Remote Estimation over Packet-Dropping Wireless Channels with Partial State Information

Ioannis Tzortzis, Evagoras Makridis, Charalambos D. Charalambous, and Themistoklis Charalambous

***Abstract*— In this paper, we address the problem of designing an optimal transmission policy for remote state estimation over packet-dropping wireless channels with imperfect channel state information. We consider a setup where a smart sensor performs state estimation of a linear time-invariant (LTI) dynamical system using a Kalman filter. The resulting state estimate obtained by the smart sensor at each discrete-time step, is transmitted over the wireless channel to a remote estimator. To balance the trade-off between information freshness and reliability, we employ a Hybrid Automatic Repeat reQuest (HARQ) protocol at the smart sensor which has imperfect channel state information in the form of acknowledgment feedback signal received by the remote estimator after its attempt to decode the information packets. We formulate this problem as a finite horizon Partially Observable Markov Decision Process (POMDP) with an augmented state-space that incorporates both the Age of Information (AoI) and the unknown channel state. By defining an information state that serves as a sufficient statistic for optimal decision-making by the smart sensor, we derive the dynamic programming equations for evaluating the optimal policy. This policy is computed numerically using the point-based value iteration algorithm.**

***Index Terms*— Remote state estimation, HARQ, POMDP, partial channel state information**

I. Introduction

The penetration of smart sensor technology into mission-critical applications such as industrial automation, smart grid monitoring and control, remote healthcare, and others, enabled the transition from on-device monitoring to remote estimation and monitoring. Unlike traditional sensors that merely collect raw data, smart sensors combine sensing with integrated computational and communication capabilities. This allows smart sensors to additionally process the data locally and transmit only useful information to a remote monitor or decision-maker [1]–[3]. In mission-critical applications, smart sensors can effectively reduce the amount of data that need to be sent over the network towards the destined receiver, with the aim of reliably communicating pre-processed information in a timely manner.

In networked control systems, smart sensors often employ data fusion and state estimation algorithms like Kalman filters to pre-estimate the states of a dynamical system [4]–[7], and communicate them to the remote estimator over the network. This approach offers superior state estimation accuracy compared to traditional sensors that merely transmit raw measurement data, especially under unreliable and constrained communication channels [7]. However, communicating critical information over unreliable wireless channels can lead to severe communication impediments such as delays or even loss of information, which subsequently affect the performance of the remote estimator. These impediments arise due to channel fading, noise, interference and path loss, affecting the quality of the communication channel.

Transmitting fresh information may be more beneficial when the probability of success is equal or even higher than retransmitting previously failed information. This is a well known trade-off between information freshness and reliability [8]–[10]. However, transmitters (smart sensors) have no full information regarding the quality of the channel that could allow them to decide whether to transmit new or retransmit an old information. Instead, they may only receive an acknowledgment feedback signal via *(Hybrid) Automatic Repeat reQuest* protocols, to determine whether the transmitted information has been successfully received without errors. Using this protocol, the transmitter combines error correction techniques and retransmissions to ensure reliable delivery of data over noisy channels. Upon the reception of acknowledgment feedback signals (`ACK`/`NACK`), it updates a retransmission counter, and decides whether to discard a previously failed packet and transmit a fresh packet, or retransmit a previously failed one. However, relying only on the acknowledgement feedback to infer the quality of the channel, corresponds to an inexact and possibly delayed indicator of the *channel state information* (CSI).

In [11] the authors studied the stability condition for remote estimation over Markovian packet-dropping channels as introduced by Gilbert [12] and Elliott [13], when transmitting the raw measurements instead of the pre-processed state estimate. Unlike [11], the results in [4], [5], [14], [15], considered the remote estimation problem over packet-dropping communication channels with Markov states, when transmitting the state estimate. In particular, these works proposed remote state estimation methods where the smart sensor can decide whether to retransmit a previously failed state estimate of higher *age of information*[1] (AoI) or to send a fresh estimate with lower probability of successful reception.

The aforementioned works are developed under the assumption that both the smart sensor and the remote estimator have perfect knowledge of the channel state. However,

The authors are with the Department of Electrical and Computer Engineering, School of Engineering, University of Cyprus, 1678 Nicosia, Cyprus. E-mails: {`tzortzis.ioannis`, `makridis.evagoras`, `chadcha`, `charalambous.themistoklis`}`@ucy.ac.cy`.
T. Charalambous is also a Visiting Professor at the Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University, 02150 Espoo, Finland.

[1] Age of information is a measure of the freshness of information in a communication system that quantifies the time elapsed since the most recently received piece of information at the receiver.

in many cases, such an assumption might not hold due to rapidly changing communication channels, interference and noise, uncertainty, limited sensing range, etc. In this work, we focus on designing optimal transmission policies for remote state estimation over packet-dropping wireless channels with imperfect channel state information. The main contributions of this work are the following:

1) We formulate the problem of designing an optimal transmission policy over a wireless channel with partial CSI as a finite horizon POMDP. Our approach, uses a combined state-space that incorporates both the Age of Information (AoI) and the unknown channel state.
2) We identify an information state which satisfies the Markov property and serves as a sufficient statistic for the smart sensor's optimal decision making policy. Additionally, we derive the recursive relation for updating this information state.
3) We solve the optimal decision problem via dynamic programming, and numerically approximate the optimal transmission policy via the point-based value iteration algorithm (PBVI).

The remainder of this paper is organized as follows. In Section II, the system model is introduced and the problem of interest is discussed. In Section III, the problem is formulated as a POMDP, and the recursive relation of the information state is derived. In Section IV, the solution of the optimal decision problem is provided via a dynamic programming approach, and the PBVI algorithm is employed as a numerical approximation to the optimal decision policy. Finally, in Section V numerical experiments are given to illustrate the behavior of the proposed solution.

II. System Model

A. Dynamical System

Consider a discrete-time LTI dynamical system

$$x_k = Ax_{k-1} + w_{k-1}, \tag{1a}$$
$$y_k = Cx_k + v_k, \tag{1b}$$

where k is the discrete time index, $x_k \in \mathbb{R}^n$ is the (unknown) state vector process, $y_k \in \mathbb{R}^m$ is the measurement vector process, $w_k \in \mathbb{R}^n$ is the process noise, $v_k \in \mathbb{R}^m$ is the measurement noise, and $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times n}$ are known matrices. It is assumed that $x_0, w_0, w_1, \ldots, v_0, v_1, \ldots$ are mutually independent and Gaussian random variables

$$x_0 \sim N(0, \Sigma_0), \quad w_k \sim N(0, R_w), \quad v_k \sim N(0, R_v)$$

where $\Sigma_0 \succcurlyeq 0$, $R_w \succcurlyeq 0$, and $R_v \succ 0$.

