Optimal Power Management in Wireless Control Systems

Konstantinos Gatsis, Alejandro Ribeiro and George J. Pappas

***Abstract*—This paper considers the control of a linear plant when plant state information is being transmitted from a sensor to the controller over a wireless fading channel. The power allocated to these transmissions determines the probability of successful packet reception and is allowed to adapt online to both channel conditions and plant state in order to conserve sensor's energy resources. The goal is to design plant input and transmit power management policies that minimize an infinite horizon cost combining power expenses and the conventional linear quadratic regulator control cost. A restricted information structure is identified allowing the separate designs of plant inputs and transmit powers. After the separation the optimal plant control policy is shown to be the standard LQR controller. The optimal communication policy follows from a Markov decision process problem minimizing the transmit power at the sensor and the state estimation error at the controller. The qualitative features of the optimal power adaptation to channel and plant are examined for general forward error correcting codes. In the particular case of capacity achieving codes conventional event-triggered policies are recovered, where the decision is whether to transmit or not. Approximate dynamic programming is employed to derive a family of tractable suboptimal communication policies, exhibiting the same qualitative features as the optimal one. The performance of our suboptimal policies is shown in simulations and the advantages are contrasted to other simple transmission policies.**

I. Introduction

The networked control systems studied in this paper are characterized by the separation of sensing and actuation in different physical devices with control loops involving the communication of plant state information over a wireless channel. When sensor and controller communicate over a wireless channel the cost of controlling the plant gets mixed with the cost of sending plant state information from the sensor to the controller. The more information the sensor conveys the more precise actuation becomes, but the resulting increase in power consumption at the sensor leads to rapid depletion of its energy resources. It is therefore apparent that a tradeoff emerges between plant performance and power consumption. To quantify this tradeoff we study the problem of selecting plant inputs and power management policies that minimize a joint cost that accounts for the plant regulation cost and the cost of conveying information from the sensor to the controller.

A. Related literature

Early works on networked control systems ignore the cost of conveying information and focus their analysis on the performance of control loops when various communication effects are taken into account, see e.g., [2]–[5] and references therein. These works examine packet-based communication over analog erasure channels, analyzing necessary and sufficient requirements for stability and designing controllers and estimators to counteract random packet drops and delays. Alternatively, analog channels can be modeled as input-output systems and channel randomness can be treated as stochastic model uncertainty [6]. Controllers under the latter framework can be synthesized using robust control techniques to handle additive noise channels with signal-to-noise ratio (SNR) constraints [7]. A different set of issues arise when loops close over digital channels with data-rate constraints. In such case quantization effects become important and apart from the controller design an efficient encoding/decoding scheme is required [8]. Fundamental limits like the minimum bit rate for stabilization are also known; see, e.g., [9].

In other networked control architectures communication is not treated as a limitation but becomes an active part of the design. Typically, these setups depart from the classic periodic communication paradigm, leading to frameworks such as event-triggered sampling [10], [11], control [12], [13] or self-triggered control [14]. The underlying concept in these contributions is to prolong the time elapsed between successive transmissions or input updates as long as some Lyapunov-like plant performance criterion is satisfied. Such schemes exhibit in general an average communication/update rate lower than periodic schemes that attain similar plant performance. However, communication costs are not explicitly accounted for in the triggering design.

Communication costs are explicitly modeled in the context of remote state estimation in [15]–[18]. In this framework a sensor measuring the plant state decides whether to transmit its value to an estimator or not and each transmission incurs a fixed cost. The overall goal is to minimize the estimation error cost and the communication penalties aggregated over time. The optimal communication is event-triggered [15], similar to, e.g., [12], [13], meaning that transmissions are triggered when the estimation error exceeds a threshold. Computing the optimal transmission-triggering sets is not tractable, motivating the development of suboptimal event-triggered schemes [17], [18]. Related contributions consider plant and communication controllers jointly optimal with respect to a linear quadratic and communication cost assuming again a fixed cost per transmission [19], [20]. The problem turns out to be more complex than the case of simple state estimation but a separation principle can be established [20], and optimal inputs and schedules can be found by dynamic programming for

Work in this paper is supported by NSF CNS-0931239. The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, 200 South 33rd Street, Philadelphia, PA 19104. Email: {kgatsis, aribeiro, pappasg}@seas.upenn.edu. Part of the results in this paper were submitted to the 2013 American Control Conference [1].

a finite horizon. The characterization and determination of jointly optimal plant and transmission policies in this context is otherwise open. We note however that in the case of control over digital channels some aspects of jointly optimal encoder and controller design have recently been studied [21].

B. Contributions and summary

Instead of accounting for communication cost in terms of transmissions, in this paper we are interested in the allocation of the actual resource used by the wireless sensors to communicate, namely, transmitted power. This perspective permits the incorporation of fading effects in the wireless channel and provides the flexibility in power allocation to protect some transmissions more than others. Fading refers to large unpredictable variations in wireless channel transferences whose mitigation involves extensive use of power adaptation to channel conditions [22, Chapters 3,4]. Besides counteracting fading effects power adaption may be also helpful in closed-loop control to, e.g., increase the likelihood of successful packet decoding when the plant state deviates from target. In this paper transmit powers are allowed to adapt to both, the fading channel realization and the plant state. The allocated power and the realization of the fading channel determine the likelihood of successful packet decoding at the receiver by a known complementary error function (Section II). This communication model has been used for state estimation in sensor networks [23], [24] and can be regarded as a generalization of the erasure channel with i.i.d. dropouts of, e.g., [2]–[5], since here the probability of packet drops is actively controlled by an online transmit power adaptation policy.

Given our general communication model, we are interested in the trade-off between closed-loop plant performance and power resources. To this end we combine the transmit power with a conventional linear quadratic regulator (LQR) cost to form an aggregate infinite horizon cost that we seek to minimize through proper joint selection of plant and power control policies (Section II-A). For the novel problem formulation proposed we begin by identifying restricted information structures that permit decoupling of plant input and power control policies (Section II-B). For this particular information structure the usual LQR control law becomes optimal at the controller side while the optimal communication policy at the transmitter follows from a Markov decision process (MDP) formulation accounting for transmit power and the state estimation error at the controller (Section III). The optimal power control policy is then expressed in terms of a value function solving the MDP problem (Section IV). While this does not allow computation of optimal policies it does allow us to understand the qualitative characteristics of the optimal resource allocation.

In contrast to the work in, e.g., [15]–[17], [19] where transmission is based just on plant state, the availability of channel state information at the transmitter leads to new insights on the optimal communication policy. In particular when channel gain is low or estimation error small no transmission is triggered, since it would be costly or unnecessary respectively. On the other hand, similar to the above work, there is an event/set of plant and channel states where transmission is triggered, but the optimal power allocation on this event still needs to adapt to the channel and plant states. Alternatively, our power management policy can be viewed as a 'soft' version of the event-triggered paradigm of, e.g., [12], as instead of just deciding whether to transmit or not we select how much power to allocate to the transmission attempt. This interpretation is further fostered by the realization that conventional event-triggered policies emerge as the optimal communication when the sensor uses capacity achieving forward error correcting (FEC) codes (Section IV-A).

Finally since optimal communication is not computationally tractable we devise suboptimal power control policies using approximate dynamic programming, in particular rollout algorithms (Section V). These policies maintain the same qualitative characteristics as the optimal policies for general error correcting or capacity achieving codes, and this is verified in numerical simulations (Section VI). Our rollout policies are shown to have significant performance benefits compared to other simple policies that adapt only to channel conditions and not the plant state, such as the ones proposed in [23], [24]. We close the paper with conclusions and suggestions on how the adopted model can be used for more complex wireless communication/control design problems in future work (Section VII).

Notation: Let $\mathcal{N}_{\mu,\Sigma}$ denote the n-dimensional Gaussian distribution with mean μ and covariance Σ. For a square matrix $M \in \mathbb{R}^{n\times n}$ let $\lambda_{\max}(M), \lambda_{\min}(M)$ denote respectively the largest and smallest eigenvalues in magnitude. For compactness a set of variables $\{x_k, x_{k+1}, \ldots, x_{k+t}\}$ is denoted by $x_{k:k+t}$. Subscripts of variables as in x_k, x_{k+1} denote discrete time. When time index k is clear from the context, subscripts are omitted and the respective variables are denoted as x, x_+.

II. Problem Formulation

We consider the architecture shown in Fig. 1 deployed to control a discrete-time linear time-invariant plant described by the difference equation

$$x_{k+1} = Ax_k + Bu_k + w_k,\ k \geq 0, \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the plant's state with x_0 given, $u_k \in \mathbb{R}^m$ the driving input, and $\{w_k, k \geq 0\}$ is the process noise composed of independent identically distributed (i.i.d) n-dimensional Gaussian random variables $w_k \sim \mathcal{N}_{0,W}$ with zero mean and covariance W. We assume the plant is unstable ($\lambda_{\max}(A) > 1$) but that (A, B) is stabilizable.

The wireless control system considered in this paper includes a sensor/transmitter collecting state measurements x_k that it communicates with power $p_k \in [0, p_{\max}]$ over a wireless fading channel with coefficient h_k. At the other side of the channel the receiver/controller uses the received information to determine a control input u_k that it feedbacks into the plant. The effects of state quantization and transmission delays are considered negligible and are thus ignored henceforth.

Due to propagation effects the channel gain h_k changes unpredictably [22, Chapter 3]. We adopt the standard block fading model of wireless communications whereby channels

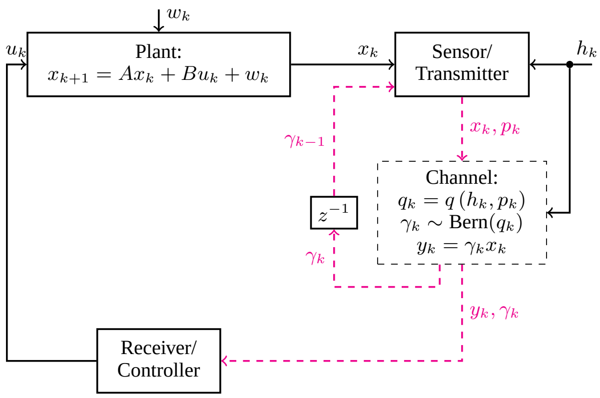

Fig. 1. Wireless control system architecture. A sensor measures the plant and wireless fading channel states x_k, h_k respectively and transmits with power p_k. Messages are successfully decoded at the controller with probability q_k that depends on the channel state h_k and the power p_k. The sensor receives acknowledgments with a one-step delay.

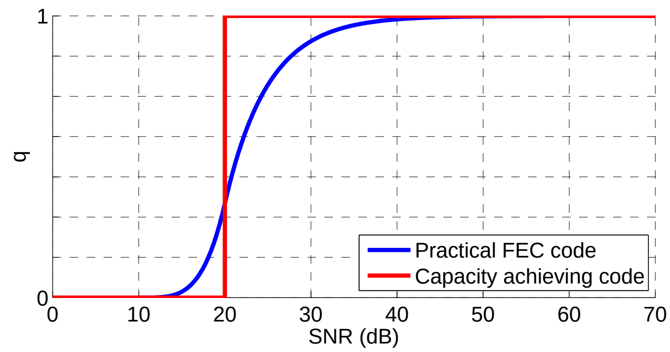

Fig. 2. Complementary error function for FEC and capacity achieving codes. The probability of successful decoding q for a practical FEC code is a sigmoid function of the received SNR $= h\,p/N_0$, while for a capacity achieving code a threshold value SNR$_0$ determines whether a packet is successfully received.

$\{h_k, k \geq 0\}$ are modeled as i.i.d. random variables taking values in the positive reals $\mathbb{R}_+$ according to some known distribution m_H and are independent of the plant process noise $\{w_k, k \geq 0\}$. We make the technical assumption that the distribution m_H of the channel state has a probability density function on $\mathbb{R}_+$. To allow for transmissions adapted to the current channel conditions the transmitter has access to the channel state information h_k before transmitting at time k – the development is equally valid if estimates are available in lieu of h_k as discussed in Remark 2.

At the controller side the received signal includes the information bearing signal and additive white Gaussian noise (AWGN). The noise power is denoted by N_0 and the power of the information bearing signal is the product $h_k\, p_k$. Assuming the receiver also has channel state information, successful decoding of the transmitted packet is determined by the signal to noise ratio (SNR) at the receiver defined as SNR$_k := h_k p_k / N_0$. More precisely, given the particular type of modulation and FEC code used, the SNR determines the probability of successful detection q_k. To keep the analysis general we define a generic complementary error function

$$q_k = q\,(h_k, p_k)\,, \tag{2}$$

mapping SNR$_k := h_k p_k / N_0$ to the probability q_k. We assume that $q\,(h, p)$ is a known increasing function of the product $h\,p$ - see Remark 1.

