2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Optimal Scheduling Policies for Remote Estimation of Autoregressive Markov Processes over Time-Correlated Fading Channel

Manali Dutta and Rahul Singh

***Abstract*— We consider the problem of optimally scheduling transmissions for remote estimation of a discrete-time autoregressive Markov process that is driven by white Gaussian noise. A sensor observes this process, and then decides to either encode the current state of this process into a data packet and attempts to transmit it to the estimator over an unreliable wireless channel modeled as a Gilbert-Elliott channel [1] [2] [3], or does not send any update. Each transmission attempt consumes λ units of transmission power, and the remote estimator is assumed to be linear. The channel state is revealed only via the feedback (ACK/NACK) of a transmission, and hence the channel state is not revealed if no transmission occurs. The goal of the scheduler is to minimize the expected value of an infinite-horizon cumulative discounted cost, in which the instantaneous cost is composed of the following two quantities: (i) squared estimation error, (ii) transmission power. We posed this problem as a partially observable Markov decision process (POMDP), in which the scheduler maintains a belief about the current state of the channel, and makes decisions on the basis of the current value of the error $e(t)$ (defined in (6)), and the belief state. To aid its analysis, we introduce an easier-to-analyze "folded POMDP." We then analyze this folded POMDP and show that there is an optimal scheduling policy that has threshold structure, i.e. for each value of the error e, there is a threshold $b^\star(e)$ such that when the error is equal to e, this policy transmits only when the current belief state is greater than $b^\star(e)$.**

***Index Terms*— Remote estimation, Gilbert-Elliott channel, partially observable Markov decision process (POMDP), threshold-type policy.**

I. Introduction

A. Literature Overview

In distributed networked control systems (NCS), several network nodes are connected through a communication network, which enables them to exchange information and collaborate to achieve a common goal [4], [5]. In such a control architecture, decision-making is decentralized since each node can communicate only with its neighbors, and makes decisions based on its own local information. Such systems have gained widespread interest in recent years due to their ability to enable remote control and monitoring of physical systems. They are used in various fields, including industrial automation, robotics, and transportation systems. Remote state estimation is one of the fundamental problems in NCS. Such a system is comprised of a sensor that observes an underlying process, encodes its observations into data packets, and then transmits it over a communication channel to a remote estimator that has a different location. We will be exclusively interested in the case where the wireless medium is used for carrying out these transmissions. Since wireless devices are typically battery-operated, and transmissions consume energy, it is not efficient for the sensor to continually transmit the observations. Hence, in order to strike a balance between the communication cost and estimation error, sensors typically employ scheduling policies that make dynamic decisions regarding whether or not to send a packet, based on the information available with them.

In this work, we consider a remote estimator that attempts to estimate the state of a Markovian source in real-time. [6]–[10] consider the remote estimation problem for systems in which the communication channels are ideal and packet transmissions are always successful. [6] imposes constraints on the number of transmissions, while [7] fixes the estimator to be "Kalman-like" and then optimizes the scheduling policy. It shows that the optimal scheduling decisions are solely a function of the current value of the state estimation error. [8] does not impose any conditions on the structure of the scheduling policy or the estimator, and uses majorization theoryin order to show that a threshold-type communication policy at the sensor, and a Kalman-like estimator are jointly optimal. [9] shows that the structure of optimal communication and estimation policies derived in [8] continue to hold when additionally there are energy constraints on the transmitter. [10] also derives jointly optimal scheduling policy and estimator for the average cost problem by viewing it as a limiting case of the discounted cost problem in the limit the discount factor approaches unity. Policies that transmit only when the current value of the estimation error is greater than a threshold are also called event-triggered communication policies, and are studied in [11]–[14] for error-free communication channels.

Transmissions using the wireless medium are unreliable due to factors such as environmental conditions, interference, blockages. [15]–[17] study remote estimation over wireless medium by modeling packet losses to be random and i.i.d. across times and show the existence of an optimal policy that has a threshold structure with respect to the estimation error. A more realistic way to model the wireless fading channel is to model it using a finite state Markov chain [18]. In this work, we model the unreliable wireless channel as a Gilbert-Elliott channel [1], which is a Markovian channel in which the channel state can assume two values. At each discrete time, the channel is either in a good state and packet transmissions are successful, or it is in a bad state and any

This work was partially supported by the SERB Grant SRG/2021/002308.
The authors are with the Department of Electrical and Communication Engineering, Indian Institute of Science, Bengaluru, Karnataka 560012, India (e-mail: manalidutta@iisc.ac.in and rahulsingh@iisc.ac.in)

2023 62nd IEEE Conference on Decision and Control (CDC) | 979-8-3503-0124-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/CDC49753.2023.10384144

attempted transmission fails. The works [2], [19], [20] study remote estimation over Markovian channels. [19] derives transmission power control and remote estimation policies that are jointly optimal. It assumes that the channel state is instantaneously known to the sensor and estimator. The problem is formulated as a partially observable Markov decision problem (POMDP), with a belief over the common information available with the sensor and the estimator, and it is shown that the optimal transmission strategy has a threshold structure with respect to the belief state. A model similar to [19] is considered in [2] and [20], but with the difference that the channel state is known perfectly to the sensor with a delay of one unit. The optimality of a transmission policy that is of threshold-type with respect to the error is shown. However, obtaining a perfect knowledge of the channel state is difficult due to the complexity involved in measuring the characteristics of the communication channel.

B. Contributions

We address the problem of optimally scheduling transmissions to a remote linear estimator when the transmitter does not employ a probing mechanism to continually sense the channel state, and hence gets to know the channel state only via acknowledgments sent by the estimator when there is a transmission attempt. If there is no transmission attempt, then the current state is not known. The underlying process at the sensor which is being estimated is an autoregressive (AR) Markov process [20], [21], and our objective is to minimize the infinite-horizon cumulative expected discounted cost. Our main contributions are as follows:

1) We pose the problem faced by the sensor as a dynamic optimization problem that involves minimizing an infinite-horizon cumulative expected value of a discounted cost that consists of i) the squared estimation error and, ii) the transmission power. We show that this can be formulated as a POMDP [22] in which the state comprises of (a) the "belief state," i.e. the conditional probability (conditioned on the information available with the sensor) that the channel state is good, (b) the current value of the error.
2) Since our POMDP involves a one-stage cost function that is unbounded, it is not obvious that the value iteration algorithm [23], [24] can be used to solve the POMDP. We show that, under mild assumptions (Assumption 1) on the AR process and the Markovian transition probabilities of the channel (14), the value iteration algorithm converges and yields an optimal policy.
3) We then introduce a certain "folded POMDP," that is much easier to analyze since the error in this folded POMDP is always positive, and show that it is equivalent to the original POMDP, i.e. we can recover an optimal policy for the original problem by solving the folded POMDP. The concept of "folding a Markov decision process (MDP)" was introduced in [25]. However, since in our setup the channel state is not known by the sensor we cannot use the results of [25].
4) We then derive novel structural results for the original POMDP by analyzing this folded POMDP. Specifically, we show that the optimal transmission strategy exhibits a threshold structure with respect to the belief state. Since POMDPs are PSPACE hard [26], such a result reduces the policy search space. Though there is an extensive literature on structural results for POMDPs [22], we cannot use these as they mostly restrict the state-space to be a simplex. Notably, [3] considers the problem of age minimization, and derives structural results for a POMDP in which the state-space is the Cartesian product of a simplex with the natural numbers.

