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 a sensor transmits plant state information over a wireless fading channel to a controller physically separated from the sensor. The power allocated to these transmissions determines the probability of successful reception and is adapted to channel and plant state in order to conserve the sensor's energy resources. Our goal is to design plant control and power management policies to minimize an infinite horizon cost combining power consumption with the conventional linear quadratic regulator control cost. A method to separate the designs of plant inputs and transmitting powers is provided. The resulting optimal controller is the standard LQR control law while the optimal communication policy follows from a Markov decision process problem accounting for power at the sensor and state estimation error at the controller. The features of the optimal power management for general forward error correcting are examined qualitatively. In the particular case of transmissions protected with capacity achieving codes, conventional event-triggered policies are recovered, where the decision is whether to transmit or not. Further a suboptimal communication policy is computed using approximate dynamic programming and its behavior is validated in simulations and contrasted to other simple transmission policies.

I. INTRODUCTION

The Networked Control Systems (NCS) studied in this paper involve communication of plant state information from sensor to controller over a wireless channel. The more information the sensor conveys the more precise actuation becomes, but the resulting increase in transmitted power rapidly depletes the sensor's energy resources. To quantify the emerging tradeoff between plant performance and power consumption, we study plant input design and power control that minimize a joint cost accounting for both plant regulation and communication penalty.

Early works on NCS ignore communication costs and focus their analysis on the performance of control loops when various communication effects are taken into account, see e.g., [1]–[3] and references therein. Fundamental limits like the minimum bit rate for stabilization over bandlimited channels are also known; e.g., [4]. Efforts to regulate the communication cost include the notions of event-triggered [5] or self-triggered [6] control. In both cases the idea is to prolong the interval between successive communications by avoiding transmission as long as a plant performance level is guaranteed. Such triggering rules implicitly reduce the communication cost but transmission expenses are not explicitly accounted.

Explicit account of communication cost in the context of remote state estimation, when plant control is not part of the

design, can be found in [7]–[9]. The sensor measures the plant state and decides whether to transmit its value to an estimator or not. Transmissions incur a fixed cost and the overall goal is to minimize the combination of an estimation error cost and the communication penalties. Related contributions consider jointly optimal plant and communication controllers using again a fixed cost per transmission [10]. The problem is more complex than the case of simple state estimation but a separation principle for finite horizon costs can be established [11], and optimal plant and transmission controls can be found by dynamic programming.

In this paper instead of transmissions we penalize the resources used by the sensor to communicate, that is the transmitted power. This modeling also allows to take into account fading effects in the wireless channel. Power is allowed to adapt to plant state and channel fading and affects the likelihood of successful packet decoding through a known complementary error function (Section II). This communication framework generalizes the well-studied erasure model with i.i.d. dropouts (see e.g. [1], [2]) as the probability of packet drop is now controlled by the selected transmitted power. Quantization effects are ignored in this paper but have been introduced for remote state estimation in [12]. By proper joint selection of plant and power control policies we seek to minimize an aggregate infinite horizon cost combining power consumption with a linear quadratic regulator (LQR) cost (Section II-A).

A restricted information structure is identified (Section II-B), by which the usual LQR control law becomes optimal while the optimal communication policy follows from a Markov decision process (MDP) formulation involving transmitted power and the state estimation error at the controller (Section III). We leverage this separation principle to express optimal power control policies in terms of a value function (Section IV). While this does not allow computation of optimal policies it does provide a qualitative characterization. Our work can be considered as a generalization of the work in [7]–[10] where instead of scheduling based on plant state we allow for power control policies adapted to the channel and plant state. Moreover, conventional event-triggered policies reminiscent of [5] where one just decides whether to transmit or not emerge as the optimal communication strategy at the theoretical limit if the sensor uses capacity achieving forward error correcting (FEC) codes (Section IV-A). Further, suboptimal power control policies are derived using a rollout algorithm (Section V) and numerical simulations show how they adapt power to plant and channel state to regulate estimation error at the controller (Section VI). We close the paper with conclusions and suggestions for further work (Section VII).
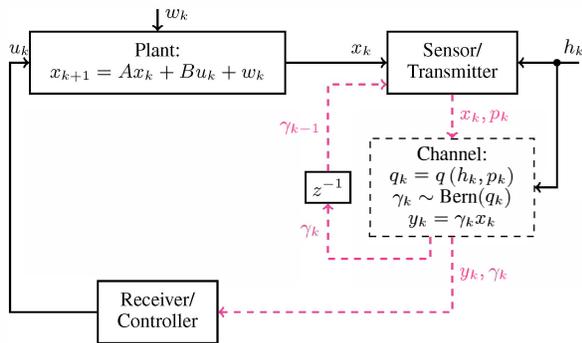
Fig. 1. Wireless control system architecture. A sensor measures the plant and channel states x_k, h_k and transmits with power p_k over a wireless fading channel. Messages are successfully decoded at the controller with probability q_k depending on h_k and p_k. The sensor receives acknowledgments with a one-step delay.

II. PROBLEM FORMULATION

Consider the architecture shown in Fig. 1 controlling a discrete-time linear time-invariant plant described by

$$x_{k+1} = Ax_k + Bu_k + w_k, \; k \geq 0, \qquad (1)$$

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) Gaussian random variables $w_k \in \mathbb{R}^n$ with zero mean and covariance W. We assume the plant is unstable but controllable.

The NCS considered in this paper includes a sensor/transmitter collecting state observations x_k that it communicates with power $p_k \in [0, p_{\max}]$ through a wireless fading channel with coefficient h_k. At the other side of the channel the receiver/controller uses the received information to determine the control input u_k applied to the plant.

