Opportunistic sensor scheduling in wireless control systems

Konstantinos Gatsis, Miroslav Pajic, Alejandro Ribeiro, and George J. Pappas

***Abstract*— We consider a wireless control system where multiple power-constrained sensors transmit plant output measurements to a controller over a shared wireless medium. A centralized scheduler grants channel access to a single sensor on each time step. Assuming an a priori designed controller, we design scheduling and transmit power policies that opportunistically adapt to the random wireless channel conditions experienced by each sensor. The objective is to obtain a stable system, by minimizing the expected decrease rate of a given Lyapunov function, while respecting the sensors' power constraints. We develop an online optimization algorithm based on the random channel sequence observed during execution which converges almost surely to the optimal protocol design.**

I. Introduction

Wireless control systems in, e.g., industrial or building automation applications, often involve sensing and actuating devices at different physical locations that communicate control-relevant information over shared wireless mediums. Scheduling access to the medium is critical to avoid interferences between transmissions but also affects the overall control performance. Previous work in wired and/or wireless networked control systems, focused on deriving stability conditions under given scheduling protocols – see, e.g., [1]–[3]. The typical approach is to convert the system in some form of a switching system whence stability properties can be derived [2], often in conjunction with other network phenomena such as communication delays, uncertain communication times, and/or packet drops.

Beyond the question of stability, the problem of designing schedulers suitable for control applications has also been addressed. The proposed protocols can be generally classified as either fixed or dynamic. Typical examples of the first type are periodic protocols, i.e., repeating in a predefined sequence (e.g., round-robin). Fixed protocols leading to stability [4], controllability and observability [5], or minimizing linear quadratic objectives [6] have been proposed. Deriving otherwise optimal scheduling sequences is recognized as a hard combinatorial problem [7]. Dynamic scheduler design constitutes a different approach where based on the current plant/control system states, informally speaking, the subsystem with the largest state discrepancy is scheduled to communicate. Examples of such dynamic schedulers can be found in, e.g., [2], [8]–[10].

In this paper we focus on scheduling for wireless control systems and, in contrast to the above approaches, we examine how scheduling can opportunistically exploit the varying channel conditions on the shared wireless medium. In previous work we have shown that such channel-aware designs can be utilized when scheduling independent control tasks whose performance requirements translate to different channel utilization demands [11]. Here we consider the problem of scheduling plant output measurements from sensors to controller when the sensors have limited power resources (Section II). These power resources can be used during transmission to counteract channel fading effects and obtain a higher decoding probability at the receiver/controller [12]. However, the channel fading conditions that a sensor experiences change randomly over time and also differ among sensors [13, Ch. 14]. Hence dynamically assigning access to the sensor currently experiencing, e.g., the most favorable conditions, can save up power. On the other hand, scheduling should lead to a closed loop control system with stability guarantees.

We formulate the design of channel-aware scheduling and power allocation protocols in a stochastic optimization framework (Section II-A), where a protocol is feasible if the sensors' power constraints are met. The objective is to optimize a closed-loop stability margin measured as the decrease rate of a given Lyapunov function, in expectation over the random channel conditions. In Section III we present an optimization algorithm based on the Lagrange dual problem. The algorithm does not require prior knowledge of the channel distribution, and it can be implemented online based on a random observed channel sequence. We show that the algorithm converges almost surely to a feasible protocol, which additionally leads to a stable system if the system is stabilizable with respect to the selected Lyapunov function. We conclude with numerical simulations and a discussion on our results.

Notation: A set of variables $a_0, a_1, \dots a_k$ is denoted compactly as $a_{0:k}$. We denote by $\geq, \succeq, \succ$ the comparison with respect to the cones of the real m-dimensional nonnegative orthant $\mathbb{R}_+^m$, of the real $n \times n$ symmetric positive semi-definite matrices S_+^n, and of the real $n \times n$ symmetric positive definite matrices S_{++}^n respectively. For a matrix M we denote by $\|M\|$ the Frobenius norm.

II. Problem Formulation

We consider the wireless control architecture of Fig. 1 where m sensors measuring plant outputs communicate over a shared wireless medium to the system controller. To avoid interferences between transmissions, a centralized scheduler guarantees that at most one sensor is *scheduled* to access the medium at each time step. Due to uncertainties in the wireless channel, which we will be model in detail next, the transmitted sensor measurements might get lost. We indicate

This work was supported in part by NSF CNS-0931239, and by TerraSwarm, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA. The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, 200 South 33rd Street, Philadelphia, PA 19104. Email: {kgatsis, pajic, aribeiro, pappasg}@seas.upenn.edu.