Notation: Let $\mathbb{N}, \mathbb{R}_+$ and $\mathbb{R}_-$ denote the set of natural numbers, non-negative and non-positive real numbers, respectively, and let $\mathbb{Z}_{\geq 0} := \{0\} \cup \mathbb{N}$. For a sigma-algebra $\mathcal{F}$, $\mathbb{E}(\cdot|\mathcal{F})$ denotes the conditional expectation with respect to $\mathcal{F}$. $\mathcal{N}(\mu, \sigma^2)$ is the Gaussian distribution with mean μ and variance σ^2, and $\delta_x(\cdot)$ is the delta function with unit mass at x. Due to space limitations, we have moved proofs of some results to the detailed technical report [27].

II. Problem Formulation

Consider a networked system comprising of a sensor and a remote estimator. The sensor observes an AR process $\{x(t)\}$ that evolves as follows, $x(0) = 0$ and,

$$x(t+1) = ax(t) + w(t),\ t = 0, 1, 2, \ldots,$$

where $a, x(t) \in \mathbb{R}$, $w(t)$ is an i.i.d. Gaussian noise process that satisfies $w(t) \sim \mathcal{N}(0, 1)$. Sensor encodes its observations into data packets and transmits these to a remote estimator via an unreliable wireless channel. We denote the state of the channel at time t by $c(t) \in \{0, 1\}$. $c(t) = 0$ denotes that the channel is in "bad state" and any transmissions are unsuccessful, while $c(t) = 1$ denotes that any packet transmitted at t will be delivered to the estimator. The channel has a *memory*, and hence we assume that $\{c(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ is a Markov process with parameters,

$$p_{01} := \mathbb{P}(c(t+1) = 1 | c(t) = 0), \quad (1)$$

$$p_{11} := \mathbb{P}(c(t+1) = 1 | c(t) = 1), \quad (2)$$

where $p_{01}, p_{11} \in (0, 1)$. Let $u(t) \in \{0, 1\}$ denote the decision made by the sensor regarding whether ($u(t) = 1$) or not ($u(t) = 0$) to attempt a packet transmission at time t. We assume that each transmission attempt consumes λ units of power/resource. Let $y(t)$ denote the output of the channel, i.e. the observation made by the estimator at time t. Let $\hat{x}(t)$ denote the state of the estimator, or equivalently the point estimate made by the estimator. It evolves as follows,

$$\hat{x}(t) = \begin{cases} a\hat{x}(t-1) & \text{if } y(t) = \Xi, \\ y(t) & \text{otherwise,} \end{cases} \quad (3)$$

where $y(t) = \Xi$ denotes that no packet was received by the estimator, either because no transmission was carried out, or because the channel state was bad. Scheduler does not observe the channel state $c(t)$. However, if there is a successful transmission at t, then the estimator sends an acknowledgment to the sensor. Hence, if $u(t) = 1$, then the channel state $c(t)$ at t is known to the sensor at time $t+1$, or we say that upon transmitting a packet the scheduler gets to "probe" the channel. Let $z(t)$ be the delayed channel state information delivered to the sensor upon a transmission attempt made at $t-1$. Note that no channel information is delivered if no transmission attempt is made. The scheduler has access to $\{z(s)\}_{s=0}^{t-1}$ and also $\{u(s)\}_{s=0}^{t-1}, \{x(s)\}_{s=0}^{t}$ while making the decision at time t. The goal of the scheduler is to choose $\{u(t)\}_{t\in\mathbb{Z}_{\geq 0}}$ so as to minimize the expected value of the cumulative errors, as well as keep the cumulative transmission power utilized at minimal level. Hence, the goal of the scheduler at the sensor is to dynamically make the decisions $\{u(t)\}_{t\in\mathbb{Z}_{\geq 0}}$ so as solve,

$$\min_{\phi} \ \mathbb{E}_{\phi}\left(\sum_{t=0}^{\infty}\beta^t\big((x(t)-\hat{x}(t))^2+\lambda u(t)\big)\right), \quad (4)$$

where $\beta \in (0,1)$ is the discount factor, $\phi = \{\phi_t\}_{t\in\mathbb{Z}_{\geq 0}}$ is a measurable policy that for each time t maps the history $\mathcal{F}_t := \sigma(\{x(s), u(s), z(s)\}_{s=0}^{t-1}, x(t))$ to decision $u(t)$, and $\mathbb{E}_{\phi}$ denotes that the expectation is taken w.r.t. the measure induced by the policy ϕ.

III. POMDP Formulation

Note that while solving (4), the channel state is not completely observed by the scheduler. Let $b(t) := \mathbb{E}(c(t)|\mathcal{F}_t)$ be its estimate of the current channel state at time t. This can be updated recursively using the ACK/NACK as follows,

$$b(t+1) = \begin{cases} p_{11} & \text{if } u(t)=1, c(t)=1,\\ p_{01} & \text{if } u(t)=1, c(t)=0,\\ \mathcal{T}(b(t)) & \text{if } u(t)=0, \end{cases} \quad (5)$$

where for $x \in \mathbb{R}$, we define $\mathcal{T}(x) := xp_{11} + (1-x)p_{01}$. Let $\{e(t)\}_{t\in\mathbb{Z}_{\geq 0}}$ be the "error process"[1] which evolves as follows, $e(0) = 0$, and for $t \geq 0$,

$$e(t+1) = \begin{cases} ae(t)+w(t) & \text{if } u(t)c(t)=0,\\ w(t) & \text{if } u(t)c(t)=1. \end{cases} \quad (6)$$

For the purpose of solving (4), we pose it as the following POMDP [21], [22],

$$\min_{\phi} \ \mathbb{E}_{\phi}\left(\sum_{t=0}^{\infty}\beta^t(d(e(t),b(t),y(t),u(t)))\right), \quad (7)$$

in which the system state at time t is given by $(e(t), b(t), y(t))$, where $e(t) \in \mathbb{R}, b(t) \in [0,1]$ and $y(t) \in \mathbb{R}\cup\{\Xi\}$, and $b(t), e(t)$ evolve according to (5) and (6), respectively. $u(t) \in \{0,1\}$, and the instantaneous cost incurred at time t is equal to $d(e(t), b(t), y(t), u(t))$, where

$$d(e,b,y,u) := \begin{cases} e^2+\lambda u & \text{if } y = \Xi,\\ \lambda & \text{if } y \neq \Xi. \end{cases} \quad (8)$$

[1] $e(t)$ should not be confused with the estimation error

Policy ϕ chooses $u(t)$ on the basis of the operational history $\{(e(s), b(s))\}_{s=0}^{t} \cup \{y(s)\}_{s=0}^{t-1}$. Next, we consider the following simpler POMDP with a reduced state-space,

$$\min_{\phi}\mathbb{E}_{\phi}\left(\sum_{t=0}^{\infty}\beta^t\tilde{d}(e(t),b(t),u(t))\right), \quad (9)$$

where $b(t)$ and $e(t)$ evolve according to (5) and (6) respectively, $u(t) \in \{0,1\}$, and the instantaneous cost function $\tilde{d}$ is given by,

$$\tilde{d}(e,b,u) := \begin{cases} a^2e^2+1 & \text{if } u=0,\\ (1-b)(a^2e^2+1)+\lambda & \text{if } u=1, \end{cases} \quad (10)$$

and ϕ chooses $u(t)$ on the basis of $\{(e(s), b(s))\}_{s=0}^{t}$. [27] shows that (9)-(10) is equivalent to (7)-(8), and hence a policy that is optimal for (9) also solves (7). Henceforth, we will focus exclusively on solving (9).