B. Smart Sensor

Define the *a priori* and *a posterior* state estimates

$$\hat{x}^s_{k|k-1} := \mathbb{E}[x_k | y_{1:k-1}],$$
$$\hat{x}^s_{k|k} := \mathbb{E}[x_k | y_{1:k}],$$

and their corresponding error covariances

$$P^s_{k|k-1} := \mathbb{E}[(x_k - \hat{x}^s_{k|k-1})(x_k - \hat{x}^s_{k|k-1})^T | y_{1:k-1}],$$
$$P^s_{k|k} := \mathbb{E}[(x_k - \hat{x}^s_{k|k})(x_k - \hat{x}^s_{k|k})^T | y_{1:k}].$$

where $y_{1:t} := \{y_1, \ldots, y_t\}$, $t \in \mathbb{N}$, denotes the sequence of noisy observations up to time t. The Kalman filter is given by the following prediction-correction equations

$$\hat{x}^s_{k|k-1} = A\hat{x}^s_{k-1|k-1}, \tag{2a}$$
$$P^s_{k|k-1} = AP^s_{k-1|k-1}A^T + R_w, \tag{2b}$$
$$K_k = P^s_{k|k-1}C^T(CP^s_{k|k-1}C^T + R_v)^{-1}, \tag{2c}$$
$$\hat{x}^s_{k|k} = \hat{x}^s_{k|k-1} + K_k(y_k - C\hat{x}^s_{k|k-1}), \tag{2d}$$
$$P^s_{k|k} = P^s_{k|k-1} - K_kCP^s_{k|k-1}, \tag{2e}$$

which are computed for each time step $k = 1, 2, \ldots$. Next, we make the following assumption.

Assumption 1: Stabilizability and detectability conditions hold, so that, the error covariance matrix of the smart sensor estimate $\hat{x}^s_{k|k}$ has reached its steady-state, that is, $\lim_{k\to\infty} P^s_{k|k} = P^s$, where P^s is the unique stabilizing solution of the algebraic Riccati equation.

The above assumption is justified by the fact that for stable systems with bounded process and measurement noise the Kalman filter's error covariance converges to its steady-state value exponentially fast. With the assumption that the error covariance matrix has reached its steady-state, the smart sensor at each time k must decide whether to transmit a new piece of information, *i.e.,* its current local state estimate $\hat{x}^s_{k|k}$, or retransmit old information, *i.e.,* an old local state estimate $\hat{x}^s_{k-\tau_k|k-\tau_k}$, as follows

$$\hat{x}^s_k = \begin{cases} \hat{x}^s_{k|k}, & \text{if } \alpha_k = 1, \\ \hat{x}^s_{k-1}, & \text{if } \alpha_k = 0, \end{cases} \tag{3}$$

where $\alpha_k \in \{0,1\}$ denotes the transmission action of the smart sensor at time k, and represents either the transmission of a new packet, *i.e.,* $\alpha_k = 1$, or the retransmission of a previously failed packet, *i.e.,* $\alpha_k = 0$.

The transmission action of the smart sensor is the design parameter we aim to design in this work.

C. Wireless Channel

The wireless channel is considered as lossy and is modeled as a packet-erasure channel with ideal acknowledgment feedback. The quality of the channel can be modeled as a random process that varies with time in a correlated manner, although it is assumed constant within each time slot.

The state of the channel at time slot k could be in one of the n_c states that describe its quality, *i.e.,* $s^c_k \in \mathcal{S}^c \triangleq \{1, 2, \ldots, n_c\}$. Thus, depending on the quality of the channel, a packet that is transmitted by the smart sensor to the remote estimator might arrive in error. In such a case, and with the remote estimator being incapable of correcting the error with the aid of the HARQ protocol, we say that a packet error occurred.

D. Remote Estimator

We assume that the remote estimator is equipped with a HARQ protocol aiming at correcting possible errors that occurred due to channel fading. To indicate whether the packet sent by the smart sensor has been received by the remote estimator with or without errors, we introduce the random variable $z_k \in \mathcal{Z} \triangleq \{0,1\}$. That is, when a packet loss occurs at time k, the remote estimator sends a negative acknowledgment (NACK) back to the smart sensor, *i.e.,* $z_k = 0$, while if the packet has been received without errors at time k, the remote estimator sends an acknowledgment (ACK), $z_k = 1$. Then, according to z_k, the remote state estimate $\hat{x}_k$ and its corresponding error covariance matrix P_k are given by

$$\hat{x}_k = \begin{cases} \hat{x}_k^s, & \text{if } z_k = 1, \\ A\hat{x}_{k-1}, & \text{otherwise,} \end{cases} \tag{4}$$

and

$$P_k = \begin{cases} P_{k|k}^s, & \text{if } z_k = 1, \\ AP_{k-1}A^T + R_w, & \text{otherwise.} \end{cases} \tag{5}$$

E. Problem Setup

The ultimate goal of the smart sensor is to minimize the mean squared error and energy consumption by optimally deciding between transmitting new information or re-transmitting previously failed information. Given that the exact channel state is unknown, the smart sensor relies its decision based on its belief about the channel state. This belief is updated over time based on partial information inferred from an acknowledgment feedback signal received from the remote estimator, see Fig.1.

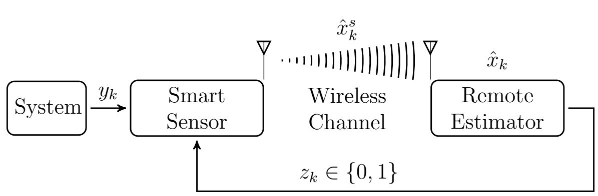

Fig. 1. Remote state estimation scheme.