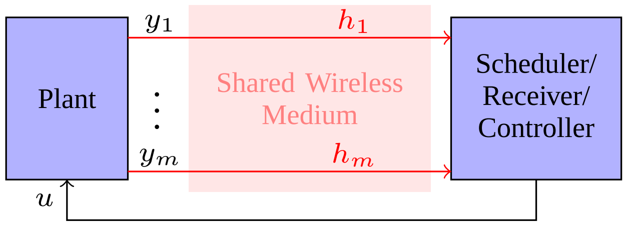

Fig. 1. Sensor scheduling in a wireless control architecture. Each sensor i measures and transmits a plant output y_i to a centralized controller over a shared wireless medium. A scheduler implemented at the receiver/controller opportunistically selects which sensor accesses the channel at each time step based on the random wireless channel conditions $h_1, \dots, h_m$ experienced by the sensors.

with $\gamma_{i,k} \in \{0,1\}$ the event that sensor i is scheduled at the discrete time step k *and* the respective transmission is successful. Let also $\gamma_{0,k} \in \{0,1\}$ denote the event that no sensor transmits successfully at time k, so that $\sum_{i=0}^{m} \gamma_{i,k} = 1$ for all k since these events are disjoint.

Let $x_k \in \mathbb{R}^n$ denote the overall state of the plant and control system before transmission at time k. System evolution from x_k to x_{k+1} depends on whether a transmission occurs at time k and which of the sensors transmits. Suppose the system follows linear dynamics denoted by $A_i \in \mathbb{R}^{n\times n}$ if sensor i transmits successfully ($\gamma_{i,k} = 1$), and $A_0 \in \mathbb{R}^{n\times n}$ when no sensor transmits ($\gamma_{0,k} = 1$). We describe then the system evolution by the switched linear discrete time system

$$x_{k+1} = \sum_{i=0}^{m} \gamma_{i,k} \, A_i x_k + w_k. \tag{1}$$

with w_k modeling an independent identically distributed (i.i.d.) noise process with mean zero and covariance $W \succeq 0$. An example of such a setup follows.

Example 1. Consider a linear continuous time plant

$$\begin{aligned} \dot{x} &= A_p x + B_p u + w, \\ y &= Cx, \end{aligned} \tag{2}$$

perturbed by some white noise process w. Each output $y_i(t)$ of the vector output $y(t) \in \mathbb{R}^m$ is measured by a wireless sensor i, for $i = 1, \dots, m$. Also consider a continuous time dynamic controller

$$\begin{aligned} \dot{z} &= A_c z + B_c \hat{y}, \\ u &= Fz + G\hat{y}, \end{aligned} \tag{3}$$

designed for desirable performance when fed with the plant output $\hat{y} = y$. Due to the wireless sensor communication the controller has access to a perturbed version $\hat{y}$ of the real output y. If at most one sensor measurement can be received on discrete time steps t_k, a standard convention [1], [2] is to update the received output at the controller and hold the remaining ones constant, i.e.,

$$\hat{y}(t_k) = \sum_{i=1}^{m} I_{ii} \left[\gamma_{i,k} C x(t_k) + (1-\gamma_{i,k})\hat{y}(t_{k-1})\right], \tag{4}$$

where I_{ii} is a square matrix with (i,i) being the only non-zero element and equal to 1. If communication is periodic with period T_s the closed loop system can be transformed (see, e.g., [1], [2]) by augmenting the state space as

$$\begin{bmatrix} x_{k+1} \\ z_{k+1} \\ \hat{y}_k \end{bmatrix} = \exp\left(\begin{bmatrix} A_p & B_p F & B_p G \\ 0 & A_c & B_c \\ 0 & 0 & 0 \end{bmatrix} T_s \right) \sum_{i=0}^{m} \gamma_{i,k} \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ I_{ii} C & 0 & I - I_{ii} \end{bmatrix} \begin{bmatrix} x_k \\ z_k \\ \hat{y}_{k-1} \end{bmatrix} + \begin{bmatrix} w_k \\ 0 \\ 0 \end{bmatrix} \tag{5}$$

where $I_{00} \equiv 0$, and index k corresponds to the variables before transmission at time t_k. This is a system of the form (1). The dual architecture where a controller distributes plant inputs to a set of actuators can be similarly formulated. $\square$