Due to propagation effects the channel coefficient h_k changes unpredictably [13, Ch. 3]. We adopt a block fading model whereby channels $\{h_k, k \geq 0\}$ are modeled as i.i.d. random variables of some known distribution m_H on $\mathbb{R}^+$, independent of the noise process $\{w_k, k \geq 0\}$, with the technical assumption that m_H has a probability density function. The transmitter measures the value of h_k before each transmission, e.g. by a pilot signal, however perfect channel knowledge is not necessary (see Remark 1).

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 $h_k p_k$. Successful decoding of the transmitted packet is determined by the signal to noise ratio (SNR) at the receiver defined as $\text{SNR}_k := h_k p_k / N_0$. 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

$$q_k = q(h_k, p_k), \qquad (2)$$

mapping SNR_k to the probability q_k. We assume that $q(h, p)$ is a known increasing function of the product $h\,p$.
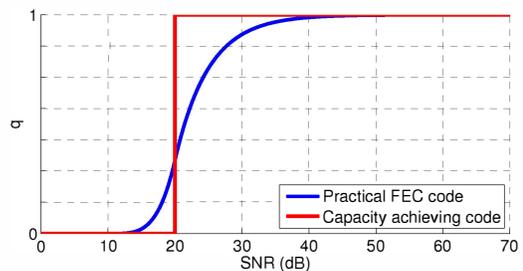

Fig. 2. Complementary error function for practical FEC and capacity achieving codes. The probability of successful decoding q for a FEC code is a sigmoid function of the received $\text{SNR} = h\,p/N_0$, while for a capacity achieving code decoding depends on a threshold value SNR_0.

Considering packet decoding as a part of the communication process, we can model communication as a sequence of successful decoding indicator variables γ_k taking value $\gamma_k = 1$ when information is correctly decoded and $\gamma_k = 0$ otherwise. Variables $\gamma_k \sim \text{Bern}(q_k)$ are Bernoulli distributed with time-varying success probabilities q_k. Ignoring quantization errors, the controller receives a signal $y_k = \gamma_k x_k$ and feedbacks γ_k 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. We assume the receiver also gets γ_k so that the cases $x_k = 0$ and $\gamma_k = 0$ are distinguishable.

In the next section we present the formal problem statement of control inputs u_k and transmitting powers p_k design based on information available at time k.

Remark 1. Perfect channel knowledge is not required in our framework. If h_k is only an estimate of the fading coefficient the complementary error function (2) captures the uncertainty over channel realization as well. 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 [14], [15]. The typical shape of $q(h_k, p_k)$ is a sigmoid function of $h_k p_k$ with exponential tails as shown in Fig. 2. In the theoretical limit correct decoding depends on the channel capacity $C_k = W \log_2(1 + \text{SNR}_k)$, with W the channel bandwidth. Packet transmitted at a rate smaller than C_k bits per second are almost surely successfully decoded, and almost surely incorrectly decoded otherwise. Thus, we can write the successful decoding probability as the indicator

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

for some threshold SNR_0. With α samples per second and β bits per sample x_k we require a transmission rate of $\alpha\beta$ bits per second so $\text{SNR}_0 = 2^{\alpha\beta/W} - 1$. This limit requires an infinitely long code with prohibitive delay, but delays are ignored in this paper. Our interest in (3) is conceptual as event-triggered communication [7] will be shown to arise from the use of capacity achieving codes – see Section IV-A. The form of (3) is shown in Fig. 2. $\square$

A. Joint optimal design of plant and power control

To formulate the joint design of plant controller and power management we introduce an equivalent architecture
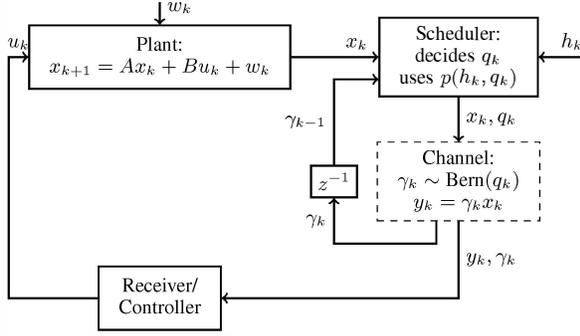
Fig. 3. Equivalent wireless control system architecture. A scheduler decides the successful decoding probability q_k and transmits the state measurement x_k with the required power $p_k = p(h_k, q_k)$.

resembling those considered in, e.g., [7], [10]. In view of (2), choosing p_k is equivalent to choosing the desired probability of successful delivery q_k at time k and transmitting with the minimum required power to achieve this q_k, namely

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

We now interpret q_k as the decision variable with $p(h_k, q_k)$ the cost of decision q_k. This equivalent NCS architecture is shown in Fig. 3 where a scheduler responsible for deciding q_k replaces the sensor/transmitter of Fig 1.

The assumed monotonicity of the function $q(h, p)$ on the product $h\,p$ implies that the $p(h, q)$ is increasing in q and decreasing in h. Using maximum power $p_{\max}$, the transmitter can achieve a maximum success 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. For all channel realizations h, the function $p(h, q)$ in (4) is continuous in the successful decoding probability variable q.

Assumption 2. For all channel realizations h, the largest achievable successful decoding probability $q_{\max}(h)$ satisfies $q_{\max}(h) > q_{crit} := 1 - 1/\max_i |\lambda_i(A)|^2$.

Assumption 1 is of a technical nature and will be used in Theorem 2. The value q_{crit} in Assumption 2 is the necessary and sufficient success probability of i.i.d. Bernoulli γ_k for stability of the estimation error in (11) (see e.g. [1, Th. 2]). In our case success probabilities are controlled and we require q_{crit} to be feasible for all channel realizations h.

