Transmit or Retransmit: A Tradeoff in Networked Control of Dynamical Processes over Lossy Channels with Ideal Feedback

Touraj Soleymani, *IEEE*, *Member*, John S. Baras, *IEEE, Life Fellow*, and Deniz Gündüz, *IEEE*, *Fellow*

***Abstract*—We study networked control of a dynamical process over a lossy channel with a hybrid automatic repeat request protocol that connects a sensor to an actuator. The dynamical process is modeled by a Gauss–Markov process, and the lossy channel by a packet-erasure channel with ideal feedback. We suppose that data is communicated in the format of packets with negligible quantization error. In such a networked control system, whenever a packet loss occurs, there exists a tradeoff between transmitting new sensory information with a lower success probability and retransmitting previously failed sensory information with a higher success probability. In essence, an inherent tradeoff between freshness and reliability. To address this tradeoff, we consider a linear-quadratic-regulator performance index, which penalizes state deviations and control efforts over a finite horizon, and jointly design optimal encoding and decoding policies for the encoder and the decoder, which are collocated with the sensor and the actuator, respectively. Our emphasis here lies specifically on designing switching and control policies, rather than error-correcting codes. We show that the optimal encoding policy is a threshold switching policy and the optimal decoding policy is a certainty-equivalent control policy. In addition, we determine the equations that the encoder and the decoder need to solve in order to implement the optimal policies. More specifically, we show that the encoder must solve the Kalman filtering equations, a mismatch linear equation, and a Bellman optimality equation, while the decoder must solve a linear filtering equation and an algebraic Riccati equation.**

***Index Terms*—communication channels, dynamical processes, freshness, erasure channels, hybrid automatic repeat request (HARQ), networked control, optimal policies, packet loss, team decision making, reliability, retransmission.**

I. Introduction

NETWORKED CONTROL systems are distributed feedback systems where the underlying components, i.e., sensors, actuators, and controllers, are connected to each other via communication channels [2]. In these systems, *wireless communication* can play an important role due to various key reasons [3]. First, wireless communication eliminates the need for extensive wiring infrastructure, reducing installation costs of networked control systems and enhancing their flexibility by enabling effortless reconfiguration of control devices. This flexibility is especially advantageous in applications such as industrial automation and smart grids. Second, wireless communication facilitates the placement of control devices in remote or hard-to-reach locations, expanding the accessibility and the coverage of networked control systems. This capability is particularly valuable in infrastructure systems such as environmental monitoring and emergency response systems. Third, wireless communication can easily accommodate the addition or the removal of control devices, providing networked control systems with the ability to scale according to requirements. This scalability is indeed beneficial in environments with dynamic specifications or rapidly changing network topologies such as smart cities and autonomous vehicles.

Notwithstanding the advantages, wireless communication presents some challenges. In particular, wireless channels, which serve to close the feedback control loops in networked control systems, are prone to noise. A direct consequence of the channel noise in real-time tasks is information loss, which can severely degrade the performance of the underlying system or can even lead to instability. It is known that reliable communication close to the capacity limit can be attained with error correction subject to infinite delay or with persistent retransmission based on feedback. In networked control, where data is real-time, any delay more than a certain threshold is typically intolerable. Moreover, retransmission of stale information may not be favorable if fresh information can be transmitted with the same success rate instead. In this situation, the adoption of a *hybrid automatic repeat request (HARQ) protocol*, which integrates error correction with retransmission, seems to be highly promising. Note that an HARQ protocol is able to effectively increase the successful detection probability of a retransmission by combining multiple copies from previously failed transmissions [4]–[8]. Despite the substantial research conducted on HARQ for enhancing transmission reliability in wireless communication systems, its application in networked control systems, however, has received very limited attention in the literature.

In this article, we study networked control of a dynamical process over a lossy channel with an HARQ protocol that connects a sensor to an actuator. We suppose that data is communicated in the format of packets with negligible quantization error. Note that, although the exact relationship between the probability of successful packet detection and the number of retransmissions in an HARQ protocol varies depending on the channel conditions and the adopted HARQ scheme, it has been observed that the probability of successful packet detection generally increases with each retransmission. This suggests that, whenever a packet loss occurs in our networked

Touraj Soleymani and Deniz Gündüz are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, United Kingdom (e-mails: touraj@imperial.ac.uk, d.gunduz@imperial.ac.uk). John S. Baras is with the Institute for Systems Research, University of Maryland College Park, MD 20742, USA (e-mail: baras@umd.edu). This article was presented in part at the *IEEE International Symposium on Information Theory* [1].

control system, there exists a tradeoff between transmitting new sensory information with a lower success probability and retransmitting previously failed sensory information with a higher success probability. In essence, an inherent tradeoff between *freshness* and *reliability*. To address this tradeoff, we consider a performance index that penalizes state deviations and control efforts over a finite horizon, and jointly design optimal encoding and decoding policies for the encoder and the decoder, which are collocated with the sensor and the actuator, respectively.

A. Related Work

Previous research in the literature has investigated the severe effects of packet loss on stability and the performance of state estimation and feedback control [9]–[22]. The majority of works have considered independent and identically distributed (i.i.d.) erasure channels [9]–[14]. In the seminal work in [9], mean-square stability of Kalman filtering over an i.i.d. erasure channel was studied, where it was proved that there exists a critical point for the packet error rate above which the expected estimation error covariance is unbounded. This work was extended to feedback control in [10], where it was shown that there exists a separation between estimation and control when packet acknowledgment is available. Moreover, several works have employed Gilbert–Elliott channels to capture the temporal correlation of wireless channels [15]–[18]. Notably, peak-covariance stability of Kalman filtering over a Gilbert–Elliott channel was addressed in [15], where it was proved that there exists a critical region defined by the recovery and failure rates outside which the expected prediction error covariance is unbounded. The corresponding feedback control problem was considered in [16], where it was shown that the separation principle still holds when packet acknowledgment is available. Furthermore, a number of works have employed fading channels in order to take into account the time variation of the strengths of wireless channels [19]–[21]. In particular, mean-square stability of Kalman filtering over a fading channel with correlated gains was investigated in [19], where a sufficient condition that ensures the exponential boundedness of the expected estimation error covariance was established. Stabilization of a dynamical process in the robust mean-square stability sense over a fading channel was also studied in [20], where a controller with the largest stability margin was designed. Finally, it is worth mentioning that state estimation and feedback control of a Gauss–Markov process over an erasure channel was investigated in [22], where it was demonstrated that transmitting the minimum mean-square-error (MMSE) state estimate at the encoder at each time leads to the maximal information set for the decoder, and a necessary and sufficient condition for the packet-loss probability of an i.i.d. erasure channel that guarantees the boundedness of the expected estimation error covariance was found.

Earlier research has also considered status updating and state estimation over communication channels with retransmission [23]–[34]. In particular, status updating over a binary erasure channel was studied in [23], where the tradeoff between the protection afforded by additional redundancy and the decoding delay associated with longer codewords was analyzed, and the optimal codeword length for a few transmission schemes was obtained. This problem was further examined in [24] for two specific HARQ protocols based on finite-blocklength information-theoretic bounds, where it was shown that there exists an optimal blocklength minimizing the average age and the average peak age of information. Status updating over a noisy channel with reactive and proactive HARQ protocols was investigated in [25], where unified closed-form expressions for the average age and the average peak age of information were derived. Status updating over a binary erasure channel with two specific HARQ protocols was introduced in [26], where the effect of these schemes on the transmission time of data was analyzed. Status updating with the HARQ protocols used in [26] subject to random updates was studied in [27], where it was proved that the optimal encoding policy discards the newly generated update and does not preempt the current one. Status updating over an erasure channel with an HARQ protocol subject to a frequency constraint was studied in [29], where the structure of the optimal signal-independent encoding policy was determined and a reinforcement learning algorithm was proposed for the case in which the channel statistics are unknown. This work was extended in [30] to a broadcast channel setting with multiple receivers. More recently, state estimation over an erasure channel with an HARQ protocol was studied in [32]–[34]. In these works, loss functions were expressed as nonlinear functions of the age of information, and optimal signal-independent encoding policies were obtained. Note that the above studies directly rely on the age of information, a metric that quantifies the time elapsed since the generation of the last successful delivery at each time. Although the age of information is an appropriate instrument for shaping the information flow in many networked real-time systems, it has been shown in [35]–[40] that, for networked control systems, more comprehensive instruments such as the value of information, which measures the difference between the benefit and the cost of a piece of information at each time, are required, and that strategies that depend purely on the age of information can lead to a degradation in the system performance.

B. Contributions and Outline

In this study, we aim to contribute to the field of *semantic*[1] *communications* [41], [42], a new communication paradigm that focuses on exchanging the most significant part of data by considering its contextual relevance. In particular, we explore and assess the relative significance of two types of data, namely new sensory information and previously failed sensory information, sequentially in advanced communication networks, with a specific emphasis on the regulation of dynamical systems. Our main contributions are as follows:

(i) We propose a novel framework for networked control of a dynamical process over a lossy channel with an HARQ protocol, where the objective is to minimize state deviations and control efforts over a finite horizon.

[1]The word "semantics" is etymologically derived from the ancient Greek word "semantikos", which means "significant".

The dynamical process is modeled by a Gauss–Markov process, and the lossy channel by a packet-erasure channel with ideal feedback. We formulate the problem mathematically, and derive the structural properties of the optimal encoding and decoding policies. Our emphasis here lies specifically on designing switching and control policies, rather than error-correcting codes.

(ii) We show that the optimal encoding policy is a threshold switching policy that depends on the system's dynamics and its realizations, and the optimal decoding policy is a certainty-equivalent control policy with a switching filter that depends on the encoder's decision. In addition, we determine the equations that the encoder and the decoder need to solve in order to implement the optimal policies. More specifically, we show that the encoder must solve the Kalman filtering equations, a mismatch linear equation, and a Bellman optimality equation, while the decoder must solve a linear filtering equation and an algebraic Riccati equation.

(iii) We show how the structural properties of the optimal encoding and decoding policies alter when an automatic repeat request (ARQ) protocol is used instead of an HARQ protocol. In this case, we prove that, as expected, retransmitting previously failed measurements is suboptimal, and corroborate that the optimal encoding and decoding policies match the prior findings in the literature (see e.g., [22], [43]). Finally, we conduct a numerical performance comparison between the optimal policies with HARQ and ARQ protocols to validate our theoretical results.

Note that our study differs from [9]–[22], which investigate the impact of channel conditions, given a stability constraint or a performance index, in a setting where, at each time, the encoder transmits new sensory information with some success probability. We here examine the impact of an HARQ protocol, given a performance index, in a setting where, at each time, the encoder decides whether to transmit new sensory information with a lower success probability or retransmit previously failed sensory information with a higher success probability. Besides, our study diverges from [23], [24], [26]–[30], [32], [33], which look for an optimal signal-independent codeword length or an optimal signal-independent encoding policy by considering a performance index for status updating or state estimation, which is expressible in terms of the age of information. We here formulate a dynamic team game with two decision makers, and search for the optimal signal-dependent encoding and decoding policies simultaneously by considering a well-established performance index for feedback control, which cannot be expressed solely in terms of the age of information.

The article is organized as follows. We formulate the problem of interest in Section II. We present our main theoretical results in Section III, and provide the derivation of these results in Section IV. Then, we present a numerical example in Section V. Finally, we conclude the article and discuss future research in Section VI.

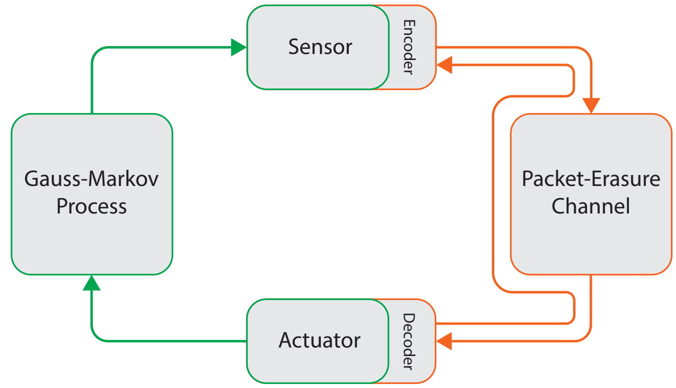

Fig. 1: Networked control of a Gauss–Markov process over a packet-erasure channel with an HARQ protocol. The encoder, collocated with the sensor, and the decoder, collocated with the actuator, are the decision makers. An ideal feedback channel connects the decoder to the encoder.