In this work the system dynamics are given, i.e., a controller has been already designed, and we focus on designing the wireless communication (scheduling and associated transmit power) which affects the transmission indicators $\gamma_{i,k}$. We describe the wireless channel conditions for link i, between sensor i and the controller, at time k by the channel fading coefficient $h_{i,k}$ that sensor i experiences if it transmits at time k. Due to propagation effects, the channel fading states $h_{i,k}$ change unpredictably [13, Ch. 3] and take values in a subset $\mathcal{H} \subseteq \mathbb{R}_+$ of the positive reals. Channel states $h_{i,k}$ change not only over time k but also between sensors i. We group $h_{i,k}$ for $1 \le i \le m$ at time k in a vector $h_k \in \mathcal{H}^m$, and we adopt a block fading model whereby h_k are random variables independent across time slots k and identically distributed with a multivariate distribution ϕ on $\mathcal{H}^m$. Channel states are also independent of the plant process noise w_k. We assume that h_k are available before transmission – see Remark 1 for a practical implementation. We make the following technical assumption to avoid a degenerate channel distribution, but otherwise no other prior information about the channel distribution will be needed for the communication design in this paper.

Assumption 1. The joint distribution ϕ of channel states h_k has a probability density function on $\mathcal{H}^m$.

If sensor i is scheduled to transmit at time k it selects a transmit power level $p_{i,k} \in [0, p_{\max}]$. Channel fading and transmit power affect the probability of successful decoding of the message at the receiver. In particular, given the forward error-correcting code (FEC) in use, the probability q that a packet is successfully decoded is a function of the received signal-to-noise ratio (SNR). The SNR is proportional to the received power level expressed by the product $h\,p$ of the channel fading state and the allocated transmit power. Overall we express the probability of success by some given relationship of the form $q(h_{i,k}, p_{i,k})$ – see [12] for more details on this model. An illustration of this relationship is shown in Fig. 2. The assumptions on the form of the function $q(hp)$ are the following.

Assumption 2. The function $q(.)$ as a function of the product $r = h\,p$ for $r \ge 0$ satisfies:

(a) $q(0) = 0$,

(b) $q(r)$ is continuous, and strictly increasing when $q(r) > 0$, i.e., for any $r' > r$ it holds that $q(r') > q(r)$,

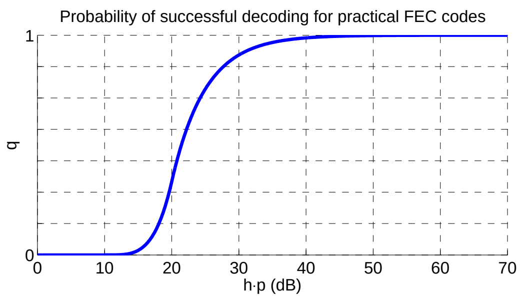

Fig. 2. Complementary error function for practical FEC codes. The probability of successful decoding q for a FEC code is a sigmoid function of the received SNR $\sim h\,p$.

(c) for any $\mu \geq 0$ and for almost all values $h \in \mathcal{H}$ the set $\operatorname{argmin}_{0\leq p\leq p_{\max}} p - \mu q(h\,p)$ is a singleton.

Parts (a),(b) of this assumption state that the probability of successful decoding $q(h\,p)$ will be zero when the received power level $h\,p$ is small, and it becomes positive $q(h\,p) > 0$ and strictly increasing for larger values of $h\,p$. Part (c) is more stringent, stating essentially that $q(h\,p)$ cannot behave linearly in p for a range of channel values h. As shown in Fig. 2 for cases of practical interest $q(h\,p)$ has a sigmoid form and all the above requirements are expected to hold.

Before transmission, a scheduler selects which sensor will access the channel. We allow for randomized scheduling and we denote with $\alpha_{i,k}$ the probability that sensor i is selected at time k. For simplicity we require that exactly one sensor is scheduled, meaning that $\sum_{i=1}^{m} \alpha_{i,k} = 1$. Hence the scheduling decision is a probability vector of the form

$$\alpha_k \in \Delta_m = \left\{ \alpha \in \mathbb{R}^m : \alpha \geq 0, \sum_{i=1}^{m} \alpha_i = 1 \right\}, \qquad (6)$$

