chooses the estimate $\hat{X}_t$ as follows:

$$\hat{X}_t = g_t(I_t^2) = g_t(S_{0:t}, Y_{0:t}) \tag{5}$$

where g_t is called the *estimation rule* at time t. The collection $\boldsymbol{g} := (g_0, g_1, \dots)$ for all time is called the *estimation strategy*.

The collection $(\boldsymbol{f}, \boldsymbol{g})$ is called a *communication strategy*.

C. Performance Measures and Problem Formulation

At each time t, the system incurs two costs: a transmission cost $\lambda(U_t)$ and a distortion or estimation error $d(X_t, \hat{X}_t)$. Thus, the per-step cost is

$$c(X_t, U_t, \hat{X}_t) = \lambda(U_t) + d(X_t, \hat{X}_t).$$

We assume that $\lambda(u)$ is (weakly) increasing in u with $\lambda(0) = 0$ and $\lambda(u_{\max}) < \infty$. For the autoregressive source model, we assume that the distortion is given by $d(X_t - \hat{X}_t)$, where $d(\cdot)$ is even and quasi-convex with $d(0) = 0$.

We are interested in the following optimization problems.

Problem 1 (Finite horizon): In the model described above, identify a communication strategy $(\boldsymbol{f}^*, \boldsymbol{g}^*)$ that minimizes the total cost given by

$$J_T(\boldsymbol{f}, \boldsymbol{g}) := \mathbb{E}\left[\sum_{t=0}^{T-1} c(X_t, U_t, \hat{X}_t)\right]. \tag{6}$$

Problem 2 (Infinite horizon): In the model described above, given a discount factor $\beta \in (0, 1]$, identify a communication strategy $(\boldsymbol{f}^*, \boldsymbol{g}^*)$ that minimizes the total cost given as follows.

1) For $\beta \in (0, 1)$,

$$J_\beta(\boldsymbol{f}, \boldsymbol{g}) := (1 - \beta)\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t c(X_t, U_t, \hat{X}_t)\right]. \tag{7}$$

2) For $\beta = 1$,

$$J_1(\boldsymbol{f}, \boldsymbol{g}) := \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=0}^{T-1} c(X_t, U_t, \hat{X}_t)\right]. \tag{8}$$

Remark 1: In the model, it has been assumed that whenever the transmitter transmits (i.e., $U_t \neq 0$), it sends the source realization uncoded. This is without loss of generality because the channel input alphabet is the same as the source alphabet and the channel is symmetric. For such models, coding does not improve performance [26].

Problems 1 and 2 are decentralized stochastic control problems. The main conceptual difficulty in solving such problems is that the information available to the decision makers and, hence, the domain of their strategies grow with time, making the optimization problem combinatorial. One could circumvent this issue by identifying suitable *information states* at the decision makers that do not grow with time. In the following section, we discuss one such method to establish the structural results.

III. Main Results for Finite State Markov Sources

A. Structure of Optimal Communication Strategies

We establish two types of structural results. First, we use *person-by-person approach* to show that $(X_{0:t-1}, U_{0:t-1})$ is irrelevant at the transmitter (see Lemma 1); then, we use the common information approach of [27] and establish a belief state for the common information $(S_{0:t}, Y_{0:t})$ between the transmitter and the receiver (see Theorem 1).

Lemma 1: For any estimation strategy of the form (5), there is no loss of optimality in restricting attention to transmission strategies of the form

$$U_t = f_t(X_t, S_{0:t-1}, Y_{0:t-1}). \tag{9}$$

The proof proceeds by establishing that the process $\{X_t, S_{0:t-1}, Y_{0:t-1}\}_{t\geq 0}$ is a controlled Markov process controlled by $\{U_t\}_{t\geq 0}$. See Appendix A for details.

For any strategy $\boldsymbol{f}$ of the form (9) and any realization $(s_{0:T}, y_{0:T})$ of $(S_{0:T}, Y_{0:T})$, define $\varphi_t : \mathbb{X} \to \mathbb{U}$ as,

$$\varphi_t(x) = f_t(x, s_{0:t-1}, y_{0:t-1}), \quad \forall x \in \mathbb{X}.$$

Furthermore, define conditional probability measures π_t^1 and π_t^2 on $\mathbb{X}$ as follows: for any $x \in \mathbb{X}$

$$\pi_t^1(x) := \mathbb{P}^{\boldsymbol{f}}(X_t = x \mid S_{0:t-1} = s_{0:t-1}, Y_{0:t-1} = y_{0:t-1}),$$
$$\pi_t^2(x) := \mathbb{P}^{\boldsymbol{f}}(X_t = x \mid S_{0:t} = s_{0:t}, Y_{0:t} = y_{0:t}).$$

We call π_t^1 the *pre-transmission belief* and π^2 the *post-transmission belief*. Note that when $(S_{0:T}, Y_{0:T})$ are random variables, then π_t^1 and π_t^2 are also random variables (taking values in $\Delta(\mathbb{X})$), which we denote by Π_t^1 and Π_t^2.

For the ease of notation, define $B(\pi^1, s, \varphi)$ as follows:

$$\begin{aligned} B(\pi^1, s_t, \varphi) &:= \mathbb{P}(Y_t = \mathfrak{E} \mid S_{0:t} = s_{0:t}, Y_{0:t-1} = y_{0:t-1}) \\ &= \sum_{x_t \in \mathbb{X}} \pi^1(x_t) p(s_t, \varphi(x_t)). \end{aligned} \tag{10}$$

Furthermore, define $\pi^1|_{\varphi_t, s}$ as follows:

$$\pi^1|_{\varphi, s}(x) := \frac{\pi^1(x) p(s, \varphi(x))}{B(\pi^1, s, \varphi)}. \tag{11}$$

Then, using Bayes' rule one can show the following.

Lemma 2: Given any transmission strategy $\boldsymbol{f}$ of the form (9):

1) There exists a function F^1 such that

$$\pi_{t+1}^1 = F^1(\pi_t^2) = \pi_t^2 P. \tag{12}$$

2) There exists a function F^2 such that

$$\pi_t^2 = F^2(\pi_t^1, s_t, \varphi_t, y_t) = \begin{cases} \delta_{y_t}, & \text{if } y_t \in \mathbb{X} \\ \pi_t^1|_{\varphi_t, s_t}, & \text{if } y_t = \mathfrak{E}. \end{cases} \tag{13}$$

Note that in (12), we are treating π_t^2 as a row vector and in (13), δ_{y_t} denotes a Dirac measure centered at y_t. The update equations (12) and (13) are standard nonlinear filtering equations. See supplementary material for proof.

Theorem 1: In Problem 1 with finite state Markov source, we have the following conditions.

1) *Structure of optimal strategies:* There is no loss of optimality in restricting attention to transmission and estimation strategies of the form

$$U_t = f_t^*(X_t, S_{t-1}, \Pi_t^1), \tag{14}$$

$$\hat{X}_t = g_t^*(\Pi_t^2). \tag{15}$$

2) *Dynamic program:* Let $\Delta(\mathbb{X})$ denote the space of probability distributions on $\mathbb{X}$. Define value functions $V_t^1 : \Delta(\mathbb{X}) \times \mathbb{S} \to \mathbb{R}$ and $V_t^2 : \Delta(\mathbb{X}) \times \mathbb{S} \to \mathbb{R}$ as follows: for any $s_t \in \mathbb{S}$

$$V_{T+1}^1(\pi_t^1, s_t) = 0, \tag{16}$$

and for $t \in \{T, \ldots, 0\}$

$$V_t^1(\pi_t^1, s_t) = \min_{\varphi_t : \mathbb{X} \to \mathbb{U}} \{\Lambda(\pi_t^1, \varphi_t) + H_t(x_t, \pi_t^1, s_t, \varphi_t)\}, \tag{17}$$

$$V_t^2(\pi_t^2, s_t) = \min_{\hat{x} \in \mathbb{X}} D(\pi_t^2, \hat{x}) + V_{t+1}^1(\pi_t^2 P, s_t), \tag{18}$$

where

$$\Lambda(\pi^1, \varphi) := \sum_{x \in \mathbb{X}} \lambda(\varphi(x)) \pi^1(x),$$
$$\begin{aligned} H_t(x, \pi^1, s, \varphi) &:= B(\pi^1, s, \varphi) V_t^2(\delta_x, s) \\ &\quad + (1 - B(\pi^1, s, \varphi)) V_t^2(\pi^1|_{\varphi, s}, s), \end{aligned}$$
$$D(\pi^2, \hat{x}) := \sum_{x \in \mathbb{X}} d(x, \hat{x}) \pi^2(x).$$

Let $\Psi_t(s, \pi^1)$ denote the argmin of the right-hand side of (17) and $g_t^*(\pi_t^2) := \arg\min_{\hat{x} \in \mathbb{X}} D(\pi^2, \hat{x})$. Then, the optimal transmission strategy is given by

$$f_t^*(\cdot, s, \pi_t^1) = \Psi_t(s, \pi_t^1)$$

and the optimal estimation strategy is given by g_t^*.

The proof follows from the *common information approach* [27]. See Appendix B for details.

Remark 2: The first term in (17) is the expected communication cost, the second term is the expected cost-to-go. The first term in (18) is the expected distortion and the second term is the expected cost-to-go.

Remark 3: In (17) we use $\min$ instead of $\inf$ for the following reasons. Let Φ denote the set of functions from $\mathbb{X}$ to $\mathbb{U}$, which is equal to $\prod_{x \in \mathbb{X}} \mathbb{U}$ (since $\mathbb{X}$ is finite). When $\mathbb{U}$ is finite, Φ is also finite and, thus, we can use $\min$ in (17). When $\mathbb{U}$ is uncountable, Φ is a product of compact sets and, hence, is compact and, thus, we can use $\min$ in (17).

Remark 4: Note that the dynamic program in Theorem 1 is similar to a dynamic program for a partially observable Markov decision process with finite state space and finite or uncountable action space (see Remark 3). Thus, the dynamic program can be extended to infinite horizon discounted cost model after verifying standard assumptions. However, doing so does not provide any additional insight, so we do not present infinite horizon results for this model. We will do so for the autoregressive source model later in the paper, where we provide an algorithm to find the optimal time-homogeneous strategy for infinite horizon criteria.

IV. Main Results for Autoregressive Sources

A. Structure of Optimal Strategies for Finite Horizon Model

We start with a change of variables. Define a process $\{Z_t\}_{t\geq 0}$ as follows: $Z_0 = 0$ and for $t \geq 0$

$$Z_t = \begin{cases} aZ_{t-1}, & \text{if } Y_t = \mathfrak{E} \\ Y_t, & \text{if } Y_t \in \mathbb{X}. \end{cases}$$

Next, define processes $\{E_t\}_{t\geq 0}$, $\{E_t^+\}_{t\geq 0}$, which we call the *error processes* and $\{\hat{E}_t\}_{t\geq 0}$ as follows:

$$E_t := X_t - aZ_{t-1}, \quad E_t^+ := X_t - Z_t, \quad \hat{E}_t := \hat{X}_t - Z_t.$$

The processes $\{E_t\}_{t\geq 0}$ and $\{E_t^+\}_{t\geq 0}$ are related as follows: $E_0 = 0$, $E_0^+ = 0$, and for $t \geq 0$

$$E_t^+ = \begin{cases} E_t, & \text{if } Y_t = \mathfrak{E} \\ 0, & \text{if } Y_t \in \mathbb{X} \end{cases} \quad \text{and} \quad E_{t+1} = aE_t^+ + W_t. \tag{19}$$

The above dynamics may be rewritten as

$$E_{t+1} = \begin{cases} aE_t + W_t, & \text{if } Y_t = \mathfrak{E} \\ W_t, & \text{if } Y_t \neq \mathfrak{E}. \end{cases} \tag{20}$$

Since $X_t - \hat{X}_t = E_t^+ - \hat{E}_t$, we have that $d(X_t - \hat{X}_t) = d(E_t^+ - \hat{E}_t)$. Thus, with this change of variables, the per-step cost may be written as $\lambda(U_t) + d(E_t^+ - \hat{E}_t)$.

Note that Z_t is a deterministic function of $Y_{0:t}$. Hence, at time t, Z_{t-1} is measurable at the transmitter and, thus, E_t is measurable at the transmitter. Moreover, at time t, Z_t is measurable at the receiver.

Lemma 3: For any transmission and estimation strategies of the form (9) and (5), there exists an equivalent transmission and estimation strategy of the form

$$U_t = \tilde{f}_t(E_t, S_{0:t-1}, Y_{0:t-1}), \tag{21}$$

$$\hat{X}_t = \tilde{g}_t(S_{0:t}, Y_{0:t}) + Z_t. \tag{22}$$

Moreover, for any transmission and estimation strategies of the form (21) to (22), there exist transmission and estimation strategies of the form (9) and (5) that are equivalent.