Considering packet decoding as a part of the communication process, we can model communication as a sequence of indicator variables γ_k taking value $\gamma_k = 1$ when information is successfully decoded and $\gamma_k = 0$ otherwise. Variables $\gamma_k \sim$ Bern(q_k) are Bernoulli distributed with time-varying success probabilities q_k. With this communication model the controller receives the output of the decoding process which we model by the signal $y_k = \gamma_k x_k$. We further assume that the controller also gets γ_k so that it can distinguish between the cases $x_k = 0$ and $\gamma_k = 0$. Packet receipt acknowledgment γ_k is also sent to the sensor as provided by 802.11 and TCP protocols. We assume lossless acknowledgments, so that the sensor knows what information is received at the controller.

The problem addressed in this paper is the joint design of the control inputs u_k and the transmit powers p_k. The control input u_k is determined by the received information $y_{0:k}$, $\gamma_{0:k}$. The power p_k is determined as a function of the plant state measurements $x_{0:k}$, the observed channel realizations $h_{0:k}$, and the controller acknowledgments $\gamma_{0:k}$. Informally, to conserve power at the sensor side we want to transmit information only when the state x_k deviates from its desired value or when the channel realization h_k is favorable. In the first case transmission is necessary to keep the plant under control. In the latter case the transmission cost is minimal. A formal problem specification is presented in the next section after the following remarks.

Remark 1. The error profiles $1 - q\,(h_k, p_k)$ of particular FEC codes are difficult to determine analytically but can be measured in actual or simulated experiments [25], [26]. The typical shape of $q\,(h_k, p_k)$ is a sigmoid function of $h_k\, p_k$ with exponential tails as depicted in Fig. 2. In the theoretical limit, correct decoding depends on the channel capacity $C_k = W \log_2(1 + \text{SNR}_k)$, where W is the channel bandwidth. If the packet is transmitted at a rate smaller than C_k bits per second it is almost surely successfully decoded, and it is almost surely incorrectly decoded otherwise. Thus, we can write the successful decoding probability as the indicator function

$$q\,(h_k, p_k) = \mathbb{I}\left(\frac{h_k p_k}{N_0} \geq \text{SNR}_0\right), \tag{3}$$

for some constant SNR$_0$. Determining the threshold SNR$_0$ requires specification of the sampling rate α and quantization resolution β of the state x_k. With α samples per second and β bits per sample we require a transmission rate of $\alpha\beta$ bits per second. The SNR threshold is then given by SNR$_0 = 2^{\alpha\beta/W} - 1$. Our interest in (3) is conceptual as it will allow us to recover results in event-triggered communication [15] as arising from the use of capacity achieving codes – see Section IV-A. The form of (3) is shown in Fig. 2.

Remark 2. The assumption that channel state information (CSI) is available at the transmitter is typical in modern wireless communication setups [22, Chapter 9]. To measure the wireless channel conditions a short pilot signal of fixed

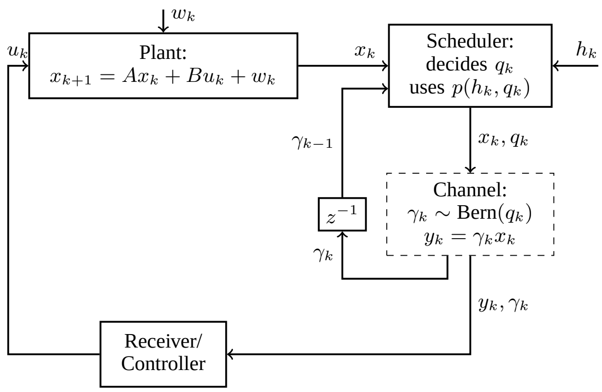

Fig. 3. Equivalent wireless control system architecture. A scheduler decides the successful decoding probability q_k and transmits the state measurement x_k with the required power $p_k = p(h_k, q_k)$. The controller receives the message with probability q_k.

power can be sent from the transmitter and then the fading characteristics can be estimated at the receiver and sent back to the transmitter by utilizing the reverse channel. Although accurate CSI is difficult to acquire at the transmitter side, our development is still valid if channel estimates are available in lieu of the actual channel value h_k. Reinterpreting h_k as an estimate of the fading coefficient the complementary error function $q(h, p)$ in (2) captures not only the success of decoding but also the uncertainty over the real channel gain. It suffices to integrate $q(h, p)$ with respect to the conditional distribution of the channel realization given the estimate.

Remark 3. There is a distinction to be made between errors that are detected by the receiver and errors that are undetected and may confuse the controller. The model here handles the former and ignores the latter. This is justified because practical communication schemes include the use of cyclic redundancy checks (CRC) for error detection that can drive the probability of undetected errors to very small values [27]. The use of simple CRCs reduces the probability of undetected errors to 10^{-3}, while longer codes can reduce this probability to 10^{-7}.

A. Joint design of plant and power control

To formulate the joint design of plant controller and power management we introduce an equivalent architecture. In view of (2), choosing p_k is equivalent to choosing the desired probability of successful decoding q_k at time k and transmitting with the minimum required power to achieve this q_k, namely

$$p_k = p(h_k, q_k) := \inf\left\{0 \leq p \leq p_{\max} : q(h_k, p) \geq q_k\right\}. \quad (4)$$

We can therefore interpret q_k as our decision variable with $p(h_k, q_k)$ denoting the cost of selecting transmission success probability q_k. This leads to the equivalent control system architecture shown in Fig. 3 where a scheduler block responsible for deciding q_k replaces the sensor/transmitter block of Fig 1. Our formulation generalizes the simple transmit-or-not decision as considered in, e.g., [15].

We note for future reference that the assumed monotonicity of the function $q(h, p)$ on the product $h\,p$ implies that the power function $p(h, q)$ is increasing in q and decreasing in h. Using maximum power $p_{\max}$, the transmitter can achieve a maximum successful decoding probability $q_{\max}(h) := q(h, p_{\max})$ for a given channel state h. Therefore, the decision variables q_k belong in the interval $[0, q_{\max}(h_k)]$. We also make the following assumptions.

Assumption 1. The maximum achievable successful decoding probability $q_{\max}(h)$ satisfies

$$\mathbb{E}_h q_{\max}(h) > q_{crit} := 1 - 1/\lambda_{\max}(A)^2, \quad (5)$$

where expectation is taken over the channel distribution m_H.

Assumption 2. For any channel realization h, the function $p(h, q)$ in (4) is continuous in the successful decoding probability variable q.

Assumption 1 is essentially a stability condition, stating that transmitter has enough power to keep the plant state bounded in second moment, as we discuss later after (17), and it will be used to establish our main Theorems 1 and 2. Assumption 2 is of a technical nature and will be used in Theorem 2.

In the architecture of Fig. 3 the communication decision q_k is chosen as a function of the information available at the sensor, while the plant control signal u_k is a function of the information available at the controller. These choices are in general allowed to be randomized. The sequence $\pi := \{q_0, q_1, \ldots\}$, or equivalently the power allocation $\{p_0, p_1, \ldots\}$, is termed the communication policy, whereas the sequence $\theta := \{u_0, u_1, \ldots\}$ denotes the control policy. With fixed policies π, θ, all random variables are defined on an appropriate product space and have a measure that we denote as $\mathbb{P}^{\pi,\theta}$. We use $\mathbb{E}^{\pi,\theta}$ to signify integration with respect to $\mathbb{P}^{\pi,\theta}$, which we simplify to $\mathbb{E}$ when not leading to confusion. We remark that sensor and controller know each other's policy.

The policy pair (π, θ) incurs a control cost and a communication cost. As a control cost we adopt the standard linear quadratic regulator cost

$$J^N_{\text{LQR}}(\pi, \theta) := \mathbb{E}^{\pi,\theta} \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k, \quad (6)$$

for some pair of matrices $R > 0$ and $Q \geq 0$, with $(A, Q^{1/2})$ detectable. The communication cost is given by the expected power consumption

$$J^N_{\text{PWR}}(\pi, \theta) := \mathbb{E}^{\pi,\theta} \sum_{k=0}^{N-1} p(h_k, q_k). \quad (7)$$

To quantify the tradeoff between plant performance and power consumption we combine the LQR cost in (6) and the power cost in (7) into the limit aggregate cost

$$J(\pi, \theta) := \limsup_{N\to\infty} 1/N \left[J^N_{\text{LQR}}(\pi, \theta) + \lambda J^N_{\text{PWR}}(\pi, \theta)\right], \quad (8)$$

for some positive constant $\lambda > 0$. Our goal is to design plant and power control policies θ and π respectively that minimize the joint cost (8). These policies depend on what information is available to the sensor and controller. The specific information structure considered in this paper is introduced in the following section.

B. Information structure

Denote as O_k the information known at the controller side at time k just before deciding the input u_k. This information includes the given initial plant state x_0, the history of decoding success variables $\gamma_{0:k}$ and the decoded signals $y_{0:k}$, as well as the previously chosen control inputs $u_{0:k-1}$, i.e.,

$$O_k := \{x_0, \gamma_{0:k}, y_{0:k}, u_{0:k-1}\}. \tag{9}$$

Then the control input u_k is chosen as a function of O_k, or more formally, measurable with respect to the σ-field generated by O_k.

Given the possibility of lost packets as indicated by $\gamma_k = 0$, the controller has partial information on the plant state x_k. It is then of importance to study the MMSE estimate $\mathbb{E}^{\pi,\theta}(x_k|O_k)$. This estimation is complicated by the fact that the event $\gamma_k = 0$ possibly contains information about the state x_k through the dependence of the probability q_k on the value of x_k – see Remark 4. To avoid this complication we discard the information given by events of the form $\gamma_k = 0$. Formally, define $\tau_k := \max\{0 \le l \le k : \gamma_l = 1\}$ as the time of the last successful transmission by time k and define the sequence of σ-fields

$$G_k := \{x_0, \gamma_{0:\tau_k}, y_{0:\tau_k}, u_{0:k-1}\}. \tag{10}$$

with $G_0 = \{x_0\}$. When $\gamma_k = 1$, G_k coincides with O_k. When $\gamma_k = 0$, G_k only contains information received till the last successful transmission which occurred at time $\tau_k < k$.

We restrict attention to control policies θ selecting inputs u_k as functions of G_k, possibly randomized, and denote the set of all such policies by Θ. Unlike $\mathbb{E}^{\pi,\theta}(x_k|O_k)$, the state MMSE estimate $\hat{x}_k := \mathbb{E}^{\pi,\theta}(x_k|G_k)$ with respect to G_k is easy to compute. When $\gamma_k = 1$ the state $x_k = y_k$ becomes known at the receiver side. When $\gamma_k = 0$ no new information becomes available and $\hat{x}_k$ is obtained by propagating $\hat{x}_{k-1}$ through the plant's dynamics in (1). Putting these two cases together yields

$$\hat{x}_k := \mathbb{E}^{\pi,\theta}(x_k|G_k) = \begin{cases} y_k & \text{if } \gamma_k = 1, \\ A\hat{x}_{k-1} + Bu_{k-1} & \text{if } \gamma_k = 0 \end{cases}, \tag{11}$$

with $\hat{x}_0 = x_0$ since the initial state is given.

At the other side of the link at time k the sensor/transmitter has access to the channel realization h_k and the plant state x_k which allows selection of the successful transmission probability q_k to depend on the values of both of h_k, x_k. This affects the controller design however, because when the controller decides u_{k-1} to control x_k, it should consider the indirect effect on q_k. This information structure renders the joint communication and control co-design problem hard to analyze. To overcome this, we restrict transmission policies to depend on the channel state h_k and the information about plant state x_k that the controller does not know. More precisely consider the difference between the sensor measurement x_k and the controller's estimate $\hat{x}_k$ by (11) if the kth packet is *not* successfully decoded, that is

$$\varepsilon_k := x_k - (A\hat{x}_{k-1} + Bu_{k-1}), \tag{12}$$

with $\varepsilon_0 := 0$. Observe that the term in the parenthesis is known to the sensor since by the acknowledgment mechanism the controller's previous estimate $\hat{x}_{k-1}$ and input u_{k-1} can be replicated at the sensor. Alternatively the terms ε_k can be viewed as the innovations of the controller's estimate (11) when a new message is received.