We now describe the controlled transition probabilities associated with (9)-(10). Let $p(e_+, b_+ \mid e, b; u)$ denote the transition density function from the current state (e,b) to the next state (e_+, b_+) under the application of the action u. Consider the following two possibilities for u:

Case i) $u = 0$: Then the state at the next step (e_+, b_+) has the following density,

$$p(e_+, b_+ \mid e, b; 0) = \\ \exp(-(e_+-ae)^2/2)\delta_{p_{11}b+p_{01}(1-b)}(b_+). \quad (11)$$

Case ii) $u = 1$: The density function of the resulting joint distribution of (e_+, b_+) is as follows,

$$p(e_+, b_+ \mid e, b; 1) = b\exp(-e_+^2/2)\ \delta_{p_{11}}(b_+) \\ +(1-b)\exp(-(e_+-ae)^2/2)\ \delta_{p_{01}}(b_+). \quad (12)$$

A. Value Iteration

We now show that under mild assumptions on the system parameters, value iteration algorithm can be used to solve (9). Value iteration algorithm is popularly used in order to solve MDPs. However, in order that we can use it to solve (9), we need to verify whether our POMDP satisfies certain conditions [24, p. 46]. This is done below.

Let $J^{(\beta)}(e, b; \phi)$ be the β-discounted cost (9) incurred by ϕ when the system starts in state $(e,b) \in \mathbb{R}\times[0,1]$,

$$J^{(\beta)}(e,b;\phi) := \mathbb{E}_{\phi}\left(\sum_{t=0}^{\infty}\beta^t\tilde{d}(e(t),b(t),u(t))\right). \quad (13)$$

Assumption 1: The Markovian channel probabilities and the system parameter a satisfy the following condition

$$a^2(1-p_{01}) < 1. \quad (14)$$

Following result is shown in [27].

Lemma 3.1: Consider the POMDP (9), and let Assumption 1 hold. The following properties hold:

P1. The one-stage cost function $\tilde{d}(e,b,u)$ (10) is continuous, non-negative, and inf-compact on $(\mathbb{R}\times[0,1]\times\{0,1\})$.

P2. The transition kernels $\{P(\cdot, u, \cdot)\}_{u\in\{0,1\}}$ that describe the transition probabilities which result when control u is applied, are strongly continuous.

P3. There exists a policy ϕ for which $J^{(\beta)}(e, b; \phi) < \infty$ for each $(e, b) \in \mathbb{R} \times [0, 1]$.

The above result allows us to use value iteration. This is shown next. We begin by describing these iterations. Let $V_n^{(\beta)}$ denote the value function at stage $n \in \mathbb{Z}_{\geq 0}$ of the value iterations [24]. We have the following for all $(e, b) \in \mathbb{R} \times [0, 1]$,

$$V_{n+1}^{(\beta)}(e, b) = \min_{u\in\{0,1\}} Q_{n+1}^{(\beta)}(e, b; u), \tag{15}$$

where,

$$Q_{n+1}^{(\beta)}(e, b; 0) := (a^2e^2 + 1) + \beta\mathbb{E}\Big[V_n^{(\beta)}(ae + w, \mathcal{T}(b))\Big],$$
$$Q_{n+1}^{(\beta)}(e, b; 1) := (1 - b)(a^2e^2 + 1) + \lambda +$$
$$\beta\mathbb{E}\Big[bV_n^{(\beta)}(w, p_{11}) + (1 - b)V_n^{(\beta)}(ae + w, p_{01})\Big],$$

with,

$$V_0^{(\beta)}(e, b) = 0. \tag{16}$$

Let $V^{(\beta)}(e, b)$ denote the β-discounted cost function for the POMDP (9) , i.e.,

$$V^{(\beta)}(e, b) := \min_{\phi} J^{(\beta)}(e, b; \phi), \tag{17}$$

where $J^{(\beta)}(e, b; \phi)$ is given by (13).

The following proposition introduces the optimality equation for $V^{(\beta)}$ and shows the convergence of value iteration method to $V^{(\beta)}$. It follows from [24, Lemma 4.2.8, Theorem 4.2.3], and involves P1.-P3. derived in Lemma 3.1.

Proposition 3.1: Consider the POMDP (9) that satisfies Assumption 1. Then,

a) Value iteration algorithm (15)-(16) converges to $V^{(\beta)}$ (17), i.e. for each $(e, b) \in \mathbb{R} \times [0, 1]$,

$$\lim_{n\to\infty} V_n^{(\beta)}(e, b) = V^{(\beta)}(e, b).$$

b) Value function $V^{(\beta)}$ (17) satisfies the following equation, for each $(e, b) \in \mathbb{R} \times [0, 1]$,

$$V^{(\beta)}(e, b) = \min_{u\in\{0,1\}} Q^{(\beta)}(e, b; u), \tag{18}$$

where,

$$Q^{(\beta)}(e, b; 0) = (a^2e^2 + 1) + \beta\mathbb{E}\Big[V^{(\beta)}(ae + w, \mathcal{T}(b))\Big],$$
$$Q^{(\beta)}(e, b; 1) = (1 - b)(a^2e^2 + 1) + \lambda$$
$$+ \beta\mathbb{E}\Big[bV^{(\beta)}(w, p_{11}) + (1 - b)V^{(\beta)}(ae + w, p_{01})\Big].$$

c) There exists an optimal stationary deterministic policy that implements the minimizer of the right-hand side of (18) for each state $(e, b) \in \mathbb{R} \times [0, 1]$.

B. Folding the POMDP

We now construct an equivalent "folded POMDP" with a state-space $\mathbb{R}_+ \times [0, 1]$, such that it suffices to study this POMDP in lieu of the original POMDP that has the state-space $\mathbb{R} \times [0, 1]$. Specifically, the error of the folded POMDP does not take negative values, in contrast to the original POMDP in which the error takes both nonnegative and negative values. Consequently, while analyzing the set of optimal policies, it is convenient to work with the folded POMDP rather than the original POMDP. The work [25] introduces the concept of a folded MDP. More specifically, for MDPs in which the state-space is $\mathbb{R}$, it shows that under certain conditions on the transition probability kernel and the instantaneous cost function, one can construct an equivalent MDP, called the folded MDP that has a state-space $\mathbb{R}_+$ and is easier to study. However, the framework of [25] cannot be used in order to study POMDPs. Hence, we now utilize the structure of POMDP (9) to introduce the folded POMDP. Before discussing this construction, we present a structural property of the value function $V^{(\beta)}$ of the POMDP (9).