The proof is given in Appendix C.

An implication of Lemma 3 is that we may assume that the transmitter transmits E_t and the receiver estimates

$$\hat{E}_t = \hat{X}_t - Z_t = \tilde{g}_t(S_{0:t}, Y_{0:t}).$$

For this model, we can further simplify the structures of optimal transmitter and estimator as follows.

Theorem 2: In Problem 1 with first-order autoregressive source, we have the following properties.

1) *Structure of optimal estimation strategy:* At each time t, there is no loss of optimality in choosing the estimates $\{\hat{E}_t\}_{t\geq 0}$ as

$$\hat{E}_t = 0,$$

or, equivalently, choosing the estimates $\{\hat{X}_t\}_{t\geq 0}$ as: $\hat{X}_0 = 0$, and for $t > 0$

$$\hat{X}_t = \begin{cases} a\hat{X}_{t-1}, & \text{if } Y_t = \mathfrak{E} \\ Y_t, & \text{if } Y_t \in \mathbb{R}. \end{cases} \tag{23}$$

2) *Structure of optimal transmission strategy:* There is no loss of optimality in restricting attention to transmission strategies of the form

$$U_t = \tilde{f}_t(E_t, S_{t-1}). \tag{24}$$

3) *Dynamic programming decomposition:* Recursively define the following value functions—for any $e \in \mathbb{R}$ and $s \in \mathbb{S}$,

$$J_{T+1}(e, s) = 0, \tag{25}$$

and for $t \in \{T, \dots, 0\}$

$$J_t(e, s) = \min_{u\in\mathbb{U}} \bar{H}_t(e, s, u), \tag{26}$$

where

$$\bar{H}_t(e, s, u) = \lambda(u) + \sum_{s'\in\mathbb{S}} Q_{ss'}p(s', u)d(e) + \mathbb{E}[J_{t+1}(E_{t+1}, S_t) \mid E_t = e, S_{t-1} = s, U_t = u].$$

Let $\tilde{f}_t^*(e, s)$ denote the argmin of the right-hand side of (26). Then, the transmission strategy $\tilde{\boldsymbol{f}}^* = (\tilde{f}_0^*, \dots, \tilde{f}_T^*)$ is optimal.

See Appendix D for the proof.

B. Monotonicity and Quasi Convexity of the Optimal Solution

For autoregressive sources we can establish monotonicity and quasi-convexity of the optimal solution. To that end, let us assume the following.

Assumption 1: The channel transition matrix Q is stochastic monotone, i.e., for all $i, j \in \{1, \dots, n\}$ such that $i > j$ and for any $\ell \in \{0, \dots, n-1\}$,

$$\sum_{k=\ell+1}^{n} Q_{ik} \geq \sum_{k=\ell+1}^{n} Q_{jk}.$$

Theorem 3: For any $t \in \{0, \dots, T\}$, we have the following.

1) For all $s \in \mathbb{S}$, $J_t(e, s)$ is even and quasi-convex in e.

Furthermore, under Assumption 1.

2) For every $e \in \mathbb{X}$, $J_t(e, s)$ is decreasing in s.
3) For every $s \in \mathbb{S}$, the transmission strategy $\tilde{f}_t(e, s)$ is even and quasi-convex in e.

Sufficient conditions under which the value function and the optimal strategy are even and quasi-convex are identified in [28, Th. 1]. Properties 1 and 3 follow because the model satisfies these sufficient conditions. Property 2 follows from standard stochastic monotonicity arguments. The details are presented in the supplementary material.

An immediate consequence of Theorem 3 is the following.

Corollary 1: Suppose that Assumption 1 is satisfied and $\mathbb{U}$ is finite set given by $\mathbb{U} = \{0, u^{(1)}, \dots, u^{(m)}\}$. For any $i \in \{0, 1, \dots, m\}$, define[3]

$$k_t^{(i)}(s) := \inf\{e \in \mathbb{R}_{\geq 0} : \tilde{f}_t(e, s) = u^{(i)}\}.$$

For ease of notation, define $k_t^{(m+1)}(s) = \infty$.

Then, for any $s \in \mathbb{S}$, $i \in \{0, \dots, m\}$ and $|e| \in [k_t^{(i)}(s), k_t^{(i+1)}(s))$, the optimal strategy is a threshold-based strategy given as follows:

$$\tilde{f}_t(e, s) = u^{(i)}. \tag{27}$$

Some Remarks:

1) It can be shown that under the optimal strategy, Π_t^2 is symmetric and unimodal (SU) (see Definition 1) around $\hat{X}_t$ and, therefore, Π_t^1 is SU around $a\hat{X}_{t-1}$. Thus, the transmission and estimation strategies in Theorem 2 depend

[3]Note that $k_t^{(0)}(s) = 0$ and Theorem 3 implies $k_t^{(i)}(s) \leq k_t^{(i+1)}(s)$ for any $i \in \{0, 1, \dots, m\}$.

on the pre- and post-transmission beliefs only through their means.

2) Since the distortion function is even and quasi-convex, we can write the threshold conditions

$$k_t^{(i-1)}(s) \leq |e| < k_t^{(i)}(s)$$

in (27) as

$$d(k_t^{(i-1)}(s)) \leq d(e) < d(k_t^{(i)}(s)).$$

Thus, if we define distortion levels $D_t^{(i)}(s) = \big[d(k_t^{(i-1)}(s)), d(k_t^{(i)}(s))\big)$, then we can say that the optimal strategy is to transmit at power level $u_{(i)}$ if $E_t \in D_t^{(i)}(S_{t-1})$.

3) When $Y_t = \mathfrak{E}$, the update of the optimal estimate is same as the update equation of Kalman filter. For this reason, we refer to the estimation strategy (23) as a *Kalman filter like estimator*.

C. Generalization to Infinite Horizon Model

Given a communication strategy $(\boldsymbol{f}, \boldsymbol{g})$, let $D_\beta^{(f,g)}(e,s)$ and $P_\beta^{(f,g)}(e,s)$ denote, respectively, the expected distortion and expected transmitted power when the system starts in state (e,s), i.e., for $\beta \in (0,1)$,

$$D_\beta^{(f,g)}(e,s) := (1-\beta)\mathbb{E}^{(f,g)}\left[\sum_{t=0}^{\infty} \beta^t d(E_t) \,|\, E_0 = e, S_{-1} = s\right],$$

$$P_\beta^{(f,g)}(e,s) := (1-\beta)\mathbb{E}^{(f,g)}\left[\sum_{t=0}^{\infty} \beta^t \lambda(U_t) \,|\, E_0 = e, S_{-1} = s\right],$$

and for $\beta = 1$,

$$D_1^{(f,g)}(e,s) := \lim_{T\to\infty} \frac{1}{T}\mathbb{E}^{(f,g)}\left[\sum_{t=0}^{T-1} d(E_t) \,|\, E_0 = e, S_0 = s\right],$$

$$P_1^{(f,g)}(e,s) := \lim_{T\to\infty} \frac{1}{T}\mathbb{E}^{(f,g)}\left[\sum_{t=0}^{T-1} \lambda(U_t) \,|\, E_0 = e, S_0 = s\right].$$

Then, the performance of the strategy $(\boldsymbol{f}, \boldsymbol{g})$ when the system starts in state (e,s) is given by

$$J_\beta^{(f,g)}(e,s) := D_\beta^{(f,g)}(e,s) + P_\beta^{(f,g)}(e,s).$$

The structure of optimal estimator, as established in Theorem 2, continues to hold for the infinite horizon setup as well. Thus, we can restrict attention to Kalman filter like estimator given by (23) and look at the problem of finding the best response transmission strategy. This is a single agent stochastic control problem. If the per-step distortion is unbounded, then we need the following assumption—which implies that there exists a strategy whose performance is bounded—for the infinite horizon problem to be meaningful.

Assumption 2: Let $f^{(0)}$ denote the transmission strategy that always transmits at power level $u_{\max}$ and g^* denote the Kalman filter like strategy given by (23). Then, for given $\beta \in (0,1]$, and for all $e \in \mathbb{X}$ and $s \in \mathbb{S}$, $D_\beta^{(f^{(0)},g)}(e,s) < \infty$.

Assumption 2 is always satisfied if $d(\cdot)$ is bounded. For $\beta = 1$, $d(e) = e^2$, and $\mathbb{S} = \{0,1\}$, the condition $a^2(1-Q_{00}) < 1$ is sufficient for Assumption 2 to hold (see [20, Th. 8] and [22, Corollary 12]). Similar sufficient conditions are given in [23, Th. 1] for vector-valued Markov source processes with a Markovian packet-drop channel.

We now state the main theorem of this section.

Theorem 4: In Problem 2 with first-order autoregressive processes under Assumption 2, we have the following properties.

1) *Structure of optimal estimation strategy:* The time-homogeneous strategy $\tilde{\boldsymbol{g}}^* = \{\tilde{g}^*, \tilde{g}^*, \dots\}$, where $\tilde{g}^*$ is given by (23), is optimal.

2) *Structure of optimal transmission strategy:* There is no loss of optimality in restricting attention to time-homogeneous transmission strategies of the form

$$U_t = \tilde{f}_\beta(E_t, S_{t-1}).$$

3) *Dynamic programming decomposition:* For $\beta \in (0,1)$, let J_β be the smallest bounded solution of the following fixed point equation: for all $e \in \mathbb{R}$ and $s \in \mathbb{S}$

$$J_\beta(e,s) = \min_{u\in\mathbb{U}} \bar{H}_\beta(e,s,u), \tag{28}$$

where

$$\bar{H}_\beta(e,s,u) = (1-\beta)\lambda(u) + \sum_{s'\in\mathbb{S}} Q_{ss'} p(s',u) d(e) + \beta\mathbb{E}[J_\beta(E_{t+1}, S_t) \,|\, E_t = e, S_{t-1} = s, U_t = u].$$

Let $\tilde{f}_\beta^*(e,s)$ denote the argmin of the right-hand side of (28). Then, the transmission $\tilde{\boldsymbol{f}}_{\boldsymbol{\beta}}^* = (\tilde{f}_\beta^*, \tilde{f}_\beta^*, \dots)$ is optimal.

4) *Results for $\beta = 1$:* Let $\tilde{f}_1^*$ be any limit point of $\{\tilde{f}_\beta^*\}_{\beta\in(0,1)}$ as $\beta \uparrow 1$. Then, $\tilde{f}_1^*$ is optimal strategy for Problem 2 with $\beta = 1$.

The proof is given in Appendix E.

Remark 5: We are not asserting that the dynamic program (28) has a *unique* fixed point. To make such an assertion, we would need to check the sufficient conditions for Banach fixed point theorem. These conditions [29] are harder to check than the sufficient conditions (P1)–(P3) of Proposition 2 that we verify in Appendix E.

Corollary 2: The monotonicity properties of Theorem 3 hold for the infinite horizon value function J_β and transmission strategy $\tilde{f}_\beta$ as well.

An immediate consequence of Corollary 2 is the following.

Corollary 3: Suppose that Assumption 1 is satisfied and $\mathbb{U}$ is finite set given by $\mathbb{U} = \{0, u^{(1)}, \dots, u^{(m)}\}$. For any $i \in \{0,1,\dots,m\}$, define[4]

$$k_\beta^{(i)}(s) := \inf\{e \in \mathbb{R}_{\geq 0} : \tilde{f}_(e,s) = u^{(i)}\}.$$

For ease of notation, define $k_\beta^{(m+1)}(s) = \infty$.

Then, the optimal strategy is a threshold-based strategy given as follows: for any $s \in \mathbb{S}$, $i \in \{0,\dots,m\}$ and $|e| \in \big[k_\beta^{(i)}(s), k_\beta^{(i+1)}(s)\big)$

$$\tilde{f}_\beta(e,s) = u^{(i)}. \tag{29}$$

[4]Note that $k_\beta^{(0)}(s) = 0$ and Corollary 2 implies $k_\beta^{(i)}(s) \leq k_\beta^{(i+1)}(s)$ for any $i \in \{0,1,\dots,m\}$.

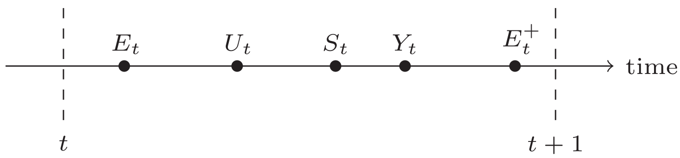

Fig. 2. Timeline for pre- and post-transmission error state.