We restrict then information at the sensor side to the set F_k defined as a collection of the channel history $h_{0:k}$, the history of innovations $\varepsilon_{0:k}$, and past decisions $q_{0:k-1}$, i.e.,

$$F_k := \{\varepsilon_{0:k}, h_{0:k}, q_{0:k-1}\}. \tag{13}$$

Let us also add a technical requirement that the sensor selects maximum transmit power $p_{\max}$ when the innovation ε_k gets too large and the channel h_k is favorable. In particular consider a positive constant $L > 0$, and a threshold on channel values $h_t \in \mathbb{R}_+$ where a positive success probability can be achieved $q_{\max}(h_t) > 0$ that also satisfies

$$\int_{h \ge h_t} q_{\max}(h)\, dm_H(h) > q_{crit}, \tag{14}$$

with q_{crit} given in (5). Such a channel threshold exists by Assumption 1. We consider then communication policies π selecting decoding success q_k as functions of F_k for each k, possibly randomized, and also satisfying $q_k \in Q(\varepsilon_k, h_k)$ where

$$Q(\varepsilon, h) := \begin{cases} q_{max}(h) & \text{if } \|\varepsilon\| \ge L \text{ and } h \ge h_t \\ [0, q_{\max}(h)] & \text{otherwise} \end{cases}. \tag{15}$$

We denote the set of all such policies with Π. The technical power saturation requirement is inconsequential as we may pick L arbitrarily large, and will be used to prove Proposition 2 and Theorem 2 in the sequel. Similar requirements have been introduced in [15], [16], however our setup is further complicated by the availability of the random channel states.

The proposed information structure is depicted in Fig. 4. The sensor block is split into a pre-processor and a scheduler. The pre-processor based on the sample x_k and the acknowledgment γ_{k-1} computes and feeds ε_k to the scheduler who, upon measuring the channel h_k decides the transmission success probability q_k while incurring power cost $p(h_k, q_k)$. Our goal in this paper is to study policies $\pi \in \Pi$ and $\theta \in \Theta$ that are optimal with respect to the joint objective (8), that is

$$\underset{\pi \in \Pi, \theta \in \Theta}{\text{minimize}} \quad J(\pi, \theta). \tag{16}$$

In particular, the next section shows that the information structure we introduced allows a separate design of the optimal communication and control policies. We then leverage this result to study optimal communication policies in Section IV and to develop tractable suboptimal policies in Section V.

Remark 4. If the controller uses the complete information O_k to estimate x_k, the optimal plant estimate is not $\hat{x}_k$ given by (11) anymore. When a sequence of packet drops $\gamma_k = \gamma_{k-1} = \ldots = 0$ is observed, and since the communication policy is known, the controller should consider the possibility that the sensor did not actually transmit anything, and this could in general give indirect information about the expected value of x_k. This issue is further discussed in [19], [20]. We note that the restriction to G_k in (10) is not necessary for the separation results of Section III, but it is necessary to obtain

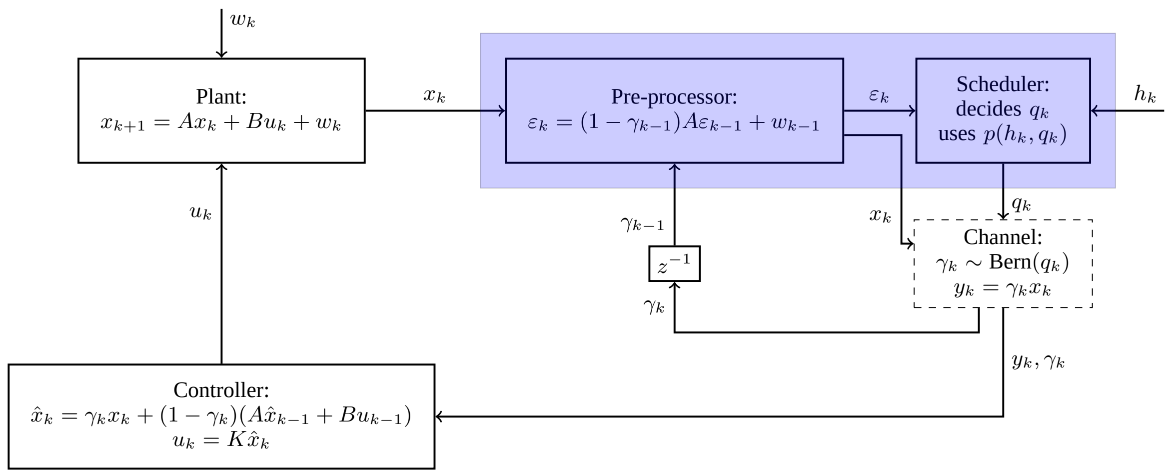

Fig. 4. Wireless control system with a restricted information structure. The sensor consists of two blocks. A pre-processor computes the error ε_k given the measurement x_k and the acknowledgment γ_{k-1}. A scheduler decides q_k based on ε_k and the channel state h_k, and transmits x_k with the required power $p_k = p(h_k, q_k)$. The controller receives the message with probability q_k, computes the state estimate $\hat{x}_k$ and provides input u_k to the plant.

linear dynamics of the estimation error e_k and the related ε_k as described next by (17) and (18) respectively.

III. Separation of designs

In this section we show that with the imposed restrictions on the information available at sensor and controller the control law $\theta \in \Theta$ and the communication policy $\pi \in \Pi$ can be designed separately. In particular the control policy has no effect on the estimation process at the receiver and by utilizing a separation principle the optimal controller becomes the standard linear quadratic one.

Let us denote the difference between the plant state and the estimate kept at the controller by $e_k := x_k - \hat{x}_k$ and its covariance as seen at the controller by $\Sigma_k := \mathbb{E}^\pi \left[e_k e_k^T \,\middle|\, G_k \right]$. The estimation error dynamics can be found by subtracting (11) from the system dynamics (1) to get

$$e_k = (1-\gamma_k)(Ae_{k-1} + w_{k-1}), \tag{17}$$

with $e_0 = 0$ since x_0 is given. Stabilizability of estimation error is guaranteed by Assumption 1. Indeed if transmitter were to use maximum power all the time the dynamics in (17) become a jump linear system since γ_k are Bernoulli with constant probability equal to the left hand side of (5). Then condition (5) is sufficient for bounded second moment as, e.g., in [2, Theorem 2]. It is also tight in the sense that estimation error becomes unstable if $\mathbb{E}_h q_{\max}(h) < q_{crit}$.

Turning our attention to the innovation substituting x_k by (1) in the definition of ε_k in (12) gives $\varepsilon_k = Ae_{k-1} + w_{k-1}$. The term e_{k-1} equals $(1-\gamma_{k-1})\varepsilon_{k-1}$ as seen by (17), therefore ε_k evolves according to

$$\varepsilon_k = (1-\gamma_{k-1})A\varepsilon_{k-1} + w_{k-1}, \tag{18}$$

with initial value $\varepsilon_0 = 0$. The following proposition establishes a separation principle in our restricted information structure setup, stating that the control action has no effect on the quality of the future estimates at the controller.

Proposition 1. *Consider any communication policy π selecting successful decoding probabilities q_k as functions of F_k given in (13), possible randomized, with ε_k defined in (12) and channel states h_k independently drawn from a distribution m_H. Then at any step k the distributions of the future processes $\{\varepsilon_\ell, q_\ell, \gamma_\ell, e_\ell, \ell > k\}$ given G_k do not depend on the chosen control policy $\theta \in \Theta$.*

Proof: First note that the processes $\{w_k, h_k, k \geq 0\}$ are by assumption independent of any other process. Then we follow an induction argument to prove the claim. At $k = 0$, ε_0 is equal to 0, q_0 depends only on h_0 and ε_0, γ_0 is an independent Bernoulli with success q_0, and e_0 is also 0 since x_0 is initially known. Consider then a time k with a given G_k, the corresponding estimation error e_k given G_k having zero mean and covariance Σ_k, and a control input u_k that is a function of G_k as described by the control policy θ. The term ε_{k+1} equals $Ae_k + w_k$, as indicated by the arguments preceding (18), which given G_k has mean 0 and covariance $A\Sigma_k A^T + W$. The choice $q_{k+1} \in F_{k+1}$ by construction depends on past variables in F_k which by causality do not depend on the action u_k, as well as the new variables $\varepsilon_{k+1}, h_{k+1}$ which are also independent of u_k. Also the distribution of $\gamma_{k+1} \sim$ Bern(q_{k+1}) only depends on the distribution of q_{k+1}, and the same holds for e_{k+1} which equals $(1-\gamma_{k+1})\varepsilon_{k+1}$ again by the arguments preceding (18). To sum up all variables $\varepsilon_{k+1}, q_{k+1}, \gamma_{k+1}, e_{k+1}$ given G_k do not depend on u_k. ■

The intuition behind this proposition is that the effect of control inputs is subtracted from x_k when forming the innovation terms ε_k in (12) that are fed to the communication policy π. Similar separation results based on innovation terms have been utilized in setups where the sensor just decides whether to transmit or not [19], even though this need not be optimal [20], as well as in encoder/decoder design for digital channels [9]. The above proposition restates the separation principle for our problem of power selection in the presence of channel state information in addition to plant measurements.

Since the power cost $J^N_{\text{PWR}}(\pi,\theta)$ in (7) only depends on pairs (q_k,h_k), the above proposition shows that the control policy θ has no effect on the power cost. Thus we can rewrite the objective in (8) as

$$J(\pi,\theta)=\limsup_{N\to\infty}\frac{1}{N}J^N_{\text{LQR}}(\pi,\theta)+\lambda\limsup_{N\to\infty}\frac{1}{N}J^N_{\text{PWR}}(\pi). \tag{19}$$

This means that the optimal control policy $\theta\in\Theta$ for a given communication policy $\pi\in\Pi$ is the one minimizing the limit LQR cost. It turns out that the above proposition can help establish a stronger result, that the form of the optimal controller does not depend on the communication policy, leading to a stronger separability than what follows from (19).

Indeed for any finite horizon by the above separation principle standard dynamic programming arguments show that the optimal control law is given by the standard LQR one, and this has been shown when e.g. the sensor just decides whether to transmit or not [19], [20]. We are interested however in the infinite horizon problem. Formally the controller has partial state information. However the setup differs from, e.g., the standard problem of Gaussian observation noise and Kalman filtering, where the estimation error covariance Σ_k is shown to converge to some limit and the system is assumed to start at time $k=0$ with this limit estimation error. In our setup whenever a packet is successfully decoded the estimation error is reset to zero, otherwise it grows, so for the general communication policies $\pi\in\Pi$ under consideration it is not clear whether some limit covariance exists. Alternatively the following proposition shows that estimation errors admit a uniform bound in second moment.

Proposition 2. *Suppose Assumption 1 holds. Then there exists a finite positive constant M such that for any communication policy $\pi\in\Pi$ selecting successful decoding probabilities q_k with respect to F_k given in (13), possibly randomized, satisfying the additional restriction $q_k\in Q(\varepsilon_k,h_k)$ given by (15), and for every $k=0,1,\ldots$, it holds that*

$$\mathbb{E}^\pi e_k^Te_k\le M. \tag{20}$$

Proof: See Appendix A ■

With this bound on expected magnitude of estimation error established, uniform over k and over any policy $\pi\in\Pi$, the following theorem shows that the optimal control law for the average infinite horizon case is indeed the standard steady-state LQR one.

Theorem 1 (Optimal control policy). *Consider the wireless control system of Fig. 4 with any communication policy $\pi:=\{q_0,q_1,\ldots\}\in\Pi$ selecting successful decoding probabilities q_k as functions of F_k given in (13), possibly randomized, with innovation terms ε_k as defined in (12) and channel states h_k independently drawn from a distribution m_H, satisfying the additional restriction $q_k\in Q(\varepsilon_k,h_k)$ given by (15). Suppose Assumption 1 holds. Then for any control policy $\theta:=\{u_0,u_1,\ldots\}\in\Theta$ composed of inputs u_k as possibly randomized functions of G_k in (10) such that*

$$\lim_{N\to\infty}1/N\ \mathbb{E}^{\pi,\theta}x_N^Tx_N=0, \tag{21}$$

the joint objective $J(\pi,\theta)$ described by (6) - (8) satisfies

$$J(\pi,\theta)\ge Tr(PW)+\limsup_{N\to\infty}\frac{1}{N}\mathbb{E}^\pi\sum_{k=0}^{N-1}e_k^T\tilde{P}e_k+\lambda p(h_k,q_k) \tag{22}$$

where P is the solution to the standard algebraic Riccati equation $P=A^TPA+Q-A^TPB(R+B^TPB)^{-1}B^TPA$ for the system in (1) and the linear quadratic regulator cost (LQR) in (6), and the matrix $\tilde{P}$ is defined as

$$\tilde{P}:=A^TPA+Q-P. \tag{23}$$

Moreover, the minimum value in (22) is achieved for the control policy

$$u_k=K\hat{x}_k, \tag{24}$$

with $\hat{x}_k$ the state estimate described in (11) and the steady state LQR gain $K:=-(R+B^TPB)^{-1}B^TPA$.

Proof: See Appendix B ■

The statement of Theorem 1 determines the optimal control policy θ as the conventional LQR controller in (24), which is shown in Fig. 4. The optimal cost given in (22) equals a constant $Tr(PW)$ and a limit average sum term that only depends on the communication policy $\pi\in\Pi$. This result shows that the optimal communication policy needs to jointly regulate the power consumption at the sensor and the weighted estimation error $e_k^T\tilde{P}e_k$ at the controller side.