Given scheduling $\alpha_k \in \Delta^m$, power allocation $p_k \in [0, p_{\max}]^m$, and channel state $h_k \in \mathcal{H}^m$, we can model the transmission events $\gamma_{i,k}$ as Bernoulli random variables with

$$\mathbb{P}[\gamma_{i,k} = 1 \,|\, h_k, \alpha_k, p_k] = \alpha_{i,k}\, q(h_{i,k}, p_{i,k}) \qquad (7)$$

This expression states that the probability that sensor i successfully transmits equals the probability that i is scheduled to transmit *and* the message is correctly decoded at the receiver. The event that no sensor transmits happens with probability

$$\mathbb{P}[\gamma_{0,k} = 1 \,|\, h_k, \alpha_k, p_k] = 1 - \sum_{i=1}^{m} \alpha_{i,k}\, q(h_{i,k}, p_{i,k}), \qquad (8)$$

which is the complement of the probability that some sensor transmits.

Our goal is to design scheduling and power allocation protocols that exploit the random channel conditions on the shared wireless medium in order to make an efficient use of the sensors' power resources and lead to a stable control system. The exact problem specification is presented next.

A. Communication design specification

We consider scheduling and power variables α_k, p_k that adapt to the current channel states h_k, so they can be expressed as mappings $\alpha_k = \alpha(h_k)$, $p_k = p(h_k)$ of the form

$$\mathcal{A} = \{\alpha : \mathcal{H}^m \mapsto \Delta_m\},\ \mathcal{P} = \{p : \mathcal{H}^m \mapsto [0, p_{\max}]^m\}. \qquad (9)$$

Since channel states h_k are i.i.d. over time k these mappings do not need to change over time. Substituting $\alpha(.), p(.)$ in our communication model (7), the expected probability of successful transmission for a sensor i at time k becomes

$$\begin{aligned}\mathbb{P}(\gamma_{i,k} = 1) &= \mathbb{E}_{h_k} \left\{ \mathbb{P}[\gamma_{i,k} = 1 \,\big|\, h_k, \alpha(h_k), p(h_k)] \right\} \\ &= \mathbb{E}_h \alpha_i(h)\, q(h_i, p_i(h)), \qquad (10)\end{aligned}$$

where in the last equality we dropped the index of the channel variable h_k since they are i.i.d. with distribution ϕ over time k, meaning that the probabilities in (10) become constant for all k. Similarly by (8) we have

$$\mathbb{P}(\gamma_{0,k} = 1) = 1 - \sum_{i=1}^{m} \mathbb{E}_h \alpha_i(h)\, q(h_i, p_i(h)). \qquad (11)$$

The goal of the communication design is to make an efficient use of the power resources available at the sensors while ensuring that the resulting control system is stable. In particular suppose each sensor i has a power budget b_i and we require that the expected power consumption induced by the communication design at each slot k is limited to

$$\mathbb{E}_h \alpha_i(h) p_i(h) \leq b_i, \quad \text{for all } i = 1, \ldots, m. \qquad (12)$$

The expectation on the left hand side is with respect to the channel distribution $h_k \sim \phi$ and accounts for the consumed transmit power whenever sensor i is scheduled.

Next we motivate the control system stability specification. Under the described communication design the transmission sequence $\{\gamma_{i,k}, 0 \leq i \leq m, k \geq 0\}$ is independent of the system state x_k. The resulting system (1) becomes a random jump linear system with i.i.d. jumps since probabilities $\mathbb{P}(\gamma_{i,k} = 1)$ are constant over time k. Necessary and sufficient stability conditions for such systems are known.

Theorem 1. *[14, Cor. 1] Consider system (1) with constant probabilities $\mathbb{P}(\gamma_{i,k} = 1)$ for all k. Then the system is mean square stable, i.e., there exist $x_\infty \in \mathbb{R}^n$ and $X_\infty \in \mathbb{S}^n_+$ such that*

$$\lim_{k\to\infty} \|\mathbb{E} x_k - x_\infty\| = 0, \text{ and } \lim_{k\to\infty} \|\mathbb{E} x_k x_k^T - X_\infty\| = 0 \qquad (13)$$

if and only if there exists a matrix $P \in S^n_{++}$ satisfying

$$\sum_{i=0}^{m} \mathbb{P}(\gamma_{i,k} = 1) A_i^T P A_i \prec P. \qquad (14)$$

The intuition behind the theorem is that for fixed probabilities $\mathbb{P}(\gamma_{i,k} = 1)$ a Lyapunov-like function $V(x) = x^T P x$, $x \in \mathbb{R}^n$ decreases in expectation at each step. In particular, (14) is equivalent to

$$\begin{aligned}\mathbb{E}\left[V(x_{k+1}) \,\big|\, x_k\right] &= \sum_{i=0}^{m} \mathbb{P}(\gamma_{i,k} = 1) x_k^T A_i^T P A_i x_k + Tr(PW) \\ &< V(x_k) + Tr(PW) \qquad (15)\end{aligned}$$

holding for any $x_k \in \mathbb{R}^n$, where the first equality follows from (1). Motivated by this observation about stability, we pose the problem of designing wireless communication variables that make the decrease rate in (15) as low as possible.