Upon the reception of the feedback signal, the smart sensor is able to track the time elapsed since the last successful reception of information by the remote estimator. We denote this variable at time k with $s_k^r \in \mathcal{S}^r \triangleq \{0,1,\ldots,n_r\}$, $n_r < \infty$, which is given by

$$s_k^r = \begin{cases} 1, & \text{if } \alpha_{k-1} = 1 \wedge z_{k-1} = 0, \\ s_{k-1}^r + 1, & \text{if } \alpha_{k-1} = 0 \wedge z_{k-1} = 0, \\ 0, & \text{if } z_{k-1} = 1, \end{cases} \tag{6}$$

for $k \geq 0$ and with $s_0^r = 0$. Clearly, when $s_k^r = 0$ then only $\alpha_k = 1$ can be taken. This is natural since $s_k^r = 0$ represents the state where a fresh packet is to be transmitted. Moreover, when $s_{k-1}^r = n_r$, then we require $s_k^r = 0$ and $\alpha_k = 1$, since the number of retransmission attempts is reached. To indicate that the AoI s_k^r is a function of the previous state s_{k-1}^r, the action α_{k-1} and the observation z_{k-1}, we write

$$s_k^r = L_{k-1}^r(s_{k-1}^r, z_{k-1}, \alpha_{k-1}). \tag{7}$$

We assume that the packet error probability reduces exponentially with the number of retransmissions $r \in \mathcal{S}^r$, based on the model in [16], *i.e.,*

$$g(r,j) = q_j \lambda^r \tag{8}$$

for $\lambda \in (0,1)$, and q_j being the packet error probability of a freshly transmitted packet when in channel $j \in \mathcal{S}^c$. With smaller parameter λ, the probability of error decreases due to the ability of HARQ to decode the received packet correctly. With $\lambda = 1$, the HARQ protocol reduces to the classical ARQ protocol without combining, where the probability of packet error is not reduced, but it is instead constant for all subsequent retransmissions.

The AoI variable s_k^r, allows the smart sensor to determine the current state estimate of the remote estimator at time k, where from (4) and (5), we get

$$\hat{x}_k = \begin{cases} \hat{x}_k^s, & \text{if } s_k^r = 0, \\ A^{s_k^r}\hat{x}_{k-s_k^r}, & \text{otherwise,} \end{cases} \tag{9}$$

$$P_k = \begin{cases} P_{k|k}^s = P^s, & \text{if } s_k^r{=}0, \\ A^{s_k^r} P^s (A^{s_k^r})^T + \sum\limits_{j=0}^{s_k^r - 1} A^j R_w (A^j)^T, & \text{otherwise,} \end{cases} \tag{10}$$

III. Problem Formulation

A. Finite horizon POMDP

In this section, the partially observable Markov decision process is used for modeling the smart sensor optimal transmission problem over imperfect channel state information. The precise POMDP formulation is as follows.

A discrete-time POMDP on a finite horizon is a collection

$$(\mathcal{S}, \mathcal{A}, \mathcal{Z}, T(\alpha), O(\alpha), c_k(\alpha), c_N) \tag{11}$$

consisting of the following elements:

a) The state space $\mathcal{S} := \mathcal{S}^c \times \mathcal{S}^r$ where $\mathcal{S}^c := \{1,2,\ldots,n_c\}$ denotes the set of possible channel states, and $\mathcal{S}^r := \{0,1,\ldots,n_r\}$ denotes the set of possible AoI values. Each $s := (s^c, s^r) \in S$ represents the system being in channel state $s^c \in \mathcal{S}^c$ and having an AoI $s^r \in \mathcal{S}^r$.

b) The action space $\mathcal{A} := \{0,1\}$. Each $\alpha \in \mathcal{A}$ represents the action that the smart sensor can take.

c) The observation space $\mathcal{Z} := \{0,1\}$. Each $z \in \mathcal{Z}$ provides information on whether or not the package has been successfully received by the remote estimator.

d) The transition probability matrix $T(\alpha)$ defined by

$$\begin{aligned} T(\alpha) &:= \Pr(s_{k+1}|s_k, \alpha_k) \\ &= \Pr(s_{k+1}^c, s_{k+1}^r | s_k^c, s_k^r, \alpha_k) \\ &= \Pr(s_{k+1}^c | s_k^c, s_{k+1}^r, \alpha_k) \Pr(s_{k+1}^r | s_k^c, s_k^r, \alpha_k) \\ &\stackrel{(a)}{=} \Pr(s_{k+1}^c | s_k^c) \Pr(s_{k+1}^r | s_k^c, s_k^r, \alpha_k) \end{aligned} \tag{12}$$

where (a) stems from the fact that the channel state s_{k+1} depends only on the current state s_k. Let $T^c = [\tau^c_{i,j}]$, $i, j \in \mathcal{S}^c$, be the channel state transition probability matrix with

$$\tau^c_{i,j} = \Pr(s^c_{k+1} = j | s^c_k = i) \tag{13}$$

denoting the (i, j)–th element of T^c. Let $T^r(\alpha) = [\tau^r_{i,\ell,q}(\alpha)]$, $i \in \mathcal{S}^c$, $\ell, q \in \mathcal{S}^r$, $\alpha \in \mathcal{A}$, be the AoI state transition probability matrix with

$$\tau^r_{i,\ell,q}(\alpha) = \Pr(s^r_{k+1} = q | s^c_k = i, s^r_k = \ell, \alpha_k = \alpha) \tag{14}$$

denoting the (i, ℓ, q)–th element of $T^r(\alpha)$.

e) The observation probability matrix $O(\alpha) = [o_{i,j,m}(\alpha)]$, $i \in \mathcal{S}^c$, $j \in \mathcal{S}^r$, $m \in \mathcal{Z}$, $\alpha \in \mathcal{A}$, defined by

$$O(\alpha) := \Pr(z_k | s_k, \alpha_k) = \Pr(z_k | s^c_k, s^r_k, \alpha_k) \tag{15}$$

with

$$o_{i,j,m}(\alpha) = \Pr(z_k = m | s^c_k = i, s^r_k = j, \alpha_k = \alpha) \tag{16}$$

denoting the (i, j, m)–th element of $O(\alpha)$.

f) The cost function $c_k(s, \alpha)$ incurred at time k when the state is $s \in \mathcal{S}$ and action $\alpha \in \mathcal{A}$ is applied.

g) The terminal cost function $c_N(s)$ incurred at time $k = N$ when the state is $s \in \mathcal{S}$.

The POMDP defined by (11) represents a partially observable model in which the augmented system state $s_k \in \mathcal{S}$ consists of both known and unknown components. In particular, the AoI $s^r \in \mathcal{S}^r$ is the known part of the system state while the channel state $s^c \in \mathcal{S}^c$ is the unknown (unobserved) part of the system state. Hence, the smart sensor must infer the unknown part of the system state through some indirect observations at each time k. The observations $z \in \mathcal{Z}$ provide insight into the hidden channel state $s^c \in \mathcal{S}^c$ based on whether the package was successfully received by the remote estimator.