In Fig. 3 the decision q_k is a causal function of information available at the sensor, while the plant control signal u_k is a causal function of the information available at the controller. The sequence $\pi := \{q_0, q_1, \ldots\}$, or equivalently the power allocation $\{p_0, p_1, \ldots\}$, is termed the communication policy, whereas $\theta := \{u_0, u_1, \ldots\}$ denotes the control policy. With fixed policies π, θ all random variables are defined on an appropriate probability space. Let the measure be $\mathbb{P}^{\pi,\theta}$ and $\mathbb{E}^{\pi,\theta}$ the integration with respect to that, 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 and a communication cost. Let the control cost be of a standard LQR form

$$J_{\text{LQR}}^N(\pi, \theta) := \mathbb{E}^{\pi,\theta} \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k, \tag{5}$$

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

$$J_{\text{PWR}}^N(\pi, \theta) := \mathbb{E}^{\pi,\theta} \sum_{k=0}^{N-1} p(h_k, q_k). \tag{6}$$

To quantify the tradeoff between plant performance and power consumption we combine the LQR cost in (5) and the power cost in (6) into the limit aggregate cost

$$J(\pi, \theta) := \lim_{N \to \infty} 1/N \left[J_{\text{LQR}}^N(\pi, \theta) + \lambda J_{\text{PWR}}^N(\pi, \theta) \right], \tag{7}$$

for some weight $\lambda > 0$. The problem of joint design of plant and power control asks for policies π and θ that minimize (7). These 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/receiver at time k just before deciding the input u_k. This information includes the history of transmission success variables $\gamma_{0:k} := \{\gamma_0, \ldots, \gamma_k, \}$ and received signals $y_{0:k} := \{y_0, \ldots, y_k, \}$, i.e.,

$$O_k := \{\gamma_{0:k}, y_{0:k}\}. \tag{8}$$

Given the possibility of lost packets, the controller has partial information on the plant state x_k so we study the MMSE estimate $\mathbb{E}^{\pi,\theta}(x_k|O_k)$. This estimation is difficult because the event $\gamma_k = 0$ contains information about the state x_k through the dependence of the probability q_k on its value. To avoid this complication we discard the information given by events of the form $\gamma_k = 0$. Formally, define $\tau_k := \sup\{l \leq k : \gamma_l = 1\}$ as last successful transmission time by time k with the convention that $\tau_0 := 0$, and define the sequence of σ-fields

$$G_k := O_{\tau_k} = O_{\sup\{l \leq k : \gamma_l = 1\}}. \tag{9}$$

When $\gamma_k = 1$, G_k coincides with O_k. When $\gamma_k = 0$, G_k only contains information available until the last successful transmission which occurred at time $\tau_k < k$.

We restrict attention to control policies u_k that are functions of G_k, or more formally measurable with respect to the σ-field G_k, and denote the set of all such policies as 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 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). Thus the estimate evolves by

$$\hat{x}_k := \mathbb{E}^{\pi,\theta}(x_k|G_k) = \begin{cases} y_k & \text{if } \gamma_k = 1, \\ A\hat{x}_{k-1} + Bu_{k-1} & \text{if } \gamma_k = 0 \end{cases} \tag{10}$$
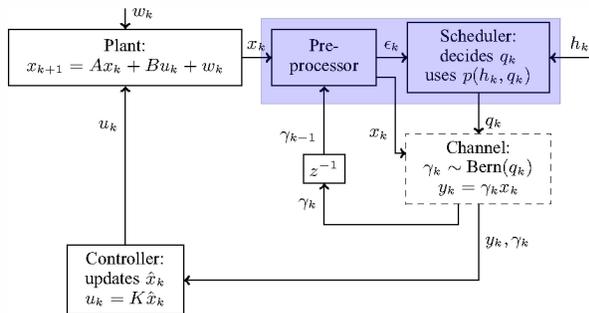
Fig. 4. Wireless control system with restricted information structure. A pre-processor computes the error ϵ_k given x_k and the acknowledgment γ_{k-1}. A scheduler decides q_k based on ϵ_k and the channel h_k, and transmits x_k with the required power $p(h_k, q_k)$. The controller receives the message with probability q_k, updates the estimate $\hat{x}_k$ and provides input u_k.

and the corresponding estimation error $e_k := x_k - \hat{x}_k$ by

$$e_k = (1 - \gamma_k)(Ae_{k-1} + w_{k-1}), \qquad (11)$$

with $e_0 = 0$ since we assumed x_0 is given.

The sensor at time k has access to the channel realization h_k and the plant state x_k which allows selection of the success probability q_k to depend on x_k. This however affects the controller design, because when the controller decides u_{k-1} to control x_k, it should consider the indirect effect on q_k. This information structure makes the joint design problem in (7) hard to analyze. To overcome this we adopt [11] and restrict transmission policies to depend on the channel state h_k and the sensor's belief on what the future estimation error at the controller side will be. In particular consider by (10) the value of the estimation error e_k if the kth packet is *not* successfully decoded, that is

$$\epsilon_k := x_k - (A\hat{x}_{k-1} + Bu_{k-1}). \qquad (12)$$

Observe that ϵ_k is known to the sensor, since x_k is measured and by the acknowledgment mechanism the controller's previous estimate $\hat{x}_{k-1}$ and input u_{k-1} can also be computed. Then $e_k = (1 - \gamma_k)\epsilon_k$, so the sensor predicts that the estimation error e_k is $e_k = 0$ with probability q_k and $e_k = \epsilon_k$ with probability $1 - q_k$. We restrict information at the sensor side to the set F_k defined as the channel history $h_{0:k} := \{h_0, \ldots, h_k, \}$ and the history of error beliefs $\epsilon_{0:k} := \{\epsilon_0, \ldots, \epsilon_k, \}$, i.e.,