C. Preliminaries

In the sequel, the sets of real numbers and non-negative integers are denoted by $\mathbb{R}$ and $\mathbb{N}$, respectively. For $x, y \in \mathbb{N}$ and $x \leq y$, the set $\mathbb{N}_{[x,y]}$ denotes $\{z \in \mathbb{N} | x \leq z \leq y\}$. The sequence of all vectors x_t, $t = p, \dots, q$, is represented by $\boldsymbol{x}_{p:q}$. For matrices X and Y, the relations $X \succ 0$ and $Y \succeq 0$ denote that X and Y are positive definite and positive semi-definite, respectively. The logical AND and the logical OR are represented by $\wedge$ and $\vee$, respectively. The indicator function of a subset $\mathcal{A}$ of a set $\mathcal{X}$ is denoted by $\mathbb{1}_{\mathcal{A}} : \mathcal{X} \to \{0, 1\}$. The product operator $\prod_{t=p}^{q} X_t$, where X_t is a matrix, is defined according to the order $X_p \cdots X_q$, and is equal to one when $q < p$. The probability measure of a random variable x is represented by $\mathsf{P}(x)$, its probability density or probability mass function by $\mathsf{p}(x)$, and its expected value and covariance by $\mathsf{E}[x]$ and $\mathsf{cov}[x]$, respectively. We will adopt stochastic kernels to represent decision policies. Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ and $(\mathcal{Y}, \mathcal{B}_{\mathcal{Y}})$ be two measurable spaces. A Borel measurable stochastic kernel $\mathsf{P} : \mathcal{B}_{\mathcal{Y}} \times \mathcal{X} \to [0, 1]$ is a mapping such that $\mathcal{A} \mapsto \mathsf{P}(\mathcal{A}|x)$ is a probability measure on $(\mathcal{Y}, \mathcal{B}_{\mathcal{Y}})$ for any $x \in \mathcal{X}$, and $x \mapsto \mathsf{P}(\mathcal{A}|x)$ is a Borel measurable function for any $\mathcal{A} \in \mathcal{B}_{\mathcal{Y}}$.

II. Networked Control of a Dynamical Process over a Lossy Channel with an HARQ Protocol

Consider a networked control system composed of a dynamical process, a sensor with an encoder, an actuator with a decoder, and a lossy channel with an HARQ protocol that connects the sensor to the actuator (see Fig. 1). At each time k, a message containing a new measurement, represented by $\breve{x}_k$, or a previously failed measurement, represented by $\breve{x}_{k'}$ for $k' < k$, can be transmitted or retransmitted over the channel from the sensor to the actuator, where an actuation input (i.e., control input), represented by a_k, should be computed causally in real time and over a finite time horizon N. We assume that time is discretized into equal time slots, and the measurement $\breve{x}_k$ is chosen such that the best MMSE state estimate is achieved at the controller.

The lossy channel is modeled as a packet-erasure channel with ideal feedback. The decision of the encoder at time k, denoted by $u_k \in \{\mathsf{tx}, \mathsf{rtx}\}$, can be either transmitting a new measurement (i.e., $u_k = \mathsf{tx}$) or retransmitting a previously failed measurement (i.e., $u_k = \mathsf{rtx}$). The packet loss in the channel (i.e. failure in detection) is modeled by a random variable $\gamma_k \in \{0, 1\}$ such that $\gamma_k = 0$ if a packet loss occurs at time k, and $\gamma_k = 1$ otherwise. Let us introduce the variable τ_k such that if the communication at time $k-1$ failed, τ_k is equal to the time elapsed since the last tx decision was made; otherwise, $\tau_k = 0$. Note that τ_k satisfies the recursive relation

$$\tau_k = \begin{cases} 1, & \text{if } u_{k-1} = \mathsf{tx} \wedge \gamma_{k-1} = 0, \\ \tau_{k-1} + 1, & \text{if } u_{k-1} = \mathsf{rtx} \wedge \gamma_{k-1} = 0, \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

for $k \in \mathbb{N}_{[0,N]}$ with initial condition $\tau_0 = 0$. The channel satisfies the input-output relation

$$z_{k+1} = \begin{cases} \check{x}_k, & \text{if } u_k = \mathsf{tx} \wedge \gamma_k = 1, \\ \check{x}_{k-\tau_k}, & \text{if } u_k = \mathsf{rtx} \wedge \gamma_k = 1, \\ \mathfrak{F}, & \text{otherwise} \end{cases} \tag{2}$$

for $k \in \mathbb{N}_{[0,N]}$ with $z_0 = \mathfrak{F}$ by convention, where z_k is the output of the channel and $\mathfrak{F}$ represents the occurrence of a packet loss. Let ω_k represent the number of communication attempts before time k associated with the previously failed measurement $\check{x}_{k-\tau_k}$. It is not difficult to observe that ω_k is in fact equal to τ_k. Let the packet error rate at time k associated with a measurement after s retransmissions be denoted by $\lambda_k(s)$. Then, the packet error rate for communication of the new measurement $\check{x}_k$ at time k is $\lambda_k(0)$, and that for communication of the previously failed measurement $\check{x}_{k-\tau_k}$ at time k is $\lambda_k(\omega_k)$. It is assumed that the packet error rate $\lambda_k(\omega_k)$ for any ω_k and $k \in \mathbb{N}_{[0,N]}$ is a random variable forming a Markov chain; the packet error rates $\lambda_k(0)$ and $\lambda_k(\omega_k)$ are estimated perfectly and known at the encoder at each time k; the random variables γ_k for $k \in \mathbb{N}_{[0,N]}$ are mutually independent given their respective packet error rates; signaling effect [44] and measurement quantization error [2] are negligible; a transmitted or retransmitted measurement is received after one-step delay, which is fixed and independent of the packet content; and packet acknowledgments are sent back from the decoder to the encoder via an ideal feedback channel.

Remark 1: Note that the packet error rate $\lambda_k(\omega_k)$ at time k associated with a previously failed measurement after ω_k retransmissions depends on ω_k, the channel conditions, and the particular HARQ scheme used for combining multiple copies from previously failed messages (see e.g., [7] for how to estimate the packet error rate empirically when an HARQ protocol is used). It has been observed that, in any reasonable HARQ scheme, $\lambda_k(\omega_k)$ is generally non-increasing in ω_k, i.e., $\lambda_k(\omega_k^1) > \lambda_k(\omega_k^2)$ for all $\omega_k^1 < \omega_k^2$. Furthermore, note that standard HARQ schemes impose a finite limit on the maximum number of retransmissions permitted, i.e., $\omega_k \leq \omega_{\max}$. This limit is taken into account in our study.

Remark 2: Although the main focus of this study is on a lossy channel with an HARQ protocol, as described above, we will utilize our findings to draw conclusions also about a lossy channel with an ARQ protocol. It is important to note that unlike an HARQ protocol, which incorporates both retransmission and error correction, an ARQ protocol relies solely on retransmission. In particular, in an ARQ protocol, failed messages are discarded, and each retransmission is decoded as a new message. As a result, when an ARQ protocol is adopted instead of an HARQ protocol, the packet error rate for communication of the new measurement $\check{x}_k$ at time k is $\lambda_k(0)$, and that for communication of the previously failed measurement $\check{x}_{k-\tau_k}$ at time k is also $\lambda_k(0)$.

The dynamical process, equipped with a sensor and an actuator with computational capabilities, is modeled as a partially observable Gauss–Markov process. This process satisfies the state and output equations

$$x_{k+1} = A_k x_k + B_k a_k + w_k, \tag{3}$$

$$y_k = C_k x_k + v_k \tag{4}$$

for $k \in \mathbb{N}_{[0,N]}$ with initial condition x_0, where $x_k \in \mathbb{R}^n$ is the state of the process, $A_k \in \mathbb{R}^{n \times n}$ is the state matrix, $B_k \in \mathbb{R}^{n \times m}$ is the input matrix, $a_k \in \mathbb{R}^m$ is the actuation input, $w_k \in \mathbb{R}^n$ is a Gaussian white noise with zero mean and covariance $W_k \succ 0$, $y_k \in \mathbb{R}^p$ is the output of the sensor, $C_k \in \mathbb{R}^{p \times n}$ is the output matrix, and $v_k \in \mathbb{R}^p$ is a Gaussian white noise with zero mean and covariance $V_k \succ 0$. It is assumed that the initial condition x_0 is a Gaussian vector with mean m_0 and covariance M_0, and the random variables x_0, w_t, and v_s for $t, s \in \mathbb{N}_{[0,N]}$ are mutually independent.

Remark 3: Note that the adopted process model represents a wide class of physical systems. Such a Gauss–Markov model has been adopted extensively in control and communication, as studying this basic model can provide a foundation for development of more sophisticated systems. Moreover, the time-varying nature of the model allows approximation of nonlinear systems around their nominal trajectories, and the partially observable nature of it respects the fact that in reality only a noisy version of the output of the sensor can be observed.

The information sets of the encoder and the decoder at time k can be represented by

$$\mathcal{I}_k^e = \left\{ y_t, z_t, \lambda_t, a_{t-1}, u_{t-1}, \gamma_{t-1} \middle| t \in \mathbb{N}_{[0,k]} \right\}, \tag{5}$$

$$\mathcal{I}_k^d = \left\{ z_t, \lambda_t, a_{t-1}, u_{t-1}, \gamma_{t-1} \middle| t \in \mathbb{N}_{[0,k]} \right\} \tag{6}$$

for $k \in \mathbb{N}_{[0,N]}$, respectively. At each time k for $k \in \mathbb{N}_{[0,N]}$, the encoder must decide about u_k and the decoder about a_k based on the Borel measurable stochastic kernels $\mathsf{P}(u_k|\mathcal{I}_k^e)$ and $\mathsf{P}(a_k|\mathcal{I}_k^d)$, respectively. A coding policy profile (π, μ), consisting of an encoding (i.e., switching) policy π and a decoding (i.e., control) policy μ, is considered admissible if $\pi = \{\mathsf{P}(u_k|\mathcal{I}_k^e) | k \in \mathbb{N}_{[0,N]}\}$ and $\mu = \{\mathsf{P}(a_k|\mathcal{I}_k^d) | k \in \mathbb{N}_{[0,N]}\}$. We would like to identify the best possible solution, denoted as $(\pi^\star, \mu^\star)$, to the stochastic optimization problem

$$\underset{\pi \in \mathcal{P}, \mu \in \mathcal{M}}{\text{minimize}} \ \Upsilon(\pi, \mu) \tag{7}$$

subject to the lossy channel model in (2), and the dynamical process model in (3) and (4), where $\mathcal{P}$ and $\mathcal{M}$ are the sets of admissible encoding policies and admissible decoding policies, respectively, and

$$\Upsilon(\pi,\mu) := \frac{1}{N+1}\,\mathsf{E}\left[\sum_{k=0}^{N+1} x_k^T Q_k x_k + \sum_{k=0}^{N} a_k^T R_k a_k\right] \tag{8}$$

for $Q_k \succeq 0$ and $R_k \succ 0$ as weighting matrices.

Remark 4: The optimization problem in (7) is a team decision-making problem with a loss function that is expressed in terms of state deviations and control efforts over a finite time horizon. This loss function, which is widely used in the literature, measures the performance of feedback control. The main challenge in solving the optimization problem in (7) is that optimal encoding and decoding policies must be designed jointly based on an information structure that is specified by $(\mathcal{I}_k^e, \mathcal{I}_k^d)$. In the subsequent section, we will proceed to characterize a globally optimal solution $(\pi^\star, \mu^\star)$. This optimal solution will play a crucial role in determining the fundamental performance limit of the networked control system under consideration. In fact, by finding a globally optimal solution, we can establish the best achievable system performance.

III. Main Results: Structural Properties of Optimal Encoding and Decoding Policies

In this section, we focus on presenting our theoretical findings. However, before delving into the details, it is essential to first establish several key definitions. We say a policy profile $(\pi^\star, \mu^\star)$ associated with the loss function $\Upsilon(\pi,\mu)$ is globally optimal (i.e., team optimal) if

$$\Upsilon(\pi^\star,\mu^\star) \le \Upsilon(\pi,\mu), \text{ for all } \pi \in \mathcal{P}, \mu \in \mathcal{M}.$$

Note that globally optimal solutions express a stronger solution concept than Nash equilibria.