Suppose a quadratic Lyapunov function $V(x) = x^T P x$, $x \in \mathbb{R}^n$, with $P \in S^n_{++}$, is fixed. We are interested in channel-aware scheduling and power allocation variables (cf. (9)) that minimize the Lyapunov decrease rate in (15) and also meet the power budgets (12). This is a stochastic optimization problem of the form

Optimal scheduling and power allocation design

$$\underset{r,\, \alpha\in\mathcal{A},\, p\in\mathcal{P}}{\textbf{minimize}} \quad r^2 \tag{16}$$

$$\textbf{subject to} \quad \mathbb{E}_h \sum_{i=1}^m \alpha_i(h) p_i(h) \leq b_i, \quad i = 1, \dots, m \tag{17}$$

$$D_0 - \sum_{i=1}^m \mathbb{E}_h \alpha_i(h)\, q(h_i, p_i(h)) D_i \preceq rP \tag{18}$$

where for compactness we defined

$$D_0 = A_0^T P A_0, \quad D_i = A_0^T P A_0 - A_i^T P A_i, \tag{19}$$

for $i = 1, \dots, m$. The semidefinite constraint (18) follows from (14) by substituting the probabilities (10), (11) induced by the communication design and introducing an auxiliary variable r for the Lyapunov decrease rate (or increase if $r > 1$). The objective in (16) is an increasing function of r, so that the optimal rate is as small as possible, and for convenience is chosen to be strictly convex. For technical reasons we keep an implicit constraint $0 \leq r \leq r_{\max}$, which is not restrictive. The left hand side in (18) will always be bounded since the terms in expectations are probabilities (bounded by 1).

Finally we note that problem (16) is always strictly feasible. Consider for instance $p \equiv 0$ and $r \geq 0$ sufficiently large so that both (17) and (18) hold with strict inequality. We denote then the optimal value of the problem by P^* and an optimal solution by $r^*, \alpha^*(.), p^*(.)$. Even though the problem is infinite-dimensional and non-convex in general, in the following section we present an algorithm based on the Lagrange dual problem which converges to the optimal solution. Moreover the algorithm does not require any prior knowledge of the channel distribution, but can be implemented using the channel states measured online during execution.

Remark 1. The centralized channel-aware scheduler of the multiple access channel architecture in Fig. 1 can be implemented as follows. Channel conditions on each wireless link can be measured by pilot signals sent from the sensors to the receiver/controller at each time step before the scheduling decision. Depending on the measured channel states, the scheduler at the receiver selects and notifies via the reverse channel a sensor to transmit. Channel state information can also be passed this way back to the selected sensor, which accordingly adapts its transmit power. □

III. Optimal scheduling and power allocation

In this section we present an algorithm that converges to the optimal channel-aware scheduling and power allocation policy. The algorithm employs the Lagrange dual problem of (16) and exploits the fact that there is no duality gap. Moreover, the algorithm can be implemented online based on a random channel sequence and converges almost surely to the optimal operating point with respect to (16).

To define the Lagrange dual problem of (16) consider nonnegative dual variables $\nu \in \mathbb{R}^m_+$ corresponding to each of the m power capacity constraints in (17), and a symmetric positive semidefinite matrix $\Lambda \in \mathbb{S}^n_+$ corresponding to the semidefinite constraint (18). The Lagrangian is written as

$$L(r, \alpha, p, \nu, \Lambda) = r^2 + \sum_{i=1}^m \nu_i [\mathbb{E}_h \alpha_i(h) p_i(h) - b_i]$$
$$+ Tr(\Lambda [D_0 - \sum_{i=1}^m \mathbb{E}_h \alpha_i(h)\, q(h_i, p_i(h)) D_i - rP]), \tag{20}$$

while the dual function is defined as

$$g(\nu, \Lambda) = \min_{r,\, \alpha\in\mathcal{A},\, p\in\mathcal{P}} L(r, \alpha, p, \nu, \Lambda). \tag{21}$$