V. Computing Optimal Thresholds for Autoregressive Sources With Finite Actions

Suppose the power levels are finite and given by

$$\mathbb{U} = \{0, u^{(1)}, \dots, u^{(m)}\}, \quad m \in \mathbb{Z}_{>0}$$

with $u^{(i)} < u^{(i+1)}$ and $i \in \{0, 1, \dots, m-1\}$. From Corollary 3, we know that the optimal strategy for Problem 2 is a time-homogeneous threshold-based strategy of the form (27). Let $\mathbf{k}$ denote the thresholds $\{k_\beta^{(i)}(s)\}$ and $f^{(\mathbf{k})}$ denote the strategy (29). In this section, we first derive formulas for computing the performance of a general threshold-based strategy $f_\beta^{(\mathbf{k})}$ of the form (27) and then propose a stochastic approximation based algorithm to identify the optimal thresholds.

It is conceptually simpler to work with a *post-decision model* where the *pre-decision state* is E_t and the *post-decision state* is E_t^+ given by (19). The timeline of the various system variables is shown in Fig. 2. In this model, the per-step cost is given by $\lambda(U_t) + d(E_t^+)$.[5]

A. Performance of an Arbitrary Threshold-Based Strategy

For $\beta \in (0, 1]$, pick a reference channel state $s^\circ \in \mathbb{S}$. Given an arbitrary threshold-based strategy $f_\beta^{(\mathbf{k})}$, suppose the system starts in state $(E_{-1}^+, S_{-1}) = (0, s^\circ)$ and follows strategy $f_\beta^{(\mathbf{k})}$. Then, the process $\{(E_t^+, S_t)\}_{t\ge 0}$ is a Markov process. Let $\tau^{(0)} = 0$ and for $n \in \mathbb{Z}_{>0}$ let

$$\tau^{(n)} := \{t > \tau^{(n-1)} : (E_{t-1}^+, S_{t-1}) = (0, s^\circ)\}$$

denote the *stopping times* when the Markov process $\{(E_t^+, S_t)\}_{t\ge 0}$ revisits $(0, s^\circ)$. We say that the Markov process *regenerates* at times $\{\tau^{(n)}\}_{n\in\mathbb{Z}_{\ge 0}}$ and refer to the interval $\{\tau^{(n)}, \dots, \tau^{(n-1)}\}$ as the nth regenerative cycle.

Define the following parameter.

1) $L_\beta^{(\mathbf{k})}$ is the expected cost during a regenerative cycle, i.e.,

$$L_\beta^{(\mathbf{k})} := \mathbb{E}\left[\sum_{t=0}^{\tau^{(1)}-1} \beta^t (\lambda(U_t) + d(E_t^+)) \,\middle|\, E_{-1}^+ = 0, S_{-1} = s^\circ\right]. \quad (30)$$

2) $M_\beta^{(\mathbf{k})}$ is the expected time during a regenerative cycle, i.e.,

$$M_\beta^{(\mathbf{k})} := \mathbb{E}\left[\sum_{t=0}^{\tau^{(1)}-1} \beta^t \,\middle|\, E_{-1}^+ = 0, S_{-1} = s^\circ\right]. \quad (31)$$

[5]From Theorem 2, we have that $\hat{E}_t = 0$. Thus, $d(E_t^+ - \hat{E}_t) = d(E_t^+)$.

Using ideas from renewal theory, we have the following.

Theorem 5: For any $\beta \in (0, 1]$, the performance of threshold-based strategy $f_\beta^{(\mathbf{k})}$ is given by

$$C_\beta^{(\mathbf{k})} := C_\beta(f_\beta^{(\mathbf{k})}, g^*) = \frac{L_\beta^{(\mathbf{k})}}{M_\beta^{(\mathbf{k})}}. \quad (32)$$

See Appendix F for the proof.

B. Necessary Condition for Optimality

In order to find the optimal threshold, we first observe the following.

Lemma 4: For any $\beta \in (0, 1]$, $L_\beta^{(\mathbf{k})}$, and $M_\beta^{(\mathbf{k})}$ are differentiable with respect to $\mathbf{k}$. Consequently, $C_\beta^{(\mathbf{k})}$ is also differentiable.

The proof of Lemma 4 follows from first principles using an argument similar to that in the supplementary material for [14].

Let $\nabla_\mathbf{k} L_\beta^{(\mathbf{k})}$, $\nabla_\mathbf{k} M_\beta^{(\mathbf{k})}$, and $\nabla_\mathbf{k} C_\beta^{(\mathbf{k})}$ denote the derivatives of $L_\beta^{(\mathbf{k})}$, $M_\beta^{(\mathbf{k})}$, and $C_\beta^{(\mathbf{k})}$, respectively. Then, a sufficient condition for optimality is the following.

Proposition 1: A necessary condition for thresholds $\mathbf{k}^*$ to be optimal is that $N_\beta^{(\mathbf{k}^*)} = 0$, where

$$N_\beta^{(\mathbf{k}^*)} := M_\beta^{(\mathbf{k})} \nabla_\mathbf{k} L_\beta^{(\mathbf{k})} - L_\beta^{(\mathbf{k})} \nabla_\mathbf{k} M_\beta^{(\mathbf{k})}.$$

Proof: The result follows from observing that $\nabla_\mathbf{k} C_\beta^{(\mathbf{k})} = N_\beta^{(\mathbf{k})}/(M_\beta^{(\mathbf{k})})^2$.

Remark 6: If $C_\beta^{(\mathbf{k})}$ is convex in $\mathbf{k}$, then the condition in Proposition 1 is also sufficient for optimality. Based on numerical calculations we have observed that $C_\beta^{(\mathbf{k})}$ is convex in $\mathbf{k}$ but we have not been able to prove it analytically.

C. Stochastic Approximation Algorithm to Compute Optimal Thresholds

In this section, we present an iterative algorithm based on simultaneous perturbation and renewal Monte Carlo (RMC) [9], [30] to compute the optimal thresholds. We present this algorithm under the following assumption.

Assumption 3: There exists a $K \in \mathbb{X}_{\ge 0}$ such that for the optimal transmission strategy

$$k^{(m-1)}(s) \le K \quad \forall s \in \mathbb{S}.$$

Remark 7: Assumption 3 is equivalent to stating that for each channel state s, there is a state $e^*(s)$ such that for all $e \ge e^*(s)$, the optimal transmission strategy transmits at the maximum power level $u_{\max}$ in state (e, s). Assumption 3 is similar to the channel saturation assumption in [13, Assumption 1] and [12, Remark 5].

Under Assumption 3, there is no loss of optimality in restricting attention to threshold strategies in the set

$$\mathcal{K} := \{\mathbf{k} : k^{(i)}(s) \le K, \forall s \in \mathbb{S}, i \in \{0, \dots, m-1\}\}.$$

The main idea behind the RMC is as follows. Given a threshold $\mathbf{k}$, consider the following sample-path-based unbiased

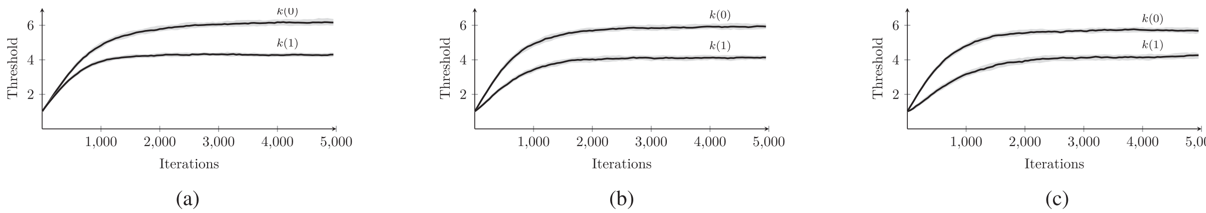

Fig. 3. Thresholds versus iterations for different values of (q_r, q_f). The experiment is repeated 100 times. The bold lines represent the sample median and the shaded regions represent the 1st and 3rd quartiles across the runs. (a) $(q_r, q_f) = (0.1, 0.1)$. (b) $(q_r, q_f) = (0.1, 0.3)$. (c) $(q_r, q_f) = (0.1, 0.5)$.

estimators of $L_\beta^{(\mathbf{k})}$ and $M_\beta^{(\mathbf{k})}$:

$$\widehat{\mathsf{L}}_\beta^{(\mathbf{k})} := \frac{1}{N}\sum_{n=0}^{N-1}\sum_{t=\tau^{(n)}}^{\tau^{(n+1)}-1}\beta^{t-\tau^{(n)}}\Big[\lambda(U_t) + d(E_t^+)\Big], \quad (33)$$

$$\widehat{\mathsf{M}}_\beta^{(\mathbf{k})} := \frac{1}{N}\sum_{n=0}^{N-1}\sum_{t=\tau^{(n)}}^{\tau^{(n+1)}-1}\beta^{t-\tau^{(n)}}, \quad (34)$$

where N is a large nonnegative integer.

Then, *simultaneous perturbation* based unbiased estimators of $\nabla_{\mathbf{k}} L_\beta^{(\mathbf{k})}$ and $\nabla_{\mathbf{k}} M_\beta^{(\mathbf{k})}$ are given by

$$\widehat{\nabla}\mathsf{L}_\beta^{(\mathbf{k})} = \delta(\widehat{\mathsf{L}}_\beta^{(\mathbf{k}+c\delta)} - \widehat{\mathsf{L}}_\beta^{(\mathbf{k}-c\delta)})/2c, \quad (35)$$

$$\widehat{\nabla}\mathsf{M}_\beta^{(\mathbf{k})} = \delta(\widehat{\mathsf{M}}_\beta^{(\mathbf{k}+c\delta)} - \widehat{\mathsf{M}}_\beta^{(\mathbf{k}-c\delta)})/2c, \quad (36)$$

where δ is an appropriately chosen random variable having the same dimension as $\mathbf{k}$ and c is a small positive constant. Typically, all components of δ are chosen independently as either $\mathrm{Rademacher}(\pm 1)$ [31], [32] or $\mathrm{Normal}(0,1)$ [33], [34].

If the estimates $(\widehat{\mathsf{L}}_\beta^{(\mathbf{k})}, \widehat{\mathsf{M}}_\beta^{(\mathbf{k})})$ and $(\widehat{\nabla}\mathsf{L}_\beta^{(\mathbf{k})}, \widehat{\nabla}\mathsf{M}_\beta^{(\mathbf{k})})$ are generated from independent sample paths, then the unbiasedness and independence of these estimates imply that

$$\widehat{\mathsf{N}}_\beta^{(\mathbf{k})} := \widehat{\mathsf{M}}_\beta^{(\mathbf{k})}\widehat{\nabla}\mathsf{L}_\beta^{(\mathbf{k})} - \widehat{\mathsf{L}}_\beta^{(\mathbf{k})}\widehat{\nabla}\mathsf{M}_\beta^{(\mathbf{k})}$$

is an unbiased estimator of $N_\beta^{(\mathbf{k})}$.

Then, the RMC algorithm to compute the optimal threshold is as follows.

1) Let $j = 0$ and pick any initial guess $\mathbf{k}_0$.
2) For sufficiently large number of iterations do
   a) Generate $\widehat{\mathsf{L}}_\beta^{(\mathbf{k}_j)}$ and $\widehat{\mathsf{M}}_\beta^{(\mathbf{k}_j)}$ according to (33) and (34).
   b) Sample a random direction δ.
   c) Generate $\widehat{\nabla}\mathsf{L}_\beta^{(\mathbf{k}_j)}$ and $\widehat{\nabla}\mathsf{M}_\beta^{(\mathbf{k}_j)}$ according to (35)–(36).
   d) Compute:
   $$\widehat{\mathsf{N}}_\beta^{(\mathbf{k}_j)} := \widehat{\mathsf{M}}_\beta^{(\mathbf{k}_j)}\widehat{\nabla}\mathsf{L}_\beta^{(\mathbf{k}_j)} - \widehat{\mathsf{L}}_\beta^{(\mathbf{k}_j)}\widehat{\nabla}\mathsf{M}_\beta^{(\mathbf{k}_j)}.$$
   e) Update:
   $$\mathbf{k}_{j+1} = [\mathbf{k}_j - \alpha_j \widehat{\mathsf{N}}_\beta^{(\mathbf{k}_j)}]_{\mathcal{K}}$$
   where $[\cdot]_{\mathcal{K}}$ denotes projection on to the set $\mathcal{K}$ and $\{\alpha_j\}$ is a sequence of learning rates that satisfy $\sum_{j=1}^\infty \alpha_j = \infty$ and $\sum_{j=1}^\infty \alpha_j^2 < \infty$.
   f) Set $j = j + 1$.

We assume the following (which is a standard assumption for stochastic approximation algorithms, see e.g., [34, Assumption 5.6]).

Assumption 4: The set of globally asymptotically stable equilibrium of the differential equation $d\mathbf{k}/dt = -N_\beta^{(\mathbf{k})}$ is compact.

Theorem 6: Consider the sequence of iterates $\{\mathbf{k}_j\}_{j\geq 0}$ obtained by the RMC algorithm described previously. Let $\mathbf{k}^*$ be any limit point of $\{\mathbf{k}_j\}_{j\geq 0}$. Then, under Assumptions 3 and 4, $N_\beta^{(\mathbf{k}^*)} = 0$ and, therefore, $\nabla_{\mathbf{k}} C_\beta^{(\mathbf{k}^*)} = 0$.