Proposition 3.2: The functions $Q^{(\beta)}(\cdot, b), V^{(\beta)}(\cdot, b)$ for the POMDP (9) are even, i.e., we have the following for all $(e, b) \in \mathbb{R} \times [0, 1], u \in \{0, 1\}$,

$$Q^{(\beta)}(e, b; u) = Q^{(\beta)}(|e|, b; u), V^{(\beta)}(e, b) = V^{(\beta)}(|e|, b).$$

Proof: Please see [27] for the proof. ■

In what follows, we will use $\tilde{\phi}, \tilde{u}, \tilde{e}$ and $\tilde{b}$ to denote a policy, control, "error" and "belief state," respectively for the folded POMDP.

Definition 3.1 (Folded POMDP): Given (9), its folded version is a POMDP on the state-space $\mathbb{R}_+ \times [0, 1]$, control space $\{0, 1\}$, and has the transition density function $\tilde{p}$ as follows,

$$\tilde{p}\Big(\tilde{e}_+, \tilde{b}_+ \mid \tilde{e}, \tilde{b}; \tilde{u}\Big) = p\Big(\tilde{e}_+, \tilde{b}_+ \mid \tilde{e}, \tilde{b}; \tilde{u}\Big)$$
$$+ p\Big(-\tilde{e}_+, \tilde{b}_+ \mid \tilde{e}, \tilde{b}; \tilde{u}\Big), \tag{19}$$

where $\tilde{e}, \tilde{e}_+ \in \mathbb{R}_+$, $\tilde{b}, \tilde{b}_+ \in [0, 1]$ and $\tilde{u} \in \{0, 1\}$. The instantaneous cost function $\tilde{d}$ (10), and the objective (9) remain the same as that of the original one, with $e(t), u(t)$ replaced by $\tilde{e}(t), \tilde{u}(t)$ respectively.

For ease of notation we denote:

$$\psi(v) := \exp(-v^2/2), \tilde{\psi}(v, s) := \psi(v - s) + \psi(v + s).$$

We next show the equivalence of the POMDP (9) and its folded version. We begin by discussing few properties of the folded POMDP. Let $\tilde{J}^{(\beta)}(\tilde{e}, \tilde{b}; \tilde{\phi})$ and $\tilde{V}^{(\beta)}(\tilde{e}, \tilde{b})$ denote the β-discounted cost and β-discounted value function given the initial state $(\tilde{e}, \tilde{b})$, respectively, of the folded POMDP. These are analogous to (13) and (17), respectively, of the POMDP (9). We can show that under Assumption 1, the folded POMDP $(\mathbb{R}_+ \times [0, 1], \{0, 1\}, \tilde{p}, d)$ also satisfies the properties P1.-P3. stated in Lemma 3.1. The proof is similar to that of Lemma 3.1, which deals with the POMDP (9). Thus, value iteration can be used to solve the folded POMDP too. Let $\tilde{V}_n^{(\beta)}$ denote the iterates during stage n of the value iteration algorithm [24] when it is applied to solve the folded POMDP. We have the following for all $(\tilde{e},\tilde{b}) \in \mathbb{R}_+ \times [0,1]$,

$$\tilde{V}_{n+1}^{(\beta)}\big(\tilde{e},\tilde{b}\big) = \min_{\tilde{u}\in\{0,1\}} \tilde{Q}_{n+1}^{(\beta)}\big(\tilde{e},\tilde{b};\tilde{u}\big), \tag{20}$$

where, $\tilde{Q}_{n+1}^{(\beta)}\big(\tilde{e},\tilde{b};0\big)$ is as follows,

$$\begin{aligned}
&\tilde{Q}_{n+1}^{(\beta)}\big(\tilde{e},\tilde{b};0\big) = (a^2\tilde{e}^2+1) + \beta \\
&\times \int_{\mathbb{R}_+} \tilde{p}\Big(\tilde{e}_+,\mathcal{T}\big(\tilde{b}\big) \mid \tilde{e},\tilde{b};0\Big)\tilde{V}_n^{(\beta)}\Big(\tilde{e}_+,\mathcal{T}\big(\tilde{b}\big)\Big)\,d\tilde{e}_+, \quad (21)\\
&= (a^2\tilde{e}^2+1) \\
&+\beta\int_{\mathbb{R}_+} \tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}_n^{(\beta)}\Big(\tilde{e}_+,\mathcal{T}\big(\tilde{b}\big)\Big)\,d\tilde{e}_+, \quad (22)
\end{aligned}$$

where (22) follows from (19) and the definition of transition density in (11).

While for $\tilde{u}=1$ we get,

$$\begin{aligned}
&\tilde{Q}_{n+1}^{(\beta)}\big(\tilde{e},\tilde{b};1\big) \\
&= (1-b)(a^2\tilde{e}^2+1) + \lambda + \\
&+\beta\Bigg[\tilde{b}\int_{\mathbb{R}_+} \tilde{p}\big(\tilde{e}_+,p_{11} \mid \tilde{e},\tilde{b};1\big)\tilde{V}_n^{(\beta)}(\tilde{e}_+,p_{11})\,d\tilde{e}_+ \\
&+(1-\tilde{b})\int_{\mathbb{R}_+} \tilde{p}\big(\tilde{e}_+,p_{01} \mid \tilde{e},\tilde{b};1\big)\tilde{V}_n^{(\beta)}(\tilde{e}_+,p_{01})\,d\tilde{e}_+\Bigg] \\
&= (1-b)(a^2\tilde{e}^2+1) + \lambda \\
&+\beta\tilde{b}\int_{\mathbb{R}_+} 2\psi(\tilde{e}_+)\tilde{V}_n^{(\beta)}(\tilde{e}_+,p_{11})\,d\tilde{e}_+ \\
&+\beta(1-\tilde{b})\int_{\mathbb{R}_+} \tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}_n^{(\beta)}(\tilde{e}_+,p_{01})\,d\tilde{e}_+, \quad (23)
\end{aligned}$$

where (23) follows from (12) and (19). Initialization is as follows, for each $(\tilde{e},\tilde{b}) \in \mathbb{R}_+ \times [0,1]$,

$$\tilde{V}_0^{(\beta)}\big(\tilde{e},\tilde{b}\big) = 0. \tag{24}$$

We have the following properties for the folded POMDP, analogous to the results for the original POMDP shown in Proposition 3.1. These follow from [24, Theorem 4.2.3],

a) The value iteration algorithm with iterates $\tilde{V}_n^{(\beta)}$ converges to $\tilde{V}^{(\beta)}$, i.e. for each $(\tilde{e},\tilde{b}) \in \mathbb{R}_+ \times [0,1]$,

$$\lim_{n\to\infty} \tilde{V}_n^{(\beta)}\big(\tilde{e},\tilde{b}\big) = \tilde{V}^{(\beta)}\big(\tilde{e},\tilde{b}\big). \tag{25}$$

b) The value function $\tilde{V}^{(\beta)}$ satisfies the following optimality equation, for each $(\tilde{e},\tilde{b}) \in \mathbb{R}_+ \times [0,1]$,

$$\tilde{V}^{(\beta)}\big(\tilde{e},\tilde{b}\big) = \min_{\tilde{u}\in\{0,1\}} \tilde{Q}^{(\beta)}\big(\tilde{e},\tilde{b};\tilde{u}\big), \tag{26}$$