Observe that as per (11) and (17) it holds that $e_k=(1-\gamma_k)\varepsilon_k$. Also $\mathbb{E}^\pi[\gamma_k|F_k]=\mathbb{P}^\pi[\gamma_k=1|F_k]=q_k$ and $\varepsilon_k\in F_k$. So we can write

$$\mathbb{E}^\pi[e_k^T\tilde{P}e_k|F_k]=\mathbb{E}^\pi[(1-\gamma_k)\varepsilon_k^T\tilde{P}\varepsilon_k|F_k]=(1-q_k)\varepsilon_k^T\tilde{P}\varepsilon_k, \tag{25}$$

and taking the expectation in both sides gives

$$\mathbb{E}^\pi[e_k^T\tilde{P}e_k]=\mathbb{E}^\pi[(1-q_k)\varepsilon_k^T\tilde{P}\varepsilon_k]. \tag{26}$$

Substituting the expression (26) into the second summand of (22) it follows that the optimal communication policy $\pi\in\Pi$ is the one that achieves the infimum cost

$$J^*_{\text{COMM}}:=\inf_{\pi\in\Pi}\lim_{N\to\infty}\frac{1}{N}\mathbb{E}^\pi\sum_{k=0}^{N-1}c(\varepsilon_k,h_k,q_k), \tag{27}$$

where we define the cost-per-stage to be

$$c(\varepsilon,h,q):=(1-q)\varepsilon^T\tilde{P}\varepsilon+\lambda p(h,q). \tag{28}$$

The difference between the sum in (22) and the objective in (27) is that in the former e_k is not known at the sensor at time k, while ε_k in the latter is. This way (27) takes the form of a Markov decision process (MDP) problem with an infinite horizon average cost criterion. The state of the problem at time k is the pair $(\varepsilon_k,h_k)\in\mathbb{R}^n\times\mathbb{R}_+$ and the available action is $q_k\in Q(\varepsilon_k,h_k)$ by (15). The state transition probabilities can be obtained from (18) and are given by

$$\begin{aligned}&\mathbb{P}(\varepsilon^+,h^+|\varepsilon,h,q)\\&\quad=\left[q\,\mathcal{N}_{0,W}(\varepsilon^+)+(1-q)\,\mathcal{N}_{A\varepsilon,W}(\varepsilon^+)\right]\ m_H(h^+).\end{aligned} \tag{29}$$

Here ε,h and ε^+,h^+ denote the current and next states respectively, and q the current action. When q is chosen at state (ε, h), a variable $\gamma \sim \text{Bern}(q)$ is drawn. By (18) on the event $\gamma = 1$, $\varepsilon^+ = w \sim \mathcal{N}_{0,W}$, while on the event $\gamma = 0$, $\varepsilon^+ = A\varepsilon + w$ with $w \sim \mathcal{N}_{0,W}$, which is equivalent to $\varepsilon^+ \sim \mathcal{N}_{A\varepsilon,W}$. Since h^+ is independent of $\varepsilon, h, \varepsilon^+$, its distribution m_H appears as a product in (29). We denote $\mathbb{E}\left[\varepsilon^+, h^+ \,\middle|\, \varepsilon, h, q\right]$ the integration with respect to the above transition probability measure.

To sum up, we have exploited the proposed decoupling information structure to determine the optimal control policy as the standard LQR control input. We proceed in the following section to show that an optimal communication policy exists and we characterize its main features in the case of general FEC codes and in the special case of capacity achieving codes.

Remark 5. The technical condition (21) for the controller in Theorem 1 can be viewed as an additional stability condition requiring that the norm of the plant state grows at a sub-linear rate. Such conditions appear in general in optimal control problems with average cost, see e.g. [28, Vol.II, p.254-5], and have also been used in average linear quadratic problems [28, Vol.II, p.272-3]. This technical condition may potentially be relaxed by employing a different proof technique.

IV. Optimal Communication Policy

Exploiting the MDP formulation of (27) we can show that optimal communication policies for the co-design problem in (16) exist. This existence result provides a characterization of these policies from which we infer the general features of optimal transmit powers p_k and corresponding successful decoding probabilities q_k as a function of innovation terms ε_k and channel realizations h_k.

In general the existence of optimal policies for average infinite-horizon MDPs on Borel spaces requires some technical conditions [29]. In our case restriction to communication policies $\pi \in \Pi$ that uniformly satisfy (15) guarantee existence, as the following theorem shows, and a useful characterization of this policy is provided.

Theorem 2 (Optimal communication policy)**.** *Consider the Markov decision process with optimal cost as in (27), state transition probabilities as in (29), and actions restricted to $q_k \in Q(\varepsilon_k, h_k)$ with $Q(\varepsilon, h)$ abiding to (15). If Assumptions 1 and 2 hold true there exists a function $V : \mathbb{R}^n \times \mathbb{R}_+ \mapsto \mathbb{R}$ such that for all $\varepsilon \in \mathbb{R}^n$ and $h \in \mathbb{R}_+$ it satisfies*

$$V(\varepsilon, h) = \min_{q \in Q(\varepsilon,h)} \{c(\varepsilon, h, q) - J^*_{COMM} + \mathbb{E}\left[V(\varepsilon^+, h^+) \,\middle|\, \varepsilon, h, q\right]\}. \qquad (30)$$

The optimal communication cost can be written as $J^*_{COMM} = \mathbb{E}_{w,h} V(w, h)$, where $\mathbb{E}_{w,h}$ denotes integration with respect to the product measure $\mathcal{N}_{0,W} \times m_H$. The optimal communication policy π^* achieving the minimum cost can be written as a function of the error and channel states at time k, $q^*_k = q^*(\varepsilon_k, h_k)$, and is the one achieving the minimum in the right hand side of (30), i.e.

$$q^*(\varepsilon, h) := \operatorname*{argmin}_{q \in Q(\varepsilon,h)} \{c(\varepsilon, h, q) - J^*_{COMM} + \mathbb{E}\left[V(\varepsilon^+, h^+) \,\middle|\, \varepsilon, h, q\right]\}. \qquad (31)$$

Proof: See Appendix C. ■

The theorem states that the optimal communication policy exists, is deterministic, and also stationary in the sense that q^*_k adapts only to the current state (ε_k, h_k) and not the complete history F_k in (13). The optimal policy is described by (31) in terms of a function $V(\varepsilon, h)$ that solves (30). Note that this function is unique up to a constant. Related characterizations of optimal policies when the decision is whether to transmit or not have appeared in [15], [16]. Our setup however differs since the decision is on the transmit power and this depends on the random wireless channel state. The proof of the theorem relies on constructing a Lyapunov-like function that is common for all policies $\pi \in \Pi$ and applying the MDP results of [30]. This methodology has been used in [16], however a refined construction is required here to take into account the random channel states as well.

An informal interpretation of the theorem and the condition (30) based on finite state spaces [28] is the following. The Markov chain induced by any stationary policy $q_k = q(\varepsilon_k, h_k)$ is (positive) recurrent. Fix some state $(\hat{\varepsilon}, \hat{h})$, and then the optimal cost J^*_{COMM} in (27) of any stationary policy can be expressed as the optimal expected cost gathered starting from $(\hat{\varepsilon}, \hat{h})$ till the first return to $(\hat{\varepsilon}, \hat{h})$, divided by the expected number of steps this transition takes. On the other hand (30) has exactly the form of standard Bellman equation but for a relative cost per stage $c(\varepsilon, h, q) - J^*_{\text{COMM}}$, indicating how far we are from the optimal average cost per stage. The function $V(\varepsilon, h)$ expresses the expected relative cost gathered starting from state (ε, h) and following the optimal policy till the first return to $(\hat{\varepsilon}, \hat{h})$. Thus the term $V(\varepsilon^+, h^+)$ on the right hand side of (30) refers to this cost evaluated at the next state (ε^+, h^+). Bellman's equation states that the optimal choice q at every step minimizes the sum of the current-stage relative cost $c(\varepsilon, h, q) - J^*_{\text{COMM}}$ and the expected future relative cost $\mathbb{E}\left[V(\varepsilon^+, h^+) \,\middle|\, \varepsilon, h, q\right]$. The minimization over the current action q gives again the value $V(\varepsilon, h)$ at the current state, as in the left hand side of (30).

In principle one can find $V(\varepsilon, h)$ using value iteration or policy iteration algorithms which involve iterative application of (30) [29]. This procedure is, however, computationally onerous as each iteration requires minimizing the right hand side of (30) for all possible state pairs $(\varepsilon, h) \in \mathbb{R}^n \times \mathbb{R}_+$. Nevertheless, (30) still gives qualitative information on the optimal policy.

Let us ignore the case $\|\varepsilon\| \geq L, h \geq h_t$ in (15) as it is irrelevant for the following discussion. Integrating $V(\varepsilon, h)$ with respect to the transition (29) gives

$$\mathbb{E}\left[V(\varepsilon^+, h^+) | \varepsilon, h, q\right] = q\mathbb{E}_{w,h^+} V(w, h^+) + (1-q)\mathbb{E}_{w,h^+} V(A\varepsilon + w, h^+). \qquad (32)$$

We substitute this and the cost-per-stage $c(\varepsilon, h, q)$ defined by (28), and the expression $J^*_{\text{COMM}} = \mathbb{E}_{w,h} V(w, h)$ in the minimization of (31), and upon reordering terms, the optimal communication policy can be written as

$$q^*(\varepsilon, h) = \operatorname*{argmin}_{q \in [0, q_{max}(h)]} \lambda p(h, q) + (1-q) R(\varepsilon), \qquad (33)$$

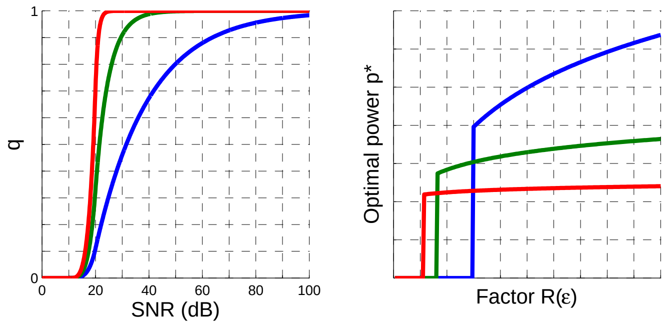

Fig. 5. Optimal power allocation for FEC codes with different complementary error functions. The optimal transmit power p^* is plotted as a function of the factor $R(\varepsilon)$ for a fixed channel state h using FEC codes with different q-SNR characteristics. When the q-SNR curve becomes steeper, the optimal power allocation resembles a step function.

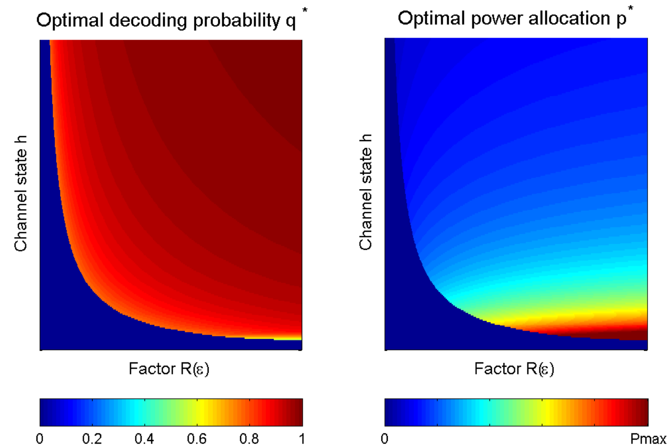

Fig. 6. Optimal decoding probability and power allocation for a FEC code. Color intensity indicates the magnitude of optimal decoding probability q^* and optimal transmit power p^* as functions of the factor $R(\varepsilon)$ and the channel state h.

where for convenience we defined the function

$$R(\varepsilon) := \mathbb{E}_{w,h}\left[V(A\varepsilon + w, h) - V(w,h)\right] + \varepsilon^T \tilde{P}\varepsilon. \qquad (34)$$

The optimal policy $q^*(\varepsilon, h)$ depends on the shape of the function $p(h,q)$. In general it takes values anywhere in the interval $[0, q_{max}(h)]$. The optimal power allocation can be found by converting (33) to power by (2), (4), and is described by

$$p^*(\varepsilon, h) := \operatorname*{argmin}_{p\in[0,p_{max}]} \lambda p + (1 - q(h,p))R(\varepsilon). \qquad (35)$$

Despite the fact that $V(\varepsilon, h)$ and $R(\varepsilon)$ are hard to compute, the above expression is an important characterization of the optimal power allocation. It provides a tool for qualitative analysis of different FEC codes in wireless control systems. We illustrate this in Fig. 5 where we examine how the q-SNR relationship of a FEC code affects the optimal power allocation. For simplicity we assume a fixed channel state h and we plot p^* in Fig. 5 as a function of $R(\varepsilon)$. In all cases, when the error penalty $R(\varepsilon)$ is below some threshold, the best option is to not transmit. Above the threshold, the optimal power increases with $R(\varepsilon)$. For powerful FEC codes characterized by a steep q-SNR relationship, close to the theoretical limit in (3), the optimal power allocation resembles a step function, since the probability of successful decoding becomes practically one for large powers. For fat q-SNR tails, this behavior deteriorates as the sensor needs to transmit with higher power to achieve a larger q.