We define two distinct value functions $V_k^e(\mathcal{I}_k^e)$ and $V_k^d(\mathcal{I}_k^d)$ associated with the loss function $\Upsilon(\pi,\mu)$ as

$$V_k^e(\mathcal{I}_k^e) := \min_{\pi\in\mathcal{P}:\mu=\mu^\star} \mathsf{E}\left[\sum_{t=k+1}^{N} s_t^T \Lambda_t s_t \Big| \mathcal{I}_k^e\right], \tag{9}$$

$$V_k^d(\mathcal{I}_k^d) := \min_{\mu\in\mathcal{M}:\pi=\pi^\star} \mathsf{E}\left[\sum_{t=k}^{N} s_t^T \Lambda_t s_t \Big| \mathcal{I}_k^d\right] \tag{10}$$

for $k \in \mathbb{N}_{[0,N]}$ given a policy profile $(\pi^\star,\mu^\star)$, where $s_t = a_t + (B_t^T S_{t+1} B_t + R_t)^{-1} B_t^T S_{t+1} A_t x_t$ and $\Lambda_t = B_t^T S_{t+1} B_t + R_t$, and $S_t \succeq 0$ obeys the algebraic Riccati equation

$$S_t = Q_t + A_t^T S_{t+1} A_t - A_t^T S_{t+1} B_t \times \left(B_t^T S_{t+1} B_t + R_t\right)^{-1} B_t^T S_{t+1} A_t \tag{11}$$

for $t \in \mathbb{N}_{[0,N]}$ with initial condition $S_{N+1} = Q_{N+1}$. Note that this definition takes into account the fact that the encoder's decision at time k can affect the cost function only from time $k+1$ onward.

In addition, we define the innovation at the encoder $\nu_k := y_k - C_k\,\mathsf{E}[x_k|\mathcal{I}_{k-1}^e]$, the estimation error at the encoder based on the conditional mean $\check{e}_k := x_k - \mathsf{E}[x_k|\mathcal{I}_k^e]$, the estimation error at the decoder based on the conditional mean $\hat{e}_k := x_k - \mathsf{E}[x_k|\mathcal{I}_k^d]$, the estimation mismatch based on the conditional means $\tilde{e}_k := \mathsf{E}[x_k|\mathcal{I}_k^e] - \mathsf{E}[x_k|\mathcal{I}_k^d]$, and the value residual from the perspective of the encoder $\Delta_k := (\lambda_k(\omega_k) - \lambda_k(0))\,\mathsf{E}[V_{k+1}(\mathcal{I}_{k+1}^e)|\mathcal{I}_k^e, \gamma_k = 0] - (1-\lambda_k(0))\,\mathsf{E}[V_{k+1}(\mathcal{I}_{k+1}^e)|\mathcal{I}_k^e, u_k = \mathsf{tx}, \gamma_k = 1] + (1-\lambda_k(\omega_k))\,\mathsf{E}[V_{k+1}(\mathcal{I}_{k+1}^e)|\mathcal{I}_k^e, u_k = \mathsf{rtx}, \gamma_k = 1]$, for $k \in \mathbb{N}_{[0,N]}$. Note that ν_k, $\check{e}_k$, $\hat{e}_k$, $\tilde{e}_k$, and Δ_k are all computable at the encoder at each time k.

Our main theoretical results are presented in the next theorems, which specify the structural properties of the encoding policy $\pi^\star$ and the decoding policy $\mu^\star$ of a globally optimal solution $(\pi^\star,\mu^\star)$.

Theorem 1: The optimal encoding policy $\pi^\star$ in networked control of a Gauss–Markov process over a packet-erasure channel with an HARQ protocol is the threshold switching policy

$$u_k^\star = \begin{cases} \mathsf{tx}, & \text{if } \omega_k > \omega_{\max} \vee \tau_k = 0 \vee \Omega_k \ge 0, \\ \mathsf{rtx}, & \text{otherwise} \end{cases} \tag{12}$$

along with $\check{x}_k = \mathsf{E}[x_k|\mathcal{I}_k^e]$ for $k \in \mathbb{N}_{[0,N]}$, where $\Omega_k = (\lambda_k(\omega_k) - \lambda_k(0))\tilde{e}_k^T A_k^T \Gamma_{k+1} A_k \tilde{e}_k + (1-\lambda_k(\omega_k))\varepsilon_k^T \Gamma_{k+1} \varepsilon_k + \Delta_k$, $\Gamma_k = A_k^T S_{k+1} B_k (B_k^T S_{k+1} B_k + R_k)^{-1} B_k^T S_{k+1} A_k$, and $\varepsilon_k = \sum_{t=0}^{\tau_k - 1} (\prod_{t'=0}^{t} A_{k-t'}) K_{k-t} \nu_{k-t}$. This encoding policy can be expressed at each time k as a function of $\tilde{e}_k$, $\boldsymbol{\nu}_{k-\tau_k+1:k}$, τ_k, and λ_k. To apply this encoding policy, the following equations need to be solved online:

$$\check{x}_k = A_{k-1}\check{x}_{k-1} + B_{k-1} a_{k-1} + K_k \nu_k, \tag{13}$$

$$P_k = \left(\left(A_{k-1} P_{k-1} A_{k-1}^T + W_{k-1}\right)^{-1} + C_k^T V_k^{-1} C_k\right)^{-1}, \tag{14}$$

$$\tilde{e}_k = \mathbb{1}_{u_{k-1}=\mathsf{tx} \wedge \gamma_{k-1}=1} K_k \nu_k + \mathbb{1}_{u_{k-1}=\mathsf{rtx} \wedge \gamma_{k-1}=1} \sum_{t=0}^{\tau_{k-1}} \left(\prod_{t'=1}^{t} A_{k-t'}\right) K_{k-t} \nu_{k-t} + \mathbb{1}_{\gamma_{k-1}=0}\left(A_{k-1}\tilde{e}_{k-1} + K_k \nu_k\right) \tag{15}$$

for $k \in \mathbb{N}_{[1,N]}$ with initial conditions $\check{x}_0 = m_0 + K_0 \nu_0$, $P_0 = (M_0^{-1} + C_0^T V_0^{-1} C_0)^{-1}$, and $\tilde{e}_0 = K_0 \nu_0$, where $K_k = P_k C_k^T V_k^{-1}$.

Proof: See Section IV. ■

Theorem 2: The optimal decoding policy $\mu^\star$ in networked control of a Gauss–Markov process over a packet-erasure channel with an HARQ protocol is the certainty-equivalent control policy

$$a_k^\star = -L_k \hat{x}_k \tag{16}$$

for $k \in \mathbb{N}_{[0,N]}$, where $L_k = (B_k^T S_{k+1} B_k + R_k)^{-1} B_k^T S_{k+1} A_k$. To apply this decoding policy, the following equation needs to *be solved online:*

$$\begin{aligned}\hat{x}_k = &\mathbb{1}_{u_{k-1}=\mathsf{tx}\wedge\gamma_{k-1}=1}\Big(A_{k-1}\check{x}_{k-1}+B_{k-1}a_{k-1}\Big)\\&+\mathbb{1}_{u_{k-1}=\mathsf{rtx}\wedge\gamma_{k-1}=1}\Bigg(\Bigg(\prod_{t=1}^{\tau_{k-1}+1}A_{k-t}\Bigg)\check{x}_{k-\tau_{k-1}-1}\\&+\sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t}A_{k-t'}\Bigg)B_{k-t-1}a_{k-t-1}\Bigg)\\&+\mathbb{1}_{\gamma_{k-1}=0}\Big(A_{k-1}\hat{x}_{k-1}+B_{k-1}a_{k-1}\Big)\end{aligned}\tag{17}$$

for $k\in\mathbb{N}_{[1,N]}$ with initial condition $\hat{x}_0=m_0$, where $\hat{x}_k=\mathsf{E}[x_k|\mathcal{I}^d_k]$.

Proof: See Section IV. ■

Remark 5: The structural results in Theorems 1 and 2 certify the existence of a globally optimal solution composed of a threshold switching policy that depends on the system's dynamics and its realizations, and a certainty-equivalent control policy with a switching filter that depends on the encoder's decision. Note that these policies can be designed completely separately. Moreover, the results assert that the optimal encoding policy transmits the encoder's current MMSE state estimate, i.e., $\check{x}_k$, only if $\omega_k>\omega_{\max}\vee\tau_k=0\vee\Omega_k\geq 0$, and retransmits the encoder's previously failed MMSE state estimate, i.e., $\check{x}_{k-\tau_k}$, otherwise; and that the optimal decoding policy is constructed by inserting the decoder's current MMSE state estimate, i.e., $\hat{x}_k$, into the corresponding optimal state-feedback control policy.

Remark 6: Note that the computational complexity of the solution $(\pi^\star,\mu^\star)$ is the same as that of solving (9). As we will prove, $V_k(\mathcal{I}^e_k)$ is a function of $\tilde{e}_k$, $\boldsymbol{\nu}_{k-\tau_k+1:k}$, τ_k, and λ_k. Now, if τ_k is bounded by $\tau_{\max}$, and $\tilde{e}_k$, ν_t for $t\in\mathbb{N}_{[k-\tau_k+1,k]}$, and λ_k are discretized in grids with c_1^n, c_2^p, and c_3 points, respectively, and the expected value is expressed based on a weighted sum of c_4 samples, the computational complexity is then $\mathcal{O}(Nc_1^nc_2^{p\times(\tau_{\max}-1)}c_3c_4\tau_{\max})$. This can be expensive especially when n and p are large. In practice, the value residual Ω_k can be approximated based on the one-step lookahead algorithm (see, e.g., [45]). Applying this procedure, we find $\Omega_k\simeq(\lambda_k(\omega_k)-\lambda_k(0))\tilde{e}_k^TA_k^T\Gamma_{k+1}A_k\tilde{e}_k+(1-\lambda_k(\omega_k))\varepsilon_k^T\Gamma_{k+1}\varepsilon_k$, which is quadratic in terms of $\tilde{e}_k$ and ε_k. Following the definitions of $\tilde{e}_k$ and ε_k, we observe that, while these variables are influenced by the age of information at the decoder, they cannot be expressed solely in terms of it, as they in general depend on the system's dynamics and its realizations.

The next propositions, which are direct applications of the above theorems, show how the structural properties of the optimal encoding and decoding policies alter when an ARQ protocol is used instead of an HARQ one.

Proposition 1: The optimal encoding policy $\pi^\star$ in networked control of a Gauss–Markov process over a packet-erasure channel with an ARQ protocol is the uniform transmission policy

$$u_k^\star=\mathsf{tx}\tag{18}$$

along with $\check{x}_k=\mathsf{E}[x_k|\mathcal{I}^e_k]$ for $k\in\mathbb{N}_{[0,N]}$. To apply this encoding policy, the following equations need to be solved online:

$$\check{x}_k=A_{k-1}\check{x}_{k-1}+B_{k-1}a_{k-1}+K_k\nu_k,\tag{19}$$

$$P_k=\Big(\Big(A_{k-1}P_{k-1}A_{k-1}^T+W_{k-1}\Big)^{-1}+C_k^TV_k^{-1}C_k\Big)^{-1}\tag{20}$$

for $k\in\mathbb{N}_{[1,N]}$ with initial conditions $\check{x}_0=m_0+K_0\nu_0$ and $P_0=(M_0^{-1}+C_0^TV_0^{-1}C_0)^{-1}$, where $K_k=P_kC_k^TV_k^{-1}$.

Proof: See Section IV. ■

Proposition 2: The optimal decoding policy $\mu^\star$ in networked control of a Gauss–Markov process over a packet-erasure channel with an ARQ protocol is the certainty-equivalent control policy

$$a_k^\star=-L_k\hat{x}_k\tag{21}$$

for $k\in\mathbb{N}_{[0,N]}$, where $L_k=(B_k^TS_{k+1}B_k+R_k)^{-1}B_k^TS_{k+1}A_k$. To apply this decoding policy, the following equation needs to be solved online:

$$\hat{x}_k=A_{k-1}(\gamma_{k-1}\tilde{e}_{k-1}+\hat{x}_{k-1})+B_{k-1}a_{k-1}\tag{22}$$

for $k\in\mathbb{N}_{[1,N]}$ with initial condition $\hat{x}_0=m_0$, where $\hat{x}_k=\mathsf{E}[x_k|\mathcal{I}^d_k]$.

Proof: See Section IV. ■

Remark 7: The results of Propositions 1 and 2 indicate that, as expected, retransmitting previously failed measurements is indeed suboptimal when an ARQ protocol is adopted instead of an HARQ protocol, i.e., when $\lambda_k(\omega_k)=\lambda_k(0)$. This finding suggests that the adoption of an ARQ protocol does not provide any advantage for networked control systems. Note that, in this case, the optimal encoding policy transmits the encoder's current MMSE state estimate, i.e., $\check{x}_k$, uniformly; and the optimal decoding policy is constructed based on the decoder's current MMSE state estimate, i.e., $\hat{x}_k$, which obeys simple dynamics. These results match the prior findings in the networked control literature (see e.g., [22], [43]).