The smart sensor must make optimal decisions on whether to transmit new information or re-transmit old information, based on the available information it has about the current system state. To specify the information available to the smart sensor, let us define the set

$$I_k := \{s^r_0, z_0, \dots, z_{k-1}, \alpha_0, \dots, \alpha_{k-1}\},\ k = 1, 2, \dots, N.$$

Remark 1: Notice from (6) that given the information history I_k, then the AoI state $s^r_k \in \mathcal{S}^r$ can be calculated for all k. It follows that, for each k, I_k conveys the same amount of information as

$$\bar{I}_k := \{z_0, \dots, z_{k-1}, \alpha_0, \dots, \alpha_{k-1}, s^r_0, \dots, s^r_k\}. \tag{17}$$

Consider the set of all policies compatible with I_k. For any $g = \{g_0, g_1, \dots, g_{N-1}\}$, $g_k : I_k \longmapsto \mathcal{A}$ is feasible if

$$\alpha_k = g_k(I_k) \in \mathcal{A}, \quad \forall I_k,\ k = 1, 2, \dots, N-1. \tag{18}$$

Let G denote the set of all feasible policies.

The next step in solving a POMDP is to define an information state, also referred to as the belief state. The information state at time k, denoted as π_k, is the posterior distribution over the unknown component of the system state conditioned on the available information, that is,

$$\pi_k(j|I_k) := \Pr(s^c_k = j | I_k), \quad \forall j \in \mathcal{S}^c. \tag{19}$$

Using Bayes' rule, the recursive relation of the information state from time k to time $k+1$ is given by (see Appendix A)

$$\begin{aligned}\pi_{k+1}(j|I_{k+1}) &= \Pr(s^c_{k+1} = j | I_{k+1}) \\ &= \frac{\sum_{i \in \mathcal{S}^c} \tau^{r,z}_{i,\ell,q,m}(\alpha) \tau^c_{i,j} \pi_k(i|I_k)}{\sum_{i \in \mathcal{S}^c} \tau^{r,z}_{i,\ell,q,m}(\alpha) \pi_k(i|I_k)},\end{aligned} \tag{20}$$

with

$$\begin{aligned}\tau^{r,z}_{i,\ell,q,m}(\alpha) &= \Pr(s^r_{k+1} = q | s^c_k = i, s^r_k = \ell, z_k = m, \alpha_k = \alpha) \\ &\quad \times \Pr(z_k = m | s^c_k = i, s^r_k = \ell, \alpha_k = \alpha).\end{aligned} \tag{21}$$

The first term on the right side is given by (30), while the second term is given by (31). To indicate that the information state is a function of the prior distribution π_k, the action α_k and the observation z_k, we write

$$\pi_{k+1}(I_{k+1}) = L_k(\pi_k(I_k), z_k, \alpha_k) \tag{22}$$

where $\pi_k(I_k) = (\pi_k(1|I_k), \dots, \pi_k(n_c|I_k))$ is an n_c–dimensional row vector.

Remark 2: By utilizing the fact that (a) the information state π_k provides a sufficient statistic for the information available to the smart sensor, that is, $\alpha_k = g_k(I_k) = g_k(\pi_k) \in \mathcal{A}$, and (b) for any fixed sequence of actions $\{\alpha_0, \dots, \alpha_k\}$, the sequence $\{\pi_0, \dots, \pi_k\}$ is a Markov process, that is, $\Pr(\pi_{k+1} | \pi_0, \dots, \pi_k, \alpha_k) = \Pr(\pi_{k+1} | \pi_k, \alpha_k)$, then the POMDP (11) is equivalent to the corresponding fully observed Markov decision process [17], [18].

B. Performance Criterion

Each policy $g \in G$ incurs a series of costs

$$c_k(s_k, \alpha_k) = \mathrm{Tr}(P_k) + \epsilon_j(\alpha), \tag{23}$$

where P_k is specified by (10), and $\epsilon_j(\alpha)$ denotes the energy consumption modeled as a deterministic function of the channel state $s^c_k = j \in S^c$ and action $\alpha_k = \alpha \in \mathcal{A}$.

Remark 3: In HARQ protocols, retransmissions often involve transmitting a packet at a higher power level due to additional encoding to improve the chances of decoding the message at the receiver, especially when previous attempts have failed due to channel conditions. This attempt raises the energy cost ϵ_j for retransmitted packets compared to a simple new packet transmission.

Define the N–stage expected cost by

$$J(g) := \mathbb{E}^g \Big[\sum_{k=0}^{N-1} c_k(s_k, \alpha_k) + c_N(s_N) \Big], \tag{24}$$

where $\mathbb{E}^g[\cdot]$ indicates the dependence of the expectation operation on the policy $g \in G$, and c_N denotes the terminal cost, a given deterministic function $c_N(s_N) := \mathrm{Tr}(P_N)$.

The optimal decision problem is that of selecting the admissible policy $g \in G$ such that the performance criterion (24) is minimized. In particular, the optimal decision problem is to choose $g^* \in G$ such that

$$J(g^*) = \inf_{g \in G} J(g) = J^*. \tag{25}$$

A policy $g^* \in G$ that satisfies (25) is called an optimal policy, and the corresponding $J^*(\cdot)$ is called the minimum cost.

IV. Problem Solution

In this section, we provide via dynamic programming the solution of the optimal decision problem (25).

Let $V_k(\alpha^g_{[k,N-1]}, I_k)$ denote the value function on the time horizon $k, k+1, \ldots, N$ given an optimal policy g_t^*, $t = 1, 2, \ldots, k-1$ defined by

$$V_k(\alpha^g_{[k,N-1]}, I_k) = \mathbb{E}\Big[\sum_{t=k}^{N-1} c_t(s_t^g, \alpha_t^g) + c_N(s_N^g)|I_k\Big] \tag{26}$$

where $\alpha^g_{[k,N-1]}$ denotes the restriction of policies in $[k, N-1]$. Let us also define the probability row vector $\pi = (\pi(1), \ldots, \pi(n_c)) \in \Pi_1(\mathcal{S}^c)$ where

$$\Pi_1(\mathcal{S}^c) := \{\pi \in \mathbb{R}^{n_c} : \sum_{i \in \mathcal{S}^c} \pi(i) = 1, \pi(i) \geq 0, \forall i \in \mathcal{S}^c\}.$$

The main theorem of this paper follows.