where,

$$\begin{aligned}
&\tilde{Q}^{(\beta)}\big(\tilde{e},\tilde{b};0\big) = (a^2\tilde{e}^2+1) + \beta \\
&\times \int_{\mathbb{R}_+} \tilde{p}\Big(\tilde{e}_+,\mathcal{T}\big(\tilde{b}\big) \mid \tilde{e},\tilde{b};0\Big)\tilde{V}^{(\beta)}\Big(\tilde{e}_+,\mathcal{T}\big(\tilde{b}\big)\Big)\,d\tilde{e}_+,
\end{aligned}$$

and,

$$\begin{aligned}
&\tilde{Q}^{(\beta)}\big(\tilde{e},\tilde{b};1\big) = (1-b)(a^2\tilde{e}^2+1) + \lambda \\
&+\beta\Bigg[\tilde{b}\int_{\mathbb{R}_+} \tilde{p}\big(\tilde{e}_+,p_{11} \mid \tilde{e},\tilde{b};1\big)\tilde{V}_n^{(\beta)}(\tilde{e}_+,p_{11})\,d\tilde{e}_+ \\
&+(1-\tilde{b})\int_{\mathbb{R}_+} \tilde{p}\big(\tilde{e}_+,p_{01} \mid \tilde{e},\tilde{b};1\big)\tilde{V}_n^{(\beta)}(\tilde{e}_+,p_{01})\,d\tilde{e}_+\Bigg]
\end{aligned}$$

c) There exists an optimal stationary deterministic policy that implements the minimizer of the right-hand side of (26) for each state $(\tilde{e},\tilde{b}) \in \mathbb{R}_+ \times [0,1]$.

The following is proved in [27], and shows the equivalence of the folded POMDP and the original one (9).

Proposition 3.3: The functions $\tilde{Q}^{(\beta)},\tilde{V}^{(\beta)}$ corresponding to the folded POMDP match with $Q^{(\beta)},V^{(\beta)}$ of the original POMDP on $\mathbb{R}_+ \times [0,1]$, i.e., for all $(e,b) \in \mathbb{R} \times [0,1]$ and $u \in \{0,1\}$, we have,

$$Q^{(\beta)}(e,b;u) = \tilde{Q}^{(\beta)}(|e|,b;u), V^{(\beta)}(e,b) = \tilde{V}^{(\beta)}(|e|,b).$$

IV. Structural Results

Definition 4.1 (Threshold-type Policy): We say that a scheduling policy for the folded POMDP $\tilde{\phi}:\mathbb{R}_+ \times [0,1] \mapsto \{0,1\}$ is of threshold type if for each $\tilde{e} \in \mathbb{R}_+$, there exists a threshold $b^\star(\tilde{e})$ such that it transmits in state $(\tilde{e},\tilde{b}) \in \mathbb{R}_+ \times [0,1]$ only if $\tilde{b} \geq b^\star(\tilde{e})$. Similarly, a policy for (9) has a threshold structure if it transmits in state $(e,b) \in \mathbb{R} \times [0,1]$ only when $b \geq b^\star(e)$.

The following is commonly assumed about the Gilbert-Elliott channels [3], [28], and we will require this while analyzing properties of the optimal policy.

***Assumption** 2:* The Markovian channel parameters (1), (2) satisfy $p_{11} \geq p_{01}$.

We now show that the optimal policy of the folded POMDP has a threshold-type structure.

Theorem 4.1: Consider the folded POMDP $(\mathbb{R}_+ \times [0,1],\{0,1\},\tilde{p},d)$. Its value function $\tilde{V}^{(\beta)}$ satisfies the following properties:

(A) For each $\tilde{b}$, the function $\tilde{V}^{(\beta)}\big(\cdot,\tilde{b}\big)$ is non-decreasing (with regards to $\tilde{e}$).

(B) For each $\tilde{e}$, the function $\tilde{V}^{(\beta)}(\tilde{e},\cdot)$ is non-increasing (with respect to $\tilde{b}$).

(C) For beliefs $x,y,z,\tilde{b}$ such that $x \geq y$ and $z = \tilde{b}x + (1-\tilde{b})y$, we have,

$$\begin{aligned}
&(1-\tilde{b})\lambda + \tilde{b}\tilde{V}^{(\beta)}(\tilde{e},x) \\
&\quad + (1-\tilde{b})\tilde{V}^{(\beta)}(\tilde{e},y) \geq \tilde{V}^{(\beta)}(\tilde{e},z). \quad (27)
\end{aligned}$$

(D) For each $\tilde{e} \in \mathbb{R}_+$, there exists a threshold $\tilde{b}^\star(\tilde{e})$ such that it is optimal to transmit only when $\tilde{b} \geq \tilde{b}^\star(\tilde{e})$. Thus, the optimal strategy corresponding to $\tilde{V}^{(\beta)}$ exhibits a threshold structure.

Proof: We will prove (A)-(D) for the iterates $\tilde{V}_n^{(\beta)}(\tilde{e},\tilde{b}), n \in \mathbb{Z}_{\geq 0}$ in (20). We will show this via induction. The result would then follow from (25), since we have $\lim_{n\to\infty}\tilde{V}_n^{(\beta)}(\tilde{e},\tilde{b}) = \tilde{V}^{(\beta)}(\tilde{e},\tilde{b})$.

Since $\tilde{V}_0^{(\beta)}\big(\tilde{e},\tilde{b}\big) \equiv 0$ (24), (A)-(D) hold for $n=0$. Next, assume that (A)-(C) hold for $k=1,2,\ldots,n$. The proof is divided into four steps. We will firstly show that the threshold property (D) holds for $k=n+1$, and then show (A)-(C) also hold for $k=n+1$.

Step I: (D) holds for step $n+1$: We have $\tilde{V}^{(\beta)}_{n+1}(\tilde{e},\tilde{b}) = \min_{\tilde{u}\in\{0,1\}} \tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};\tilde{u})$. Firstly, note that $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};1)$ is a linear function of $\tilde{b}$ by the definition of $\tilde{Q}^{(\beta)}_{n+1}$ in (23). We will now show that $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};0)$ is concave in $\tilde{b}$. Note that $\tilde{V}^{(\beta)}_{n}(\tilde{e},\tilde{b})$ is concave with respect to $\tilde{b}$ [29], so that for $\alpha\in[0,1]$ and beliefs $\tilde{b}_1,\tilde{b}_2\in[0,1]$, we have,

$$\begin{aligned}
&\int_{\mathbb{R}_+}\tilde{\psi}(\tilde{e}_+,a\tilde{e})\Big[\alpha\tilde{V}^{(\beta)}_n\Big(\tilde{e}_+,\mathcal{T}(\tilde{b}_1)\Big)\\
&+(1-\alpha)\tilde{V}^{(\beta)}_n\Big(\tilde{e}_+,\mathcal{T}(\tilde{b}_2)\Big)\Big]\,d\tilde{e}_+\\
&\geq\int_{\mathbb{R}_+}\tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n\Big(\tilde{e}_+,\mathcal{T}\big(\alpha\tilde{b}_1+(1-\alpha)\tilde{b}_2\big)\Big)\,d\tilde{e}_+\\
&=\int_{\mathbb{R}_+}\tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n\Big(\tilde{e}_+,\alpha\mathcal{T}(\tilde{b}_1)+(1-\alpha)\mathcal{T}(\tilde{b}_2)\Big)\,d\tilde{e}_+,
\end{aligned}$$

where the last equality follows from simple algebraic manipulations. Concavity of $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\cdot;0)$ then follows from (22).