Then in Fig. 6 we present qualitative plots of the optimal decoding probability q^* and optimal transmit power p^* as functions of both the factor $R(\varepsilon)$ and the channel state h for a given q-SNR characteristic. The blue region indicate the event where no transmission occurs. This happens if channel gain h is low, where transmission is costly, or if error ε has a low penalty, meaning that there is no need to update the receiver's estimate. We note that this no-transmission region becomes larger if one increases the power scaling factor λ. Outside this region a transmission occurs and transmit power adapts to both channel and error states. In principle when channel gain h is high, small amounts of power suffice. For intermediate values of channel h power takes a large range of values depending on the error as well. Overall this optimal power management displays different features from the standard ”0-1” event-triggered transmission paradigms, as in e.g., [15] or [12]. It can be though as a ’soft” version of these policies since the power decision ranges between $[0, p_{max}]$, or equivalently the decoding q between $[0, q_{max}(h)]$. Finally we note that the transmit power/estimation problem has also been studied in the very recent works [23], [24], however the former allows only power adaptation to channel and does not adapt to plant state, while the latter does not provide the important qualitative characterization we discuss here and the connections with the event-triggered paradigm.

A. *Optimal solution for capacity achieving codes*

Consider now the case of capacity achieving codes. By (3), at time k the transmitter needs to use either $p_k = 0$, i.e. not transmitting, or $p_k = p_0/h_k$ with $p_0 := N_0\text{SNR}_0$, which certainly guarantees correct packet delivery. Any other power allocation is unfavorable. However the instantaneous power is bounded by $p_k \le p_{max}$, so the sensor can transmit only when $p_0/h_k \le p_{max}$, or equivalently when the channel state exceeds $h_k \ge p_0/p_{max}$.

In this case we are looking again for a randomized policy, i.e. a distribution on the two power options $\{0, p_0/h_k\}$ when $h_k \ge p_0/p_{max}$. With a slight abuse of notation we denote $q_k \in [0,1]$ the probability of choosing power p_0/h_k. Then when $h_k \ge p_0/p_{max}$ the transmitter draws independent $\gamma_k \sim \text{Bern}(q_k)$ and transmits with power $p_k = \gamma_k\, p_0/h_k$. The decoding success at the receiver is given by the same γ_k. The expected power consumption becomes

$$\mathbb{E}\sum_{k=0}^{N-1} p_k = \mathbb{E}\sum_{k=0}^{N-1} q_k \frac{p_0}{h_k}\,\mathbb{I}\left(h_k \ge \frac{p_0}{p_{max}}\right). \qquad (36)$$

Observe that this is of the form as the expected power consumption of the original problem given in (7) with $p(h,q) = q\,p_0/h\,\mathbb{I}\,(h \ge p_0/p_{max})$. Then the statements of the results so far hold for the capacity achieving codes as well. For this special case for $p(h,q)$ however the minimization in (33) becomes linear in q, and the optimal communication policy is deterministic,

$$q^{CA}(\varepsilon,h) := \begin{cases} 0 & \text{if } \; h\,R(\varepsilon) \leq \lambda\, p_0 \text{ or } h \leq p_0/p_{max} \\ 1 & \text{otherwise} \end{cases}, \tag{37}$$

or in terms of power

$$p^{CA}(\varepsilon,h) := \begin{cases} 0 & \text{if } \; h\,R(\varepsilon) \leq \lambda\, p_0 \text{ or } h \leq p_0/p_{max} \\ p_0/h & \text{otherwise} \end{cases}. \tag{38}$$

This is an event-triggered transmission scheme along the lines of, e.g., [15], except that now the decision is also affected by the current channel state h apart from the error ε. This deterministic policy was expected as the limit behavior of powerful FEC codes in Fig. 5. The region of the plant/channel state space $\mathbb{R}^n \times \mathbb{R}_+$ outside of which it is optimal for the sensor to transmit is described in (38) as $h\,R(\varepsilon) > \lambda p_0$ and $h \geq p_0/p_{max}$. Qualitatively the condition $h\,R(\varepsilon) > \lambda p_0$ shows that when the channel is in a good state, transmitting is worthy since it does not cost much, while when a measure $R(\varepsilon)$ of the error is large, it is necessary to transmit in order to reset it to zero. This region gets larger when p_0 increases, since successful transmission in this case requires more power, or when λ increases, since then power penalty becomes more important.

In the following section we present a simple computable approximation of the above optimal communication policies, which we examine with simulations in Section VI.

V. A ROLLOUT COMMUNICATION POLICY

The optimal communication policy $q^*(\varepsilon,h)$ is described by Theorem 2 in terms of the relative value function $V(\varepsilon,h)$. The practical value of this characterization is limited because determination of $V(\varepsilon,h)$ is not computationally tractable in general. The purpose of this section is to show how approximate dynamic programming techniques can be used to devise approximations of $V(\varepsilon,h)$ leading to tractable suboptimal policies.

As shown in Theorem 2, the optimal communication policy $q^*(\varepsilon,h)$ given by (31) is to choose the current decision q that minimizes a combination of the current cost $c(\varepsilon,h,q)$ and the *optimal* expected future cost $\mathbb{E}\left[V(\varepsilon^+,h^+)|\varepsilon,h,q\right]$. However the function $V(\varepsilon,h)$ is not available, so modeling the optimal future cost is not possible.

Suppose on the other hand that some communication policy π suboptimal in general is available, for which the corresponding relative value function $V^\pi(\varepsilon,h)$ is known at all state pairs (ε,h). If we assume that at all future time steps the sensor/transmitter employs this given policy π we can model the expected future cost induced by π as $\mathbb{E}\left[V^\pi(\varepsilon^+,h^+)|\varepsilon,h,q\right]$. Then the optimal current action selected with respect to this suboptimal assumption on future communication decisions is described by

$$q^{\text{roll}}(\varepsilon,h) := \underset{q\in[0,q_{max}(h)]}{\operatorname{argmin}} \; c(\varepsilon,h,q) + \mathbb{E}\left[V^\pi(\varepsilon^+,h^+)|\varepsilon,h,q\right]. \tag{39}$$

This approximation defines a rollout algorithm [28, Vol. I]. If the suboptimal policy π is not far from the optimal policy the rollout transmission success probability $q^{\text{roll}}(\varepsilon,h)$ is close to the optimal $q^*(\varepsilon,h)$.

To find a family of policies with computable relative value function suppose we adapt $q_k = q(h_k)$ to the current channel state h_k but not to the innovation ε_k. Policies of this form have been proposed in prior works [23], [24]. Since channel states are independent of ε_k the policy $q(h)$ results in successful packet decodings with expected probability $\bar{q} := \mathbb{E}_h q(h)$ implying that the communication success indicator variable is $\gamma_k \sim \text{Bern}(\bar{q})$. The expected power consumption at every stage is also constant given by $\mathbb{E}_h p(h,q(h))$. Thus the cost of this policy $q(h)$ for the MDP problem in (27) becomes

$$J^{q(\cdot)} := \limsup_{N\to\infty} \frac{1}{N}\mathbb{E}\sum_{k=0}^{N-1}(1-\bar{q})\varepsilon_k^T\tilde{P}\varepsilon_k + \lambda\mathbb{E}_h p(h,q(h)). \tag{40}$$

For any policy of the form $q_k = q(h_k)$ the corresponding relative value function $V^{q(\cdot)}(\varepsilon,h)$ can be determined in closed form as stated in the following theorem that also provides an explicit expression for the cost $J^{q(\cdot)}$.

Theorem 3 (Cost of channel-adaptive communication policies)**.** *Consider the Markov decision process with state pair (ε,h) and state transition probabilities as in (29). Consider policies $q(\cdot)$ for which the success transmission probability is selected as a function $q(h)$ independent of the innovation terms ε. For any policy of this form satisfying $\bar{q} := \mathbb{E}_h q(h) > q_{crit}$ for the critical probability q_{crit} of Assumption 1, the cost $J^{q(\cdot)}$ in (40) becomes*

$$J^{q(\cdot)} = Tr(\tilde{P}E) + \lambda\mathbb{E}_h p(h,q(h)), \tag{41}$$

where the matrix E is the unique solution of

$$E = (1-\bar{q})(AEA^T + W). \tag{42}$$

Furthermore, the relative value function $V^{q(\cdot)}$ is given by

$$V^{q(\cdot)}(\varepsilon,h) = \frac{1-q(h)}{1-\bar{q}}\varepsilon^T H\varepsilon + \lambda p(h,q(h)), \tag{43}$$

where the matrix H is the unique solution of

$$H = (1-\bar{q})(A^T HA + \tilde{P}). \tag{44}$$

Proof: See Appendix D. ∎

Theorem 3 provides an explicit formula for a family of relative value functions $V^{q(\cdot)}(\varepsilon,h)$ that can be used in the rollout algorithm in (39). Substituting (43) into (39) and removing constants from the resulting expression we find the rollout policy

$$q^{\text{roll}}(\varepsilon,h) := \underset{q\in[0,q_{max}(h)]}{\operatorname{argmin}} \; \lambda p(h,q) + (1-q)\frac{\varepsilon^T H\varepsilon}{1-\bar{q}}. \tag{45}$$

Computing such policies is easy. Given the parameter $\bar{q}$ that models the suboptimal future actions, we can compute H by (44) and then solve (45) given the function $p(h,q)$. Observe that (45) is of the same form as the optimal communication policy (33) except that the optimal unknown function $R(\varepsilon)$ is replaced by the quadratic form $\varepsilon^T H\varepsilon/(1-\bar{q})$. Since the rollout policy is suboptimal the quadratic can be viewed as an approximation of the function $R(\varepsilon)$. As a side note, the rollout policy need not satisfy the technical requirement $q^{\text{roll}}(\varepsilon, h) \in Q(\varepsilon, h)$ of (15).

For the particular case of a capacity achieving FEC we can repeat the analysis in Section IV-A to modify (45) and obtain the explicit (suboptimal) policy

$$q^{\text{roll},CA}(\varepsilon, h) := \begin{cases} 0 & \text{if } h\,\dfrac{\varepsilon^T H \varepsilon}{1-\bar{q}} \le \lambda\, p_0 \text{ or } h \le \dfrac{p_0}{p_{max}} \\ 1 & \text{otherwise} \end{cases} . \quad (46)$$

Again we managed to approximate the unknown function $R(\varepsilon)$ in (37) by a quadratic that we can compute by (44). This gives us an explicit event triggered communication policy, where the events depend on the current values of the channel state h and the error ε.

A question that arises is how suboptimal is the performance of these policies compared to the optimal cost J^*_{COMM}. Unfortunately since the rollout is a heuristic it is not easy to characterize the optimality gap. It is however guaranteed to perform not worse than the reference policy, and in many practical problems the improvement is significant [28, Vol. I]. In the following section we simulate the constructed policies and characterize numerically the improvement to the reference policies adapting to channel only.

VI. Simulations

We begin by presenting simulations of the rollout algorithm for capacity achieving codes given in (46). We assume the channel state distribution to be exponential with mean 0.5. The plant is given by

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 0.8 \end{bmatrix}, \; B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \quad (47)$$

The rest of the parameters are $Q = W = I$, $R = 1$, $\lambda = 50$, $p_0 = 1$. As a reference communication policy we use $q(h) = \mathbb{I}(h \ge h_t)$ which transmits whenever the channel state is above some threshold value h_t that induces $\bar{q} \approx 0.79$. The simulations of the rollout policy reveal a dramatic decrease in the empirical rate of transmissions $q_{emp} = 1/N \sum_{k=0}^{N-1} \gamma_k \approx 0.37$, which is also much lower than the minimum non-adaptive policy $q_{crit} = 0.75$ that would keep the error stable. Similarly, the empirical cost $J_{emp} = 1/N \sum_{k=0}^{N-1} e_k^T \tilde{P} e_k + \lambda \gamma_k p_0 / h_k \approx 56$ decreased compared to the reference $J^{q(\cdot)} \approx 124$.

The event-triggered nature of the rollout policy in the case of capacity achieving codes is captured in Fig. 7 where we plot the two plant states along with the channel, $|\varepsilon_{k,1}|, h_k$ and $|\varepsilon_{k,2}|, h_k$ during the simulation. Blue points indicate the decision not to transmit, $q_k = 0$, while red are the points where $q_k = 1$. When the channel fading coefficient h_k is low, the sensor avoids transmission as it requires large power consumption. The rollout policy is also adapted to the plant structure. The error state $\varepsilon_{k,1}$ is related to the unstable eigenvalue of A, so the sensor always decides to transmit when this state is far from 0. The hyperbolic shape of the $|\varepsilon_{k,1}|, h_k$ plot was expected by the form of the rollout algorithm in (46).