Theorem 1: Define recursively the functions $V_k(\pi, s^r)$, $0 \leq k \leq N$, $\pi \in \Pi_1(\mathcal{S}^c)$, $s^r \in \mathcal{S}^r$ by

$$V_N(\pi, s^r) = \mathbb{E}[c_N(s_N^c, s^r)|\pi_N(I_N) = \pi, s_N^r = s^r], \tag{27a}$$

$$V_k(\pi, s^r) = \min_{\alpha_k \in \mathcal{A}} \mathbb{E}[c_k(s_k^c, s^r, \alpha_k) \tag{27b}$$
$$+V_{k+1}(L_k(\pi, z_k, \alpha_k), L_k^r(s^r, z_k, \alpha_k))|\pi_k(I_k){=}\pi, s_k^r{=}s^r].$$

1) Let $g \in G$. Then,

$$V_k(\pi_k(I_k), s_k^r) \leq V_k(\alpha^g_{[k,N-1]}, I_k). \tag{28}$$

2) Let $g \in G$ be such that for all $\pi \in \Pi_1(\mathcal{S}^c)$, $g_k(\pi)$ achieves the infimum in (27b). Then, g is optimal and

$$V_k(\pi_k(I_k), s_k^r) = V_k(\alpha^g_{[k,N-1]}, I_k). \tag{29}$$

The expectation in the right-hand side of (27a) is given by

$$\sum_{j \in \mathcal{S}^c} c_N(j, s^r) \Pr(j|\pi_N(I_N) = \pi, s_N^r = s^r)$$
$$= \sum_{j \in \mathcal{S}^c} c_N(j, s^r) \Pr(j|\pi_N(I_N) = \pi) = \sum_{j \in \mathcal{S}^c} c_N(j, s^r)\pi(j)$$

where the first equality follows since $\pi_N(I_N) = \pi$ summarizes all the information of I_N, including the information of $s_N^r = s^r$ (i.e., see Remark 1). Similarly, the expectation in the right-hand side of (27b) is given by

$$\sum_{j \in \mathcal{S}^c} c_k(j, s^r, \alpha)\pi(j) + \sum_{m \in \mathcal{Z}} V_{k+1}(L_k(\pi, m, \alpha), L_k^r(s^r, m, \alpha))$$
$$\times \Pr(m|\pi_k(I_k) = \pi, s_k^r = s^r, \alpha_k = \alpha)$$

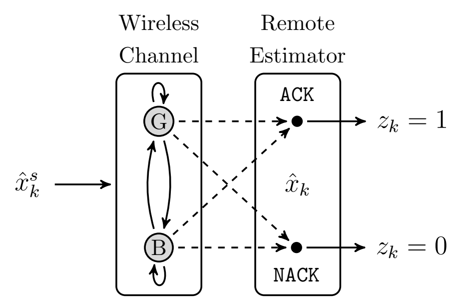

Fig. 2. Gilbert-Elliot channel with a "good" (G) and "bad" (B) state, each one corresponding to a certain probability of packet error.

where the conditional probability of $z_k = m$ given $\pi_k(I_k) = \pi$, $s_k^r = s^r$, and $\alpha_k = \alpha$, is

$$\Pr(z_k = m|\pi_k(I_k) = \pi, s_k^r = s^r, \alpha_k = \alpha)$$
$$= \sum_{i \in \mathcal{S}^c} \Pr(z_k = m|s_k^c = i, s_k^r = s^r, \alpha_k = \alpha)$$
$$\times \Pr(s_k^c = i|\pi_k(I_k) = \pi, s_k^r = s^r, \alpha_k = \alpha)$$
$$= \sum_{i \in \mathcal{S}^c} o_{i,s^r,m}(\alpha)\pi(i).$$

Proof: See Appendix B. ■

It is well-known that solving a POMDP exactly is computationally challenging. For this reason, in this work the point-based value iteration (PBVI) algorithm [19] is employed to numerically approximate the optimal policy. The PBVI algorithm, using a finite set of representative belief points $\mathcal{B}_\pi$, approximates the value function and iteratively solves the dynamic programming equations (27) for each AoI state. At each iteration of the PBVI algorithm, belief states are updated recursively generating child belief states for each observation and action, capturing how the probability of channel states evolves over time based on AoI dynamics and new observations. This recursive process allows the PBVI algorithm to compute an optimal transmission policy while considering both the AoI and channel state uncertainty, resulting in a policy that balances between the freshness of information and the reliability of the channel state estimation.

In the next section, the effectiveness of the proposed solution is illustrated via numerical experiments.

V. Numerical Experiments

Consider the discrete-time LTI dynamical system in (1) with

$$A = \begin{bmatrix} 0.9974 & 0.0539 \\ -0.1078 & 0.1591 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

The smart sensor employs the Kalman filter defined in (2) with

$$R_w = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.25 \end{bmatrix}, \quad R_v = 0.05,$$

and at each time step k it transmits either a fresh packet with its current local state estimate $\hat{x}^s_{k|k}$, or retransmits an old state estimate $\hat{x}^s_{k-\tau_k|k-\tau_k}$, according to the action α_k, as described in (3). In this example we model the channel, over which

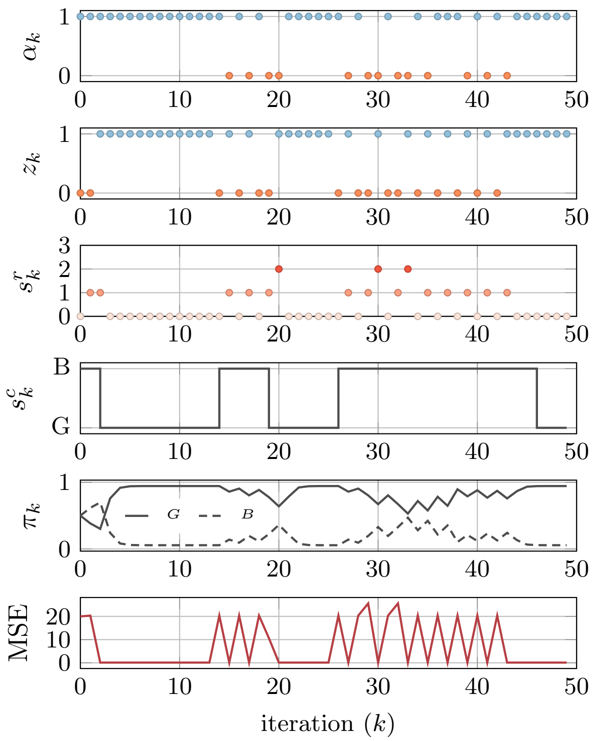

Fig. 3. Example of HARQ under T_1^c with $\lambda = 0.5$ using the PBVI policy.