IV. Derivation of Main Results

This section is dedicated to the derivation of the main results, presented in Section III. Note that, given any optimal switching policy implemented, the optimal value that minimizes the mean square error at the decoder at time k is the conditional mean $\mathsf{E}[x_k|\mathcal{I}^d_k]$. Besides, the conditional mean $\mathsf{E}[x_k|\mathcal{I}^e_k]$ combines all current and previous outputs of the sensor that are accessible to the encoder at time k. This implies that if this new message is transmitted by the encoder at time k, from the MMSE perspective, the decoder is able to develop a state estimate upon the successful receipt of the message at time $k+1$ that would be the same if it had all the previous outputs of the sensor until time k, which is the best possible case for the decoder. Moreover, given an HARQ protocol, when the last transmitted message fails, aside from transmitting a new message, the encoder at time k has an additional choice to retransmit a previously failed message. This message, following the definition of τ_k, is in fact $\mathsf{E}[x_{k-\tau_k}|\mathcal{I}^e_{k-\tau_k}]$. Therefore, without loss of optimality, the encoder at each time k should decide to transmit either $\check{x}_k = \mathsf{E}[x_k|\mathcal{I}_k^e]$ or $\check{x}_{k-\tau_k} = \mathsf{E}[x_{k-\tau_k}|\mathcal{I}_{k-\tau_k}^e]$.

We first characterize in the next two lemmas the recursive equations that the optimal estimators at the encoder and the decoder must satisfy.

Lemma 1: The optimal estimator minimizing the MMSE at the encoder satisfies the recursive equations

$$\check{x}_k = m_k + K_k\Big(y_k - C_k m_k\Big), \tag{23}$$

$$m_k = A_{k-1}\check{x}_{k-1} + B_{k-1}a_{k-1}, \tag{24}$$

$$P_k = \Big(M_k^{-1} + C_k^T V_k^{-1} C_k\Big)^{-1}, \tag{25}$$

$$M_k = A_{k-1}P_{k-1}A_{k-1}^T + W_{k-1} \tag{26}$$

for $k \in \mathbb{N}_{[1,N]}$ *with initial conditions* $\check{x}_0 = m_0 + K_0(y_0 - C_0 m_0)$ *and* $P_0 = (M_0^{-1} + C_0^T V_0^{-1} C_0)^{-1}$*, where* $\check{x}_k = \mathsf{E}[x_k|\mathcal{I}_k^e]$*,* $m_k = \mathsf{E}[x_k|\mathcal{I}_{k-1}^e]$*,* $P_k = \mathsf{cov}[x_k|\mathcal{I}_k^e]$*,* $M_k = \mathsf{cov}[x_k|\mathcal{I}_{k-1}^e]$*, and* $K_k = P_k C_k^T V_k^{-1}$*.*

Proof: Given the information set $\mathcal{I}_k^e$, the optimal estimator at the encoder must satisfy the Kalman filter equations for the conditional mean and the conditional covariance. For detailed derivation of these equations see e.g., [46]. ■

Lemma 2: The optimal estimator minimizing the MMSE at the decoder satisfies the recursive equation

$$\begin{aligned}\hat{x}_k = {} & \mathbb{1}_{u_{k-1}=\mathsf{tx}\wedge\gamma_{k-1}=1}\Big(A_{k-1}\check{x}_{k-1} + B_{k-1}a_{k-1}\Big) \\ & + \mathbb{1}_{u_{k-1}=\mathsf{rtx}\wedge\gamma_{k-1}=1}\Bigg(\Bigg(\prod_{t=1}^{\tau_{k-1}+1} A_{k-t}\Bigg)\check{x}_{k-\tau_{k-1}-1} \\ & + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)B_{k-t-1}a_{k-t-1}\Bigg) \\ & + \mathbb{1}_{\gamma_{k-1}=0}\Big(A_{k-1}\hat{x}_{k-1} + B_{k-1}a_{k-1}\Big)\end{aligned} \tag{27}$$

for $k \in \mathbb{N}_{[1,N]}$ *with initial condition* $\hat{x}_0 = m_0$*, where* $\hat{x}_k = \mathsf{E}[x_k|\mathcal{I}_k^d]$*.*

Proof: Writing x_k in terms of x_{k-1} based on (3), and taking expectation given the information set $\mathcal{I}_k^d$, we obtain

$$\mathsf{E}\Big[x_k|\mathcal{I}_k^d\Big] = A_{k-1}\,\mathsf{E}\Big[x_{k-1}|\mathcal{I}_k^d\Big] + B_{k-1}a_{k-1} \tag{28}$$

for $k \in \mathbb{N}_{[1,N]}$ as $\mathsf{E}[w_{k-1}|\mathcal{I}_k^d] = 0$. If $\gamma_{k-1} = 0$, regardless of u_{k-1}, we have $z_k = \mathfrak{F}$. Note that $\mathsf{E}[x_{k-1}|\mathcal{I}_k^d] = \mathsf{E}[x_{k-1}|\mathcal{I}_{k-1}^d, z_k = \mathfrak{F}, \lambda_k, a_{k-1}, u_{k-1}, \gamma_{k-1} = 0] = \mathsf{E}[x_{k-1}|\mathcal{I}_{k-1}^d] = \hat{x}_{k-1}$. Therefore, using (28), when $\gamma_{k-1} = 0$, regardless of u_{k-1}, we get

$$\mathsf{E}\Big[x_k|\mathcal{I}_k^d\Big] = A_{k-1}\hat{x}_{k-1} + B_{k-1}a_{k-1} \tag{29}$$

for $k \in \mathbb{N}_{[1,N]}$.

However, if $u_{k-1} = \mathsf{tx}$ and $\gamma_{k-1} = 1$, we have $z_k = \check{x}_{k-1}$. In this case, we get $\mathsf{E}[x_{k-1}|\mathcal{I}_k^d] = \mathsf{E}[x_{k-1}|\mathcal{I}_{k-1}^d, z_k = \check{x}_{k-1}, \lambda_k, a_{k-1}, u_{k-1} = \mathsf{tx}, \gamma_{k-1} = 1] = \mathsf{E}[x_{k-1}|\check{x}_{k-1}, P_{k-1}] = \check{x}_{k-1}$ as $\{\check{x}_{k-1}, P_{k-1}\}$ is a sufficient statistic of $\mathcal{I}_k^d$ with respect to x_{k-1}. Hence, using (28), when $u_{k-1} = \mathsf{tx}$ and $\gamma_{k-1} = 1$, we get

$$\mathsf{E}\Big[x_k|\mathcal{I}_k^d\Big] = A_{k-1}\check{x}_{k-1} + B_{k-1}a_{k-1} \tag{30}$$

for $k \in \mathbb{N}_{[1,N]}$.

Furthermore, writing x_k in terms of $x_{k-\tau_{k-1}-1}$ based on (3), and taking expectation given the information set $\mathcal{I}_k^d$, we obtain

$$\begin{aligned}\mathsf{E}\Big[x_k|\mathcal{I}_k^d\Big] = {} & \Bigg(\prod_{t=1}^{\tau_{k-1}+1} A_{k-t}\Bigg)\,\mathsf{E}\Big[x_{k-\tau_{k-1}-1}|\mathcal{I}_k^d\Big] \\ & + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)B_{k-t-1}a_{k-t-1}\end{aligned} \tag{31}$$

for $k \in \mathbb{N}_{[1,N]}$ as $\sum_{t=0}^{\tau_{k-1}}(\prod_{t'=1}^{t} A_{k-t'})\,\mathsf{E}[w_{k-t-1}|\mathcal{I}_k^d] = 0$. If $u_{k-1} = \mathsf{rtx}$ and $\gamma_{k-1} = 1$, we have $z_k = \check{x}_{k-\tau_{k-1}-1}$. In this case, we get $\mathsf{E}[x_{k-\tau_{k-1}-1}|\mathcal{I}_k^d] = \mathsf{E}[x_{k-\tau_{k-1}-1}|\mathcal{I}_{k-1}^d, z_k = \check{x}_{k-\tau_{k-1}-1}, \lambda_k, a_{k-1}, u_{k-1} = \mathsf{rtx}, \gamma_{k-1} = 1] = \mathsf{E}[x_{k-\tau_{k-1}-1}|\check{x}_{k-\tau_{k-1}-1}, P_{k-\tau_{k-1}-1}] = \check{x}_{k-\tau_{k-1}-1}$ as $\{\check{x}_{k-\tau_{k-1}-1}, P_{k-\tau_{k-1}-1}\}$ is a sufficient statistic of $\mathcal{I}_k^d$ with respect to $x_{k-\tau_{k-1}-1}$. Hence, using (31), when $u_{k-1} = \mathsf{rtx}$ and $\gamma_{k-1} = 1$, we get

$$\begin{aligned}\mathsf{E}\Big[x_k|\mathcal{I}_k^d\Big] = {} & \Bigg(\prod_{t=1}^{\tau_{k-1}+1} A_{k-t}\Bigg)\check{x}_{k-\tau_{k-1}-1} \\ & + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)B_{k-t-1}a_{k-t-1}\end{aligned} \tag{32}$$

for $k \in \mathbb{N}_{[1,N]}$.

We obtain (27) by combining (29), (30), and (32). Note that the initial condition is $\mathsf{E}[x_0] = m_0$ because no measurement is available at the decoder at time $k = 0$. ■

Lemma 3: The estimation error at the decoder satisfies the recursive equation

$$\begin{aligned}\hat{e}_k = {} & \mathbb{1}_{u_{k-1}=\mathsf{tx}\wedge\gamma_{k-1}=1}\Big(A_{k-1}\check{e}_{k-1} + w_{k-1}\Big) \\ & + \mathbb{1}_{u_{k-1}=\mathsf{rtx}\wedge\gamma_{k-1}=1}\sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)w_{k-t-1} \\ & + \mathbb{1}_{\gamma_{k-1}=0}\Big(A_{k-1}\hat{e}_{k-1} + w_{k-1}\Big)\end{aligned} \tag{33}$$

for $k \in \mathbb{N}_{[1,N]}$ *with initial condition* $\hat{e}_0 = x_0 - m_0$*, where* $\hat{e}_k = x_k - \mathsf{E}[x_k|\mathcal{I}_k^d]$

Proof: Using (3), we can write x_k in terms of x_{k-1} as

$$x_k = A_{k-1}x_{k-1} + B_{k-1}a_{k-1} + w_{k-1} \tag{34}$$

for $k \in \mathbb{N}_{[1,N]}$ with initial condition x_0, and in terms of $x_{k-\tau_{k-1}-1}$ as

$$\begin{aligned}x_k = {} & \Bigg(\prod_{t=1}^{\tau_{k-1}+1} A_{k-t}\Bigg)x_{k-\tau_{k-1}-1} \\ & + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)B_{k-t-1}a_{k-t-1} \\ & + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)w_{k-t-1}\end{aligned} \tag{35}$$

for $k \in \mathbb{N}_{[1,N]}$ with initial condition x_0.

Moreover, from the law of total probability, $\mathsf{p}(u_{k-1} = \mathsf{tx} \wedge \gamma_{k-1} = 1) + \mathsf{p}(u_{k-1} = \mathsf{rtx} \wedge \gamma_{k-1} = 1) + \mathsf{p}(\gamma_{k-1} =$

$0) = 1$. Thus, using (34) and (35), we can get

$$x_k = \mathbb{1}_{u_{k-1}=\mathsf{tx} \wedge \gamma_{k-1}=1}\Big(A_{k-1}x_{k-1} + B_{k-1}a_{k-1} + w_{k-1}\Big) + \mathbb{1}_{u_{k-1}=\mathsf{rtx} \wedge \gamma_{k-1}=1}\Bigg(\Bigg(\prod_{t=1}^{\tau_{k-1}+1} A_{k-t}\Bigg)x_{k-\tau_{k-1}-1} + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)B_{k-t-1}a_{k-t-1} + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)w_{k-t-1}\Bigg) + \mathbb{1}_{\gamma_{k-1}=0}\Big(A_{k-1}x_{k-1} + B_{k-1}a_{k-1} + w_{k-1}\Big) \tag{36}$$

for $k \in \mathbb{N}_{[1,N]}$ with initial condition x_0.