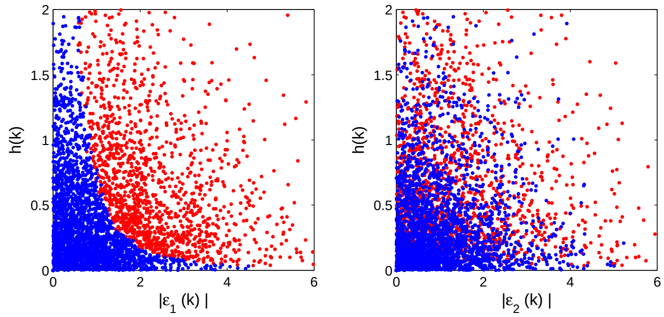

Fig. 7. Simulation results of the rollout policy for a capacity achieving code. The points $|\varepsilon_{k,1}|, h_k$ and $|\varepsilon_{k,2}|, h_k$ are plotted respectively, with blue indicating the decision to not transmit, $q_k = 0$, and red to transmit, $q_k = 1$.

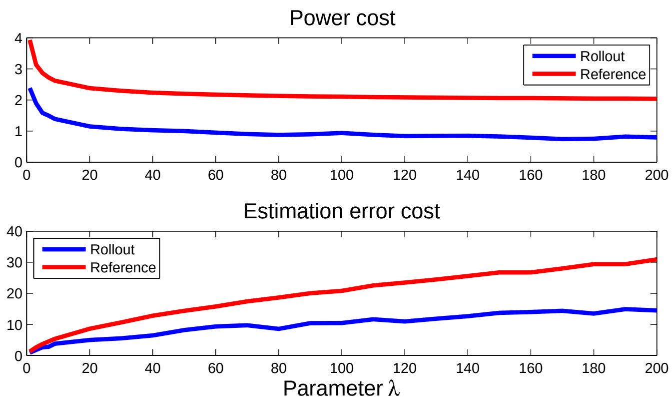

Fig. 8. Comparison of the optimal threshold policy and the resulting rollout policy performance for a capacity achieving code. The power and estimation error costs of the two policies are plotted for different values of the weighting factor λ. The estimation error cost for the reference increases at a faster rate than the rollout.

In contrast, such a correlation between the error state $\varepsilon_{k,1}$ and the decision to transmit is not clear. Even when $\varepsilon_{k,2}$ takes large values, the sensor might choose not to transmit. The reason is that this state's dynamics are related to a stable eigenvalue, so informally it will remain bounded even if the sensor takes no action. More precisely, as long as the sensor keeps $\varepsilon_{k,1}$ bounded, $\varepsilon_{k,1}$ will also be bounded.

Next, for the plant and channel described above we compare the performance of the rollout algorithm with that of the reference policy we used to compute the value function in (43). For different values of λ we find the optimal threshold policy $q(h) = \mathbb{I}(h \ge h_t)$, i.e. the one that minimizes (41), and the corresponding rollout policy in (46). The resulting power cost $1/N \sum_{k=0}^{N-1} \gamma_k p_0 / h$ and estimation cost $1/N \sum_{k=0}^{N-1} e_k^T \tilde{P} e_k$ that we got from simulating the rollout algorithm are plotted separately in Fig. 8 along with the costs of the reference policy. As λ increases the power consumption decreases, since it is penalized more in the aggregate cost (27), and the decrease rate is similar for the rollout and the reference policies. On the other hand, when λ increases the estimation cost increases, since the sensor decides to transmit less often. However the increase for the rollout policy is slower than that of the reference policy. The reason is that the reference only adapts to the channel, avoiding transmissions when the channel state

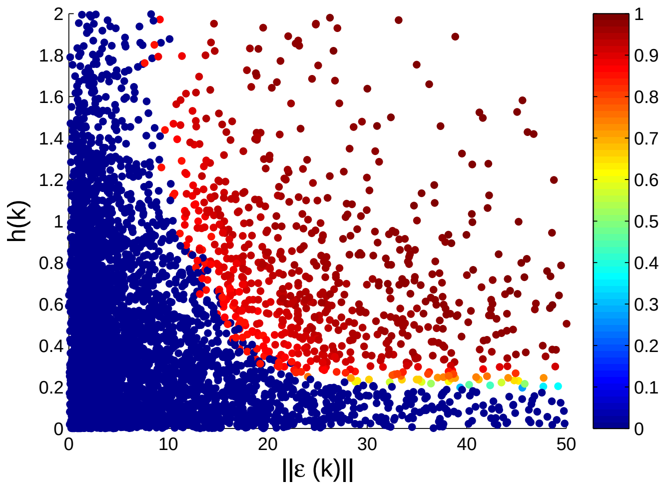

Fig. 9. Simulation results of the rollout policy for a FEC code. The points $\|\varepsilon_k\|, h_k$ during the simulation are plotted with colors denoting the magnitude of the chosen decoding probability q_k. Practically q_k takes values either 0 or close to 1.

h_k is low. The rollout algorithm adapts not only to the channel, but also to the error ε_k. By transmitting only when ε_k is large, it results in only a moderate increase in the estimation cost without sacrificing too much power.

Finally, we simulate the rollout algorithm in (45) for the above plant and channel model when a FEC code is employed. The probabilities q_k of successful decoding that the rollout provided during the simulation are plotted in Fig. 9 on $\|\varepsilon_k\|, h_k$ axes (compare with the optimal policy in Fig. 6). Unlike the capacity achieving codes, the decisions q_k take values smaller than 1. However, due to the sigmoid form of the q-SNR characteristic of the FEC code, we observe q_k that are practically either 0 or very close to 1, especially when the channel state is good (h_k large). For low channel fading gain h_k there is a very high power penalty if the sensor wants to transmit with high success probability. In this case, the rollout policy is either to not transmit ($q_k = 0$) or transmit with a success probability q_k very close to 1. In general, we observe that the points in the plot are accumulated at the region where the error $\|\varepsilon_k\|$ is small. The reason is that when the error gets larger, q_k is chosen close to 1 by the communication policy, so a successful transmission $\gamma_k = 1$ occurs with very high probability and resets the error.

VII. Conclusions and Future Work

In this paper we examined a control system with a wireless fading channel between the sensor and the controller. The sensor adapts transmit power to plant and channel states and affects the probability of successful decoding at the controller. For the problem of co-designing transmit powers and control inputs to minimize an average LQR and power cost a method to separate the two designs is provided, leading to the standard LQR controller. The optimal power allocation is then characterized qualitatively for general FEC codes and capacity achieving codes. Tractable suboptimal policies are derived and their performance is compared with alternative policies in simulations.

The design of near-optimal and computationally efficient policies for the general co-design framework requires further research. For example the performance of power management policies adapting to the plant state directly, not the estimation innovation term, needs to be evaluated. Further work includes also the incorporation of other wireless channel models, such as Markov. Moreover interference effects when multiple control loops close over the same channel need to be considered, although they are expected to be limited since sensors abstain from transmitting in unfavorable channel conditions and/or favorable plant states. Overall the proposed framework can be expanded to accommodate modeling and analysis of more complex wireless sensor & actuator networks. This unified control/wireless networking framework could lead to novel communication/control design problems, such as a control-aware network resource allocation, or a resource-aware networked controller synthesis.

Appendix

A. Proof of Proposition 2

First note that, by the same arguments we use to derive (26) later, if we condition on F_k we can rewrite

$$\mathbb{E}^\pi \, e_k^T e_k = \mathbb{E}^\pi (1 - q_k) \varepsilon_k^T \varepsilon_k. \qquad (48)$$

The uniform bound of (20) will be proved by an equivalent bound on the innovation process $\{\varepsilon_k, k \geq 0\}$. By Proposition 1 for any communication policy $\pi \in \Pi$ this process is independent of the chosen control policy $\theta \in \Theta$ and its evolution is given by (18). This evolution can be described more formally along with the i.i.d. channel process $h_k \sim m_H$ by a stochastic transition kernel given the values of ε, h and action q at each step as

$$\begin{aligned} &\mathbb{P}(\varepsilon^+, h^+ | \varepsilon, h, q) \\ &= \left[q \, \mathcal{N}_{0,W}(\varepsilon^+) + (1-q) \, \mathcal{N}_{A\varepsilon,W}(\varepsilon^+) \right] \, m_H(h^+). \end{aligned} \qquad (49)$$

This expression is included later again in (29), where its derivation is explained in detail.

The following technical lemma shows that under Assumption 1 one can construct a Lyapunov-like function common for all communication policies, satisfying explicitly the technical requirements of [30, Assumptions 3.1, 3.2]. The uniform bound (20) is a direct consequence of these requirements, while the lemma will be subsequently used to prove Theorem 2 based on the results of [30].

Lemma 1. *Suppose Assumption 1 holds and consider the innovation and channel processes $\{\varepsilon_k, h_k, k \geq 0\}$ described by the transition (49), with communication decisions satisfying $q_k \in Q(\varepsilon_k, h_k)$ given in (15). Then there exists a measurable function W on $\mathbb{R}^n \times \mathbb{R}_+$ bounded below by a constant $\gamma > 0$ such that*

$$(1 - q)\varepsilon^T \varepsilon + c \leq K W(\varepsilon, h), \qquad (50)$$

where $c \geq 0$ is some constant, for all $\varepsilon, h \in \mathbb{R}^n \times \mathbb{R}_+$, $q \in Q(\varepsilon, h)$, for some positive K. Moreover there exists a non-trivial measure ν on $\mathbb{R}^n \times \mathbb{R}_+$, a non-negative measurable *function $\phi(\varepsilon,h,q)$ for $\varepsilon,h \in \mathbb{R}^n\times\mathbb{R}_+$, $q\in Q(\varepsilon,h)$, and a positive constant $\mu<1$ such that*

(i) $\nu(W):=\int W(\varepsilon,h)d\nu(\varepsilon,h)<\infty,$

(ii) $\mathbb{P}(\varepsilon^+\in B_1, h^+\in B_2|\varepsilon,h,q)\geq \nu(B_1,B_2)\phi(\varepsilon,h,q)$ *for all measurable subsets* $(B_1,B_2)\in\mathcal{B}(\mathbb{R}^n\times\mathbb{R}_+)$,

(iii) $\mathbb{E}\left[W(\varepsilon^+,h^+)|\varepsilon,h,q\right]\leq \mu W(\varepsilon,h)+\phi(\varepsilon,h,q)\nu(W)$

(iv) $\int \phi(\varepsilon,h,q)d\nu(\varepsilon,q)>0$ *for all* $q\in Q(\varepsilon,h)$.

Proof: Let

$$\nu:=\mathcal{N}_{0,W}\times m_H, \tag{51}$$

$$\phi(\varepsilon,h,q):=q. \tag{52}$$

Let us denote the set where the choice of q is free as

$$S:=\{(\varepsilon,h)\in\mathbb{R}^n\times\mathbb{R}_+ : \|\varepsilon\|<L \text{ or } h<h_t\}. \tag{53}$$

We choose $\mu<1$ such that

$$\mu>1-q_{\max}(h_t)+q_{\max}(h_t)\nu(S), \tag{54}$$

and

$$\mu>(1-\bar q)\,\lambda_{\max}(A)^2, \tag{55}$$

where $\bar q$ denotes the integral introduced in (14),

$$\bar q:=\int_{h_t}^{+\infty} q_{\max}(h)dm_H(h). \tag{56}$$

The left hand side of (54) is less than 1 because the event S under the measure ν happens with probability less than 1 and we have assumed $q_{\max}(h_t)>0$. The left hand side of (55) is also less than 1 because of Assumption 1 and by the choice for h_t that satisfies (14).

Finally for any $L>0$ when $\|\varepsilon\|\geq L$ observe that by construction of the set $Q(\varepsilon,h)$ we can upper bound

$$1-q\leq 1-q_{\max}(h)\mathbb{I}\left(h\geq h_t\right)=:\psi(h), \tag{57}$$

where we named the quantity on the right $\psi(h)$ to be used within this proof. This inequality holds because when $h<h_t$, $q\geq 0$, and when $h\geq h_t$, $q=q_{\max}(h)$.

Then we pick

$$W(\varepsilon,h):=\psi(h)\varepsilon^T H\varepsilon+\beta\,\mathbb{I}\left(\varepsilon,h\in S\right)+\gamma, \tag{58}$$

where $\beta,\gamma>0$ are appropriate positive constants that will be designed next, and $H>0$ is a positive definite matrix satisfying

$$(1-\bar q)A^THA-\mu H=-\Theta, \tag{59}$$

for some positive definite matrix $\Theta>0$. This Lyapunov equation is feasible by our choice of μ that satisfies (55).