For convenience let us also denote the set of primal variables r, α, p that minimize the Lagrangian at ν, Λ by

$$(\mathcal{R}, \mathcal{A}, \mathcal{P})(\nu, \Lambda) = \underset{r,\, \alpha\in\mathcal{A},\, p\in\mathcal{P}}{\text{argmin}}\ L(r, \alpha, p, \nu, \Lambda). \tag{22}$$

In general this set might contain multiple solutions. We will refer to any such solution triplet as $r(\nu, \Lambda), \alpha(\nu, \Lambda), p(\nu, \Lambda)$. We define then the Lagrange dual problem as

$$D^* = \underset{\nu\in\mathbb{R}^m_+, \Lambda\in\mathbb{S}^n_+}{\text{maximize}} \quad g(\nu, \Lambda). \tag{23}$$

By standard Lagrange duality theory the dual function $g(\nu, \Lambda)$ at any point ν, Λ is a lower bound on the optimal cost P^* of problem (16), hence also $D^* \leq P^*$ (weak duality). The following proposition however, based on the results in similar stochastic optimization problems [15], establishes a strong duality result ($D^* = P^*$) and provides a relationship between the optimal primal and dual variables.

Proposition 1. *Let Assumption 1 hold, let P^* be the optimal value of the optimization problem (16) and (r^*, α^*, p^*) be an optimal solution, and let D^* be the optimal value of the dual problem (23) and ν^*, Λ^* be an optimal solution. Then*
(a) $P^* = D^*$ (strong duality)
(b) $(r^*, \alpha^*, p^*) \in (\mathcal{R}, \mathcal{A}, \mathcal{P})(\nu^*, \Lambda^*)$

Proof. As noted after problem (16), a strictly feasible solution always exists. Statement (a) under Assumption 1 and strict feasibility follows from [15, Theorem 1] where a similar optimization setup is examined. The proof is omitted due to space limitations.

To show (b) consider a primal optimal solution (r^*, α^*, p^*). This gives an optimal value P^* for problem (16). The Lagrangian in (20) at the point of optimal primal and dual variables evaluates to

$$L(r^*, \alpha^*, p^*, \nu^*, \Lambda^*) = P^* + \sum_{i=1}^m \nu_i^* s_i + Tr(\Lambda^* S) \leq P^*, \tag{24}$$

where for compactness we denote the constraint slack of (r^*, α^*, p^*), i.e., the brackets in (20), as s_i for the power constraints of i and as S for the semidefinite constraint. Since the optimal primal solution is feasible for (16), it satisfies $s_i \leq 0$ and $S \preceq 0$, and since the dual variables satisfy $\nu^* \geq 0$ and $\Lambda^* \succeq 0$, we get the last inequality in (24).

On the other hand by definition of the dual function g in (21) at the point ν^*, Λ^* we have that

$$L(r^*, \alpha^*, p^*, \nu^*, \Lambda^*) \geq g(\nu^*, \Lambda^*) = P^* \tag{25}$$

where for the last equality we used the fact that $g(\nu^*, \Lambda^*) = D^*$ by dual optimality, and $D^* = P^*$ by part (a). Combining (24) and (25) we conclude that all the included inequalities hold with equality. Then (25) holding with equality shows that r^*, α^*, p^* yield an optimal value for the Lagrangian at ν^*, Λ^*, and verifies (b) by definition (22). □

Interestingly this proposition states that strong duality holds despite the fact that (16) is not convex. More importantly, as we follow next, it suggests the possibility of developing an algorithm to find the optimal dual variables ν^*, Λ^*, and then via (b) recover the optimal primal variables r^*, α^*, p^*. To prepare for the algorithm, note that the Lagrangian in (20) can be equivalently written as

$$L(r, \alpha, p, \nu, \Lambda) = r^2 - r\,Tr(\Lambda P) + Tr(\Lambda D_0) - \nu^T b + \mathbb{E}_h \sum_{i=1}^{m} \alpha_i(h)\ \left[\nu_i p_i(h) - Tr(\Lambda D_i)\, q(h_i, p_i(h))\right] \tag{26}$$

By this expression finding the primal Lagrange optimizers in (21)-(22) is easy. By strict convexity and differentiability with respect to r, the minimizer $r(\nu, \Lambda)$ is unique and equals

$$r(\nu, \Lambda) = \min\{\, 1/2\, Tr(\Lambda P),\, r_{\max}\} \tag{27}$$

where we enforced the implicit constraint $0 \leq r \leq r_{\max}$.