$$F_k := \{\epsilon_{0:k}, h_{0:k}\}. \qquad (13)$$

Communication policies are restricted to be measurable with respect to F_k. Call the set of all such policies Π. This information structure is depicted in Fig. 4. The sensor block is split into a pre-processor and a scheduler. The preprocessor gets the sample x_k and the acknowledgment γ_{k-1} 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 $\theta \in \bullet$ that minimize (7), that is

$$J^* := \inf_{\pi \in \Pi, \theta \in \bullet} J(\pi, \theta). \qquad (14)$$

In particular, the next section shows that the information structure we introduced allows 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.

III. SEPARATION OF DESIGNS

With the imposed restrictions on the information available at sensor and controller, the control law $\theta \in \bullet$ and the communication policy $\pi \in \Pi$ can be designed separately. To show this, let us substituting x_k by (1) in (12) to get $\epsilon_k = Ae_{k-1} + w_{k-1}$. Then by $e_{k-1} = (1 - \gamma_{k-1})\epsilon_{k-1}$,

$$\epsilon_k = (1 - \gamma_{k-1})A\epsilon_{k-1} + w_{k-1}, \qquad (15)$$

with initial value $\epsilon_0 = 0$. The key argument to why separation works is that $\{w_k, h_k, k \geq 0\}$ are chosen independently, and the processes $\{\epsilon_k, q_k, \gamma_k, e_k, k \geq 0\}$ are completely determined by the communication policy $\pi \in \Pi$ and do not depend on the control policy $\theta \in \bullet$. We show this by induction. 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. Then at time k, ϵ_k by (15) depends on $\gamma_{k-1}, \epsilon_{k-1}, w_{k-1}$ which are not affected by past control inputs. Also by the restriction $q_k \in F_k$, q_k depends on $\epsilon_{0:k}, h_{0:k}$ so it is not affected by control inputs. Finally $\gamma_k \sim \text{Bern}(q_k)$, and e_k by (11) are also independent of control inputs.

Consequently the power cost $J_{\text{PWR}}^N(\pi, \theta)$ in (6) is not affected by the control policy θ. We can rewrite then the optimum cost in (14) as

$$J^* = \inf_{\pi \in \Pi} \left\{ \inf_{\theta \in \Theta} \lim_{N \to \infty} \frac{1}{N} J_{\text{LQR}}^N(\pi, \theta) + \lambda \lim_{N \to \infty} \frac{1}{N} J_{\text{PWR}}^N(\pi) \right\} \qquad (16)$$

Using (16), optimal control policies $\theta \in \bullet$ for given communication policy $\pi \in \Pi$ can be found by the infimum of the limit LQR cost $\lim_{N \to \infty}(1/N)J_{\text{LQR}}^N(\pi, \theta)$. As it turns out, we can also prove that the form of the optimal controller does not depend on the communication policy leading to a stronger separability result than what follows from (16). The proof of this result requires the technical assumption that the limit and infimum operator can be interchanged in the optimal LQR cost for any given communication policy $\pi \in \Pi$, i.e.,

$$\inf_{\theta \in \Theta} \left\{ \lim_{N \to \infty} \frac{1}{N} J_{\text{LQR}}^N(\pi, \theta) \right\} = \lim_{N \to \infty} \inf_{\theta \in \Theta} \left\{ \frac{1}{N} J_{\text{LQR}}^N(\pi, \theta) \right\}. \qquad (17)$$

This assumption holds true in most cases of practical interest. The separability result is stated in the following theorem. [1]

Theorem 1 (Optimal control policy). *Consider the networked control system of Fig. 4 with control policies $\theta := \{u_0, u_1, \ldots\} \in \bullet$ composed of actions u_k measurable with respect to the set G_k in (9), and communications policies $\pi := \{q_0, q_1, \ldots\} \in \Pi$ selecting q_k measurable with respect*

---

[1]Due to space limitations, the proofs of the theorems are omitted in this paper and can be found in [16].

to the set F_k in (13). Further assume (17) holds. The optimal joint communication and control cost J^* by (5)-(7), (14) with respect to policies $\theta \in \Theta$ and $\pi \in \Pi$ satisfies

$$J^* = Tr(PW) + \inf_{\pi \in \Pi} \lim_{N \to \infty} \frac{1}{N} \mathbb{E}^\pi \sum_{k=0}^{N-1} e_k^T \tilde{P} e_k + \lambda J_{PWR}^N(\pi) \tag{18}$$

where P is the solution to the algebraic Riccati equation $P = A^T PA + Q - A^T PB(R + B^T PB)^{-1}B^T PA$ for the system in (1) and the linear quadratic regulator cost (LQR) in (5), and the matrix $\tilde{P}$ is defined as

$$\tilde{P} := A^T PA + Q - P. \tag{19}$$

Furthermore, the optimal control policy that achieves the infimum in (14) and (18) is the linear controller

$$u_k = -K\hat{x}_k, \tag{20}$$

with $K := (R + B^T PB)^{-1}B^T PA$ being the steady state LQR gain.

The statement in (20) of Theorem 1 determines the optimal control policy θ as the conventional LQR controller, which is shown in Fig. 4. The second summand in (18) shows that the optimal communication policy jointly regulates power consumption and a weighted controller's estimation error $e_k^T \tilde{P} e_k$.