Proof: The proof follows from [30, Corollary 1].

D. Numerical Example

Consider a real-valued autoregressive source with $a = 1$, $\mu = \mathrm{Normal}(0,1)$ and discount factor $\beta = 0.99$. The channel is a Gilbert–Elliott channel [35], [36] with state space $\mathbb{S} = \{0, 1\}$ and transition matrix $Q = \left[\begin{smallmatrix} 1-q_r & q_r \\ q_f & 1-q_f \end{smallmatrix}\right]$, where q_r and q_f are called the recovery rate and the failure rate of the channel. Note that when $q_r + q_f \leq 1$, the matrix Q is stochastic monotone and Assumption 1 is satisfied. The channel has two power levels $\mathbb{U} = \{0, 1\}$ with loss probability $p(0,0) = 1$, $p(0,1) = 0.7$, $p(1,0) = 1$, $p(1,1) = 0.2$, and transmission cost $\lambda(0) = 0$, $\lambda(1) = 100$.

We run the RMC algorithm with $s^\circ = 0$, $N = 1000$, the learning rates $\{\alpha_j\}$ chosen according to Adam [37] (with α parameter of Adam equal to 0.1 and other parameters taking their default values as stated in [37]), $\delta = \mathrm{Normal}(0,1)$ and $c = 0.1$. Note that since $m = 1$, $k^{(0)}(s) = 0$. We use $k(0)$ and $k(1)$ to denote $k^{(1)}(0)$ and $k^{(1)}(1)$. For a given choice of the channel transition matrix, the algorithm is run for 5000 iterations and the experiment is repeated 100 times. The median and first and third quartiles across multiple runs are stored.

To visualize the speed of convergence across different choices of channel transition matrices, we plot thresholds $(k(0), k(1))$ versus iterations for a few different values of (q_r, q_f) in Fig. 3. These plots show that the convergence is relatively fast. Similar qualitative behavior is observed for other choices of channel transition matrices as well.

To compare the performance of the system for different choices of the channel transition matrices, we choose the thresholds $\mathbf{k} = (k(0), k(1))$ to be the average over the last 500 iterations of the median across the 100 runs. We estimate $\hat{L}_\beta^{(\mathbf{k})}$ and $\hat{M}_\beta^{(k)}$ by Monte Carlo averaging over $N = 10^6$ renewals and compute $\hat{C}_\beta^{(\mathbf{k})}$ using Theorem 5. Optimal performance for different values of recovery rate q_r and failure rate q_f such that $q_r + q_f \leq 1$ is shown in Fig. 4. The plot shows that the performance improves with increase in the recovery rate q_r and

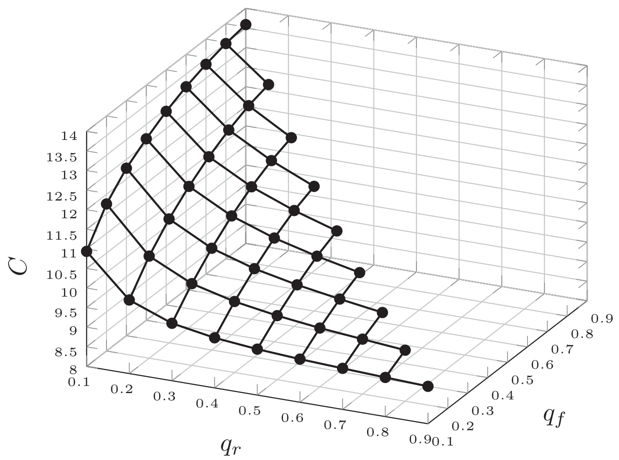

Fig. 4. Optimal performance over Gilbert–Elliott channel for different values of recovery rate q_r and failure rate q_f for $q_r + q_f \leq 1$.

deteriorates with the increase in the failure rate q_f. The code can be found in [38].

VI. Discussions

A. Comparison With the Results of [12]

Remote estimation over a packet-drop channel with Markovian state was recently considered in [12]. In [12], it is assumed that the transmitter knows the current channel state. In contrast, in our model, we assume that the receiver observes the channel state and sends it back to the transmitter. So, the transmitter has access to a one-step delayed channel state.

Ren *et al.* [12] pose the problem of identifying the optimal transmission and estimation strategies for infinite horizon average cost setup for vector-valued autoregressive sources. They identify the common information based dynamic program and identify technical conditions under which the dynamic program has a deterministic solution. The dynamic program in [12] may be viewed as the infinite horizon average cost equivalent of the finite horizon dynamic program in Theorem 1. They then show that when the source dynamics are orthogonal and the noise dynamics are isotropic, there is no loss of optimality in restricting attention to estimation strategies of the form (15) and transmission strategies of the form (14). In addition, for every $\pi \in \Delta(\mathbb{X})$ and $s \in \mathbb{S}$, $\zeta_t(\cdot) := f_t(\cdot, s, \pi)$ is symmetric and quasi-convex. This structural property of the transmitter implies that when the power levels are finite, there exist thresholds $\{k_t^i(s,\pi)\}_{i=0}^{m-1}$ such that the optimal strategy is a threshold-based strategy as follows: for any $s \in \mathbb{S}$, $\pi \in \Delta(\mathbb{X})$, $i \in \{0, \ldots, m\}$, and $|e| \in [k_t^{(i)}(s,\pi), k_t^{(i+1)}(s,\pi))$

$$f_t(e, s, \pi) = u^{(i)}. \tag{37}$$

In this paper, we follow a different approach. We investigate both finite Markov sources and first order autoregressive sources. For Markov sources, we first show that there is no loss of optimality in restricting attention to estimation strategies of the form (15) and transmission strategies of the form (14). For autoregressive sources, we show that the structure of the transmission strategies can be further simplified to (23) and (24). In addition for every $s \in \mathbb{S}$, $\varphi_t(\cdot) = \tilde{f}_t(\cdot, s)$ is symmetric and quasi-convex. This structural property of the transmitter implies that when the power levels are finite, there exist thresholds $\{k_t^i(s)\}_{i=0}^{m-1}$ such that the optimal strategy is a threshold-based strategy given by (27).

Once we restrict attention to estimation strategy of the form (23), the best response strategy at the transmitter is a centralized Markov decision process. This allows us to establish the existence of optimal deterministic strategies for both discounted and average cost infinite horizon models without having to resort to the detailed technical argument presented in [12].

Note that in the threshold-based strategies (37) identified in [12], the thresholds depend on the belief state π, while in the threshold-based strategies (27) identified in this paper, the thresholds do not depend on π. We exploit this lack of dependence on π to develop a renewal theory based method to compute the performance of a threshold based strategy. The algorithm proposed in this paper will not work for threshold strategies of the form (37) due to the dependence on π (which is uncountable).

B. Comparison With the Results of [14]

A method for computing the optimal threshold for remote estimation over noiseless communication channel (i.e., no packet drop) is presented in [14]. That method relies on computing $L_\beta^{(\mathbf{k})}$ and $M_\beta^{(\mathbf{k})}$ by solving the balance equations (which are Fredholm integral equations of the second kind) for the truncated Markov chain. When the channel is a packet-drop channel, the kernel of the Fredholm integral equation is discontinuous. Moreover, when the channel has state, the integral equation is multidimensional. Solving such integral equations is computationally difficult. The simulation-based methods presented in this paper circumvent these difficulties.

C. Special Case With Independent and Identically Distributed Packet-Drop Channels

Consider the case when the packet drops are independent and identically distributed, which can be viewed as a Markov channel with a single state (i.e., $|\mathbb{S}| = 1$). Thus, we may drop the dependence on s from the value function and the strategies. Furthermore, Assumption 1 is trivially satisfied. Thus, the result of Theorem 3 simplifies to the following.

Corollary 4: For Problem 1 with independent and identically distributed packet drops the value function $J_t(e)$ and the optimal transmission strategy $f_t(e)$ are even and quasi-convex.

The above result is same as [7, Th. 1]. Furthermore, when the power levels are finite, the optimal transmission strategy is characterized by thresholds $(k_t^{(1)}, \ldots, k_t^{(m-1)})$. For infinite horizon models the thresholds are time invariant.

In addition, the renewal relationships of Theorem 5 continues to hold. The stopping times $\{\tau^{(n)}\}_{n\geq 0}$ correspond to times of successful reception and the proposed renewal Monte Carlo algorithm is similar in spirit to [9]. Note that the algorithm proposed in [9] uses simultaneous perturbation to find the minimum of $C_\beta^{(k)}$, where $C_\beta^{(k)}$ is evaluated using renewal relationships. The algorithm proposed in this paper is different and uses simultaneous perturbations to find the roots of $N_\beta^{(\mathbf{k})}$, which coincide with the roots of $\nabla_{\mathbf{k}} C_\beta^{(\mathbf{k})}$.

It is worth highlighting that when the power levels are finite, the model can also be interpreted as a remote estimator that has the option of transmitting (at a constant power level) over one of m available independent and identically

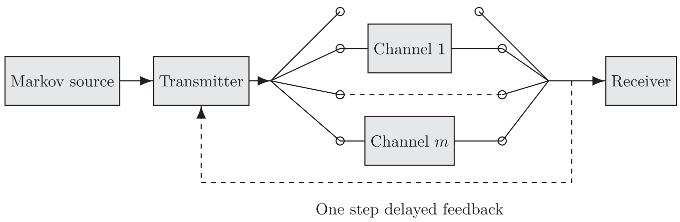

Fig. 5. Remote estimation over multiple channels.

distributed packet-drop channels, as shown in Fig 5. For channel i, $i \in \{1, \dots, m\}$, the transmission cost is $\lambda(i)$ and the drop probability is $p(i)$. We assume that the channels are ordered[6] such that $\lambda(1) \le \cdots \le \lambda(m)$ and $p(1) \ge \cdots \ge p(m)$. In addition, the sensor has the option of not transmitting, which is denoted by $i = 0$. Note that $\lambda(0) = 0$ and $p(0) = 1$. As argued previously, the optimal transmission strategy in this case is characterized by thresholds $(k_t^{(1)}, \dots, k_t^{(m-1)})$ and the sensor transmits over channel i, where i is such that $|e_t| \in [k_t^{(i)}, k_t^{(i+1)})$ (and we assume that $k_t^{(0)} = 0$ and $k_t^{(m)} = \infty$). The above result is similar in spirit to [39], which considers independent and identically distributed source and additive noise channels.

VII. Conclusion

In this paper, we study remote estimation over a Markovian channel with feedback. We assume that the channel state is observed by the receiver and fed back to the transmitter with one unit delay. In addition, the transmitter gets ACK/NACK feedback for successful/unsuccessful transmission. Using ideas from team theory, we establish the structure of optimal transmission and estimation strategies for finite Markov sources and identify a dynamic program to determine optimal strategies with that structure. We then consider first-order autoregressive sources where the noise process has unimodal and symmetric distribution. Using ideas from majorization theory, we show that the optimal transmission strategy has a monotone structure and the optimal estimation strategy is like Kalman filter.

The structural results imply that threshold-based transmitter is optimal when the power levels are finite. We provide a stochastic approximation based algorithm to compute the optimal thresholds and optimal performance. An example of a first-order autoregressive source model with Gilbert–Elliott channel is considered to illustrate the results.