Optimizing over the functions $\alpha(.), p(.)$ in (26) is also simplified because they are decoupled over channel states $h \in \mathcal{H}^m$. Power minimizers at each h are given by

$$p_i(\nu, \Lambda; h) = \underset{0 \leq p \leq p_{\max}}{\operatorname{argmin}}\ \nu_i\, p - Tr(\Lambda D_i)\, q(h_i, p), \tag{28}$$

which implies a further decoupling among sensors i – see Remark 2. Scheduling minimizers for each h in (26) are obtained as

$$\alpha(\nu, \Lambda; h) = \underset{\alpha \in \Delta^m}{\operatorname{argmin}}\ \sum_{i=1}^{m} \alpha_i\, \xi(h_i, \nu_i, \Lambda), \tag{29}$$

where

$$\xi(h_i, \nu_i, \Lambda) = \min_{0 \leq p \leq p_{\max}}\ \nu_i p - Tr(\Lambda D_i)\, q(h_i, p). \tag{30}$$

By the form of Δ^m in (6) the minimizing scheduling is deterministic. The scheduler picks with certainty the sensor with the lowest value $\xi(h_i, \nu_i, \Lambda)$ (or one of them if non-unique). This reveals the opportunistic nature of the channel-aware scheduler which, based on the current channel conditions, dynamically assigns channel access to the sensor with lowest relative value $\xi(h_i, \nu_i, \Lambda)$.

We now present an iterative algorithm to solve the dual problem. As noted earlier, this is an online algorithm depending on an observed random channel sequence. Hence the variables are indexed by real time steps $k \geq 0$. The iterative steps of the algorithm are as follows:

i) At time step k observe current channel conditions h_k, and given current dual variables ν_k, Λ_k, compute primal optimizers of the Lagrangian at h_k using (27)-(29) as

$$r_k = r(\nu_k, \Lambda_k) \tag{31}$$

$$p_{i,k} = p_i(\nu_k, \Lambda_k; h_k), \quad i = 1, \ldots, m, \tag{32}$$

$$\alpha_k = \alpha(\nu_k, \Lambda_k; h_k) \tag{33}$$

ii) Update the dual variables as

$$\nu_{i,k+1} = [\nu_{i,k} + \epsilon_k(\alpha_{i,k} p_{i,k} - b_i)]_+ \tag{34}$$

$$\Lambda_{k+1} = [\Lambda_k + \epsilon_k(D_0 - \sum_{i=0}^{m} \alpha_{i,k}\, q(h_{i,k}, p_{i,k}) D_i - r_k P)]_+ \tag{35}$$

where $[\ \ ]_+$ denotes the projection on the non-negative orthant and on the positive semidefinite cone in (34) and (35) respectively, and $\epsilon_k \geq 0$ is a step size.

The intuition behind the algorithm is that dual variables are updated in (34), (35) in a direction which in expectation is a subgradient of the dual function g. The following proposition establishes that the algorithm converges to the optimal solution for the dual of the optimal scheduling and power allocation problem.

Proposition 2. *Consider the optimization problem (16) and its dual derived in (23). Based on a sequence $\{h_k, k \geq 0\}$ of i.i.d. random variables with distribution ϕ on $\mathcal{H}^m$, let the algorithm described in steps (i)-(ii) be employed with step sizes satisfying*

$$\sum_{k=0}^{\infty} \epsilon_k^2 < \infty,\ \sum_{k=0}^{\infty} \epsilon_k = \infty. \tag{36}$$

Then almost surely with respect to $\{h_k, k \geq 0\}$ it holds

$$\lim_{k \to \infty} (\nu_k, \Lambda_k) = (\nu^*, \Lambda^*), \ \ \textit{and} \ \ \lim_{k \to \infty} g(\nu_k, \Lambda_k) = D^* \tag{37}$$

where ν^*, Λ^* is an optimal solution of the dual problem and D^* is the optimal value of the dual problem.

Proof. See Appendix A. □

Besides optimizing over dual variables, the algorithm can be interpreted as a communication protocol of how to schedule sensors and allocate transmit power, adapting online to the observed channel conditions. Since the communication protocol is designed to serve the wireless control architecture of Section II, the following theorem establishes the provided control performance guarantees.