Observe that as per (10) and (11) it holds that $e_k = (1 - \gamma_k)\epsilon_k$. Also $\mathbb{E}^\pi[\gamma_k|F_k] = \mathbb{P}^\pi[\gamma_k = 1|F_k] = q_k$ and $\epsilon_k \in F_k$. So we can write

$$\mathbb{E}^\pi[e_k^T \tilde{P} e_k|F_k] = \mathbb{E}^\pi[(1 - \gamma_k)\epsilon_k^T \tilde{P}\epsilon_k|F_k] = (1 - q_k)\epsilon_k^T \tilde{P}\epsilon_k. \tag{21}$$

and taking the expectation in both sides gives

$$\mathbb{E}^\pi[e_k^T \tilde{P} e_k] = \mathbb{E}^\pi[(1 - q_k)\epsilon_k^T \tilde{P}\epsilon_k]. \tag{22}$$

Substituting the expression (22) into the second summand of (18) it follows that the optimal communication policy $\pi \in \Pi$ is the one that achieves the infimum cost

$$J_{COM}^* := \inf_{\pi \in \Pi} \lim_{N \to \infty} \frac{1}{N} \mathbb{E}^\pi \sum_{k=0}^{N-1} c(\epsilon_k, h_k, q_k), \tag{23}$$

where we define the cost-per-stage to be

$$c(\epsilon, h, q) := (1 - q)\epsilon^T \tilde{P}\epsilon + \lambda p(h, q). \tag{24}$$

The difference between the infimum in (18) and the formulation (23) is that in the former e_k is not known at the sensor at time k, while ϵ_k in the latter is. Thus (23) takes the form of a MDP problem with an infinite horizon average cost criterion. The state of the problem at time k is the pair $(\epsilon_k, h_k) \in \mathbb{R}^n \times \mathbb{R}^+$ and the available action is $q_k \in [0, q_{\max}(h_k)]$. The state transition probabilities can be obtained from (15) and are given by

$$\mathbb{P}(\epsilon^+, h^+|\epsilon, h, q)$$
$$= [q \, \mathcal{N}_{0,W}(\epsilon^+) + (1 - q) \, \mathcal{N}_{A\epsilon, W}(\epsilon^+)] \, m_H(h^+), \tag{25}$$

where $\mathcal{N}_{\mu, \Sigma}$ is the n-dimensional Gaussian distribution with mean μ and covariance Σ. 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 (15) on the event $\gamma = 1$, $\epsilon^+ = w \sim \mathcal{N}_{0,W}$, while on the event $\gamma = 0$, $\epsilon^+ = A\epsilon + w$ with $w \sim \mathcal{N}_{0,W}$, which is equivalent to $\epsilon^+ \sim \mathcal{N}_{A\epsilon, W}$. Since h^+ is independent of ϵ, h, ϵ^+, its distribution m_H appears as a product in (25). Let us denote $\mathbb{E}[\epsilon^+, h^+ | \epsilon, h, q]$ the integration with respect to the above transition probability measure.

IV. OPTIMAL COMMUNICATION POLICY

Using (23) we can show that optimal communication policies achieving the infimum of the co-design problem in (14) exist. This existence result provides a characterization of these policies from which we infer the general features of optimal transmitted powers p_k and corresponding success probabilities q_k as a function of estimation error beliefs ϵ_k and channel realizations h_k.

Existence of optimal policies for infinite-horizon MDPs on general Borel spaces requires some technical assumptions. Here we select a sufficiently large constant L and restrict consideration to stationary policies $q_k \in Q(\epsilon_k, h_k)$ satisfying

$$Q(\epsilon, h) := \begin{cases} [0, q_{max}(h)] & \text{if } \|\epsilon\| < L, \\ q_{max}(h) & \text{otherwise,} \end{cases} \tag{26}$$

This restriction is inconsequential as we may pick L large enough so as to make its effect arbitrarily small. With this restriction we adapt the fixed point approach [17] used in [9] to determine the characterization of the optimal communication policy stated in the following theorem.

Theorem 2 (Optimal communication policy). *Consider the Markov decision process with optimal cost as in (23), state transition probabilities as in (25), and a stationary policy of the form $q_k \in Q(\epsilon_k, h_k)$ with $Q(\epsilon, h)$ abiding to (26). Under assumptions 1,2 there exists a function $V : \mathbb{R}^n \times \mathbb{R}^+ \mapsto \mathbb{R}$, with $V(0, \tilde{h}) = 0$ for some $\tilde{h} \in \mathbb{R}^+$, such that for all $\epsilon \in \mathbb{R}^n$ and $h \in \mathbb{R}^+$ it satisfies*

$$V(\epsilon, h) = \min_{q \in Q(\epsilon, h)} \{c(\epsilon, h, q) - J_{COM}^* + \mathbb{E}\left[V(\epsilon^+, h^+) | \epsilon, h, q\right]\}. \tag{27}$$

The optimal communication cost is given by $J_{COM}^* = \mathbb{E}_{w,h} V(w, h)$, where $\mathbb{E}_{w,h}$ denotes integration with respect to the product measure $\mathcal{N}_{0,W} \times m_H$. The optimal communication policy $q_k^* \in Q(\epsilon_k, h_k)$ is the one achieving the minimum in the right hand side of (27) and it can be written as

$$q^*(\epsilon, h) = \operatorname*{argmin}_{q \in [0, q_{max}(h)]} \lambda p(h, q) + (1 - q) R(\epsilon), \tag{28}$$

when $\|\epsilon\| < L$, and $q^*(\epsilon, h) = q_{max}(h)$ when $\|\epsilon\| \geq L$, where

$$R(\epsilon) := \epsilon^T \tilde{P}\epsilon + \mathbb{E}_{w,h}[V(A\epsilon + w, h) - V(w, h)]. \tag{29}$$