Appendix A
Proof of Lemma 1

Arbitrarily fix the estimation strategy $\boldsymbol{g}$ and consider the *best response* strategy at the transmitter. We will show that $\tilde{I}_t^1 := (X_t, S_{0:t-1}, Y_{0:t-1})$ is an information state at the transmitter. In particular, we will show that $\{\tilde{I}_t^1\}_{t\ge 1}$ satisfies the following properties:

$$\mathbb{P}(\tilde{I}_{t+1}^1 \mid I_t^1, U_t) = \mathbb{P}(\tilde{I}_{t+1}^1 \mid \tilde{I}_t^1, U_t) \tag{38}$$

and

$$\mathbb{E}[c(X_t, U_t, \hat{X}_t) \mid I_t^1, U_t] = \mathbb{E}[c(X_t, U_t, \hat{X}_t) \mid \tilde{I}_t^1, U_t]. \tag{39}$$

Given any realization $(x_{0:T}, s_{0:T}, y_{0:T}, u_{0:T})$ of the system variables $(X_{0:T}, S_{0:T}, Y_{0:T}, U_{0:T})$, define $i_t^1 = (x_{0:t}, s_{0:t-1}, y_{0:t-1}, u_{0:t-1})$ and $\tilde{i}_t^1 = (x_t, s_{0:t-1}, y_{0:t-1})$. Now, for any $\breve{i}_{t+1}^1 = (x_{t+1}, \breve{s}_{0:t}, \breve{y}_{0:t}) = (x_{t+1}, \breve{s}_t, \breve{y}_t, \breve{i}_t^1)$, we use the shorthand $\mathbb{P}(\breve{i}_{t+1}^1 | \tilde{i}_{0:t}^1, u_{0:t})$ to denote $\mathbb{P}(\tilde{I}_{t+1}^1 = \breve{i}_{t+1}^1 | \tilde{I}_{0:t}^1 = \tilde{i}_{0:t}^1, U_{0:t} = u_{0:t})$. Then,

$$\begin{aligned}
\mathbb{P}(\breve{i}_{t+1}^1 | i_t^1, u_t) &= \mathbb{P}(x_{t+1}, \breve{s}_t, \breve{y}_t, \breve{i}_t^1 | x_{0:t}, s_{0:t-1}, y_{0:t-1}, u_{0:t}) \\
&\overset{(a)}{=} \mathbb{P}(x_{t+1}|x_t)\mathbb{P}(\breve{y}_t|x_t, \breve{s}_t)\mathbb{P}(\breve{s}_t|s_{t-1}) \\
&\quad \times \mathbb{1}_{\{(\breve{s}_{0:t-1}, \breve{y}_{0:t-1}) = (s_{0:t-1}, y_{0:t-1})\}} \\
&= \mathbb{P}(\breve{i}_{t+1}^1 | x_t, s_{0:t-1}, y_{0:t-1}, u_t) \\
&= \mathbb{P}(\breve{i}_{t+1}^1 | \tilde{i}_t^1, u_t)
\end{aligned} \tag{40}$$

where (a) follows from the source and the channel models. This shows that (38) is true.

Now consider (39). Recall that $\hat{X}_t = g_t(I_t^2)$. Therefore, the expectation in the left-hand side of (39) depends on $\mathbb{P}(I_t^2 \mid I_t^1, U_t)$. By marginalizing (40) with respect to X_{t+1}, we get

$$\mathbb{P}(I_t^2 \mid I_t^1, U_t) = \mathbb{P}(I_t^2 \mid \tilde{I}_t^1, U_t)$$

which implies (39).

Equation (38) shows that $\{\tilde{I}_t^1\}_{t\ge 0}$ is a controlled Markov process controlled by $\{U_t\}_{t\ge 0}$. Equation (39) shows that $\tilde{I}_t^1$ is sufficient for performance evaluation. Hence, by Markov decision theory [40], there is no loss of optimality in restricting attention to transmission strategies of the form (9).

Appendix B
Proof of Theorem 1

Once we restrict attention to transmission strategies of the form (9), the information structure is partial history sharing [27]. Thus, one can use the common information approach of [27] and obtain the structure of optimal transmission strategy using this approach.

Following [27], we split the information available at each agent into a *common information* and *local information*. Common information is the information available to all decision makers in the future; the remaining data at the decision maker are the local information. Thus, at the transmitter, the common information is $C_t^1 := \{S_{0:t-1}, Y_{0:t-1}\}$ and the local information is $L_t^1 := X_t$, and at the receiver, $C_t^2 := \{S_{0:t}, Y_{0:t}\}$ and $L_t^2 = \emptyset$. The state sufficient for input output mapping of the system is (X_t, S_t). By [27, Proposition 1], we get that

$$\begin{aligned}
\Theta_t^1(x, s) &:= \mathbb{P}(X_t = s, S_t = s | C_t^1), \\
\Theta_t^2(x, s) &:= \mathbb{P}(X_t = s, S_t = s | C_t^2)
\end{aligned}$$

are sufficient statistics for the common information at the transmitter and the receiver, respectively. Now, we observe the following.

[6]The assumption is without loss of generality. In particular, if there are channels i and j such that $\lambda(i) \ge \lambda(j)$ and $p(i) \ge p(j)$, then transmission over channel j dominates the action of transmission over channel i.

1) Θ_t^1 is equivalent to (Π_t^1, S_{t-1}) and Θ_t^2 is equivalent to (Π_t^2, S_t). This is because independence of $\{X_t\}_{t\geq 0}$ and $\{S_t\}_{t\geq 0}$ implies that $\theta_t^1(x,s) = \pi_t^1(x)Q_{s_{t-1}s}$ and $\theta_t^2(x,s) = \pi_t^2(x)\mathbb{1}_{\{S_t=s\}}$.
2) The expected distortion $D(\Pi_t^2, \hat{X}_t)$ does not depend on S_t and the evolution of Π_t^2 to Π_{t+1}^1 (given by Lemma 2) does not depend on $\hat{X}_t$.

Thus, from [27, Proposition 1] we get that the optimal strategy is given by the dynamic program of (17)–(18).

Appendix C
Proof of Lemma 3

The proof relies on the fact that Z_t is a deterministic function of $Y_{0:t}$, i.e., there exists an ℓ_t such that $Z_t = \ell_t(Y_{0:t})$. We prove the two parts separately. We use the notation $\boldsymbol{f} = (f_1, \ldots, f_T)$, $\boldsymbol{g} = (g_1, \ldots, g_T)$, $\tilde{\boldsymbol{f}} = (\tilde{f}_1, \ldots, \tilde{f}_T)$, and $\tilde{\boldsymbol{g}} = (\tilde{g}_1, \ldots, \tilde{g}_T)$.

1) Given a transmission and an estimation strategy $(\boldsymbol{f}, \boldsymbol{g})$ of the form (9) and (5), define

$$\tilde{f}_t(E_t, S_{0:t-1}, Y_{0:t-1}) = f_t(E_t + a\ell_{t-1}(Y_{0:t-1}), S_{0:t-1}, Y_{0:t-1})$$

and

$$\tilde{g}_t(S_{0:t}, Y_{0:t}) = g_t(S_{0:t}, Y_{0:t}) - \ell_t(Y_{0:t}).$$

Then, by construction the strategy $(\tilde{\boldsymbol{f}}, \tilde{\boldsymbol{g}})$ is equivalent to $(\boldsymbol{f}, \boldsymbol{g})$.

2) Given a transmission and an estimation strategy $(\tilde{\boldsymbol{f}}, \tilde{\boldsymbol{g}})$ of the form (21) to (22), define

$$f_t(X_t, S_{0:t-1}, Y_{0:t-1}) = \tilde{f}_t(X_t - a\ell_{t-1}(Y_{0:t-1}), S_{0:t-1}, Y_{0:t-1})$$

and

$$g_t(S_{0:t}, Y_{0:t}) = \tilde{g}_t(S_{0:t}, Y_{0:t}) + \ell_t(Y_{0:t}).$$

Then, by construction the strategy $(\boldsymbol{f}, \boldsymbol{g})$ is equivalent to $(\tilde{\boldsymbol{f}}, \tilde{\boldsymbol{g}})$.

Appendix D
Proof of Theorem 2

A. Sufficient Statistic and Dynamic Program

Similar to the construction of a prescription for the finite state Markov sources, for any transmission strategy $\tilde{\boldsymbol{g}}$ of the form (21) and any realization $(s_{0:t-1}, y_{0:t})$ of $(S_{0:t-1}, Y_{0:t})$, define $\varphi : \mathbb{R} \to \mathbb{U}$ as

$$\varphi(e) = \tilde{f}_t(e, s_{0:t-1}, y_{0:t-1}), \forall e \in \mathbb{R}.$$

Next, redefine the pre- and post-transmission beliefs in terms of the error process. In particular, π_t^1 is the conditional probability density of E_t given $(s_{0:t-1}, y_{0:t-1})$ and π_t^2 is the conditional probability density of E_t^+ given $(s_{0:t}, y_{0:t})$.

Let $R_t = \mathbb{1}_{\{Y_t \in \mathbb{X}\}}$ and r_t denote the realization of R_t. The time evolution of π_t^1 and π_t^2 is similar to Lemma 2. In particular, we have the following lemma.

Lemma 5: Given any transmission strategy $\boldsymbol{f}$ of the form (4), the following functions exist.

1) There exists a function F^1 such that

$$\pi_{t+1}^1 = F^1(\pi_t^2) = \tilde{\pi}_t^2 \star \mu \tag{41}$$

where $\tilde{\pi}_t^2$ given by $\tilde{\pi}_t^2(e) := (1/|a|)\pi_t^2(e/a)$ is the conditional probability density of aE_t^+, μ is the probability density function of W_t and $\star$ is the convolution operation.

2) There exists a function F^2 such that for any realization $r_t \in \{0,1\}$ of R_t

$$\pi_t^2 = F^2(\pi_t^1, s_t, \varphi_t, r_t).$$

In particular,

$$\pi_t^2 = F^2(\pi_t^1, s_t, \varphi_t, r_t) = \begin{cases} \pi_t^1|_{\varphi_t, s_t}, & \text{if } r_t = 0 \\ \delta_0, & \text{if } r_t = 1. \end{cases} \tag{42}$$

The dynamic program of Theorem 1 can be rewritten in terms of the error process as follows. Consider $\mathbb{X} = \mathbb{R}$ (similar derivation holds for $\mathbb{X} = \mathbb{Z}$). Then,

$$V_{T+1}^1(\pi_t^1, s_t) = 0,$$

and for $t \in \{T, \ldots, 0\}$

$$V_t^1(\pi_t^1, s_t) = \min_{\varphi_t : \mathbb{X} \to \mathbb{U}} \Big\{ \Lambda(\pi_t^1, \varphi_t) + H_t(\pi_t^1, s_t, \varphi_t) \Big\}, \tag{43}$$

$$V_t^2(\pi_t^2, s_t) = D(\pi_t^2) + V_{t+1}^1(\pi_t^2 P, s_t), \tag{44}$$

where

$$\Lambda(\pi_t^1, \varphi_t) := \int_{\mathbb{R}} \lambda(\varphi(e))\pi_t^1(e)de,$$

$$H_t(\pi_t^1, s_t, \varphi) := B(\pi_t^1, s_t, \varphi_t)V_t^2(\delta_0, s_t) + (1 - B(\pi_t^1, s_t, \varphi_t))V_t^2(\pi^1|_{\varphi, s_t}, s_t),$$

$$D(\pi_t^2) := \min_{\hat{e} \in \mathbb{X}} \int_{\mathbb{R}} d(e - \hat{e})\pi_t^2(e)de.$$

Note that due to the change of variables, the expression for the first term of H_t does not depend on the transmitted symbol. Consequently, the expression for V_t^1 is simpler than that in Theorem 1.

Remark 8: The common-information approach for decentralized stochastic control for finite state and finite action models was presented in [27]. In general, to extend the approach to Borel state and action spaces, one needs to impose a topology on the space of prescriptions and establish an appropriate "measurable selection theorem". Such technical difficulties are not present in the dynamic program of (43) and (44) because the common information is finite[7] and, therefore, all prescriptions are measurable.

To establish the results of Theorem 2, we show that the above dynamic program satisfies a monotonicity property with respect to the partial order based on majorization. We start with some mathematical preliminaries needed to present this argument.

[7] In particular, $C_t^1 = R_{0:t-1}$ and $C_t^2 = R_{0:t}$.

B. Mathematical preliminaries

Definition 1 (Symmetric and unimodal density): A probability density function π on reals is said to be *symmetric and unimodal* (SU) around $c \in \mathbb{R}$ if for any $x \in \mathbb{R}$, $\pi(c-x) = \pi(c+x)$ and π is nondecreasing in the interval $(-\infty, c]$ and nonincreasing in the interval $[c, \infty)$.

Definition 2 (Symmetric and quasi-convex prescription): Given $c \in \mathbb{R}$, a prescription $\theta : \mathbb{R} \to \mathbb{U}$ is symmetric and quasi-convex (denoted by $\mathrm{SQC}(c)$) if $\theta(e-c)$ is even and quasi-convex.

Now, we state some properties of SU distributions.

Property 1: If π is $\mathrm{SU}(c)$, then

$$c \in \arg\min_{\hat{e} \in \mathbb{R}} \int_{\mathbb{R}} d(e-\hat{e})\pi(e)de.$$

Proof: For $c = 0$, the above property is a special case of [4, Lemma 12]. The result for general c follows from a change of variables.

Property 2: If π_t^1 is $\mathrm{SU}(0)$ and $\varphi_t \in \mathrm{SQC}(0)$, then for any $r_t \in \{0,1\}$ and $s_t \in \mathbb{S}$, $F^2(\pi_t^1, s_t, \varphi_t, r_t)$ is $\mathrm{SU}(0)$.

Proof: We prove the result for each $r_t \in \{0,1\}$ separately. Recall the update of π_t^1 given by (42).