By definition, $\hat{e}_k = x_k - \hat{x}_k$. Therefore, we obtain (33) by subtracting (27) from (36). The initial condition $\hat{e}_0$ is also obtained by subtracting $\hat{x}_0 = m_0$ from x_0. ∎

Lemma 4: The estimation mismatch satisfies the recursive equation

$$\tilde{e}_k = \mathbb{1}_{u_{k-1}=\mathsf{tx} \wedge \gamma_{k-1}=1}K_k\nu_k + \mathbb{1}_{u_{k-1}=\mathsf{rtx} \wedge \gamma_{k-1}=1}\sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)K_{k-t}\nu_{k-t} + \mathbb{1}_{\gamma_{k-1}=0}\Big(A_{k-1}\tilde{e}_{k-1} + K_k\nu_k\Big) \tag{37}$$

for $k \in \mathbb{N}_{[1,N]}$ with initial condition $\tilde{e}_0 = K_0\nu_0$, where $\tilde{e}_k = \mathsf{E}[x_k|\mathcal{I}_k^e] - \mathsf{E}[x_k|\mathcal{I}_k^d]$

Proof: From the definition of the innovation ν_k and the state estimate $\check{x}_k$, we find that

$$\nu_k = y_k - C_k\Big(A_{k-1}\check{x}_{k-1} + B_{k-1}a_{k-1}\Big) \tag{38}$$

with the exception $\nu_0 = y_0 - C_0m_0$. Note that ν_k is a white Gaussian noise with zero mean and covariance $N_k = C_kM_kC_k^T + V_k$.

Using (23), (24), and (38), we can write $\check{x}_k$ in terms of $\check{x}_{k-1}$ as

$$\check{x}_k = A_{k-1}\check{x}_{k-1} + B_{k-1}a_{k-1} + K_k\nu_k \tag{39}$$

for $k \in \mathbb{N}_{[1,N]}$ with initial condition $\check{x}_0 = m_0 + K_0\nu_0$, and in terms of $\check{x}_{k-\tau_{k-1}-1}$ as

$$\check{x}_k = \Bigg(\prod_{t=1}^{\tau_{k-1}+1} A_{k-t}\Bigg)\check{x}_{k-\tau_{k-1}-1} + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)B_{k-t-1}a_{k-t-1} + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)K_{k-t}\nu_{k-t} \tag{40}$$

for $k \in \mathbb{N}_{[1,N]}$ with initial condition $\check{x}_0 = m_0 + K_0\nu_0$.

Moreover, from the law of total probability, we have $\mathsf{p}(u_{k-1} = \mathsf{tx} \wedge \gamma_{k-1} = 1) + \mathsf{p}(u_{k-1} = \mathsf{rtx} \wedge \gamma_{k-1} = 1) + \mathsf{p}(\gamma_{k-1} = 0) = 1$. Thus, using (39) and (40), we can get

$$\check{x}_k = \mathbb{1}_{u_{k-1}=\mathsf{tx} \wedge \gamma_{k-1}=1}\Big(A_{k-1}\check{x}_{k-1} + B_{k-1}a_{k-1} + K_k\nu_k\Big) + \mathbb{1}_{u_{k-1}=\mathsf{rtx} \wedge \gamma_{k-1}=1}\Bigg(\Bigg(\prod_{t=1}^{\tau_{k-1}+1} A_{k-t}\Bigg)\check{x}_{k-\tau_{k-1}-1} + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=1}^{t} A_{k-t'}\Bigg)B_{k-t-1}a_{k-t-1} + \sum_{t=0}^{\tau_{k-1}}\Bigg(\prod_{t'=0}^{t} A_{k-t'}\Bigg)K_{k-t}\nu_{k-t}\Bigg) + \mathbb{1}_{\gamma_{k-1}=0}\Big(A_{k-1}\check{x}_{k-1} + B_{k-1}a_{k-1} + K_k\nu_k\Big) \tag{41}$$

for $k \in \mathbb{N}_{[1,N]}$ with initial condition $\check{x}_0 = m_0$.

By definition, $\tilde{e}_k = \check{x}_k - \hat{x}_k$. Therefore, we obtain (37) by subtracting (27) from (41). The initial condition $\tilde{e}_0$ is also obtained by subtracting $\hat{x}_0 = m_0$ from $\check{x}_0 = m_0 + K_0\nu_0$. ∎

We now present the proof of Theorems 1 and 2.

Proof: Let (π^o, μ^o) denote a policy profile in the set of globally optimal solutions. It is evident that this set cannot be empty. We prove that the policy profile $(\pi^\star, \mu^\star)$ in the claim is globally optimal by showing that $\Upsilon(\pi^\star, \mu^\star)$ cannot be greater than $\Upsilon(\pi^o, \mu^o)$. Our proof is structured in the following way:

$$\Upsilon(\pi^o, \mu^o) = \Upsilon(\pi^n, \mu^o) \geq \Upsilon(\pi^n, \mu^c) \geq \Upsilon(\pi^\star, \mu^\star). \tag{42}$$

In particular, we first find an innovation-based switching policy π^n such that $\Upsilon(\pi^n, \mu^o) = \Upsilon(\pi^o, \mu^o)$. Then, we derive a certainty-equivalent control policy μ^c such that $\Upsilon(\pi^n, \mu^c) \leq \Upsilon(\pi^n, \mu^o)$. Finally, we show that for the policy profile in the claim we have $\Upsilon(\pi^\star, \mu^\star) \leq \Upsilon(\pi^n, \mu^c)$. Throughout our analysis, without loss of generality, we assume that $m_0 = 0$. Similar arguments can be made for $m_0 \neq 0$ following a coordinate transformation.

In the first step of the proof, we will show that, given the control policy μ^o, we can find an innovation-based switching policy π^n that is equivalent to the switching policy π^o. Note that a switching policy in the context of our problem is innovation-based if it depends on $\boldsymbol{\nu}_{0:k}$ instead of $\boldsymbol{y}_{0:k}$ and $\boldsymbol{z}_{0:k}$ at each time k. From the definition of ν_k, we get

$$\boldsymbol{y}_k = \boldsymbol{\nu}_k + E_k\check{\boldsymbol{x}}_{k-1} + F_k\boldsymbol{a}_{k-1} \tag{43}$$

where E_k and F_k are matrices of proper dimensions. From (23) and (24), we find that

$$\check{\boldsymbol{x}}_k = G_k\boldsymbol{\nu}_k + H_k\boldsymbol{a}_{k-1} \tag{44}$$

where G_k and H_k are matrices of proper dimensions. Furthermore, from (2), we know that $\boldsymbol{z}_k$ is a function of $\check{\boldsymbol{x}}_{k-1}$, $\boldsymbol{u}_{k-1}$, and $\boldsymbol{\gamma}_{k-1}$. As a result, it is possible to write

$$\mathsf{p}_{\pi^o}(u_k|\mathcal{I}_k^e) = \mathsf{p}_{\pi^o}\Big(u_k\Big|\boldsymbol{\nu}_k, \boldsymbol{\lambda}_k, \boldsymbol{a}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{\gamma}_{k-1}\Big),$$

$$\mathsf{p}_{\mu^o}(a_k|\mathcal{I}_k^d) = \mathsf{p}_{\mu^o}\Big(a_k\Big|\boldsymbol{\nu}_{k-1}, \boldsymbol{\lambda}_k, \boldsymbol{a}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{\gamma}_{k-1}\Big).$$

Accordingly, any realizations of u_k and a_k can be expressed as $u_k = u_k(\eta_k; \boldsymbol{\nu}_k, \boldsymbol{\lambda}_k, \boldsymbol{a}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{\gamma}_{k-1})$ and $a_k = a_k(\zeta_k; \boldsymbol{\nu}_{k-1}, \boldsymbol{\lambda}_k, \boldsymbol{a}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{\gamma}_{k-1})$, respectively, where η_k

and ζ_k represent random variables that are independent of any other variables. Hence, it is possible to recursively construct π^n with $\mathsf{p}_{\pi^n}(u_k|\boldsymbol{\nu}_k, \boldsymbol{\lambda}_k, \boldsymbol{\zeta}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{\gamma}_{k-1})$ such that it is equivalent to $\mathsf{p}_{\pi^o}(u_k|\mathcal{I}_k^e)$. This proves that $\Upsilon(\pi^n, \mu^o) = \Upsilon(\pi^o, \mu^o)$. It should be emphasized that the switching policy π^n, initially constructed associated with the control policy μ^o, is now dependent solely on $\boldsymbol{\nu}_k$, $\boldsymbol{\lambda}_k$, $\boldsymbol{\zeta}_{k-1}$, $\boldsymbol{u}_{k-1}$, and $\boldsymbol{\gamma}_{k-1}$ at each time k.

In the second step of the proof, given the switching policy π^n, we will search for an optimal control policy μ^c, and prove that it is a certainty-equivalent control policy. We first present three identities. From (3), we have

$$x_{k+1}^T S_{k+1} x_{k+1} = \Big(A_k x_k + B_k a_k + w_k\Big)^T \times S_{k+1}\Big(A_k x_k + B_k a_k + w_k\Big). \quad (45)$$

From (11), we can write

$$x_k^T S_k x_k = x_k^T \Big(Q_k + A_k^T S_{k+1} A_k - L_k^T \Big(B_k^T S_{k+1} B_k + R_k\Big) L_k\Big) x_k. \quad (46)$$

Moreover, through a simple algebraic calculation, we obtain

$$x_{N+1}^T S_{N+1} x_{N+1} - x_0^T S_0 x_0 = \sum_{k=0}^{N} x_{k+1}^T S_{k+1} x_{k+1} - \sum_{k=0}^{N} x_k^T S_k x_k. \quad (47)$$

Incorporating the identities (45) and (46) into the identity (47), taking the expectation of both sides of (47), and using the facts that w_k is independent of x_k and a_k and that the terms $x_0^T S_0 x_0$ and $w_k^T S_{k+1} w_k$ are independent of the switching and control policies, we find the loss function

$$\Upsilon'(\pi, \mu) := \mathsf{E}\left[\sum_{k=0}^{N} s_k^T \Lambda_k s_k\right] \quad (48)$$

for any $\pi \in \mathcal{P}$ and for any $\mu \in \mathcal{M}$. Note that $\Upsilon'(\pi, \mu)$ is equivalent to $\Upsilon(\pi, \mu)$ in the sense that optimizing the former over (π, μ) yields the same optimal solutions as optimizing the latter over (π, μ).

Note that, at time k, the term $\mathsf{E}[\sum_{t=0}^{k-1} s_t^T \Lambda_t s_t]$ in (48) will not be affected by the control policy executed from time k onward. Associated with $\Upsilon'(\pi^n, \mu)$ for π^n that was obtained in the first step and for any $\mu \in \mathcal{M}$, we define the value function $V_k^d(\mathcal{I}_k^d)$ as

$$V_k^d(\mathcal{I}_k^d) := \min_{\mu \in \mathcal{M}} \mathsf{E}\left[\sum_{t=k}^{N} s_t^T \Lambda_t s_t \Big| \mathcal{I}_k^d\right] \quad (49)$$

for $k \in \mathbb{N}_{[0,N]}$ with initial condition $V_{N+1}^c(\mathcal{I}_{N+1}^c) = 0$. Building upon the prior findings in the literature (refer to, e.g., [22], [36], [37], [43], [47]), it becomes evident that the separation principle holds, and the minimizer in (49) is obtained by $a_k^\star = -L_k \hat{x}_k$. This establishes that $\Upsilon(\pi^n, \mu^c) \leq \Upsilon(\pi^n, \mu^o)$, and completes the proof of Theorem 2.

In the third step of the proof, we will show that $\Upsilon(\pi^\star, \mu^\star) \leq \Upsilon(\pi^n, \mu^c)$. Note that, at time k, the term $\mathsf{E}[\sum_{t=0}^{k} s_t^T \Lambda_t s_t]$ in (48) will not be affected by the switching policy executed from time k onward. Associated with $\Upsilon'(\pi^n, \mu^c)$ for any $\pi^n \in \mathcal{P}$ that is innovation-based and for μ^c that was obtained in the second step, we define the value function $V_k^e(\mathcal{I}_k^e)$ as

$$V_k^e(\mathcal{I}_k^e) := \min_{\pi^n \in \mathcal{P}} \mathsf{E}\left[\sum_{t=k+1}^{N} s_t^T \Lambda_t s_t \Big| \mathcal{I}_k^e\right] \quad (50)$$

for $k \in \mathbb{N}_{[0,N]}$ with initial condition $V_{N+1}^e(\mathcal{I}_{N+1}^e) = 0$. From the additivity of $V_k^e(\mathcal{I}_k^e)$ and by using $a_k = -L_k \hat{x}_k$, we obtain

$$\begin{aligned} V_k^e(\mathcal{I}_k^e) &= \min_{\mathsf{p}(u_k|\mathcal{I}_k^e)} \mathsf{E}\Big[\hat{e}_{k+1}^T \Gamma_{k+1} \hat{e}_{k+1} \\ &\quad + \min_{\mathsf{p}(u_{k+1}|\mathcal{I}_{k+1}^e)} \mathsf{E}\Big[\hat{e}_{k+2}^T \Gamma_{k+2} \hat{e}_{k+2} + \dots \Big| \mathcal{I}_{k+1}^e\Big] \Big| \mathcal{I}_k^e\Big] \\ &= \min_{\mathsf{p}(u_k|\mathcal{I}_k^e)} \mathsf{E}\Big[\hat{e}_{k+1}^T \Gamma_{k+1} \hat{e}_{k+1} + V_{k+1}^e(\mathcal{I}_{k+1}^e) \Big| \mathcal{I}_k^e\Big] \\ &= \min_{\mathsf{p}(u_k|\mathcal{I}_k^e)} \mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} + \mathrm{tr}(\Gamma_{k+1} P_{k+1}) \\ &\quad + V_{k+1}^e(\mathcal{I}_{k+1}^e) \Big| \mathcal{I}_k^e\Big] \end{aligned} \quad (51)$$

for $k \in \mathbb{N}_{[0,N]}$ with initial condition $V_{N+1}^e(\mathcal{I}_{N+1}^e) = 0$, where $\Gamma_k = A_k^T S_{k+1} B_k (B_k^T S_{k+1} B_k + R_k)^{-1} B_k^T S_{k+1} A_k$, and in the third equality we used the fact that, by the tower property of conditional expectations, $\mathsf{E}[\hat{e}_{k+1}^T \Gamma_{k+1} \hat{e}_{k+1} | \mathcal{I}_k^e] = \mathsf{E}[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} | \mathcal{I}_k^e] + \mathrm{tr}(\Gamma_{k+1} P_{k+1})$. We will prove by backward induction that $V_k^e(\mathcal{I}_k^e)$ can be written in terms of $\tilde{e}_k$, $\boldsymbol{\nu}_{k-\tau_k+1:k}$, τ_k, and λ_k. The claim is satisfied for time $N+1$. We assume that the claim holds at time $k+1$, and shall prove that it also holds at time k.