In the context of infinite horizon MDPs the function $V(\epsilon, h)$ in (27) is called relative value function and informally it represents the expected cost incurred starting at state (ϵ, h) and following the optimal policy in the future. Observe that V in (27) is defined only up to a constant but we fixed its value at $V(0, \tilde{h}) = 0$. In principle one can
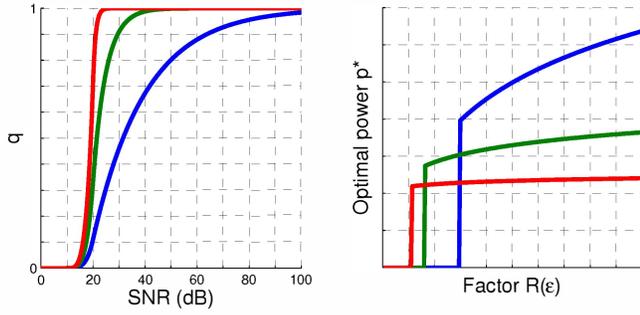
Fig. 5. Optimal transmitting power p^* for FEC codes with different complementary error functions. The optimal p^* is plotted as a function of the factor $R(\epsilon)$ for a fixed channel state h using FEC codes with different q-SNR characteristics. When the q-SNR curve becomes steeper, the optimal power allocation resembles a step function.

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

find $V(\epsilon, h)$ using, e.g., a value iteration algorithm which involves iterative application of (27) [18]. This procedure is, however, computationally onerous as each iteration requires minimizing the right hand side of (27) for all possible state pairs (ϵ, h). Nevertheless, (27) still gives qualitative information on the optimal policy.

Let us ignore the case $\|\epsilon\| \geq L$ in (26) as it is irrelevant for the following discussion. The optimal policy $q^*(\epsilon, h)$ in (28) 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 (28) to power by (2), (4), and is described by

$$p^*(\epsilon, h) := \min_{0 \leq p \leq p_{max}} \lambda p + (1 - q(h, p)) R(\epsilon). \qquad (30)$$

Despite the fact that $V(\epsilon, h)$ and $R(\epsilon)$ 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 NCS. 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 (30) as a function of $R(\epsilon)$. In all cases, when the error penalty $R(\epsilon)$ is below some threshold, the best option is to not transmit. Above the threshold, the optimal power increases with $R(\epsilon)$. 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.

In Fig. 6 we present qualitative plots of the optimal decoding probability q^* and optimal transmitting power p^* as functions of the factor $R(\epsilon)$ and the channel state h for a given sigmoid q-SNR characteristic. Blue regions correspond to no transmission. The fact that q depends on the product $h\,p$ creates a disparity in the two plots. The optimal q^* increases with the product $h\,R(\epsilon)$. On the other hand, the optimal p^* increases with $R(\epsilon)$, but the rate of increase depends on the quality of the channel h. So when h is bad, the sensor needs higher power to achieve high probability of successful
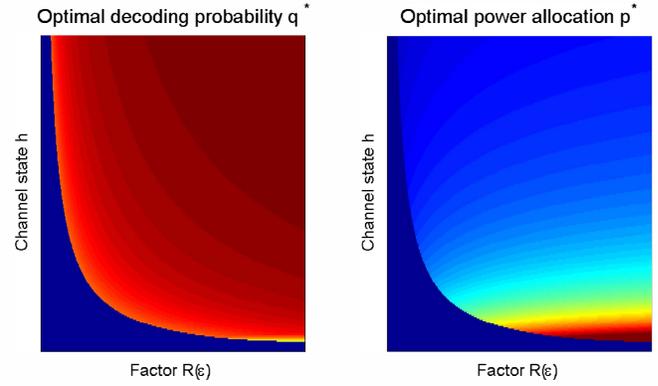
decoding at the receiver.

A. Optimal solution for capacity achieving codes

Consider now the case of capacity achieving codes. By (3), at time k the transmitter needs to allocate either $p_k = 0$ or $p_k = p_0/h_k$ with $p_0 := N_0\text{SNR}_0$, since any other choice is unfavorable. Assumption 2 in this case implies that the transmitter has enough power budget $p_{max} \geq p_0/h$ to transmit successfully for every channel state h. Suppose we are looking for a randomized policy, i.e. a distribution on the two power options, and with a slight abuse of notation we denote $q_k \in [0, 1]$ the probability of choosing power p_0/h_k. Then the transmitter draws independent $\gamma_k \sim Bern(q_k)$ 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 is

$$\mathbb{E} \sum_{k=0}^{N-1} \gamma_k\, p_0/h_k = \mathbb{E} \sum_{k=0}^{N-1} q_k\, p_0/h_k, \qquad (31)$$

which is of the form (6) with $p(h, q) = q\, p_0/h$. In this case the minimization in (28) becomes linear in q, thus the optimal communication policy is deterministic,

$$q^{CA}(\epsilon, h) := \begin{cases} 0 & \text{if } h\,R(\epsilon) \leq \lambda p_0 \\ 1 & \text{otherwise} \end{cases}, \qquad (32)$$

or in terms of power

$$p^{CA} := \begin{cases} 0 & \text{if } h\,R(\epsilon) \leq \lambda p_0 \\ p_0/h & \text{otherwise} \end{cases}. \qquad (33)$$

This is an event-triggered scheme along the lines of [7], 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 transmission-triggering event is $h\,R(\epsilon) > \lambda p_0$. Qualitatively this shows that when the channel is in a good state, transmitting is worthy since it does not cost much, while when a measure $R(\epsilon)$ of the error is large, it is necessary to transmit and reset it to zero.