Since $\lambda\geq 0$, from (22) and (23) we have that $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},0;1)\geq\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},0;0)$. Now, consider the following two possible cases depending on the relationship between $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;1)$ and $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;0)$:

Case i) $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;1)<\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;0)$: Then by the concavity of $\tilde{Q}^{(\beta)}_{n}(\tilde{e},\tilde{b};0)$ and linearity of $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};1)$ in $\tilde{b}$, there exists a unique point where the curves of $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;1)$ and $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;0)$ intersect. This intersection point corresponds to the threshold, $\tilde{b}^{\star}(\tilde{e})$, i.e. during the $n+1$-th step of the iteration, it is optimal to transmit for belief values greater than this value.

Case ii) $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;1)\geq\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;0)$: We will show that for error value equal to $\tilde{e}$, it is optimal to not transmit for any value of $\tilde{b}$. Specifically, we will prove that the curve of $\tilde{Q}^{(\beta)}_{n}(\tilde{e},\cdot;1)$ always lies above the curve of $\tilde{Q}^{(\beta)}_{n}(\tilde{e},\cdot;0)$, i.e. $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};1)\geq\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};0)$ for all $\tilde{b}\in[0,1]$. Upon substituting (22), (23) into $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;1)\geq\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},1;0)$, we obtain,

$$\begin{aligned}
&\lambda+\beta\int_{\mathbb{R}_+}2\psi(\tilde{e}_+)\tilde{V}^{(\beta)}_n(\tilde{e}_+,p_{11})\,d\tilde{e}_+\\
&\geq(a^2\tilde{e}^2+1)+\beta\int_{\mathbb{R}_+}\tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n(\tilde{e}_+,p_{11})\,d\tilde{e}_+. \quad (28)
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\tilde{Q}^{(\beta)}(\tilde{e},\tilde{b};1)-\tilde{Q}^{(\beta)}(\tilde{e},\tilde{b};0)\\
&=\lambda+\beta\tilde{b}\int_{\mathbb{R}_+}2\psi(\tilde{e}_+)\tilde{V}^{(\beta)}_n(\tilde{e}_+,p_{11})\,d\tilde{e}_+\\
&+\beta(1-\tilde{b})\int_{\mathbb{R}_+}\tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n(\tilde{e}_+,p_{01})\,d\tilde{e}_+-\tilde{b}(a^2\tilde{e}^2+1)\\
&-\beta\int_{\mathbb{R}_+}\tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n\Big(\tilde{e}_+,\mathcal{T}(\tilde{b})\Big)\,d\tilde{e}_+\\
&\geq 0,
\end{aligned}$$

where the first equality follows from the definition of $\tilde{Q}^{(\beta)}_n$ (22) and (23), while the last inequality follows from (28) and the induction hypothesis regarding property (C).

Step II: (A) holds for step $n+1$: Consider errors $\tilde{e},\tilde{e}'\in\mathbb{R}_+$ satisfying $\tilde{e}'>\tilde{e}$. We will show that $\tilde{V}^{(\beta)}_{n+1}(\tilde{e}',\tilde{b})\geq\tilde{V}^{(\beta)}_{n+1}(\tilde{e},\tilde{b})$. From (20), it suffices to show that for each value of control $\tilde{u}\in\{0,1\}$ chosen for the state $\tilde{e}'$, there exists a control $\tilde{u}'\in\{0,1\}$ under which the following holds, $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e}',\tilde{b};\tilde{u})\geq\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};\tilde{u}')$. We consider these two cases below separately.

Case i) $\tilde{u}=0$: We have,

$$\begin{aligned}
&\tilde{Q}^{(\beta)}_{n+1}(\tilde{e}',\tilde{b};0)\\
&=(a^2\tilde{e}'^2+1)+\beta\int_{\mathbb{R}_+}\psi(\tilde{e}_+,a\tilde{e}')\tilde{V}^{(\beta)}_n\Big(\tilde{e}_+,\mathcal{T}(\tilde{b})\Big)\,d\tilde{e}_+\\
&\geq(a^2\tilde{e}^2+1)+\beta\int_{\mathbb{R}_+}\psi(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n\Big(\tilde{e}_+,\mathcal{T}(\tilde{b})\Big)\,d\tilde{e}_+\\
&=\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};0),
\end{aligned}$$

where the first equality follows from the definition of $\tilde{Q}^{(\beta)}_{n+1}$ in (22), while the first inequality follows from [27].

Case ii) $\tilde{u}=1$: From (23) we have,

$$\begin{aligned}
\tilde{Q}^{(\beta)}_{n+1}(\tilde{e}',\tilde{b};1)&\geq(1-\tilde{b})(a^2\tilde{e}^2+1)+\lambda\\
&+\beta\tilde{b}\int_{\mathbb{R}_+}2\psi(\tilde{e}_+)\tilde{V}^{(\beta)}_n(\tilde{e}_+,p_{11})\,d\tilde{e}_+\\
&+\beta(1-\tilde{b})\int_{\mathbb{R}_+}\psi(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n(\tilde{e}_+,p_{01})\,d\tilde{e}_+\\
&=\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};1),
\end{aligned}$$

where the inequality follows from [27].

Step III: (B) holds for step $n+1$: Consider belief values $\tilde{b},\tilde{b}'\in[0,1]$ satisfying $\tilde{b}'\leq\tilde{b}$. We will show that $\tilde{V}^{(\beta)}_{n+1}(\tilde{e},\tilde{b}')\geq\tilde{V}^{(\beta)}_{n+1}(\tilde{e},\tilde{b})$. To prove this, we will prove that for each value of control $\tilde{u}$, we have $\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b}';\tilde{u})\geq\tilde{Q}^{(\beta)}(\tilde{e},\tilde{b};\tilde{u})$. Since $p_{11}\geq p_{01}$, we have $\mathcal{T}(\tilde{b}')\leq\mathcal{T}(\tilde{b})$. Consider the following two cases.

Case i) $\tilde{u}=0$: We have,

$$\begin{aligned}
&\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b}';0)\\
&=(a^2\tilde{e}^2+1)+\beta\int_{\mathbb{R}_+}\tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n\Big(\tilde{e}_+,\mathcal{T}(\tilde{b}')\Big)\,d\tilde{e}_+\\
&\geq(a^2\tilde{e}^2+1)+\beta\int_{\mathbb{R}_+}\tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n\Big(\tilde{e}_+,\mathcal{T}(\tilde{b})\Big)\,d\tilde{e}_+\\
&=\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b};0),
\end{aligned}$$

where the first equality follows from (22), while the inequality follows since (B) holds for n by induction hypothesis.