1) For $r_t = 0$ and a given $s_t \in \mathbb{S}$, we have that if $\varphi_t \in \mathrm{SQC}(0)$, then $p(s_t, \varphi_t(e)$ is $\mathrm{SU}(0)$ since $p(s_t, \cdot)$ is decreasing and $p(s_t, \varphi_t(e)) = p(s_t, \varphi(-e))$. Then, $\pi_t^1(e)p(s_t, \varphi_t(e))$ is $\mathrm{SU}(0)$ since the product of two $\mathrm{SU}(0)$ functions is $\mathrm{SU}(0)$. Hence, π_t^2 is $\mathrm{SU}(0)$.
2) For $r_t = 1$, $\pi_t^2 = \delta_0$, which is $\mathrm{SU}(0)$.

Property 3: If π_t^2 is $\mathrm{SU}(0)$, then $F^1(\pi_t^2)$ is also $\mathrm{SU}(0)$.

Proof: Recall that F^1 is given by (41). The property follows from the fact that convolution of symmetric and unimodal distributions is symmetric and unimodal [4].

Definition 3 (Symmetric rearrangement (SR) of a set): Let $\mathbb{A}$ be a measurable set of finite Lebesgue measure, its *symmetric rearrangement* $\mathbb{A}^\sigma$ is the open interval centered around origin whose Lebesgue measure is same as $\mathbb{A}$.

Definition 4 (Level sets of a function): Given a function $\ell : \mathbb{R} \to \mathbb{R}_{\geq 0}$, its upper-level set at level ρ, $\rho \in \mathbb{R}$, is $\{x \in \mathbb{R} : \ell(x) > \rho\}$ and its lower-level set at level ρ is $\{x \in \mathbb{R} : \ell(x) < \rho\}$.

Definition 5 (SR of a function): The *symmetric decreasing rearrangement* $\ell_\downarrow^\sigma$ of ℓ is a symmetric and decreasing function whose level sets are the same as ℓ, i.e.,

$$\ell_\downarrow^\sigma(x) = \int_{\mathbb{R}_{\geq 0}} \mathbb{1}_{\{z \in \mathbb{R} : \ell(z) > \rho\}^\sigma}(x)d\rho.$$

Similarly, the symmetric increasing rearrangement $\ell_\uparrow^\sigma$ of ℓ is a symmetric and increasing function whose level sets are the same as ℓ, i.e.,

$$\ell_\uparrow^\sigma(x) = \int_{\mathbb{R}_{\geq 0}} \mathbb{1}_{\{z \in \mathbb{R} : \ell(z) < \rho\}^\sigma}(x)d\rho.$$

Definition 6 (Majorization): Given two probability density functions ξ and π over $\mathbb{R}$, ξ *majorizes* π, which is denoted by $\xi \succeq_m \pi$, if for all $\rho \geq 0$,

$$\int_{|x| \geq \rho} \xi_\downarrow^\sigma(x)dx \geq \int_{|x| \geq \rho} \pi_\downarrow^\sigma(x)dx.$$

Definition 7 (SU-majorization): Given two probability density functions ξ and π over $\mathbb{R}$, ξ *SU majorizes* π, which we denote by $\xi \succeq_a \pi$, if ξ is SU and ξ majorizes π.

An immediate consequence of Definition 7 is the following.

Lemma 6: For any nonnegative $\mathrm{SQC}(c)$ function g, $c \in \mathbb{R}$, and given two probability density functions ξ and π over $\mathbb{R}$, such that $\xi \succeq_a \pi$ and ξ is $\mathrm{SU}(c)$, we have that for any $A \subseteq \mathbb{R}$

$$\int_A g(x)\xi(x)dx \geq \int_A g(x)\pi(x)dx.$$

Property 4: For any $\xi \succeq_a \pi$, $F^1(\xi) \succeq_a F^1(\pi)$.

This follows from [4, Lemma 10].

Recall the definition of $D(\pi^2)$ given after (44).

Property 5: If $\xi \succeq_a \pi$, then

$$D(\pi) \geq D(\pi_\downarrow^\sigma) \geq D(\xi_\downarrow^\sigma) = D(\xi).$$

Proof: The inequalities follow from [4, Lemma 11]. The last equality holds since ξ is $\mathrm{SU}(0)$ and, thus, $\xi_\downarrow^\sigma = \xi$.

Lemma 7: For any $c \in \mathbb{R}$, densities π and ξ and a prescription φ, there exists a $\mathrm{SQC}(c)$ prescription θ, $c \in \mathbb{R}$, such that for any $u \in \mathbb{U}$

$$\pi(\{e \in \mathbb{R} : \varphi(e) \leq u\}) = \xi(\{e \in \mathbb{R} : \theta(e) \leq u\}).$$

We denote such a θ by $\mathcal{T}_c^{(\pi,\xi)}\varphi$.

Proof: Let $\eta(u) := \pi(\{e \in \mathbb{R} : \varphi(e) \leq u\})$. We prove the construction separately for $\mathbb{X} = \mathbb{R}$ and $\mathbb{X} = \mathbb{Z}$. First, let us consider $\mathbb{X} = \mathbb{R}$. Let $A(u)$ be a symmetric set centered around $c \in \mathbb{R}$ such that $\xi(A(u)) = \eta(u)$. By construction, $\eta(u)$ is increasing in u and, therefore, so is $A(u)$. Define $\theta(e)$ such that $A(u)$ is the lower level set of $\theta(e)$. By construction, $\theta(e)$ is symmetric around c. Moreover, since $A(u)$ is convex, $\theta(e)$ is quasi-convex.

Now, let us consider $\mathbb{X} = \mathbb{Z}$ and $c = 0$. (The proof for general c follows from a change of variables). Let $\mathbb{U}_\varphi$ denote the set $\{\varphi(e) : e \in \mathbb{Z}\}$. Note that $\mathbb{U}_\varphi$ is a finite or a countable set, which we will index by $\{u^{(i)} ; i \in \mathbb{Z}_{\geq 0}\}$. Define $k^{(i)}$ such that

$$\xi(A^{(i-1)}) < \eta^{(i)} \leq \xi(A^{(i)})$$

where $A^{(i-1)} := \{-k^{(i-1)}, \ldots, k^{(i-1)}\}$ and $A^{(i)} := \{-k^{(i)}, \ldots, k^{(i)}\}$.

Now, for any $u^{(i)} \in \mathbb{U}_\varphi$, let $\eta^{(i)}$ denote $\eta^{(i)} := \pi(\{e \in \mathbb{Z} : \varphi(e) \leq u^{(i)}\})$. Define $\delta^{(i)} := (\eta^{(i)} - \xi(A^{(i)}))/2$. Define the prescription θ as follows:

$$\theta(e) = \begin{cases} u^{(i-1)}, & \text{if } |e| \in \left(k^{(i-1)}, k^{(i)}\right) \\ u^{(i-1)} \text{ w.p. } \delta^{(i-1)}, & \text{if } |e| = k^{(i)} \\ u^{(i)} \text{ w.p. } \delta^{(i)}, & \text{if } |e| = k^{(i)}. \end{cases}$$

Then, by construction, θ is $\mathrm{SQC}(0)$ and $\xi(\{e \in \mathbb{Z} : \theta(e) \leq u\}) = \eta(u)$.

Remark 9: Suppose $\theta = \mathcal{T}_c^{(\pi,\xi)}\varphi$. Then, the probability of using power level u at pre-transmission belief π and prescription

φ is the same as that at pre-transmission belief π and prescription θ. In particular, for all $u \in \mathbb{U}$

$$\pi(\{e \in \mathbb{X} : \varphi(e) = u\}) = \xi(\{e \in \mathbb{X} : \theta(e) = u\}).$$

Property 6: For any density ξ, π, and prescription φ let $\theta = \mathcal{T}_c^{(\pi,\xi)}\varphi$. Then, for any $s \in \mathbb{S}$

1) $B(\xi, s, \theta) = B(\pi, s, \varphi)$.
2) $\Lambda(\pi, \varphi) = \Lambda(\xi, \theta)$.

The above results follow from the definitions of B and Λ and Remark 9. See supplementary material for a detailed proof.

Property 7: For any $\xi \succeq_a \pi$, where ξ is $\mathrm{SU}(c)$, and prescription φ, let $\theta = \mathcal{T}_c^{(\xi,\pi)}\varphi$. Then, for any $s \in \mathbb{S}$

$$\xi|_{\theta,s} \succeq_a \pi|_{\varphi,s}. \tag{45}$$

Consequently, for any $s \in \mathbb{S}$ and $r \in \{0, 1\}$, we have

$$F^2(\xi, s, \theta, r) \succeq_a F^2(\pi, s, \varphi, r). \tag{46}$$

Proof: We prove the result for finite $\mathbb{U} = \{u^{(0}, \dots, u^{(m)}\}$. The result for the case when $\mathbb{U}$ is an interval follows from a discretization argument.

For any $i \in \{0, \dots, m\}$, let $A^{(i)} = \{e : \theta(e) = u^{(i)}\}$ and $B^{(i)} = \{e : \varphi(e) = u^{(i)}\}$ and $\bar{A}^{(i)} = \bigcup_{j=0}^{i} A^{(i)}$ and $B^{(i)} = \bigcup_{j=0}^{i} B^{(i)}$.

Since θ is $\mathrm{SQC}(0)$, $\bar{A}^{(i)}$ is an interval (while $\bar{B}^{(i)}$ need not be an interval). Define $a^{(i)} = \xi(\bar{A}^{(i)})$ and $b^{(i)} = \pi(\bar{B}^{(i)})$. Also define $\xi^{(i)}$ as $\xi^{(i)}(e) = \xi(e)\mathbb{1}_{\{e \in \bar{A}^{(i)}\}}/a^{(i)}$ and $\pi^{(i)} = \pi^{(i)}(e) = \pi(e)\mathbb{1}_{\{e \in \bar{B}^{(i)}\}}/b^{(i)}$. Then, by [4, Lemma 8]

$$\xi^{(i)} \succeq_a \pi^{(i)}, \quad \forall i \in \{0, \dots, m\}. \tag{47}$$

Fix an $s \in \mathbb{S}$. For ease of notation, define $p^{(i)} = p(s, u^{(i)})$. Then, we can write the following expression for $\xi\big|_{\theta,s}$:

$$\begin{aligned}
\xi\big|_{\theta,s}(e) &= \frac{1}{B(\xi, s, \theta)}\xi(e)p(s, \theta(e)) \\
&= \frac{\xi(e)}{B(\xi, s, \theta)}\sum_{i=0}^{m} p^{(i)}\mathbb{1}_{\{e \in A^{(i)}\}} \\
&\overset{(a)}{=} \frac{\xi(e)}{B(\xi, s, \theta)}\sum_{i=0}^{m-1} [p^{(i)} - p^{(i+1)}]\mathbb{1}_{\{e \in \bar{A}^{(i)}\}} + p^{(m)}\mathbb{1}_{\{e \in \bar{A}^{(m)}\}} \\
&= \frac{1}{B(\xi, s, \theta)}\sum_{i=0}^{m-1} a^{(i)}[p^{(i)} - p^{(i+1)}]\xi^{(i)} + a^{(m)}p^{(m)}\xi^{(m)}
\end{aligned} \tag{48}$$

where (a) uses the fact that $A^{(i)} = \bar{A}^{(i)} \setminus \bar{A}^{(i-1)}$. By a similar argument, we have

$$\begin{aligned}
\pi\big|_{\varphi,s}(e) &= \frac{1}{B(\pi, s, \varphi)}\sum_{i=0}^{m-1} b^{(i)}[p^{(i)} - p^{(i+1)}]\pi^{(i)} \\
&\quad + b^{(m)}p^{(m)}\pi^{(m)}.
\end{aligned} \tag{49}$$

Property 6 implies that $a^{(i)} = b^{(i)}$. The monotonicity of $p(s, u)$ implies that $p^{(i)} - p^{(i+1)} \geq 0$. Using this, and combining (48) and (49) with (47), we get (45). Equation (46) follows from (13).

C. Qualitative Properties of the Value Function and Optimal Strategy

Lemma 8: The value functions V_t^1 and V_t^2 of (43)–(44), satisfy the following property.

V1) For any $i \in \{1, 2\}$, $s \in \mathbb{S}$, $t \in \{0, \dots, T\}$, and probability densities ξ^i and π^i such that $\xi^i \succeq_a \pi^i$, we have that $V_t^i(\xi^i, s) \leq V_t^i(\pi^i, s)$.
Furthermore, for any $s \in \mathbb{S}$ and $t \in \{0, \dots, T\}$, the optimal strategy satisfies the following properties.

V2) If π_t^1 is $\mathrm{SU}(c)$, then there exists a prescription $\varphi_t \in \mathrm{SQC}(c)$ that is optimal. In general, φ_t depends on π_t^1.

V3) If π_t^2 is $\mathrm{SU}(c)$, then the optimal estimate $\hat{E}_t$ is c.