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

In the previous section we showed by Theorem 2 that the optimal communication policy $q*(\epsilon, h)$ minimizes in the right hand side of (27) including the optimal value funciton V. A rollout algorithm [19] to suboptimally choose actions relies on some reference communication policy π for which the corresponding relative value function $V^\pi(\epsilon, h)$ is known. Assuming this reference policy will be followed at all future steps, the best current action, minimizing the current stage cost and the future expected cost is

$$q_k^{\text{roll}}(\epsilon, h) := \underset{0 \le q \le q_{m \bullet x}(h)}{\operatorname{argmin}} \ c(\epsilon, h, q) + \mathbb{E}\left[V^\pi(\epsilon^+, h^+)|\epsilon, h, q\right]$$
(34)

To find a family of policies with computable relative value functions suppose we adapt $q_k = q(h_k)$ to the current channel state h_k but not to the error ϵ_k. 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 $\gamma_k \sim \text{Bern}(\bar{q})$ for all k. Since the expected power consumption $\mathbb{E}_h p(h, q(h))$ at every stage is also constant the MDP cost of this policy by (23) becomes

$$J^{q(\cdot)} := \lim_{N \to \infty} \frac{1}{N} \mathbb{E} \sum_{k=0}^{N-1} (1 - \bar{q}) \epsilon_k^T \tilde{P} \epsilon_k + \lambda \mathbb{E}_h p(h, q(h)).$$
(35)

For any policy of the form $q_k = q(h_k)$ the corresponding relative value function $V^{q(\cdot)}(\epsilon, h)$ and cost $J^{q(\cdot)}$ can be determined in closed form as stated in the following theorem.

Theorem 3 (Cost of channel-adaptive communication policies). *Consider the Markov decision process with state pair (ϵ, h) and state transition probabilities as in (25). Consider policies $q(\cdot)$ for which the success transmission probability is selected as a function $q(h)$ independent of the estimation error belief ϵ. For any such policy with $\bar{q} := \mathbb{E}_h q(h) > q_{crit}$ for the critical probability q_{crit} of Assumption 2, the cost $J^{q(\cdot)}$ in (35) becomes*

$$J^{q(\cdot)} = Tr(\tilde{P}E) + \lambda \mathbb{E}_h p(h, q(h)),$$
(36)

where the matrix E is the unique solution of

$$E = (1 - \bar{q})(AEA^T + W).$$
(37)

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

$$V^{q(\cdot)}(\epsilon, h) = \frac{1 - q(h)}{1 - \bar{q}} \epsilon^T H \epsilon + \lambda p(h, q(h)),$$
(38)

where the matrix H is the unique solution of

$$H = (1 - \bar{q})(A^T H A + \tilde{P}).$$
(39)

Theorem 3 provides an explicit formula for a family of relative value functions $V^{q(\cdot)}(\epsilon, h)$. Substituting (38) into (34) and removing constants we find the rollout policy

$$q^{\text{roll}}(\epsilon, h) := \underset{0 \le q \le q_{m \bullet x}(h)}{\operatorname{argmin}} (1 - q) \frac{\epsilon^T H \epsilon}{1 - \bar{q}} + \lambda p(h, q).$$
(40)

Given a parameter $\bar{q}$ modeling suboptimal future actions, we can compute H by (39) and then solve (40) given the
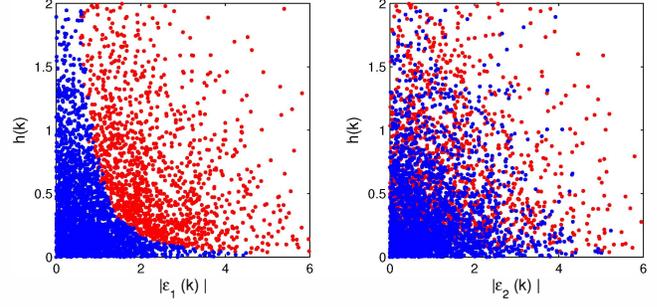

Fig. 7. Simulation points $|\epsilon_1(k)|, h(k)$ and $|\epsilon_2(k)|, h(k)$ of the rollout policy for a capacity achieving code. Blue indicates the decision to not transmit ($q_k = 0$), red to transmit ($q_k = 1$).

function $p(h, q)$. Observe that (40) is of the same form as the optimal communication policy (28) except that the optimal unknown function $R(\epsilon)$ is approximated by the suboptimal quadratic form $\epsilon^T H \epsilon / (1 - \bar{q})$.

For the case of a capacity achieving FEC we can repeat the analysis in Section IV-A to solve (40) and we obtain

$$q^{\text{roll},CA}(\epsilon, h) := \begin{cases} 0 & \text{if } \dfrac{h}{(1 - \bar{q})\lambda p_0} \epsilon^T H \epsilon \le 1 \\ 1 & \text{otherwise} \end{cases} .$$
(41)

This gives us an explicit (suboptimal) event-triggered communication policy, where events depend on current values of channel state h and error ϵ.

VI. SIMULATIONS

We simulate exponentially distributed channel fading with mean 0.5, and choose plant, cost and noise parameters as

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 0.8 \end{bmatrix}, B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, Q = R = W = I, \quad (42)$$

and $\lambda = 50$, $p_0 = 1$. For capacity achieving codes we use a reference communication policy $q(h) = \mathbb{I}(h \ge h_t)$ which transmits whenever channel state exceeds a threshold h_t inducing $\bar{q} \approx 0.79$. Simulations of the resulting rollout policy by (41) reveal a dramatic decrease in the empirical rate of transmissions $1/N \sum_{k=0}^{N-1} \gamma_k \approx 0.37$, which is also much lower than the critical $q_{crit} = 0.75$. Similarly the empirical cost $1/N \sum_{k=0}^{N-1} e_k^T \tilde{P} e_k + \lambda \gamma_k p_0 / h_k \approx 56$ decreased to a half compared to the reference $J^{q(\cdot)} \approx 124$.