By the hypothesis, $V_{k+1}^e(\mathcal{I}_{k+1}^e)$ is function of $\tilde{e}_{k+1}$, $\boldsymbol{\nu}_{k-\tau_{k+1}+2:k+1}$, τ_{k+1}, and λ_{k+1}. Note that, from (37), we can write $\tilde{e}_{k+1}$ in terms of $\tilde{e}_k$, $\boldsymbol{\nu}_{k-\tau_k+1:k+1}$, u_k, τ_k, and γ_k; from (1), we can write τ_{k+1} in terms of τ_k, u_k, and γ_k; and by the Markov property, we know that λ_{k+1} depends on λ_k. Moreover, we recall that $0 \leq \tau_{k+1} \leq \tau_k + 1$. Hence, there exists a function $g(.)$ such that

$$V_{k+1}^e(\mathcal{I}_{k+1}^e) = g\Big(\tilde{e}_k, \boldsymbol{\nu}_{k-\tau_k+1:k+1}, u_k, \tau_k, \lambda_k, \gamma_k\Big). \quad (52)$$

This implies that $\mathsf{E}[V_{k+1}^e(\mathcal{I}_{k+1}^e) | \mathcal{I}_k^e, u_k]$ is a function of $\tilde{e}_k$, $\boldsymbol{\nu}_{k-\tau_k+1:k}$, u_k, τ_k, and λ_k as ν_{k+1} and γ_k are averaged out. In addition, since we can write $\tilde{e}_{k+1}$ in terms of $\tilde{e}_k$, $\boldsymbol{\nu}_{k-\tau_k+1:k+1}$, u_k, τ_k, and γ_k, there exists a function $h(.)$ such that

$$\tilde{e}_{k+1}^T \tilde{e}_{k+1} = h\Big(\tilde{e}_k, \boldsymbol{\nu}_{k-\tau_k+1:k+1}, u_k, \tau_k, \gamma_k\Big). \quad (53)$$

This implies that $\mathsf{E}[\tilde{e}_{k+1}^T \tilde{e}_{k+1} | \mathcal{I}_k^e, u_k]$ is a function of $\tilde{e}_k$, $\boldsymbol{\nu}_{k-\tau_k+1:k}$, τ_k, and λ_k as ν_{k+1} and γ_k are averaged out. Therefore, the claim holds.

Lastly, we will need to obtain the switching condition of the optimal switching policy. Note that, according to (37), when $u_k = \mathsf{tx}$ and $\gamma_k = 1$, $\tilde{e}_{k+1}$ satisfies

$$\tilde{e}_{k+1} = K_{k+1} \nu_{k+1}. \quad (54)$$

When $u_k = \mathsf{rtx}$ and $\gamma_k = 1$, it satisfies

$$\tilde{e}_{k+1} = \sum_{t=0}^{\tau_k} \Big(\prod_{t'=1}^{t} A_{k-t'+1} \Big) K_{k-t+1} \nu_{k-t+1} = \sum_{t=0}^{\tau_k - 1} \Big(\prod_{t'=0}^{t} A_{k-t'} \Big) K_{k-t} \nu_{k-t} + K_{k+1} \nu_{k+1}. \quad (55)$$

Furthermore, when $\gamma_k = 0$, it satisfies

$$\tilde{e}_{k+1} = A_k \tilde{e}_k + K_{k+1} \nu_{k+1}. \quad (56)$$

As a result, since ν_{k+1} is a white Gaussian noise with zero mean and covariance $N_{k+1} = C_{k+1} M_{k+1} C_{k+1}^T + V_{k+1}$, from (54), (55), and (56), we can derive

$$\mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} \big| \mathcal{I}_k^e, u_k = \mathsf{tx}, \gamma_k = 1\Big] = \operatorname{tr}\Big(\Gamma_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big), \quad (57)$$

$$\mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} \big| \mathcal{I}_k^e, u_k = \mathsf{rtx}, \gamma_k = 1\Big] = \varepsilon_k^T \Gamma_{k+1} \varepsilon_k + \operatorname{tr}\Big(\Gamma_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big), \quad (58)$$

$$\mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} \big| \mathcal{I}_k^e, \gamma_k = 0\Big] = \tilde{e}_k^T A_k^T \Gamma_{k+1} A_k \tilde{e}_k + \operatorname{tr}\Big(\Gamma_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big) \quad (59)$$

where in (58) we have $\varepsilon_k = \sum_{t=0}^{\tau_k - 1} (\prod_{t'=0}^{t} A_{k-t'}) K_{k-t} \nu_{k-t}$ and used the fact that $\mathsf{E}[\varepsilon_k | \mathcal{I}_k^e] = \varepsilon_k$, and in (59) we used the fact that $\mathsf{E}[\tilde{e}_k | \mathcal{I}_k^e] = \tilde{e}_k$. Note that, since (56) holds regardless of the value of u_k, we have

$$\mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} \big| \mathcal{I}_k^e, u_k = \mathsf{tx}, \gamma_k = 0\Big] = \mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} \big| \mathcal{I}_k^e, u_k = \mathsf{rtx}, \gamma_k = 0\Big] = \mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} \big| \mathcal{I}_k^e, \gamma_k = 0\Big].$$

Employing (57), (58), and (59), and applying the law of total expectation for the quadratic mismatch terms in (51), we get

$$\begin{aligned}\mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} \big| \mathcal{I}_k^e, u_k = \mathsf{tx}\Big] &= \lambda_k(0) \tilde{e}_k^T A_k^T \Gamma_{k+1} A_k \tilde{e}_k + \lambda_k(0) \operatorname{tr}\Big(\Gamma_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big) + \big(1 - \lambda_k(0)\big) \operatorname{tr}\Big(\Gamma_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big) \\ &= \lambda_k(0) \tilde{e}_k^T A_k^T \Gamma_{k+1} A_k \tilde{e}_k + \operatorname{tr}\Big(\Gamma_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big) \end{aligned} \quad (60)$$

and

$$\begin{aligned}\mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} \big| \mathcal{I}_k^e, u_k = \mathsf{rtx}\Big] &= \lambda_k(\omega_k) \tilde{e}_k^T A_k^T \Gamma_{k+1} A_k \tilde{e}_k + \lambda_k(\omega_k) \operatorname{tr}\Big(\Gamma_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big) \\ &\quad + \big(1 - \lambda_k(\omega_k)\big) \varepsilon_k^T \Gamma_{k+1} \varepsilon_k + \big(1 - \lambda_k(\omega_k)\big) \operatorname{tr}\Big(\Gamma_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big) \\ &= \lambda_k(\omega_k) \tilde{e}_k^T A_k^T \Gamma_{k+1} A_k \tilde{e}_k + \big(1 - \lambda_k(\omega_k)\big) \varepsilon_k^T \Gamma_{k+1} \varepsilon_k + \operatorname{tr}\Big(\Gamma_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big). \end{aligned} \quad (61)$$

Besides, applying the law of total expectation for the cost-to-go terms in (51), we get

$$\begin{aligned}\mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{tx}\Big] &= \lambda_k(0)\, \mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{tx}, \gamma_k = 0\Big] \\ &\quad + \big(1 - \lambda_k(0)\big)\, \mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{tx}, \gamma_k = 1\Big] \end{aligned} \quad (62)$$

and

$$\begin{aligned}\mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{rtx}\Big] &= \lambda_k(\omega_k)\, \mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{rtx}, \gamma_k = 0\Big] \\ &\quad + \big(1 - \lambda_k(\omega_k)\big)\, \mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{rtx}, \gamma_k = 1\Big]. \end{aligned} \quad (63)$$

Observe that since no measurement can be detected correctly at time $k+1$ when $\gamma_k = 0$, in (62) and (63), we can use the equalities

$$\mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{tx}, \gamma_k = 0\Big] = \mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{rtx}, \gamma_k = 0\Big] = \mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, \gamma_k = 0\Big].$$

Inserting (60), (61), (62), and (63) in (51), we deduce that $u_k^\star = \mathsf{tx}$ if

$$\begin{aligned}&\big(\lambda_k(\omega_k) - \lambda_k(0)\big) \tilde{e}_k^T A_k^T \Gamma_{k+1} A_k \tilde{e}_k + \big(1 - \lambda_k(\omega_k)\big) \varepsilon_k^T \Gamma_{k+1} \varepsilon_k \\ &+ \big(\lambda_k(\omega_k) - \lambda_k(0)\big)\, \mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, \gamma_k = 0\Big] \\ &- \big(1 - \lambda_k(0)\big)\, \mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{tx}, \gamma_k = 1\Big] \\ &+ \big(1 - \lambda_k(\omega_k)\big)\, \mathsf{E}\Big[V_{k+1}^e(\mathcal{I}_{k+1}^e) \big| \mathcal{I}_k^e, u_k = \mathsf{rtx}, \gamma_k = 1\Big] \geq 0. \end{aligned}$$

We also know that $u_k^\star = \mathsf{tx}$ if $\omega_k > \omega_{\max}$ or $\tau_k = 0$. The condition $\omega_k > \omega_{\max}$ comes from the HARQ scheme, and the condition $\tau_k = 0$ from the fact that the previous transmission was successful. In both cases, no retransmission is taken place. This completes the proof of Theorem 1. ■

Finally, we present the proof of Propositions 1 and 2.

Proof: We will specialize the results of Theorems 1 and 2 for a packet-erasure channel with an ARQ protocol, where the packet error rate for communication of the new measurement

$\check{x}_k$ at time k is $\lambda_k(0)$, and that for communication of the previously failed measurement $\check{x}_{k-\tau_k}$ at time k is also $\lambda_k(0)$. To that end, we need to convert $\lambda_k(\omega_k)$ in the results of Theorem 1 to $\lambda_k(0)$. Aside from this observation, we will prove by backward induction that $V_k^e(\mathcal{I}_k^e)$ can be written as $V_k^e(\mathcal{I}_k^e) = \mathsf{E}_{\boldsymbol{\lambda}_{k+1:N}}[\tilde{e}_k^T \Phi_k \tilde{e}_k + \phi_k | \mathcal{I}_k^e]$, where $\Phi_k \succeq 0$ and ϕ_k is independent of $\tilde{e}_k$. The claim is satisfied for time $N+1$. We assume that the claim holds at time $k+1$, and shall prove that it also holds at time k.