First observe that $W(\varepsilon,h)\geq\gamma>0$ by construction. Then we check (50). When $\|\varepsilon\|<L$, we have

$$(1-q)\varepsilon^T\varepsilon+c\leq L^2+c\leq K(\beta+\gamma)\leq KW(\varepsilon,h), \tag{60}$$

for a sufficiently large K, where the last inequality follows from the form of $W(\varepsilon,h)$ on $\|\varepsilon\|<L$. On the other hand if $\|\varepsilon\|\geq L$, we may use (57) to upper bound

$$\begin{aligned}(1-q)\varepsilon^T\varepsilon+c&\leq\psi(h)\varepsilon^T\varepsilon+c\\ &\leq K(\psi(h)\varepsilon^TH\varepsilon+\gamma)\leq KW(\varepsilon,h),\end{aligned} \tag{61}$$

for a sufficiently large K, by our choice for the function $W(\varepsilon,h)$ when $\|\varepsilon\|\geq L$.

We proceed by showing that parts (i)-(iv) in the statement of the lemma also hold. Part (i) holds because the integral of $W(\varepsilon,h)$ with our chosen measure ν equals

$$\nu(W)=(1-\bar q)Tr(HW)+\nu(S)\beta+\gamma<\infty. \tag{62}$$

Part (ii) holds because the transition probability in (29) gives

$$\begin{aligned}&\mathbb{P}(\varepsilon^+\in B_1,h^+\in B_2|\varepsilon,h,q)\\ &=\left[q\,\mathcal{N}_{0,W}(B_1)+(1-q)\,\mathcal{N}_{A\varepsilon,W}(B_1)\right]\ m_H(B_2)\\ &\geq q\,\mathcal{N}_{0,W}(B_1)\ m_H(B_2)=\phi(\varepsilon,h,q)\nu(B_1,B_2).\end{aligned} \tag{63}$$

Part (iv) follows by our choice $\phi(\varepsilon,h,q)=q$ and the construction of the set $Q(\varepsilon,h)$ in (15) because

$$\begin{aligned}\int\phi(\varepsilon,h,q)d\nu(\varepsilon,h)&\geq\int_{\varepsilon,h\in S^c}q_{max}(h)d\nu(\varepsilon,h)\\ &=\bar q\int_{\|\varepsilon\|\geq L}d\mathcal{N}_{0,W}(\varepsilon)>0.\end{aligned} \tag{64}$$

To prove the remaining part (iii) first observe that by the transition defined in (29) and our choices for ν and ϕ we have

$$\begin{aligned}&\mathbb{E}\left[W(\varepsilon^+,h^+)|\varepsilon,h,q\right]=\phi(\varepsilon,h,q)\nu(W)\\ &+(1-q)\int W(\varepsilon^+,h^+)d\mathcal{N}_{A\varepsilon,W}(\varepsilon^+)\,d\,m_H(h^+).\end{aligned} \tag{65}$$

Substituting (65) in (iii), we only need to show that

$$(1-q)\int W(\varepsilon^+,h^+)d\mathcal{N}_{A\varepsilon,W}(\varepsilon^+)\,d\,m_H(h^+)\leq\mu W(\varepsilon,h). \tag{66}$$

Plugging the expression of $W(\varepsilon,h)$ given by (58) in the integral of the left hand side, condition (66) becomes

$$\begin{aligned}&(1-q)\left\{(1-\bar q)\left[\varepsilon^TA^THA\varepsilon+Tr(HW)\right]\right.\\ &\qquad\left.+\beta\mathcal{N}_{A\varepsilon,W}\times m_H(S)+\gamma\right\}\leq\mu W(\varepsilon,h).\end{aligned} \tag{67}$$

We can bound $\mathcal{N}_{A\varepsilon,W}\times m_H(S)\leq\nu(S)$ for any $\varepsilon\in\mathbb{R}^n$, and also $(1-q)(1-\bar q)Tr(HW)\leq Tr(HW)$. So a sufficient condition for (67) is to show that

$$\begin{aligned}&(1-q)\left\{(1-\bar q)\varepsilon^TA^THA\varepsilon+\beta\,\nu(S)+\gamma\right\}+Tr(HW)\\ &\leq\mu W(\varepsilon,h)\end{aligned} \tag{68}$$

holds for every choice of $q\in Q(\varepsilon,h)$.

Let us first study the case $\|\varepsilon\|\geq L$. Using again (57) to upper bound $1-q\leq\psi(h)$, and upon substituting $W(\varepsilon,h)$ in (68) and rearranging terms, we need to show equivalently that

$$\begin{aligned}&\psi(h)\left\{\varepsilon^T\left[(1-\bar q)A^THA-\mu H\right]\varepsilon\right.\\ &\left.+\beta\nu(S)+\gamma\right\}+Tr(HW)\leq\mu\left\{\beta\mathbb{I}\left(h<h_t\right)+\gamma\right\}\end{aligned} \tag{69}$$

By the choice of H in (59) the quadratic on the left hand side is negative definite equal to to $-\varepsilon^T\Theta\varepsilon$. And since $\|\varepsilon\|\geq L$ we can upper bound $-\varepsilon^T\Theta\varepsilon\leq-\lambda_{\min}(\Theta)L^2\leq 0$. After these, a sufficient condition for (69) is

$$Tr(HW)+\psi(h)\left\{\beta\nu(S)+\gamma\right\}\leq\mu\left\{\beta\mathbb{I}\left(h<h_t\right)+\gamma\right\} \tag{70}$$

We now take two cases for h. If $h<h_t$, the above condition (70) becomes

$$Tr(HW)+\beta\nu(S)+\gamma\leq\mu(\beta+\gamma). \tag{71}$$

On the other hand if $h \geq h_t$ we have that $q_{\max}(h) \geq q_{\max}(h_t)$ by monotonicity assumption, so we may bound $\psi(h) = 1 - q_{\max}(h) \leq 1 - q_{\max}(h_t)$. Condition (70) becomes

$$Tr(HW) + (1 - q_{\max}(h_t)) \{\beta\nu(S) + \gamma\} \leq \mu\gamma. \quad (72)$$

We pick a $\gamma >$ to satisfy (72) with equality, that is

$$\gamma = \frac{(1 - q_{\max}(h_t))\nu(S)\beta + Tr(HW)}{\mu - (1 - q_{\max}(h_t))} \quad (73)$$

where the denominator is positive by the choice of μ in (54). We will show that condition (71) also holds by an appropriate choice for β.

Let us now examine condition (68) in the case $\|\varepsilon\| < L$. Then $q \geq 0 \Rightarrow 1 - q \leq 1$ and it is sufficient for (68) to show that

$$\begin{aligned}&\sup_{\|\varepsilon\|<L} \varepsilon^T(1-\bar{q})A^T HA\varepsilon + Tr(HW) + \beta\nu(S) + \gamma \\ &\leq \mu(\beta + \gamma)\end{aligned} \quad (74)$$

where on the right hand side we lower bounded the quadratic term of W by 0. This is of the general form

$$C_2 + \beta\nu(S) + \gamma \leq \mu(\beta + \gamma) \quad (75)$$

for some constant C_2, like the left over condition (71). Plugging in (75) the chosen γ by (73) leads to a condition of the form

$$C_3 + \frac{q_{\max}(h_t)\,\nu(S)\,\mu\beta}{\mu - (1 - q_{\max}(h_t))} \leq \mu\beta, \quad (76)$$

for some constant C_3. We want this to hold for an arbitrarily large positive β because C_3 might be negative, so we need the coefficient of $\mu\beta$ on the left side to be *strictly* smaller than the coefficient of $\mu\beta$ on the right hand side. This turns out to be equivalent to

$$\mu > 1 - q_{\max}(h_t) + q_{\max}(h_t)\nu(S), \quad (77)$$

which corresponds to our choice of μ in (54). ■

Turning back to the proof of Proposition 2, combining (48) with condition (50) of the above Lemma we have that $\mathbb{E}^\pi e_k^T e_k \leq K\,\mathbb{E}^\pi W(\varepsilon_k, h_k)$, so it suffices for (20) to show that a uniform bound on the expected value of $W(\varepsilon_k, h_k)$ exists.

By result (ii) of the above lemma for $(B_1, B_2) = (\mathbb{R}^n, \mathbb{R}_+)$ we have that $\phi(\varepsilon, h, q) \leq 1/\nu(\mathbb{R}^n, \mathbb{R}_+)$. Plugging this in (iii) leads to

$$\mathbb{E}\left[W(\varepsilon^+, h^+)|\varepsilon, h, q\right] \leq \mu W(\varepsilon, h) + \frac{\nu(W)}{\nu(\mathbb{R}^n, \mathbb{R}_+)} \quad (78)$$

Iterated applications of this inequality across some policy $\pi \in \Pi$ yields

$$\mathbb{E}^\pi W(\varepsilon_k, h_k) \leq \mu^k\,\mathbb{E}\,W(\varepsilon_0, h_0) + \frac{\nu(W)}{(1-\mu)\nu(\mathbb{R}^n, \mathbb{R}_+)} \quad (79)$$

Thus since $\mu < 1$ a uniform bound on $\mathbb{E}^\pi W(\varepsilon_k, h_k)$ exists and this completes the proof.

B. Proof of Theorem 1

First note that since $\Sigma_k := \mathbb{E}^\pi\left[e_k e_k^T \,\middle|\, G_k\right]$ we have that

$$\mathbb{E}^\pi[Tr(\Sigma_k)] = \mathbb{E}^\pi e_k^T e_k. \quad (80)$$

Then under Assumption 1 Proposition 2 states that for any $\pi \in \Pi$ condition (20) holds and guarantees that both quantities in (80) are bounded uniformly over k.

To establish the optimality of the proposed control law we use the fact that the Bellman-like equation

$$\begin{aligned}&V(G_k) + Tr(PW) + Tr(\tilde{P}\Sigma_k) = \\ &\min_{u_k} \mathbb{E}^\pi\left[x_k^T Q x_k + u_k^T R u_k + V(G_{k+1})|G_k, u_k\right],\end{aligned} \quad (81)$$

is satisfied for the function

$$V(G_k) = \mathbb{E}^\pi\left[x_k^T P x_k \,\middle|\, G_k\right], \quad (82)$$

with $V(G_0) = x_0^T P x_0$, where P is the solution to the standard algebraic Riccati equation and $\tilde{P}$ is given by (23). The existence of P is guaranteed by the stabilizability of (A, B) and detectability of $(A, Q^{1/2})$.

Indeed observe that we can rewrite the term on the right hand side of (81) as

$$\begin{aligned}&\mathbb{E}^\pi\left[V(G_{k+1})|G_k, u_k\right] = \mathbb{E}^\pi\left[x_{k+1}^T P x_{k+1}|G_k, u_k\right] \\ &= \mathbb{E}^\pi\left[(Ax_k + Bu_k)^T P(Ax_k + Bu_k)|G_k, u_k\right] + Tr(PW),\end{aligned} \quad (83)$$

where the last equality follows by substituting x_{k+1} from the system equation (1). The quadratic minimization over u_k at the right hand side of (81) takes the usual form appearing in LQR problems with partial state information - see e.g. [28, Vol. II, Section 5.2]. The argument of the minimization in (81) is given by the control law (24). Straightforward substitutions show that the optimal value of the minimization equals the left hand side of (81).

The equation (81) can be used to show that the optimal control policy is (24). First iterate (81) for $k = 0, \dots, N-1$ across some control policy $\theta \in \Theta$ to get

$$\begin{aligned}&V(G_0) + N\,Tr(PW) + \mathbb{E}^\pi \sum_{k=0}^{N-1} Tr(\tilde{P}\Sigma_k) \\ &\leq J_{\text{LQR}}^N(\pi, \theta) + \mathbb{E}^{\pi,\theta} V(G_N)\end{aligned} \quad (84)$$

Dividing (84) by N and taking the limit as $N \to \infty$, the term on the left hand side tends to

$$\begin{aligned}&\limsup_{N\to\infty} 1/N\left[x_0^T P x_0 + N\,Tr(PW) + \mathbb{E}^\pi \sum_{k=0}^{N-1} Tr(\tilde{P}\Sigma_k)\right] \\ &= Tr(PW) + \limsup_{N\to\infty} 1/N\,\mathbb{E}^\pi \sum_{k=0}^{N-1} e_k^T \tilde{P} e_k\end{aligned} \quad (85)$$

where we used (80) to convert Σ_k to e_k.

Then consider the term on the right hand side of (84). Any control policy $\theta \in \Theta$ satisfying (21) also satisfies

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}^{\pi,\theta} V(G_N) = \lim_{N\to\infty} \frac{1}{N}\mathbb{E}^{\pi,\theta} x_N^T P x_N = 0 \quad (86)$$

by the form of V given in (82). Thus taking the limit as $N \to \infty$, by (86) the term on the right hand side of (84) tends to the average LQR cost. The inequality in (84) then shows that the average LQR cost of θ is larger or equal to the limit of the left hand side which was given in (85). The result given in (22) follows by including the power cost that depends only on the communication policy π as suggested by (19).