The event-triggered nature of the rollout policy (41) is captured in Fig. 7, where we plot $|\epsilon_1(k)|, h(k)$ and $|\epsilon_2(k)|, h(k)$ during simulation. Red and blue points indicate the decision to transmit $q(k) = 0$ and not $q(k) = 1$ respectively. The sensor avoids transmission at low channel fading $h(k)$ requiring large power. Transmissions also adapt to the plant structure. Error $\epsilon_1(k)$ relates to the unstable eigenvalue of A, so the sensor always transmits when $\|e\|$ is away from 0. In contrast such a correlation between transmissions and $\epsilon_2(k)$ is not clear. Even for large $\epsilon_2(k)$ the sensor chooses not to transmit. As long as $\epsilon_1(k)$ is kept bounded, $\epsilon_2(k)$ will also be bounded as it relates to a stable eigenvalue.

Next for the same parameters we compare the performance of the rollout with the reference. For different values of λ
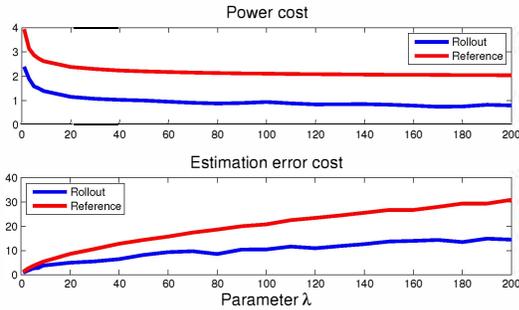
Fig. 8. Comparison of the optimal threshold policy and the resulting rollout policy performance for a capacity achieving code. Power and estimation costs are plotted for different values of the weight λ.
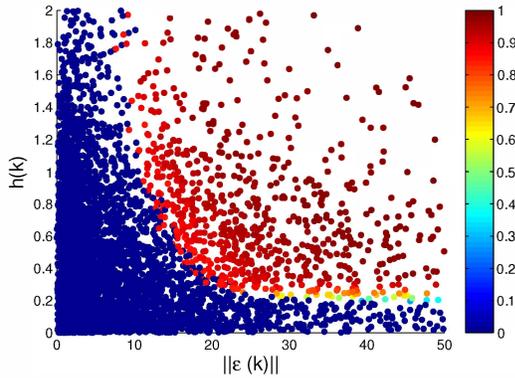

Fig. 9. Simulation points $\|\epsilon(k)\|, h(k)$ of the rollout policy for a FEC code. Colors denote the magnitude of the chosen decoding probability $q(k)$.

we find the optimal threshold policy $q(h) = \mathbb{I}(h \geq h_t)$, i.e. the one minimizing (36), and the corresponding rollout (41). The resulting power and estimation costs by simulating the rollout are plotted separately in Fig. 8 along with the costs of the reference policy. Larger λ decreases power consumption since it is penalized more in the aggregate cost (23), and the decrease rate is similar for both policies. On the other hand estimation cost increases too since the sensor transmits less often. However the increase for the rollout is slower relative to the reference. The reason is that the reference adapts only to channel so avoids transmission when channel h_k is low. The rollout adapts to both channel and error ϵ_k. Transmitting only when ϵ_k is large results in only a moderate increase in the estimation cost with a similar decrease in power.

Finally, we simulate rollout (40) for the system when a FEC code is employed. The probabilities q_k chosen by the rollout during simulation are plotted in Fig. 9 on $\|\epsilon(k)\|, h(k)$ axes (compare with the optimal policy in Fig. 6). Unlike capacity achieving codes, q_k take values smaller than 1. However, due to the sigmoid form of the q-SNR characteristic of the FEC code, q_k are practically either 0 or very close to 1, especially for high h_k. When h_k is low transmissions with high success probability require large power. Thus the rollout decides either to not transmit ($q_k = 0$) or transmit with q_k very close to 1. The points accumulate at small errors $\|\epsilon(k)\|$ because when the error gets larger, q_k is chosen close to 1 so with very high probability the error is reset.

VII. CONCLUSIONS AND FUTURE WORK

This paper examines the control of a linear plant when sensor and actuator communicate over a wireless fading channel. Fading is captured by a random channel state and the probability of a successfully decoded packet is modeled as a function of the channel state and the power used at transmission. We considered the problem of co-designing transmitting powers and control inputs that minimize an infinite horizon aggregate LQR and power cost. A method to separate the two designs is introduced and the optimal power management is characterized for general FEC codes and capacity achieving codes. Suboptimal power policies are proposed and simulated. Further work focuses on the general control/power co-design problem without separation as well as the problem of power resources management in a multi-sensor/actuator network.

REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] 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.

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

[6] M. Mazo, A. Anta, and P. Tabuada, "An ISS self-triggered implementation of linear controllers," *Automatica*, vol. 46, no. 8, pp. 1310–1314, 2010.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

[12] D. E. Quevedo, A. Ahlén, and J. Østergaard, "Energy efficient state estimation with wireless sensors through the use of predictive power control and coding," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4811–4823, 2010.

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

[14] 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.

[15] 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.

[16] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Optimal power management in wireless control systems," *Technical Report*, 2013. Available at http://www.seas.upenn.edu/~kgatsis/.

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

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

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