Proof: We proceed by backward induction. $V_{T+1}^i(\pi^1, s), i \in \{1, 2\}$ trivially satisfy (V1). This forms the basis of induction. Now assume that $V_{t+1}^1(\pi^1, s)$ also satisfies (V1). For $\xi^2 \succeq_a \pi^2$, we have that

$$\begin{aligned}
V_t^2(\pi^2, s) &= D(\pi^2) + V_{t+1}^1(F^1(\pi^2), s) \\
&\overset{(a)}{\geq} D(\xi^2) + V_{t+1}^1(F^1(\xi^2), s) = V_t^2(\xi^2, s),
\end{aligned} \tag{50}$$

where (a) follows from Properties 4 and 5 and the induction hypothesis. Equation (50) implies that V_t^2 also satisfies (V1). Now, we have

$$\begin{aligned}
&H_t(\pi^1, s, \varphi) = B(\pi^1, s, \varphi)V_t^2(\delta_0, s) \\
&\qquad + (1 - B(\pi^1, s, \varphi))V_t^2(\pi^1|_{\varphi,s}, s) \\
&\overset{(a)}{\geq} B(\xi^1, s, \theta)V_t^2(\delta_0, s) + (1 - B(\xi^1, s, \theta))V_t^2(\xi^1|_{\theta,s}, s) \\
&= H_t(\xi^1, s, \theta),
\end{aligned} \tag{51}$$

where (a) holds due to Properties 6 and 7 and (50). Then, we have from (43)

$$\begin{aligned}
V_t^1(\pi^1, s) &= \min_{\varphi:\mathbb{X}\to\mathbb{U}} \Big\{\Lambda(\pi^1, \varphi) + H_t(\pi^1, s, \varphi)\Big\} \\
&\overset{(b)}{\geq} \min_{\theta:\mathbb{X}\to\mathbb{U}} \Big\{\Lambda(\xi^1, \theta) + H_t(\xi^1, s, \theta)\Big\} \\
&= V_t^1(\xi^1, s)
\end{aligned} \tag{52}$$

where (b) holds due to Property 6 and (51) and since inequality is preserved in pointwise minimization. This completes the induction step.

In order to show (V2), note that $\pi_t^1 \succeq_a \pi_t^1$ trivially. Let φ_t be the optimal prescription at π_t^1. If $\varphi_t \in \mathrm{SQC}(c)$, then we are done. If not, define $\theta_t = \mathcal{T}_c^{(\pi_t^1,\pi_t^1)}\varphi_t$ where the transformation $\mathcal{T}_c^{(\pi_t^1,\pi_t^1)}$ is introduced in Lemma 7. Then, the argument in (52) (with ξ^1 replaced by π^1) also implies (V2). Furthermore, (V3) follows from Property 1.

D. Proof of Theorem 2

Proof of Part 1: Properties 2 and 3 imply that for all t, Π_t^2 is $\mathrm{SU}(0)$. Therefore, by Property 1, the optimal estimate $\hat{E}_t = 0$. Recall that $\hat{E}_t = \hat{X}_t - Z_t$. Thus, $\hat{X}_t = Z_t$. This proves the first part of Theorem 2.

Proof of Parts 2 and 3: Lemmas 1 and 3 imply that there is no loss of optimality in restricting attention to transmitters of the form

$$U_t = \tilde{g}_t(E_t, Y_{0:t-1}). \tag{53}$$

Part 1 implies that there is no loss of optimality in restricting attention to estimation strategies of form (23). So, we assume that the transmission and estimation strategies are of these forms.

Since the estimation strategy is fixed, Problem 1 reduces to a single agent optimization problem. $\{(E_t, S_{t-1})\}_{t\geq 0}$ is an information state for this single agent optimization problem for the following reasons.

1) Equation (20) implies that $\{E_t\}_{t\geq 0}$ is a controlled Markov process controlled by R_t. Moreover, for any realization $(e_{0:t}, u_{0:t}, s_{0:t-1})$ of $(E_{0:t}, U_{0:t}, S_{0:t-1})$, we have

$$\begin{aligned}\mathbb{P}(R_t = 0 | e_{0:t}, u_{0:t}, s_{0:t-1}) &= \sum_{s' \in \mathbb{S}} Q_{ss'} p(s', u) \\ &= \mathbb{P}(R_t = 0 | u_t, s_{t-1}).\end{aligned}$$

Combining the two we get that

$$\mathbb{P}(E_{t+1}, S_t | e_{0:t}, u_{0:t}, s_{0:t-1}) = \mathbb{P}(E_{t+1}, S_t | e_t, u_t, s_{t-1}).$$

2) Using (19), the conditional expected per-step cost may be written as

$$\begin{aligned}&\mathbb{E}[\lambda(U_t) + d(E_t^+) | e_{0:t}, u_{0:t}, s_{0:t-1}] \\ &\quad = \lambda(u_t) + \mathbb{E}[p(S_t, u) d(e) | S_{t-1} = s]. \end{aligned} \tag{54}$$

Thus, the optimization problem at the transmitter is an Markov decision process with information state (E_t, S_{t-1}). Therefore, from Markov decision theory, there is no loss of optimality in restriction attention to Markov strategies of the form $f_t(E_t, S_{t-1})$. The optimal strategies of this form are given by the dynamic program of Part 3.

Appendix E
Proof of Theorem 4

A. Proof of Part 1)

The structure of Theorem 2 holds the infinite horizon setup as well. The main idea is to use forward induction to show that the optimal estimate for the error process, $\hat{E}_t = 0$. The structure of optimal $\hat{X}_t$ is then derived by using the change of variable introduced in Section IV-A.

B. Some Preliminary Properties

We prove the following properties, which will be used to establish the existence of the solution to the dynamic program (28). Note that the per-step cost $c(e, s, u)$ given in (54) can be rewritten as $c(e, s, u) := (1-\beta)(\lambda(u) + \sum_{s_t \in \mathbb{S}} Q_{ss_t} p(s_t, u) d(e))$.

Our model satisfies the following properties ([29, Assumptions 4.2.1,4.2.2]).

Proposition 2: Under Assumption 2, for any $\lambda(\cdot) \geq 0$, the following conditions of [29] are satisfied.

P1) The per-step cost $c(e, s, u)$ is lower semicontinuous,[8] bounded from below and inf-compact[9] on $\mathbb{R} \times \{0, 1\}$, i.e., for all $e, r \in \mathbb{R}$, the set $\{u \in \mathbb{U} : c(e, s, u) \leq r\}$ is compact.

P2) For every $u \in \{0, 1\}$, the transition kernel from (E_t, S_{t-1}) to (E_{t+1}, S_t) is strongly continuous.[10]

P3) There exists a strategy for which the value function is finite.

Proof: P1) is true because of the following reasons. The action set $\mathbb{U}$ is either finite or uncountable and the per-step cost $c(\cdot, s, u)$ is continuous on $\mathbb{R}$ (and, hence, lower semicontinuous), and is nonnegative. Finally, when $\mathbb{U}$ is finite, all subsets of $\mathbb{U}$ are compact. When $\mathbb{U}$ is uncountable, all closed subsets of $\mathbb{U}$ are compact.

To check (P2), note the following fact ([29, Example C.6]).

Fact 1: Let P be a stochastic kernel and suppose that there is a σ-finite measure ϕ on $\mathbb{X}$ such that, for every $y \in \mathbb{X}$, $\mathbb{P}(\cdot|y)$ has a density $p(\cdot|y)$ with respect to ϕ, that is

$$P(B|y) = \int_B p(x|y)\phi(dx), \quad \forall B \in \mathcal{B}(\mathbb{X}), y \in \mathbb{X}.$$

If $p(x|\cdot)$ is continuous on $\mathbb{X}$ for every $x \in \mathbb{X}$, then P is strongly continuous.

P2) is true for the following reasons. Let P denote the transition kernel from (E_t, S_{t-1}) to (E_{t+1}, S_t). Then, for any Borel subset B of $\mathcal{B}(\mathbb{X})$

$$\begin{aligned}&P(E_{t+1} \in B \mid E_t = e, S_{t-1} = s, U_t = u) \\ &\quad = (1 - p(s,u)) \int_B \mu(w) dw + p(s, u) \int_B \mu(w - ae) dw.\end{aligned}$$

Then, according to Fact 1, P is strongly continuous since the real line $\mathbb{R}$ with Lebesgue measure $\phi(dx) = dx$ is σ-finite and since the density μ is continuous on $\mathbb{R}$.

P3) is true due to Assumption 2.

C. Proofs of Parts 2) and 3)

For ease of exposition, we assume $\mathbb{X} = \mathbb{R}$. Similar argument works for $\mathbb{X} = \mathbb{Z}$. We fix the optimal estimator with the Kalman filter like structure (23) and identify the best performing transmitter, which is a centralized optimization problem. For the discounted setup, one expects that the optimal solution is given by the fixed point of the dynamic program (28) [similar to (25)–(26)]. However, it is not obvious that there exists a fixed point of (28) because the distortion $d(\cdot)$ is unbounded.

Define the operator $\mathcal{B}$ given as follows, for any given $s \in \mathbb{S}$ and function: $v : \mathbb{R} \times \mathbb{S} \to \mathbb{R}$

$$\begin{aligned}[\mathcal{B}v](e, s) := \min_{u \in \mathbb{U}} \mathbb{E}[&c(e, s, u) + v(E_{t+1}, S_t) \\ &\mid E_t = e, S_{t-1} = s].\end{aligned}$$

[8]A function is lower semicontinuous if its lower level sets are closed.

[9]A function $v : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is said to be *inf-compact* on $\mathbb{X} \times \mathbb{U}$ if, for every $x \in \mathbb{X}$ and $r \in \mathbb{R}$, the set $\{u \in \mathbb{U} : v(x, u) \leq r\}$ is compact.

[10]A controlled transition probability kernel $P(\cdot \mid x, u) : \mathbb{X} \to [0, 1]$ is said to be *strongly continuous* if for any bounded measurable function v on $\mathbb{X}$ the function v' on $\mathbb{X} \times \mathbb{U}$, $v'(x, u) := \mathbb{E}^P[v \mid x, u]$, $x \in \mathbb{X}$, $u \in \mathbb{U}$, is continuous and bounded.

Then, (28) can be expressed in terms of the operator $\mathcal{B}$ as follows:

$$[\mathcal{B}J_\beta](e,s) := \min_{u\in\mathbb{U}} \mathbb{E}[c(e,s,u) + J_\beta(E_{t+1}, S_t) \mid E_t = e, S_{t-1} = s].$$

Then, the proof of Parts 2) and 3) of the theorem follows directly from [29]. In particular, we have the following proposition, where the first part follows from [29, Th. 4.2.3] and the second part follows from [29, Lemma 4.2.8].

Proposition 3: Under (P1)–(P3), there exist fixed point solutions J_β of (28). Let $J^*_\beta : \mathbb{X} \to \mathbb{R}$ denote the smallest such fixed point and $\tilde{f}^*_\beta(e,s)$ denote the argmin of the right-hand side of (28) for $J_\beta = J^*_\beta$. Then,

1) $\tilde{f}^*_\beta$ is the optimal strategy for Problem 2 with $\beta \in (0,1)$.
2) Let $J^{(0)}_\beta = 0$ and define $J^{(n+1)}_\beta := \mathcal{B}J^{(n)}_\beta$. Then, $J^*_\beta = \lim_{n\to\infty} J^{(n)}_\beta$.

D. Properties of the Value Function

We can apply the *vanishing discount approach* and show that the result for the long-term average cost is obtained as a limit of those in the discounted setup, as $\beta \uparrow 1$.

Our model satisfies the following conditions [29], [41].

Proposition 4: Under Assumption 2, for any $\lambda(\cdot) \geq 0$, the value function J_β, as given by (28), satisfies the following conditions of [29], [41]: for any $s \in \mathbb{S}$, and $\lambda \in \mathbb{R}_{\geq 0}$.

S1) There exists a reference state $e_0 \in \mathbb{X}$ and a nonnegative scalar M_s such that $J^*_\beta(e_0, s) < M_s$ for all $\beta \in (0,1)$.

S2) Define $h_\beta(e,s) = (1-\beta)^{-1}[J_\beta(e,s) - J_\beta(e_0,s)]$. There exists a function $K_s : \mathbb{X} \to \mathbb{R}$ such that $h_\beta(e,s) \leq K_s(e)$ for all $e \in \mathbb{X}$ and $\beta \in (0,1)$.

S3) There exists a nonnegative (finite) constant L_λ such that $-L_s \leq h_\beta(e,s)$ for all $e \in \mathbb{X}$ and $\beta \in (0,1)$.

Therefore, if $\tilde{f}^*_\beta$ denotes an optimal strategy for $\beta \in (0,1)$, and $\tilde{f}^*_1$ is any limit point of $\{\tilde{f}^*_\beta\}$, then $\tilde{f}^*_1$ is optimal for $\beta = 1$.

Proof: We prove the proposition for $\mathbb{X} = \mathbb{R}$. Similar argument holds for $\mathbb{X} = \mathbb{Z}$.