The final step of the proof is to show that the control policy θ^* defined by (24) gives exactly the LQR cost given in (85). This policy satisfies (81) with equality, so (84) also holds with equality. Dividing by N and taking the limit as before would prove the desired result if condition (86) also holds for θ^*.

Indeed use $u_k^* = K\hat{x}_k$ and $x_k = \hat{x}_k + e_k$ to rewrite the closed loop system equation (1) under θ^* as

$$x_{k+1} = (A + BK)\hat{x}_k + Ae_k + w_k. \tag{87}$$

Then denoting $\rho := \lambda_{\max}(A + BK)$ which is stable, $\rho < 1$, we can upper bound (83) under θ^* by

$$\begin{aligned} &\mathbb{E}^{\pi,\theta^*}\left[V(G_{k+1})|G_k\right] \\ &\leq \rho^2 \hat{x}_k^T P \hat{x}_k + Tr(A^T P A \Sigma_k) + Tr(PW) \\ &= \rho^2 V(G_k) + Tr((A^T P A - \rho^2 P)\Sigma_k) + Tr(PW) \end{aligned} \tag{88}$$

Taking expectation on both sides we have that

$$\begin{aligned} \mathbb{E}^{\pi,\theta^*} V(G_{k+1}) &\leq \rho^2 \mathbb{E}^{\pi,\theta^*} V(G_k) \\ &+ Tr((A^T P A - \rho^2 P)\, \mathbb{E}^{\pi,\theta^*} \Sigma_k) + Tr(PW) \end{aligned} \tag{89}$$

But (80) and (20) imply that $\mathbb{E}^{\pi,\theta^*}\Sigma_k$ is uniformly bounded over k so the term on the second line of (89) is bounded by some constant $\delta < \infty$. Iterating the above inequality (89) across θ^* up to $k = N$ yields

$$\mathbb{E}^{\pi,\theta^*}\left[x_N^T P x_N\right] \leq \rho^{2N} x_0^T P x_0 + \frac{\delta}{1-\rho^2} \tag{90}$$

which guarantees the limit (86) since $\rho < 1$.

C. *Proof of Theorem 2*

The proof of the theorem is a direct application of the theorems contained in [30]. For these we need to show that [30, Assumptions 3.1, 3.2, 3.4] hold in our case. In particular [30, Assumption 3.1] requires that the cost per stage is bounded $|c(\varepsilon, h, q)| \leq KW(\varepsilon, h)$ by a positive measurable function W. This is a consequence of (50) of Lemma 1, since

$$|c(\varepsilon, h, q)| \leq (1-q)\lambda_{\max}(\tilde{P})\varepsilon^T\varepsilon + \lambda p_{max} \tag{91}$$

which is of the same form as (50). Also [30, Assumption 3.2] requires exactly the conditions given in (i)-(iv) of Lemma 1. Finally [30, Assumption 3.4] requires the following for the chosen functions W and ϕ satisfying Lemma 1.

Assumption 3. For every $\varepsilon \in \mathbb{R}^n$, $h \in \mathbb{R}_+$

(i) $Q(\varepsilon, h)$ is compact,
(ii) $c(\varepsilon, h, q)$ is lower semi-continuous in $q \in Q(\varepsilon, h)$,
(iii) $\mathbb{P}(\varepsilon^+, h^+|\varepsilon, h, q)$ is strongly continuous [1] in $q \in Q(\varepsilon, h)$,
(iv) the mapping $q \to \mathbb{E}\left[W(\varepsilon^+, h^+)|\varepsilon, h, q\right]$ is continuous,
(v) $\phi(\varepsilon, h, q)$ is continuous in $q \in Q(\varepsilon, h)$.

[1]i.e. for every bounded measurable function Ψ on $\mathbb{R}^n \times \mathbb{R}_+$, the mapping $q \mapsto \mathbb{E}[\Psi(\varepsilon^+, h^+)|\varepsilon, h, q]$ is continuous

Part (i) is trivial, and part (ii) is a consequence of the continuity of $p(h, q)$ by Assumption 2. Strong continuity in part (iii) is guaranteed by the fact that the transition kernel given in (29) has a probability density function. Part (iv) holds because the transition in (29) is linear in q, and part (v) is trivial.

Having established [30, Assumptions 3.1, 3.2, 3.4], then [30, Theorems 3.5, 3.6] state that in our case the infimum J^*_{COMM} in (27) exists, there exists a function $V(\varepsilon, h)$ that satisfies (30), and the optimal policy is the minimizer of the right hand side of (30) as given in (31).

Moreover, observe that (30) still holds if we add any constant to $V(\varepsilon, h)$. So without loss of generality we may take $V(0, \hat{h}) = 0$ for some $\hat{h}$. Then if we substitute $\varepsilon = 0, h = \hat{h}$ in (30) we get

$$\begin{aligned} V(0, \hat{h}) = 0 = \min_{q \in Q(0,\hat{h})} &\Big\{ c(0, \hat{h}, q) - J^*_{\text{COMM}} \\ &+ \mathbb{E}\left[V(\varepsilon^+, h^+) \,\middle|\, 0, \hat{h}, q\right] \Big\}. \end{aligned} \tag{92}$$

Note that by (29), $\mathbb{P}(\varepsilon^+, h^+|0, \hat{h}, q) = \mathcal{N}_{0,W}(\varepsilon^+)\, m_H(h^+)$, and if we substitute $c(0, \hat{h}, q) = \lambda p(\hat{h}, q)$ by (28) in (92), we get

$$0 = \min_{q \in Q(0,\hat{h})} \left\{ \lambda p(\hat{h}, q) - J^*_{\text{COMM}} + \mathbb{E}_{w,h} V(w, h) \right\}. \tag{93}$$

The minimizer is $q = 0$ and this gives $J^*_{\text{COMM}} = \mathbb{E}_{w,h} V(w, h)$.

D. *Proof of Theorem 3*

For any channel-adaptive communication policy $q(h)$ with expected success $\bar{q}$, the estimation error e_k in (17) becomes a Markov jump linear system, with mean $\mathbb{E}(e_k) = 0$ for all $k \geq 0$ and covariance

$$\mathbb{E}(e_k e_k^T) = (1 - \bar{q})(A\, \mathbb{E}(e_{k-1} e_{k-1}^T)\, A^T + W). \tag{94}$$

Since $\bar{q} > q_{crit}$ the covariance reaches a steady state matrix E that satisfies the Lyapunov equation (42), and e_k is stable in the bounded covariance sense. The cost of such a policy is then given by (41).

The corresponding relative value function $V^{q(\cdot)}(\varepsilon, h)$ satisfies the steady state condition [c.f. (30)]

$$\begin{aligned} &V^{q(\cdot)}(\varepsilon, h) + J^{q(\cdot)} \\ &= c(\varepsilon, h, q(h)) + \mathbb{E}\left[V^{q(\cdot)}(\varepsilon^+, h^+)|\varepsilon, h, q(h)\right], \end{aligned} \tag{95}$$

where $c(\varepsilon, h, q)$ is given by (28). We need to show that $V^{q(\cdot)}(\varepsilon, h)$ in (43) satisfies condition (95). First integrate (43) with respect to (29) to get

$$\begin{aligned} &\mathbb{E}\left[V^{q(\cdot)}(\varepsilon^+, h^+)|\varepsilon, h, q(h)\right] = (1 - q(h))\varepsilon^T A^T H A \varepsilon \\ &+ Tr(HW) + \lambda \mathbb{E}_{h^+} p(h^+, q(h^+)). \end{aligned} \tag{96}$$

Using (96), the total cost $J^{q(\cdot)}$ in (41), and the cost per stage $c(\varepsilon, h, q)$ in (28), we conclude that $V^{q(\cdot)}(\varepsilon, h)$ in (43) satisfies condition (95) if the following equation holds

$$\frac{1-q(h)}{1-\bar{q}}\varepsilon^T H\varepsilon + \lambda p(h,q(h)) + Tr(\tilde{P}E) + \lambda \mathbb{E}_h p(h,q(h))$$
$$= (1-q(h))\varepsilon^T \tilde{P}\varepsilon + \lambda p(h,q(h)) + (1-q(h))\varepsilon^T A^T H A\varepsilon$$
$$+ Tr(HW) + \lambda \mathbb{E}_{h^+} p(h^+, q(h^+)). \tag{97}$$

Substituting (44) on the left hand side and canceling terms, the above condition becomes equivalent to

$$Tr(\tilde{P}E) = Tr(HW). \tag{98}$$

This is easily verified if we substitute the explicit expressions for the solutions E, H of the discrete-time Lyapunov equations (42), (44) given by

$$E = \sum_{k=0}^{\infty} (1-\bar{q})^{k+1} A^k W (A^T)^k, \tag{99}$$

$$H = \sum_{k=0}^{\infty} (1-\bar{q})^{k+1} (A^T)^k \tilde{P} A^k, \tag{100}$$

respectively. The existence and uniqueness of solutions to the Lyapunov equations (42), (44) follows by the assumption $\bar{q} > q_{crit}$.

REFERENCES

[1] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Optimal power management in wireless control systems," in *American Control Conference, 2013*. To appear.

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

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

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

[5] O. C. Imer, S. Yüksel, and T. Başar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429–1439, 2006.

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

[7] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, "Feedback stabilization over signal-to-noise ratio constrained channels," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1391–1403, 2007.

[8] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, 2004.

[9] G. Nair, F. Fagnani, S. Zampieri, and R. Evans, "Feedback control under data rate constraints: An overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[10] K. J. Astrom and B. M. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems," in *Proc. of the 41st IEEE Conference on Decision and Control*, vol. 2, pp. 2011–2016, 2002.

[11] M. Rabi, G. V. Moustakides, and J. S. Baras, "Multiple sampling for estimation on a finite horizon," in *Proc. of the 45th IEEE Conference on Decision and Control*, pp. 1351–1357, 2006.

[12] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1680–1685, 2007.

[13] M. Rabi, K. Johansson, and M. Johansson, "Optimal stopping for event-triggered sensing and actuation," in *Proc. of the 47th IEEE Conference on Decision and Control.*, pp. 3607–3612, 2008.

[14] A. Anta and P. Tabuada, "To sample or not to sample: Self-triggered control for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2030–2042, 2010.

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

[16] A. Mesquita, J. Hespanha, and G. Nair, "Redundant data transmission in control/estimation over lossy networks," *Automatica*, vol. 48, pp. 1020–1027, 2012.

[17] R. Cogill, S. Lall, and J. Hespanha, "A constant factor approximation algorithm for event-based sampling," in *Proc. of the 2007 American Control Conference*, pp. 305–311, 2007.

[18] L. Li and M. Lemmon, "Performance and average sampling period of sub-optimal triggering event in event triggered state estimation," in *Proc. of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC 2011)*, pp. 1656–1661, 2011.

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

[20] C. Ramesh, H. Sandberg, L. Bao, and K. Johansson, "On the dual effect in state-based scheduling of networked control systems," in *Proc. of the 2011 American Control Conference*, pp. 2216–2221, 2011.

[21] L. Bao, M. Skoglund, and K. H. Johansson, "Iterative encoder-controller design for feedback control over noisy channels," *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 265–278, 2011.

[22] A. Goldsmith, *Wireless communications*. Cambr. Univ. Press, 2005.

[23] D. E. Quevedo, A. Ahlén, A. S. Leong, and S. Dey, "On Kalman filtering over fading wireless channels with controlled transmission powers," *Automatica*, vol. 48, no. 7, pp. 1306–1316, 2012.

[24] A. S. Leong and S. Dey, "Power allocation for error covariance minimization in Kalman filtering over packet dropping links," in *Proc. of the 51st IEEE Conference on Decision and Control*, pp. 3335–3340, 2012.

[25] J. Zhang, K. Tan, J. Zhao, H. Wu, and Y. Zhang, "A practical SNR-guided rate adaptation," in *INFOCOM 2008. The 27th Conference on Computer Communications.*, pp. 2083–2091, 2008.

[26] N. Ploplys, P. Kawka, and A. Alleyne, "Closed-loop control over wireless networks," *IEEE Control Systems Magazine*, vol. 24, no. 3, pp. 58–71, 2004.

[27] T. K. Moon, *Error correction coding: mathematical methods and algorithms*. Wiley, 2005.

[28] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 2005.

[29] O. Hernández-Lerma and J. Lasserre, *Discrete-time Markov control processes: basic optimality criteria*. Springer New York, 1996.

[30] O. Vega-Amaya, "The average cost optimality equation: a fixed point approach," *Bol. Soc. Mat. Mexicana*, vol. 9, no. 1, pp. 185–195, 2003.