Note that, from (51) and by the hypothesis, we can write

$$\begin{aligned} V_k^e(\mathcal{I}_k^e) &= \min_{\mathsf{p}(u_k|\mathcal{I}_k^e)} \mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} + \mathrm{tr}(\Gamma_{k+1} P_{k+1}) \\ &\quad + V_{k+1}^e(\mathcal{I}_{k+1}^e) \Big| \mathcal{I}_k^e\Big] \\ &= \min_{\mathsf{p}(u_k|\mathcal{I}_k^e)} \mathsf{E}\Big[\tilde{e}_{k+1}^T \Gamma_{k+1} \tilde{e}_{k+1} + \mathrm{tr}(\Gamma_{k+1} P_{k+1}) \\ &\quad + \tilde{e}_{k+1} \Phi_{k+1} \tilde{e}_{k+1} + \phi_{k+1} \Big| \mathcal{I}_k^e\Big] \\ &= \min_{\mathsf{p}(u_k|\mathcal{I}_k^e)} \mathsf{E}\Big[\tilde{e}_{k+1}^T \Theta_{k+1} \tilde{e}_{k+1} + \theta_{k+1} \Big| \mathcal{I}_k^e\Big] \end{aligned} \tag{64}$$

where in the second equality we used the tower property of conditional expectations, and in the third equality we have $\Theta_{k+1} = \Gamma_{k+1} + \Phi_{k+1}$ and $\theta_{k+1} = \mathrm{tr}(\Gamma_{k+1} P_{k+1}) + \phi_{k+1}$. Note that θ_{k+1} is independent of u_k. Employing operations analogous to those resulted in (60) and (61), we get

$$\begin{aligned} &\mathsf{E}\Big[\tilde{e}_{k+1}^T \Theta_{k+1} \tilde{e}_{k+1} \Big| \mathcal{I}_k^e, u_k = \mathsf{tx}, \boldsymbol{\lambda}_{k+1:N}\Big] \\ &= \lambda_k(0) \tilde{e}_k^T A_k^T \Theta_{k+1} A_k \tilde{e}_k \\ &\quad + \mathrm{tr}\Big(\Theta_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big) \end{aligned} \tag{65}$$

and

$$\begin{aligned} &\mathsf{E}\Big[\tilde{e}_{k+1}^T \Theta_{k+1} \tilde{e}_{k+1} \Big| \mathcal{I}_k^e, u_k = \mathsf{rtx}, \boldsymbol{\lambda}_{k+1:N}\Big] \\ &= \lambda_k(0) \tilde{e}_k^T A_k^T \Theta_{k+1} A_k \tilde{e}_k \\ &\quad + \big(1 - \lambda_k(0)\big) \varepsilon_k^T \Theta_{k+1} \varepsilon_k \\ &\quad + \mathrm{tr}\Big(\Theta_{k+1} K_{k+1} N_{k+1} K_{k+1}^T\Big). \end{aligned} \tag{66}$$

Inserting (65) and (66) in (64), we deduce that $u_k^\star = \mathsf{tx}$ if

$$\mathsf{E}_{\boldsymbol{\lambda}_{k+1:N}}\Big[\big(1 - \lambda_k(0)\big) \varepsilon_k^T \Theta_{k+1} \varepsilon_k \Big| \mathcal{I}_k^e\Big] \geq 0.$$

This condition is always satisfied because $1 - \lambda_k(0) \geq 0$ and $\varepsilon_k^T \Theta_{k+1} \varepsilon_k \geq 0$. Consequently, we can write $V_k^e(\mathcal{I}_k^e)$ as $V_k^e(\mathcal{I}_k^e) = \mathsf{E}_{\boldsymbol{\lambda}_{k+1:N}}[\tilde{e}_k^T \Phi_k \tilde{e}_k + \phi_k | \mathcal{I}_k^e]$ with $\Phi_k = \lambda_k(0) A_k^T \Theta_{k+1} A_k$ and $\phi_k = \mathrm{tr}(\Theta_{k+1} K_{k+1} N_{k+1} K_{k+1}^T) + \theta_{k+1}$. Therefore, the claim holds. This completes the proof of Proposition 1.

Furthermore, since $u_k^\star = \mathsf{tx}$ for all $k \in \mathbb{N}_{[0,N]}$, the dynamics of the decoder's MMSE state estimate gets simplified as

$$\begin{aligned} \hat{x}_k &= \mathbb{1}_{\gamma_{k-1}=1}\Big(A_{k-1} \check{x}_{k-1} + B_{k-1} a_{k-1}\Big) \\ &\quad + \mathbb{1}_{\gamma_{k-1}=0}\Big(A_{k-1} \hat{x}_{k-1} + B_{k-1} a_{k-1}\Big) \\ &= A_{k-1}(\gamma_{k-1} \tilde{e}_{k-1} + \hat{x}_{k-1}) + B_{k-1} a_{k-1} \end{aligned}$$

for $k \in \mathbb{N}_{[1,N]}$ with initial condition $\hat{x}_0 = m_0$. This completes the proof of Proposition 2. ■

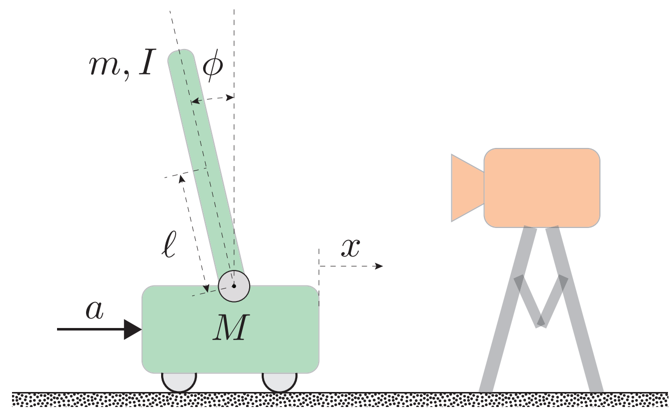

Fig. 2: An inverted pendulum on a cart with an external sensor. The sensory information is communicated to the actuation point over a wireless channel.

V. Numerical Results: An Inverted Pendulum on a Cart with an External Sensor

In this section, we present a numerical example to illustrate the theoretical results discussed in the previous sections. Consider an inverted pendulum on a cart (see Fig 2). For this dynamical process, the continuous-time equations of motion linearized around the unstable equilibrium are given by

$$(M+m)\ddot{x} + b\dot{x} - ml\ddot{\phi} = a,$$

$$(I + ml^2)\ddot{\phi} - mgl\phi = ml\ddot{x}$$

where x is the position of the cart, ϕ is the pitch angle of the pendulum, a is the force applied to the cart, $M = 0.5$ kg is the mass of the cart, $m = 0.2$ kg is the mass of the pendulum, $b = 0.1$ N/m/sec is the friction coefficient for the cart, $l = 0.3$ m is the distance from the pivot to the pendulum's center of mass, $I = 0.006$ kg.m^2 is the moment of inertia of the pendulum, and $g = 9.81$ m/s^2 is the gravity. There is an external sensor that measures the position and the pitch angle at each time. In our example, the state and output equations of the dynamical process of the form (3) and (4), the input-output relation of the channel of the form (2), and the loss function of the form (7) are specified by the state matrix $A_k = 10^{-4} \times [10000, 100, 1, 0; 0, 9982, 267, 1; 0, 0, 10016, 100; 0, -45, 3122, 10016]$, the input matrix $B_k = [0.0001; 0.0182; 0.0002; 0.0454]$, the output matrix $C_k = [1, 0, 0, 0; 0, 0, 1, 0]$, the process noise covariance $W_k = 10^{-4} \times [6, 3, 1, 6; 3, 8, 3, 4; 1, 3, 7, 6; 6, 4, 6, 31]$, the sensor noise covariance $V_k = 10^{-4} \times [20, 0; 0, 10]$ for

SWT+CE	UNI+CE	UNI+HI	UNI+ZI
24.08	25.41	28.02	28.84

TABLE I: Comparison of the performance of four different coding policy profiles: SWT+CE represents the optimal coding policy profile with an HARQ protocol, UNI+CE represents the optimal coding policy profile without an HARQ protocol, UNI+HI represents the uniform transmission policy and the hold-input control policy, and UNI+ZI represents the uniform transmission policy and the zero-input control policy. Each value indicates the associated system loss, measured in terms of Υ according to a Monte Carlo simulation.

$k \in \mathbb{N}_{[0,N]}$, the mean and the covariance of the initial condition $m_0 = [0; 0; 0.2; 0]$ and $M_0 = 10W_k$, the maximum number of retransmissions $\omega_{\max} = 1$, the packet error rates $\lambda_k(0) = 0.5$ and $\lambda_k(1) = 0.05$ for $k \in \mathbb{N}_{[0,N]}$, the weighting matrices $Q_k = \mathrm{diag}\{1, 1, 1000, 1\}$ for $k \in \mathbb{N}_{[0,N+1]}$ and $R_k = 1$ for $k \in \mathbb{N}_{[0,N]}$, and the time horizon $N = 500$.

The simulation results corresponding to a realization of the described networked control system based on the optimal coding policy profile with an HARQ protocol are illustrated in Fig. 3. In particular, in the top diagram, the solid curve represents the stage cost trajectory[2], the black dots represent the transmission time instants, the blue dots represent the retransmission time instants, and the red dots represent the packet loss time instants. In this experiment, the total number of transmissions was 404, the total number of retransmissions was 97, and the total number of packet losses was 210. In the middle diagram, the blue curve represents the cart's position trajectory and the red curve represents the cart's velocity trajectory. We can observe that the cart's position and velocity remained satisfactorily bounded within the ranges $[-0.17 \text{ m}, 2.43 \text{ m}]$ and $[-0.69 \text{ m/s}, 1.88 \text{ m/s}]$, respectively. Finally, in the bottom diagram, the blue curve represents the pendulum's pitch angle trajectory and the red curve represents the pendulum's pitch rate trajectory. We can again observe that the pendulum's pitch angle and pitch rate remained satisfactorily bounded within the ranges $[-0.27 \text{ rad}, 0.40 \text{ rad}]$ and $[-2.39 \text{ rad/s}, 1.71 \text{ rad/s}]$, respectively.

Note that, apart from a certainty-equivalent control policy, which has been shown to be optimal, one can resort to a less complex control policy such as a hold-input policy, in which the previous actuation input is executed if a packet loss occurs, or a zero-input policy, in which the actuation input is set to zero if a packet loss occurs [48]. We compared the performance of four different coding policy profiles for the above networked control system through a Monte Carlo simulation with 1000 experiments (see Table I). We found out that the system loss, measured in terms of Υ according to the Monte Carlo simulation, is equal to 24.08 for the optimal coding policy profile with an HARQ protocol, is equal to 25.41 for the optimal coding policy profile without an HARQ protocol, is equal to 28.02 when the uniform transmission policy and the hold-input control policy were adopted, and is equal to 28.84 when the uniform transmission policy and the zero-input control policy were adopted. This outcome reaffirms the effectiveness of our framework.

[2]Note that the stage cost at each time k is defined as $(x_k^T Q_k x_k + a_k^T R_k a_k)/(N+1)$.

VI. Conclusions

In this article, we developed a framework for networked control of a Gauss–Markov process over a packet-erasure channel with an HARQ protocol. We observed that, in this networked control system, whenever a packet loss occurs, there exists an inherent tradeoff between transmitting the encoder's current MMSE state estimate with a lower success probability and retransmitting the encoder's previously failed MMSE state estimate with a higher success probability. Our objective was to obtain the optimal encoding and decoding policies that minimize a linear-quadratic-regulator performance index, penalizing state deviations and control efforts over a finite horizon. We derived the structural properties of the optimal policies, and determined the equations that need to be solved for the implementation of these policies. In addition, we examined how these results change when an ARQ protocol is used instead of an HARQ one. Our results confirmed the intuition that the adoption of an HARQ protocol is essential for retransmissions to be beneficial in a networked control system. Future research should explore the application of learning algorithms for networked control systems equipped with HARQ protocols in scenarios where the parameters and the statistics of dynamical processes and communication channels are unknown.

References

[1] T. Soleymani, J. S. Baras, and D. Gündüz, "Networked control with hybrid automatic repeat request protocols," in *Proc. IEEE Int. Symp. on Information Theory*, 2024.

[2] J. Baillieul and P. J. Antsaklis, "Control and communication challenges in networked real-time systems," *Proceedings of IEEE*, vol. 95, no. 1, pp. 9–28, 2007.

[3] P. Park, S. C. Ergen, C. Fischione, C. Lu, and K. H. Johansson, "Wireless network design for control systems: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 978–1013, 2017.

[4] A. Giovanidis, G. Wunder, and J. Bühler, "Optimal control of a single queue with retransmissions: Delay-dropping tradeoffs," *IEEE Trans. on Wireless Communications*, vol. 8, no. 7, pp. 3736–3746, 2009.

[5] Y. Li, M. C. Gursoy, and S. Velipasalar, "Throughput of HARQ-IR with finite blocklength codes and QoS constraints," in *Proc. IEEE Int. Symp. on Information Theory*, pp. 276–280, 2017.

[6] X. Lagrange, "Throughput of HARQ protocols on a block fading channel," *IEEE Communications Letters*, vol. 14, no. 3, pp. 257–259, 2010.