Let $\tilde{J}^{(0)}_\beta(e,s)$ denote the value function of the "always transmit with maximum power" strategy. According to Assumption 2, $M_s = \tilde{J}^{(0)}_\beta(0,s) < \infty$. Hence, (S1) is satisfied with $e_0 = 0$ and $M_s = \tilde{J}^{(0)}_\beta(0,s)$.

Since not transmitting is optimal at state 0 (because the transmission strategy is SQC about 0), we have

$$J^*_\beta(0,s) = \beta \sum_{s_t\in\mathbb{S}} Q_{ss_t} \int_{\mathbb{R}} \mu(w) J^*_\beta(w, s_t) dw.$$

Let $\tilde{J}^{(1)}_\beta(e,s)$ denote the value function of the strategy that transmits with power level $u = \varphi(e)$ at time 0 and follows the optimal strategy from then on. Then,

$$\begin{aligned}\tilde{J}^{(1)}_\beta(e,s) &= (1-\beta)\Big[\lambda(\varphi(e)) + \sum_{s_t\in\mathbb{S}} Q_{ss_t} p(s_t, \varphi(e)) d(e)\Big] \\ &\quad + \beta \sum_{s_t\in\mathbb{S}} Q_{ss_t} \int_{\mathbb{R}} \mu(w) J^*_\beta(w, s_t) dw \\ &= (1-\beta)\Big[\lambda(\varphi(e)) + \sum_{s_t\in\mathbb{S}} Q_{ss_t} p(s_t, \varphi(e)) d(e)\Big] \\ &\quad + J^*_\beta(0,s). \end{aligned} \tag{55}$$

Since $J^*_\beta(e,s) \leq \tilde{J}^{(1)}_\beta(e,s)$ and $J^*_\beta(0,s) \geq 0$, from (55) we get that $(1-\beta)^{-1}[J^*_\beta(e,s) - J^*_\beta(0,s)] \leq \lambda(\varphi(e)) + \sum_{s_t\in\mathbb{S}} Q_{ss_t} p(s_t, \varphi(e)) d(e)$. Hence, (S2) is satisfied with $K_s(e) = \lambda(\varphi(e)) + \sum_{s_t\in\mathbb{S}} Q_{ss_t} p(s_t, \varphi(e)) d(e)$.

According to [28, Th. 1], the value function J^*_β is even and quasi-convex and, hence, $J^*_\beta(e,s) \geq J^*_\beta(0,s)$. Hence, (S3) is satisfied with $L_s = 0$.

E. Proof of Part 4)

The result for Part 4) of the theorem for $\mathbb{X} = \mathbb{Z}$ follows from [41, Th. 7.2.3] and for $\mathbb{X} = \mathbb{R}$ the result of Part 4) follows from [29, Th. 5.4.3].

Appendix F
Proof of Theorem 5

For ease of notation, we use τ instead of $\tau^{(1)}$. Let $\mathcal{F}_0$ denote the event $\{E^+_{-1} = 0, S_{-1} = s^\circ\}$ and $\mathcal{F}_\tau$ denote the event $\{E^+_{\tau-1} = 0, S_{\tau-1} = s^\circ\}$.

First note that from (31) we can write

$$\mathbb{E}[\beta^\tau \mid \mathcal{F}_0] = 1 - (1-\beta) M^{(\mathbf{k})}_\beta. \tag{56}$$

Now, consider

$$\begin{aligned} C^{(\mathbf{k})}_\beta &= (1-\beta)\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t (\lambda(U_t) + d(E^+_t)) \,\middle|\, \mathcal{F}_0\right] \\ &= (1-\beta)\mathbb{E}\left[\sum_{t=0}^{\tau-1} \beta^t (\lambda(U_t) + d(E^+_t)) \,\middle|\, \mathcal{F}_0\right] \\ &\quad + (1-\beta)\mathbb{E}\left[\beta^\tau \sum_{t=\tau}^{\infty} \beta^{t-\tau} (\lambda(U_t) + d(E^+_t)) \,\middle|\, \mathcal{F}_0\right] \\ &\overset{(a)}{=} (1-\beta) L^{(\mathbf{k})}_\beta + (1-\beta)\mathbb{E}[\beta^\tau \mid \mathcal{F}_0] \\ &\quad \times \mathbb{E}\left[\sum_{t=\tau}^{\infty} \beta^{t-\tau} (\lambda(U_t) + d(E^+_t)) \,\middle|\, \mathcal{F}_\tau\right] \\ &\overset{(b)}{=} (1-\beta) L^{(\mathbf{k})}_\beta + \Big[1 - (1-\beta) M^{(\mathbf{k})}_\beta\Big] C^{(\mathbf{k})}_\beta \end{aligned} \tag{57}$$

where the first term of (a) uses the definition of $L^{(\mathbf{k})}_\beta$, as given by (30) and the second term of (a) uses strong Markov property; (b) uses (56) and time homogeneity. Rearranging terms in (57) we get (32).

References

[1] Y. Xu and J. P. Hespanha, "Optimal communication logics in networked control systems," in *Proc. Conf. Decis. Control*, 2004, pp. 3527–3532.
[2] O. C. Imer and T. Başar, "Optimal estimation with limited measurements," in *Proc. Conf. Decis. Control*, 2005, pp. 1029–1034.
[3] M. Rabi, G. Moustakides, and J. Baras, "Adaptive sampling for linear state estimation," *SIAM J. Control Optim.*, vol. 50, no. 2, pp. 672–702, 2012.
[4] G. M. Lipsa and N. C. Martins, "Remote state estimation with communication costs for first-order LTI systems," *IEEE Trans. Autom. Control*, vol. 56, no. 9, pp. 2013–2025, Sep. 2011.
[5] A. Nayyar, T. Başar, D. Teneketzis, and V. V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Trans. Autom. Control*, vol. 58, no. 9, pp. 2246–2260, Sep. 2013.
[6] A. Molin and S. Hirche, "Event-triggered state estimation: An iterative algorithm and optimality properties," *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 5939–5946, Nov. 2017.
[7] J. Chakravorty and A. Mahajan, "Remote state estimation with packet drop," in *Proc. Workshop Distrib. Estimation Control Netw. Syst.*, Sep. 2016, pp. 7–12.
[8] G. M. Lipsa and N. C. Martins, "Optimal state estimation in the presence of communication costs and packet drops," in *Proc. Conf. Commun., Control, Comput.*, Sep. 2009, pp. 160–169.
[9] J. Chakravorty, J. Subramanian, and A. Mahajan, "Stochastic approximation based methods for computing the optimal thresholds in remote-state estimation with packet drops," in *Proc. Amer. Control Conf.*, May 2017, pp. 462–467.
[10] X. Gao, E. Akyol, and T. Başar, "Optimal communication scheduling and remote estimation over an additive noise channel," *Automatica*, vol. 88, pp. 57–69, 2018.
[11] J. Chakravorty and A. Mahajan, "Structure of optimal strategies for remote estimation over Gilbert–Elliott channel with feedback," in *Proc. Int. Symp. Inform. Theory*, Jun. 2017, pp. 1272–1276.
[12] 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," *IEEE Trans. Autom. Control*, vol. 63, no. 1, pp. 85–100, Jan. 2018.
[13] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Optimal power management in wireless control systems," *IEEE Trans. Autom. Control*, vol. 59, no. 6, pp. 1495–1510, Jun. 2014.
[14] J. Chakravorty and A. Mahajan, "Fundamental limits of remote estimation of autoregressive Markov processes under communication constraints," *IEEE Trans. Autom. Control*, vol. 62, no. 3, pp. 1109–1124, Mar. 2017.
[15] L. He, J. Chen, and Y. Qi, "Event-based state estimation: Optimal algorithm with generalized closed skew normal distribution," *IEEE Trans. Autom. Control*, vol. 64, no. 1, pp. 321–328, Jan. 2019.
[16] W. Chen, J. Wang, D. Shi, and L. Shi, "Event-based state estimation of hidden Markov models through a Gilbert–Elliott channel," *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3626–3633, Jul. 2017.
[17] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.
[18] K. Plarre and F. Bullo, "On Kalman filtering for detectable systems with intermittent observations," *IEEE Trans. Autom. Control*, vol. 54, no. 2, pp. 386–390, Feb. 2009.
[19] Y. Mo and B. Sinopoli, "A characterization of the critical value for Kalman filtering with intermittent observations," in *Proc. Conf. Decis. Control*, 2008, pp. 2692–2697.
[20] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, Apr. 2007.
[21] K. You, M. Fu, and L. Xie, "Mean square stability for Kalman filtering with Markovian packet losses," *Automatica*, vol. 47, no. 12, pp. 2647–2657, Dec. 2011.
[22] E. R. Rohr, D. Marelli, and M. Fu, "Kalman filtering with intermittent observations: On the boundedness of the expected error covariance," *IEEE Trans. Autom. Control*, vol. 59, no. 10, pp. 2724–2738, Oct. 2014.
[23] J. Wu, G. Shi, B. D. Anderson, and K. H. Johansson, "Kalman filtering over Gilbert–Elliott channels: Stability conditions and critical curve," *IEEE Trans. Autom. Control*, vol. 63, no. 4, pp. 1003–1017, Apr. 2018.
[24] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 868–886, May 1996.
[25] S. Yang, A. Kavcic, and S. Tatikonda, "Feedback capacity of finite-state machine channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 799–810, Mar. 2005.
[26] J. C. Walrand and P. Varaiya, "Optimal causal coding-decoding problems," *IEEE Trans. Inf. Theory*, vol. 29, no. 6, pp. 814–820, Nov. 1983, *arXiv:1005.2442*.
[27] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Trans. Autom. Control*, vol. 58, no. 7, pp. 1644–1658, Jul. 2013.
[28] J. Chakravorty and A. Mahajan, "Sufficient conditions for the value function and optimal strategy to be even and quasi-convex," *IEEE Trans. Autom. Control*, vol. 63, no. 11, pp. 3858–3864, Nov. 2018.
[29] O. H. Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. New York, NY, USA: Springer, 1996.
[30] J. Subramanian and A. Mahajan, "Renewal Monte Carlo: Renewal theory based reinforcement learning," in *Proc. Conf. Decis. Control*, 2019, pp. 5759–5764.
[31] 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.
[32] J. L. Maryak and D. C. Chin, "Global random optimization by simultaneous perturbation stochastic approximation," *IEEE Trans. Autom. Control*, vol. 53, no. 3, pp. 780–783, Apr. 2008.
[33] V. Katkovnik and Y. Kulchitsky, "Convergence of a class of random search algorithms," *Autom. Remote Control*, vol. 33, no. 8, pp. 1321–1326, 1972.
[34] S. Bhatnagar, H. Prasad, and L. Prashanth, Stochastic Approximation Algorithms. *London*, U.K.: Springer, 2013.
[35] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, no. 5, pp. 1253–1265, Sep. 1960.
[36] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, no. 5, pp. 1977–1997, Sep. 1963.
[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," Jan. 2017, *arXiv: 1412.6980*.
[38] "Remote estimation over a packet-drop channel with Markovian state," 2019. https://doi.org/10.24433/CO.8529792.v1
[39] X. Gao, E. Akyol, and T. Başar, "On remote estimation with multiple communication channels," in *Proc. Amer. Control Conf.*, Jul. 2016, pp. 5425–5430.
[40] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1986.
[41] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*. New York, NY, NY, USA: Wiley, 1999.

Jhelum Chakravorty (M'12) received the B.E. degree in electrical engineering from Jadavpur University, Kolkata, India, in 2007, the M.Tech. degree in systems and control engineering from the Indian Institute of Technology Bombay, Mumbai, India, in 2010, and the Ph.D. degree in electrical and computer engineering from McGill University, Montreal, QC, Canada, in 2017.

She is a Postdoctoral Fellow with the Department of Computer Science, McGill University, Montreal, Canada. Prior to the doctoral study, she worked as a Research Assistant with the Indian Institute of Technology Bombay from 2007 to 2008, and as a Research Associate with the Indian Institute of Science from 2010 to 2011. She is the recipient of the Best Student Paper Award at 6th IFAC Workshop on Distributed Estimation and Control in Networked Systems 2016. Her current research interests include reinforcement learning, decentralized stochastic control, information theory and real-time communication.

Aditya Mahajan (S'06–M'09–SM'14) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology Kanpur, Kanpur, India, in 2003, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of Michigan, Ann Arbor, MI, USA, in 2006 and 2008.

From 2008 to 2010, he was a Postdoctoral Researcher with the Yale University, New Haven, CT, USA. Since 2010, he has been with the Department of Electrical and Computer Engineering, McGill University, Montreal, Canada, where he is currently an Associate Professor.

Prof. Mahajan is the recipient of the 2015 George Axelby Outstanding Paper Award, 2014 CDC Best Student Paper Award (as supervisor), and the 2016 NecSys Best Student Paper Award (as supervisor).