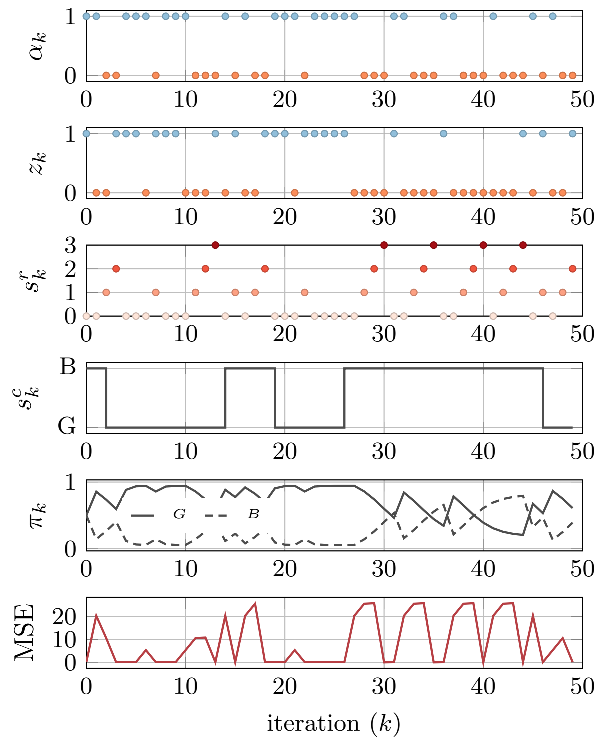

Fig. 4. Example of ARQ under T_1^c with $\lambda = 1$ using the PBVI policy.

the packets are transmitted, as a two-state time homogeneous Markov chain, as introduced by Gilbert [12] and Elliott [13]. The channel can be either in a good (G) or bad (B) state. In the results that follow, we consider two different transition matrices that denote the probabilities between the channel states, *i.e.,*

$$T_1^c = \begin{array}{cc} \scriptstyle G & \scriptstyle B \\ \end{array} \begin{bmatrix} 0.95 & 0.05 \\ 0.1 & 0.9 \end{bmatrix} \begin{array}{c} \scriptstyle G \\ \scriptstyle B \end{array}, \quad T_2^c = \begin{array}{cc} \scriptstyle G & \scriptstyle B \\ \end{array} \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \begin{array}{c} \scriptstyle G \\ \scriptstyle B \end{array}.$$

Here it is important to note that, the smart sensor is aware of the channel state transition matrix, but it does not know in which state is the channel at any time k. The probability of error of a packet transmission depends on the quality (state) of the channel. Hence, we assume that the probability of error of a fresh (new) transmission when in state G is $q_G = 0.2$, while in state B is $q_B = 0.8$. In this example, we allow the smart sensor to possibly retransmit the same packet for $n_r = 3$ times. Hence, the probability of error of a packet (re)-transmitted for $r \in \mathcal{S}^r$ times, is given by $g_j(r) = q_j \lambda^r$, for $j \in \{G, B\}$, and $\lambda \in (0, 1]$ being the parameter of HARQ combining scheme [16].

Upon the transmission of a packet by the smart sensor at time k, the remote estimator attempts to decode the packet. If the packet received at the remote estimator is decoded without errors, it sends an acknowledgment (`ACK`) back to the smart sensor, $z_k = 1$, otherwise it sends a negative acknowledgment (`NACK`), $z_k = 0$. Subsequently, the smart sensor updates s_k^r as in (6), and applies the action α_{k+1} given by the optimal transmission policy of the PBVI algorithm.

Fig.3 and Fig.4 depict the execution of the aforementioned setup using the same realization of the Markov chain transition probability matrix T_1^c, for values of $\lambda = 0.5$ and $\lambda = 1$, respectively. In both figures, we record the variables held by the smart sensor, *i.e.,* the decision a_k, the observation z_k, the AoI state s_k^r, the channel state s_k^c (which is unknown to the smart sensor), and the mean squared error (MSE) at the remote estimator.

In the following, we compare the results obtained under HARQ and ARQ protocols, as shown in Fig.3 and Fig.4.

- Under both protocols, the smart sensor's transmission policy needs to balance the trade-offs between improving the MSE by reducing the AoI and minimizing energy consumption, particularly when retransmissions are required due to the bad channel state $s_k^c = B$.
- Comparing the two protocols, we can easily observe that the main distinction lies in the plot of the AoI s_k^r. In contrast to ARQ protocol, where each packet transmission has the same success probability, HARQ protocol with parameter $\lambda = 0.5$ increases the probability of successful packet transmission with each retransmission. Combined with the smart sensor's actions α_k, HARQ protocol results in a reduced number of transmissions. This is particularly emphasized especially when the channel state is $s_k^c = B$.
- The belief state π_k at each time k is updated based on observations z_k and actions α_k. In particular, as it is illustrated in the plots of π_k, as long as a positive acknowledgment (`ACK`) is received the belief state closely follows the channel state $s_k^c = G$, reflecting high

confidence in the current channel state. However, whenever a negative acknowledgment (NACK) is received, the belief state shifts, indicating increased uncertainty about the channel state. This shift occurs because a NACK suggests a potential channel state, causing the belief state to be updated so that is reflects the probability of the channel being in the state $s_k^c = B$.

Finally, Fig.5 depicts MSE results averaged over 100 realizations of the Markov chain transition probability matrices T_1^c and T_2^c, with each result corresponding to a specific value of λ. The transition probability matrix T_2^c represents the worst-case scenario since it implies complete uncertainty about the channel state transitions. As a result, the MSE in this case is higher than the one corresponding to T_1^c.

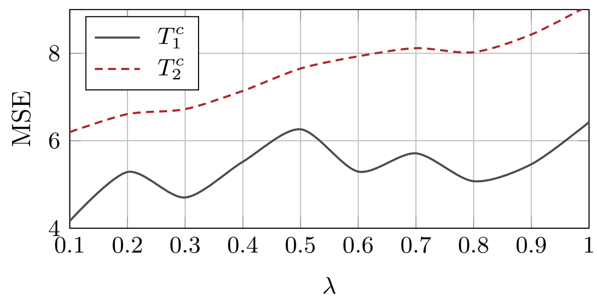

Fig. 5. Averaged MSE for different realizations of the Markov chain transition probability matrices T_1^c and T_2^c.