[7] V. Tripathi, E. Visotsky, R. Peterson, and M. Honig, "Reliability-based type II hybrid ARQ schemes," in *Proc. IEEE Int. Conf. on Communications*, vol. 4, pp. 2899–2903, 2003.

[8] P. Frenger, S. Parkvall, and E. Dahlman, "Performance comparison of HARQ with chase combining and incremental redundancy for HSDPA," in *Proc. IEEE Vehicular Technology Conf.*, vol. 3, pp. 1829–1833, 2001.

[9] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[10] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, 2007.

[11] V. Gupta, B. Hassibi, and R. M. Murray, "Optimal LQG control across packet-dropping links," *Systems & Control Letters*, vol. 56, no. 6, pp. 439–446, 2007.

[12] K. Plarre and F. Bullo, "On Kalman filtering for detectable systems with intermittent observations," *IEEE Trans. on Automatic Control*, vol. 54, no. 2, pp. 386–390, 2009.

Fig. 3: Simulation results for the networked control system illustrated in Fig. 2 when the optimal coding policy profile with an HARQ protocol is adopted. In the top diagram, the solid curve represents the stage cost trajectory, the black dots represent the transmission (TX) time instants, the blue dots represent the retransmission (RTX) time instants, and the red dots represent the packet loss (PL) time instants. In the middle diagram, the blue curve represents the cart's position trajectory and the red curve represents the cart's velocity trajectory. In the bottom diagram, the blue curve represents the pendulum's pitch angle trajectory and the red curve represents the pendulum's pitch rate trajectory.

[13] A. Chiuso, N. Laurenti, L. Schenato, and A. Zanella, "LQG-like control of scalar systems over communication channels: The role of data losses, delays and SNR limitations," *Automatica*, vol. 50, no. 12, pp. 3155–3163, 2014.

[14] S. Dey, A. Chiuso, and L. Schenato, "Feedback control over lossy SNR-limited channels: Linear encoder–decoder–controller design," *IEEE Trans. on Automatic Control*, vol. 62, no. 6, pp. 3054–3061, 2017.

[15] J. Wu, G. Shi, B. D. Anderson, and K. H. Johansson, "Kalman filtering over Gilbert–Elliott channels: Stability conditions and critical curve," *IEEE Trans. on Automatic Control*, vol. 63, no. 4, pp. 1003–1017, 2017.

[16] Y. Mo, E. Garone, and B. Sinopoli, "LQG control with Markovian packet loss," in *Proc. European Control Conf.*, pp. 2380–2385, 2013.

[17] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, 2007.

[18] 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.

[19] D. E. Quevedo, A. Ahlén, and K. H. Johansson, "State estimation over sensor networks with correlated wireless fading channels," *IEEE Trans. on Automatic Control*, vol. 58, no. 3, pp. 581–593, 2013.

[20] N. Elia, "Remote stabilization over fading channels," *Systems & Control Letters*, vol. 54, no. 3, pp. 237–249, 2005.

[21] R. Parseh and K. Kansanen, "On estimation error outage for scalar Gauss–Markov signals sent over fading channels," *IEEE Trans. on Signal Processing*, vol. 62, no. 23, pp. 6225–6234, 2014.

[22] V. Gupta, A. F. Dana, J. P. Hespanha, R. M. Murray, and B. Hassibi, "Data transmission over networks for estimation and control," *IEEE Trans. on Automatic Control*, vol. 54, no. 8, pp. 1807–1819, 2009.

[23] P. Parag, A. Taghavi, and J.-F. Chamberland, "On real-time status updates over symbol erasure channels," in *Proc. IEEE Wireless Communications and Networking Conf.*, pp. 1–6, 2017.

[24] H. Sac, T. Bacinoglu, E. Uysal-Biyikoglu, and G. Durisi, "Age-optimal channel coding blocklength for an M/G/1 queue with HARQ," in *Proc. Int. Workshop on Signal Processing Advances in Wireless Communications*, pp. 1–5, 2018.

[25] A. Li, S. Wu, J. Jiao, N. Zhang, and Q. Zhang, "Age of information with

hybrid-ARQ: A unified explicit result," *IEEE Trans. on Communications*, vol. 70, no. 12, pp. 7899–7914, 2022.

[26] R. D. Yates, E. Najm, E. Soljanin, and J. Zhong, "Timely updates over an erasure channel," in *Proc. IEEE Int. Symp. on Information Theory*, pp. 316–320, 2017.

[27] E. Najm, R. Yates, and E. Soljanin, "Status updates through M/G/1/1 queues with HARQ," in *Proc. IEEE Int. Symp. on Information Theory*, pp. 131–135, 2017.

[28] Y. Wang, S. Wu, J. Jiao, W. Wu, Y. Wang, and Q. Zhang, "Age-optimal transmission policy with HARQ for freshness-critical vehicular status updates in space–air–ground-integrated networks," *IEEE Internet of Things Journal*, vol. 9, no. 8, pp. 5719–5729, 2020.

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

[30] E. T. Ceran, D. Gündüz, and A. György, "A reinforcement learning approach to age of information in multi-user networks with HARQ," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1412–1426, 2021.

[31] S. Banerjee, S. Ulukus, and A. Ephremides, "To re-transmit or not to re-transmit for freshness," in *Proc. IEEE Symp. on Model. and Opt. in Mobile, Ad Hoc, and Wireless Net.*, pp. 516–523, 2023.

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

[33] F. Nadeem, Y. Li, B. Vucetic, and M. Shirvanimoghaddam, "Real-time wireless control with non-orthogonal HARQ," in *Proc. Globecom Workshops*, pp. 395–400, 2022.

[34] Z. An, S. Wu, Y. Wang, J. Jiao, and Q. Zhang, "HARQ based joint uplink-downlink optimal scheduling strategy for single-loop WNCS," in *Proc. IEEE Int. Conf. on Wireless Communications and Signal Processing*, pp. 7–12, 2020.

[35] T. Soleymani, J. S. Baras, and K. H. Johansson, "Relation between value and age of information in feedback control," *Age of Information: Foundations and Applications, Cambridge University Press*, pp. 286–296, 2023.

[36] T. Soleymani, J. S. Baras, S. Hirche, and K. H. Johansson, "Value of information in feedback control: Global optimality," *IEEE Trans. on Automatic Control*, vol. 68, no. 6, pp. 3641–3647, 2023.

[37] T. Soleymani, J. S. Baras, and S. Hirche, "Value of information in feedback control: Quantification," *IEEE Trans. on Automatic Control*, vol. 67, no. 7, pp. 3730–3737, 2022.

[38] T. Soleymani, *Value of Information Analysis in Feedback Control*. PhD thesis, Technical University of Munich, 2019.

[39] T. Soleymani, S. Hirche, and J. S. Baras, "Optimal self-driven sampling for estimation based on value of information," in *Proc. Int. Workshop on Discrete Event Systems*, pp. 183–188, 2016.

[40] T. Soleymani, S. Hirche, and J. S. Baras, "Optimal stationary self-triggered sampling for estimation," in *Proc. IEEE Conf. on Decision and Control*, pp. 3084–3089, 2016.

[41] E. Uysal, O. Kaya, A. Ephremides, J. Gross, M. Codreanu, P. Popovski, M. Assaad, G. Liva, A. Munari, B. Soret, T. Soleymani, and K. H. Johansson, "Semantic communications in networked systems: A data significance perspective," *IEEE Network*, vol. 36, no. 4, pp. 233–240, 2022.

[42] D. Gündüz, Z. Qin, I. E. Aguerri, H. S. Dhillon, Z. Yang, A. Yener, K. K. Wong, and C.-B. Chae, "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 1, pp. 5–41, 2023.

[43] A. Khina, V. Kostina, A. Khisti, and B. Hassibi, "Tracking and control of Gauss–Markov processes over packet-drop channels with acknowledgments," *IEEE Trans. on Control of Network Systems*, vol. 6, no. 2, pp. 549–560, 2018.

[44] T. Soleymani, J. S. Baras, and K. H. Johansson, "State estimation over delayed and lossy channels: An encoder-decoder synthesis," *IEEE Trans. on Automatic Control*, vol. 69, no. 3, pp. 1568–1583, 2024.

[45] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1. Athena Scientific, 1995.

[46] R. F. Stengel, *Optimal Control and Estimation*. Courier Corporation, 1994.

[47] V. Kostina and B. Hassibi, "Rate-cost tradeoffs in control," *IEEE Trans. on Automatic Control*, vol. 64, no. 11, pp. 4525–4540, 2019.

[48] L. Schenato, "To zero or to hold control inputs with lossy links?," *IEEE Trans. on Automatic Control*, vol. 54, no. 5, pp. 1093–1099, 2009.

Touraj Soleymani received his B.S. and M.S. degrees both in aeronautical engineering from Sharif University of Technology, Iran, in 2008 and 2011, respectively, and his Ph.D. degree in electrical and computer engineering from the Technical University of Munich, Germany, in 2019. He is currently a research associate at the Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom. He was a research associate at the School of Electrical Engineering and Computer Science, Royal Institute of Technology, Sweden, from 2019 to 2022; and was a research scholar at the Institute of Artificial Intelligence, University of Brussels, Belgium, from 2012 to 2014, at the Institute for Advanced Study, Technical University of Munich, Germany, from 2014 to 2017, and at the School of Electrical Engineering and Computer Science, Royal Institute of Technology, Sweden, from 2017 to 2019. His research interests include control, learning, communication, optimization, game theory, multi-agent systems, cyber-physical systems, and swarm robotics systems. He received the Outstanding Student Award from Sharif University of Technology in 2011, the Best Paper Award at the International Conference on Intelligent Autonomous Systems in 2014, and the Best Paper Award Finalist at the International Workshop on Discrete Event Systems in 2018.

John S. Baras received the Diploma degree in electrical and mechanical engineering from the National Technical University of Athens, Athens, Greece, in 1970, and the M.S. and Ph.D. degrees in applied mathematics from Harvard University, Cambridge, MA, USA, in 1971 and 1973, respectively. He is a Distinguished University Professor and holds the Lockheed Martin Chair in Systems Engineering, with the Department of Electrical and Computer Engineering and the Institute for Systems Research (ISR), at the University of Maryland College Park. From 1985 to 1991, he was the Founding Director of the ISR. Since 1992, he has been the Director of the Maryland Center for Hybrid Networks (HYNET), which he co-founded. His research interests include systems and control, optimization, communication networks, applied mathematics, machine learning, artificial intelligence, signal processing, robotics, computing systems, security, trust, systems biology, healthcare systems, model-based systems engineering. Dr. Baras is a Fellow of IEEE (Life), SIAM, AAAS, NAI, IFAC, AMS, AIAA, Member of the National Academy of Inventors and a Foreign Member of the Royal Swedish Academy of Engineering Sciences. Major honors include the 1980 George Axelby Award from the IEEE Control Systems Society, the 2006 Leonard Abraham Prize from the IEEE Communications Society, the 2017 IEEE Simon Ramo Medal, the 2017 AACC Richard E. Bellman Control Heritage Award, the 2018 AIAA Aerospace Communications Award. In 2016 he was inducted in the A. J. Clark School of Engineering Innovation Hall of Fame. In 2018 he was awarded a Doctorate Honoris Causa by his alma mater the National Technical University of Athens, Greece.

Deniz Gündüz received the B.S. degree in electrical and electronics engineering from Middle East Technical University (METU), Turkey, in 2002, and the M.S. and Ph.D. degrees in electrical engineering from the NYU Tandon School of Engineering (formerly Polytechnic University), USA, in 2004 and 2007, respectively. After his Ph.D. degree, he was a Post-Doctoral Research Associate at Princeton University, a Consulting Assistant Professor at Stanford University, and a Research Associate at CTTC, Barcelona, Spain. He also held visiting positions at the University of Padova from 2018 to 2020 and Princeton University from 2009 to 2012. In September 2012, he joined the Department of Electrical and Electronic Engineering, Imperial College London, UK, where he is currently a Professor of information processing. He also serves as the Deputy Head of the Intelligent Systems and Networks Group and a part-time Faculty Member at the University of Modena and Reggio Emilia, Italy. His research interests include communications and information theory, machine learning, and privacy. He is a Distinguished Lecturer of the IEEE Information Theory Society from 2020 to 2022. He was a recipient of the Consolidator in 2022 and Starting in 2016 Grants of the European Research Council (ERC), the IEEE Communications Society-Communication Theory Technical Committee (CTTC) Early Achievement Award in 2017, and several best paper awards. He is an Area Editor of the IEEE Transactions on Information Theory, the IEEE Transactions on Communications, and the IEEE Journal on Selected Areas in Communications (JSAC), Special Series on Machine Learning in Communications and Networks. He also serves as an Editor for the IEEE Transactions on Wireless Communications.