Case ii) $\tilde{u}=1$: We have,

$$\begin{aligned}
\tilde{Q}^{(\beta)}_{n+1}(\tilde{e},\tilde{b}';1)&=(1-\tilde{b}')(a^2\tilde{e}^2+1)+\lambda\\
&+\beta\int_{\mathbb{R}_+}\tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n(\tilde{e}_+,p_{01})\,d\tilde{e}_+\\
&+\beta\tilde{b}'\int_{\mathbb{R}_+}\Big(2\psi(\tilde{e}_+)\tilde{V}^{(\beta)}_n(\tilde{e}_+,p_{11})\\
&\qquad-\tilde{\psi}(\tilde{e}_+,a\tilde{e})\tilde{V}^{(\beta)}_n(\tilde{e}_+,p_{01})\Big)\,d\tilde{e}_+\\
&\geq(1-\tilde{b})(a^2\tilde{e}^2+1)+\lambda+
\end{aligned}$$

$$\beta \int_{\mathbb{R}_+} \tilde{\psi}(\tilde{e}_+, a\tilde{e}) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{01})\, d\tilde{e}_+$$
$$+\beta\tilde{b} \int_{\mathbb{R}_+} \Big(2\psi(\tilde{e}_+) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{11})$$
$$- \tilde{\psi}(\tilde{e}_+, a\tilde{e}) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{01}) \Big)\, d\tilde{e}_+$$
$$= Q_{n+1}^{(\beta)}(\tilde{e}, \tilde{b}; 1),$$

where the first equality follows from the definition of $\tilde{Q}_{n+1}^{(\beta)}$ by (23). From [27] we have, $\int_{\mathbb{R}_+} \Big(2\psi(\tilde{e}_+) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{11}) - \tilde{\psi}(\tilde{e}_+, a\tilde{e}) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{01}) \Big)\, d\tilde{e}_+ \leq 0$. Then, the last inequality follows since $\tilde{b}' \leq \tilde{b}$.

Step IV: (C) holds for $n+1$: Now, since $x \geq y$, it follows from the threshold structure of policy which is optimal at stage $n+1$, proved in Step I, that if the optimal action for state $(\tilde{e}, x)$ is to transmit, then the optimal action for state $(\tilde{e}, y)$ is also to transmit. Consider the following three possibilities:

Case i) No transmission for both $(\tilde{e}, x)$ and $(\tilde{e}, y)$: From (22) we have,

$$(1-\tilde{b})\lambda + \tilde{b}\tilde{Q}_{n+1}^{(\beta)}(\tilde{e}, x; 0) + (1-\tilde{b})\tilde{Q}_{n+1}^{(\beta)}(\tilde{e}, y; 0)$$
$$\geq (a^2\tilde{e}^2 + 1) + \beta \int_{\mathbb{R}_+} \tilde{\psi}(\tilde{e}_+, a\tilde{e}) \tilde{V}_n^{(\beta)}(\tilde{e}_+, \mathcal{T}(z))\, d\tilde{e}_+$$
$$= \tilde{Q}_{n+1}^{(\beta)}(\tilde{e}, z; 0)$$
$$\geq \tilde{V}_{n+1}^{(\beta)}(\tilde{e}, z),$$

where the first inequality follows from the induction hypothesis on property (C), while the last inequality follows from (20).

Case ii) Transmission for both the states $(\tilde{e}, x)$ and $(\tilde{e}, y)$: From (23) we have,

$$(1-\tilde{b})\lambda + \tilde{b}\tilde{Q}_{n+1}^{(\beta)}(\tilde{e}, x; 1) + (1-\tilde{b})Q_{n+1}^{(\beta)}(\tilde{e}, y; 1)$$
$$= (1-\tilde{b})\lambda + \tilde{Q}_{n+1}^{(\beta)}(\tilde{e}, z; 1)$$
$$\geq \tilde{Q}_{n+1}^{(\beta)}(\tilde{e}, z; 1)$$
$$\geq \tilde{V}_{n+1}^{(\beta)}(\tilde{e}, z),$$

where the equality follows from some simple algebraic manipulations and by the definition of $\tilde{Q}_{n+1}^{(\beta)}$ with $z = \tilde{b}x + (1-\tilde{b})y$, and the last inequality follows from (20).

Case iii) Transmission for state $(\tilde{e}, x)$, and no transmission for state $(\tilde{e}, y)$: From (22) and (23) we have,

$$(1-\tilde{b})\lambda + \tilde{b}\tilde{Q}_{n+1}^{(\beta)}(\tilde{e}, x; 1) + (1-\tilde{b})\tilde{Q}_{n+1}^{(\beta)}(\tilde{e}, y; 0)$$
$$\geq \lambda + (\tilde{b}(1-x) + (1-\tilde{b}))(a^2\tilde{e}^2 + 1)$$
$$+\tilde{b}\beta \Big[x \int_{\mathbb{R}_+} 2\psi(\tilde{e}_+) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{11})\, d\tilde{e}_+$$
$$+ (1-x) \int_{\mathbb{R}_+} \tilde{\psi}(\tilde{e}_+, a\tilde{e}) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{01})\, d\tilde{e}_+ \Big]$$
$$+(1-\tilde{b})\beta \Big[y \int_{\mathbb{R}_+} \tilde{\psi}(\tilde{e}_+, a\tilde{e}) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{11})\, d\tilde{e}_+$$

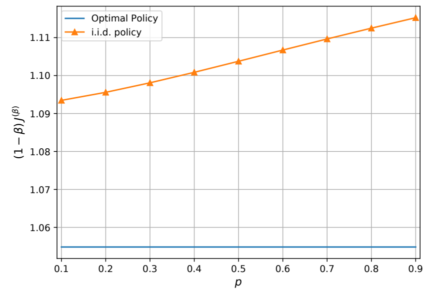

Fig. 1. Performance comparison as the transmission probability p of i.i.d. policy is varied.

$$+ (1-y) \int_{\mathbb{R}_+} \tilde{\psi}(\tilde{e}_+, a\tilde{e}) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{01})\, d\tilde{e}_+ \Big]$$
$$\geq \lambda + (\tilde{b}(1-x) + (1-\tilde{b})(1-y))(a^2\tilde{e}^2 + 1)$$
$$+\beta \Big[z \int_{\mathbb{R}_+} 2\psi(\tilde{e}_+) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{11})\, d\tilde{e}_+$$
$$+ (1-z) \int_{\mathbb{R}_+} \tilde{\psi}(\tilde{e}_+, a\tilde{e}) \tilde{V}_n^{(\beta)}(\tilde{e}_+, p_{01}) \Big]\, d\tilde{e}_+$$
$$= \tilde{Q}_{n+1}^{(\beta)}(\tilde{e}, z; 1)$$
$$\geq \tilde{V}_{n+1}^{(\beta)}(\tilde{e}, z),$$

where the first inequality holds since it was shown in Step I that $\tilde{V}_n^{(\beta)}$ is concave in b, the second inequality is shown to be true in [27], while the last inequality follows from (20). ■

As is shown below, the original POMDP (9) also admits an optimal policy that has threshold structure. The following result is an immediate consequence of Proposition 3.3 and Theorem 4.1.

Corollary 1.1: The original POMDP (9) admits an optimal policy that has a threshold structure.