VI. Conclusions

In this paper, we address the problem of designing an optimal transmission policy for remote state estimation over packet-dropping wireless channels with imperfect channel state information. We formulate this problem as a finite horizon POMDP with an augmented state-space that incorporates both the AoI and the unknown channel state. By defining an information state that serves as a sufficient statistic for optimal decision-making by the smart sensor, we derive the dynamic programming equations for evaluating the optimal policy. This policy is computed numerically using the point-based value iteration algorithm.

Appendices

A. Derivation of Equation (20)

Here, we will show the derivation of the recursive relation of the information state, as given by (20). By the definition of the information state and by Remark 1, we have that

$$\begin{aligned}\pi_{k+1}(j|I_{k+1}) &= \Pr(s_{k+1}^c = j|I_{k+1}) = \Pr(s_{k+1}^c = j|\bar{I}_{k+1}) \\ &= \Pr(s_{k+1}^c = j|\bar{I}_k, z_k = m, \alpha_k = \alpha, s_{k+1}^r = q)\end{aligned}$$

where we have assumed that during action interval k the observation under action $\alpha_k = \alpha$ was $z_k = m$. By (6), both z_k and α_k, together with $s_k^r = \ell \in \bar{I}_k$, give the AoI state $s_{k+1}^r = q$. Applying Bayes' rule, and then expanding both the numerator and denominator over channel states $s_k^c = i \in \mathcal{S}^c$, we obtain

$$\pi_{k+1}(j|I_{k+1}) = \frac{\sum_{i\in\mathcal{S}^c} \Pr(s_{k+1}^c{=}j, s_{k+1}^r = q, s_k^c{=}i, z_k{=}m|\bar{I}_k, \alpha_k{=}\alpha)}{\sum_{i\in\mathcal{S}^c} \Pr(s_{k+1}^r = q, s_k^c = i, z_k = m|\bar{I}_k, \alpha_k = \alpha)}.$$

The numerator is given by

$$\begin{aligned}&\sum_{i\in\mathcal{S}^c} \Pr(s_{k+1}^c = j, s_{k+1}^r = q, s_k^c = i, z_k = m|\bar{I}_k, \alpha_k = \alpha) \\ &= \sum_{i\in\mathcal{S}^c} \Pr(s_{k+1}^r = q, z_k = m|s_k^c = i, s_k^r = \ell, \alpha_k = \alpha) \\ &\quad \times \Pr(s_{k+1}^c = j|s_k^c = i)\Pr(s_k^c = i|I_k)\end{aligned}$$

where the first term on the right side is the joint distribution of $s_{k+1}^r \in \mathcal{S}^r$ and $z_k \in \mathcal{Z}$ given the system state $s_k \in \mathcal{S}$ and action $\alpha_k \in \mathcal{A}$, which is independent of the channel state $s_{k+1}^c \in \mathcal{S}^c$ and the information history $\bar{I}_k$ (hence, I_k as well), as given by (21). The second term on the right side is the transition probability of the channel state as given by (12), and the last term is the information state at time k with $\bar{I}_k$ replaced by I_k (since they convey the same amount of information). Similarly, the denominator is given by

$$\begin{aligned}&\sum_{i\in\mathcal{S}^c} \Pr(s_{k+1}^r = q, s_k^c = i, z_k = m|\bar{I}_k, \alpha_k = \alpha) \\ &= \sum_{i\in\mathcal{S}^c} \Pr(s_{k+1}^r = q, z_k = m|s_k^c = i, s_k^r = \ell, \alpha_k = \alpha) \\ &\quad \times \Pr(s_k^c = i|I_k).\end{aligned}$$

Thus, we have (20).

B. Proof of Theorem 1

The proof of Theorem 1 is a variation of the classical proof [20]. We prove both parts by induction. 1) For $k = N$,

$$\begin{aligned}V_N(\alpha_{[k,N-1]}^g, I_N) &= \mathbb{E}^g[c_N(s_N^c, s_N^r)|I_N] \\ &= \sum_{j\in\mathcal{S}^c} c_N(j, s^r)\Pr(s_N^c = j|I_N, s_N^r = s^r) \\ &= \sum_{j\in\mathcal{S}^c} c_N(j, s^r)\pi_N(j|I_N) \qquad (32)\end{aligned}$$

so (28) holds with equality. Assume that (28) holds for $k+1$, then

$$\begin{aligned}&V_k(\alpha_{[k,N-1]}^g, I_k) = \mathbb{E}^g\Big[c_k(s_k^c, s_k^r, \alpha_k) \\ &+ \mathbb{E}^g\Big[\sum_{t=k+1}^{N-1} c_t(s_t^r, s_t^c, \alpha_t) + c_N(s_N^r, s_N^c)|I_{k+1}\Big]|I_k\Big] \\ &\geq \mathbb{E}^g\Big[c_k(s_k^c, s_k^r, \alpha_k) + V_{k+1}(\pi_{k+1}(I_{k+1}), s_{k+1}^r)|I_k\Big] \\ &= \mathbb{E}^g\Big[\mathbb{E}^g\Big[c_k(s_k^c, s_k^r, \alpha_k) \\ &\quad + V_{k+1}(\pi_{k+1}(I_{k+1}), s_{k+1}^r)|I_k, s_k^r, \alpha_k\Big]|I_k\Big] \\ &= \mathbb{E}^g\Big[\mathbb{E}^g\Big[c_k(s_k^c, s_k^r, \alpha_k) \\ &+ V_{k+1}(L_k(\pi_k(I_k), z_k, \alpha_k), L_k^r(s_k^r, z_k, \alpha_k)) \\ &\quad |\pi_k(I_k), s_k^r, \alpha_k\Big]|I_k\Big] \\ &\geq \mathbb{E}^g\Big[V_k(\pi_k(I_k), s_k^r)|I_k\Big] = V_k(\pi_k(I_k), s_k^r)\end{aligned}$$

$$\Pr(s^r_{k+1}|s^c_k=i,s^r_k,z_k,a_k=0)=\begin{array}{c}\begin{array}{cccccccccc}(0,0)&(1,0)&\cdots&(n_r-1,0)&(n_r,0)&(0,1)&(1,1)&\cdots&(n_r-1,1)&(n_r,1)\end{array}\\ \left[\begin{array}{ccccc|ccccc}0&0&\cdots&0&p_i&0&p_i&\cdots&p_i&p_i\\0&0&\cdots&0&0&0&0&\cdots&0&0\\0&p_i&\cdots&0&0&0&0&\cdots&0&0\\ \vdots&\vdots&\ddots&\vdots&\vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&\cdots&p_i&0&0&0&\cdots&0&0\end{array}\right]\begin{array}{l}0\\1\\2\\ \vdots\\n_r\end{array}\end{array} \tag{30a}$$