V. NUMERICAL RESULTS

Throughout, we use $\beta = .99$, transmission price $\lambda = 0.65$ units, truncate the space in which $e(t)$ resides to $[-L, L]$ and discretize it with quantization width of .1, and similarly $b(t)$ is taken to be in $\{\mathcal{T}^k(p_{01})\}_{k=1}^K \cup \{\mathcal{T}^k(p_{11})\}_{k=1}^K$.

We consider an AR process with $a = 0.7$, i.i.d Gaussian noise, $w \sim \mathcal{N}(0, 1)$, and channel parameters are set to $p_{01} = 0.4, p_{11} = 0.7$. Value iteration (15)-(16) is used to get performance of optimal policy. We take $L = 0.5$ and $K = 3$. We compare the performance of our optimal policy with a policy that chooses $\{u(t)\}$ by transmitting at each time t with a probability p, independent of other times. Fig. 1 compares the performance as the transmission probability p is varied. Next, we plot their performance by varying a, p_{01}, p_{11}. Note that the average transmission energy used by an optimal policy depends upon the underlying system parameters, in order to make a fair comparison we set the transmission probability for the i.i.d. policy equal to the average energy consumption of optimal policy. Fig. 2 compares the policies as a, p_{01}, p_{11} are varied. Optimal policy is seen to outperform the naive policy.

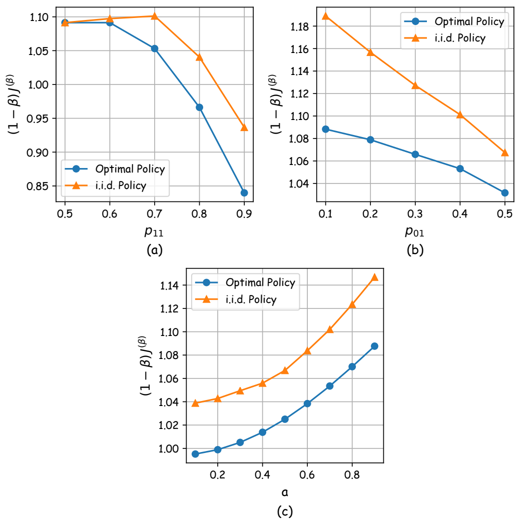

Fig. 2. Performance comparison as system parameters are varied: (a) $p_{01} = 0.4$, $a = 0.7$ while p_{11} is varied; (b) p_{01} is varied while $p_{11} = 0.7$ and $a = 0.7$; (c) $p_{01} = 0.4$ and $p_{11} = 0.7$ are fixed while a is varied.

VI. Conclusion

We consider a remote estimation problem in which the sensor observes an AR Markov process, and has to dynamically decide when to transmit updates to the estimator over a Gilbert-Elliott channel, so as to minimize a cumulative expected discounted cost that consists of estimation error and transmission power consumed. The sensor does not completely observe the channel, i.e. it obtains a delayed knowledge of the channel state only upon a transmission attempt. Even though this problem can be posed as a POMDP, its analysis is hard. We construct a simpler folded POMDP that is equivalent to the original one and derive structural results, namely that there is an optimal policy that transmits only when the belief state is greater than a certain (error-dependent) threshold. This work can be extended in multiple directions. Firstly, a simple linear estimator is used, we would like to design an estimator and scheduler that are *jointly* optimal. Secondly, since the state space is infinite, we would like to obtain an efficient algorithm to yield a good approximation to the optimal policy.

References

[1] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell system technical journal*, vol. 39, no. 5, pp. 1253–1265, 1960.

[2] J. Chakravorty and A. Mahajan, "Structure of optimal strategies for remote estimation over Gilbert-Elliott channel with feedback," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 1272–1276.

[3] G. Yao, A. Bedewy, and N. B. Shroff, "Age-optimal low-power status update over time-correlated fading channel," *IEEE Transactions on Mobile Computing*, 2022.

[4] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138–162, 2007.

[5] P. Antsaklis and J. Baillieul, "Special issue on technology of networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 5–8, 2007.

[6] O. C. Imer and T. Basar, "Optimal estimation with limited measurements," *International Journal of Systems, Control and Communications*, vol. 2, no. 1-3, pp. 5–29, 2010.

[7] Y. Xu and J. P. Hespanha, "Optimal communication logics in networked control systems," in *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, vol. 4. IEEE, 2004, pp. 3527–3532.

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

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

[10] J. Chakravorty and A. Mahajan, "Average cost optimal threshold strategies for remote estimation with communication cost," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2014, pp. 335–342.

[11] M. Rabi, G. V. Moustakides, and J. S. Baras, "Adaptive sampling for linear state estimation," *SIAM journal on control and optimization*, vol. 50, no. 2, pp. 672–702, 2012.

[12] A. Molin and S. Hirche, "An iterative algorithm for optimal event-triggered estimation," *IFAC Proceedings Volumes*, vol. 45, no. 9, pp. 64–69, 2012.

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

[14] D. Han, Y. Mo, J. Wu, S. Weerakkody, B. Sinopoli, and L. Shi, "Stochastic event-triggered sensor schedule for remote state estimation," *IEEE Transactions on Automatic Control*, vol. 60, no. 10, pp. 2661–2675, 2015.

[15] J. Chakravorty and A. Mahajan, "Remote-state estimation with packet drop," *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 7–12, 2016.

[16] G. M. Lipsa and N. C. Martins, "Optimal state estimation in the presence of communication costs and packet drops," in *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2009, pp. 160–169.

[17] J. Chakravorty, J. Subramanian, and A. Mahajan, "Stochastic approximation based methods for computing the optimal thresholds in remote-state estimation with packet drops," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 462–467.

[18] Q. Zhang and S. A. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Transactions on communications*, vol. 47, no. 11, pp. 1688–1692, 1999.

[19] 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 Transactions on Automatic Control*, vol. 63, no. 1, pp. 85–100, 2017.

[20] J. Chakravorty and A. Mahajan, "Remote estimation over a packet-drop channel with Markovian state," *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 2016–2031, 2019.

[21] P. R. Kumar and P. Varaiya, *Stochastic systems: Estimation, identification, and adaptive control*. SIAM, 2015.

[22] V. Krishnamurthy, *Partially observed Markov decision processes*. Cambridge university press, 2016.

[23] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[24] O. Hernández-Lerma and J. B. Lasserre, *Discrete-time Markov control processes: basic optimality criteria*. Springer Science & Business Media, 2012, vol. 30.

[25] J. Chakravorty and A. Mahajan, "Sufficient conditions for the value function and optimal strategy to be even and quasi-convex," *IEEE Transactions on Automatic Control*, vol. 63, no. 11, pp. 3858–3864, 2018.

[26] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of Markov decision processes," *Mathematics of operations research*, vol. 12, no. 3, pp. 441–450, 1987.

[27] M. Dutta and R. Singh, "Optimal scheduling policies for remote estimation of autoregressive Markov processes over time-correlated fading channel," 2023. [Online]. Available: https://arxiv.org/abs/2303.16285

[28] M. S. H. Abad, O. Ercetin, and D. Gündüz, "Channel sensing and communication over a time-correlated channel with an energy harvesting transmitter," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 1, pp. 114–126, 2017.

[29] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable Markov processes over a finite horizon," *Operations research*, vol. 21, no. 5, pp. 1071–1088, 1973.