$$\Pr(s^r_{k+1}|s^c_k=i,s^r_k,z_k,a_k=1)=\begin{array}{c}\begin{array}{cccccccccc}(0,0)&(1,0)&\cdots&(n_r-1,0)&(n_r,0)&(0,1)&(1,1)&\cdots&(n_r-1,1)&(n_r,1)\end{array}\\ \left[\begin{array}{ccccc|ccccc}0&0&\cdots&0&0&p_i&p_i&\cdots&p_i&p_i\\p_i&p_i&\cdots&p_i&p_i&0&0&\cdots&0&0\\0&0&\cdots&0&0&0&0&\cdots&0&0\\ \vdots&\vdots&\ddots&\vdots&\vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&\cdots&0&0&0&0&\cdots&0&0\end{array}\right]\begin{array}{l}0\\1\\2\\ \vdots\\n_r\end{array}\end{array} \tag{30b}$$

$$\Pr(z_k|s^c_k=i,s^r_k,a_k)=\begin{array}{c}\begin{array}{cccccccccc}(0,0)&(1,0)&\cdots&(n_r-1,0)&(n_r,0)&(0,1)&(1,1)&\cdots&(n_r-1,1)&(n_r,1)\end{array}\\ \left[\begin{array}{ccccc|ccccc}0&q_i\lambda&\cdots&q_i\lambda^{n_r-1}&q_i\lambda^{n_r}&q_i&q_i&\cdots&q_i&q_i\\0&1-q_i\lambda&\cdots&1-q_i\lambda^{n_r-1}&1-q_i\lambda^{n_r}&1-q_i&1-q_i&\cdots&1-q_i&1-q_i\end{array}\right]\begin{array}{l}0\ (\mathtt{NACK})\\1\ (\mathtt{ACK})\end{array}\end{array} \tag{31}$$

where the first inequality follows by induction hypothesis, and the second inequality by (27b). Thus, (28) holds for k.

2) To prove (29), assume that $g_k(\pi)$ minimizes (27b). From (32), we have showed that (29) holds for $k=N$. Assume that (29) holds for $k+1$. Then, from the proof of the first part, the first inequality becomes equality by the induction hypothesis. The second inequality becomes equality since α_k achieves the minimum. Thus, (29) holds for all k.

To show that g is optimal, we evaluate (29) at $k=0$. We have that, $V_0(\pi_0(I_0),s^r_0)=V_0(\alpha_{[0,N-1]},I_0)$. Taking expectation on both sides gives

$$\mathbb{E}[V_0(\pi_0(I_0),s^r_0)]=\mathbb{E}[V_0(\alpha_{[0,N-1]},I_0)]:=J(g).$$

On the other hand, for any $g'\in G$, setting $k=0$ in (28) and taking expectation on both sides give

$$\mathbb{E}[V_0(\pi_0(I_0),s^r_0)]\leq J(g').$$

Thus, g is optimal, and the proof is complete.

References

[1] G. Cocco, A. Munari, and G. Liva, "Remote monitoring of two-state Markov sources via random access channels: An information freshness vs. state estimation entropy perspective," *IEEE Journal on Selected Areas in Information Theory*, vol. 4, pp. 651–666, 2023.

[2] A. Munari, G. Cocco, and G. Liva, "Remote monitoring of Markov sources over random access channels: False alarm and detection probability," in *Asilomar Conference on Signals, Systems, and Computers*, pp. 64–71, IEEE, 2023.

[3] D. Li, S. Wu, Y. Wang, J. Jiao, and Q. Zhang, "Age-optimal harq design for freshness-critical satellite-iot systems," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 2066–2076, 2019.

[4] W. Liu, D. E. Quevedo, Y. Li, K. H. Johansson, and B. Vucetic, "Remote state estimation with smart sensors over Markov fading channels," *IEEE Transactions on Automatic Control*, vol. 67, no. 6, pp. 2743–2757, 2021.

[5] 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.

[6] J. Wang, X. Ren, S. Dey, and L. Shi, "Remote state estimation with usage-dependent Markovian packet losses," *Automatica*, vol. 123, p. 109342, 2021.

[7] F. Qu, A. W. Al-Dabbagh, Z. Zhao, H. Liu, and Y. Li, "Remote state estimation in networked systems: Data pre-processing or not?," *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 6, pp. 3895–3907, 2023.

[8] A. Li, S. Wu, J. Jiao, N. Zhang, and Q. Zhang, "Age of information with hybrid-arq: A unified explicit result," *IEEE Transactions on Communications*, vol. 70, no. 12, pp. 7899–7914, 2022.

[9] S. Banerjee, S. Ulukus, and A. Ephremides, "To re-transmit or not to re-transmit for freshness," in *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pp. 516–523, IEEE, 2023.

[10] T. Soleymani, J. S. Baras, and D. Gündüz, "Transmit or retransmit: A tradeoff in networked control of dynamical processes over lossy channels with ideal feedback," *IEEE Transactions on Information Theory*, 2024.

[11] 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, 2011.

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

[13] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *The Bell System Technical Journal*, vol. 42, no. 5, pp. 1977–1997, 1963.

[14] Y. Qi, P. Cheng, and J. Chen, "Optimal sensor data scheduling for remote estimation over a time-varying channel," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4611–4617, 2016.

[15] K. Huang, W. Liu, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Real-time remote estimation with hybrid arq in wireless networked control," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3490–3504, 2020.

[16] E. T. Ceran, D. Gündüz, and A. György, "Average age of information with hybrid ARQ under a resource constraint," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1900–1913, 2019.

[17] M. Aoki, "Optimal control of partially observable Markovian systems," *Journal of The Franklin Institute*, vol. 280, no. 5, pp. 367–386, 1965.

[18] K. J. Åström, "Optimal control of Markov processes with incomplete state information," *Journal of mathematical analysis and applications*, vol. 10, pp. 174–205, 1965.

[19] 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.

[20] P. Kumar and P. Varaiya, *Stochastic systems: Estimation, identification, and adaptive control*. Prentice Hall, 1986.