Theorem 2. *Consider the wireless control architecture of Fig. 1 with plant dynamics described by (1), and a given function $V(x) = x^T P x$, $P \in \mathbb{S}^n_{++}$. Consider transmission variables $\gamma_{i,k}$ described by (7), (8), depending on channel* *states $h_k \in \mathcal{H}^m$ which are i.i.d. with distribution ϕ, scheduling $\alpha_k \in \Delta_m$, and power allocation $p_k \in [0, p_{\max}]^m$. Let Assumptions 1, 2 hold. If α_k, p_k adapt to the channel sequence $h_{0:k}$ according to algorithm (31)-(35), with stepsizes ϵ_k satisfying (36), then almost surely the power consumption for each sensor i satisfies*

$$\limsup_{k\to\infty} \mathbb{E}\left[\, \alpha_{i,k} p_{i,k} \, | h_{0:k-1}\right] \leq b_i, \tag{38}$$

and the decrease rate of $V(x)$ satisfies for any $x \in \mathbb{R}^n$

$$\limsup_{k\to\infty} \mathbb{E}\left[V(x_{k+1}) \,\middle|\, x_k = x, h_{0:k-1}\right] \leq r^* V(x) + Tr(PW) \tag{39}$$

where r^* is the optimal solution of problem (16).

Proof. See Appendix B. □

According to the theorem, the protocol converges almost surely to a configuration that respects the sensors' power constraints and minimizes the decrease rate of the given Lyapunov function. This however does not a priori imply system stability. If the algorithm converges to some $r^* > 1$ then the resulting communication protocol may lead to either an unstable or a stable system. This does not contradict the necessary and sufficient stability condition of Theorem 1 which states that *some* appropriate quadratic Lyapunov function exists. The online algorithm is based on a fixed function, under which stability may not be provable. If however $r^* < 1$ then indeed stability is guaranteed (cf. Theorem 1). A necessary and sufficient condition for $r^* < 1$ is that the feasible set of problem (16) contains a point $r < 1$. We restate this observation in the following corollary.

Corollary 1. *Consider the setup of Theorem 2 and additionally suppose the optimization problem (16) contains a feasible solution with $r < 1$. Then almost surely $\mathbb{P}[\gamma_{i,k} | h_{0:k-1}]$ for $i = 1, \ldots, m$ converge to values such that system (1) is mean square stable.*

After some remarks on the structure of the communication protocol, we present numerical simulations of the online algorithm in the following section.

Remark 2. The online communication protocol implies a decentralized power allocation. In step (32), as noted in (28), the transmit power $p_{i,k}$ for sensor i, when scheduled, depends not on the whole channel vector h_k but only on the channel state $h_{i,k}$ of the respective link i, as well as on the variables $\nu_{i,k}, \Lambda_k$. Similar separability results are common in wireless communication networks [15]. From an implementation perspective, as noted in Remark 1, channel states $h_{i,k}$ can be estimated at each sensor i. The variables $\nu_{i,k}, \Lambda_k$ can be sent from the scheduler to the scheduled sensor i at each time step. As $\nu_{i,k}, \Lambda_k \to \nu_i^*, \Lambda^*$ according to Prop.2, at the limit operating point each sensor can locally store ν_i^*, Λ^* and select power according to the stored values and the current channel conditions. We note however that the scheduling variable in (33) is centralized since, as noted in (29), it depends on all dual variables and the channel states observed by all sensors. □

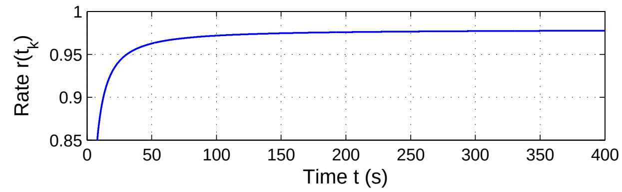

Fig. 3. Rate variable r_k during online algorithm. The variable converges to a Lyapunov decrease rate less than 1, implying mean square stability.

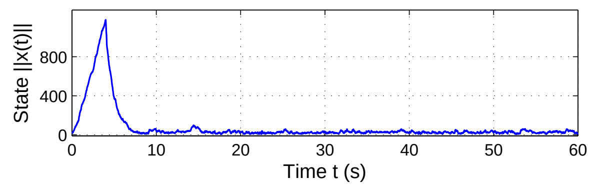

Fig. 4. Norm of system state $\|x(t)\|$ during online algorithm. The norm remains bounded, after an initial transient phase where the online algorithm has not converged to a stabilizing communication protocol.

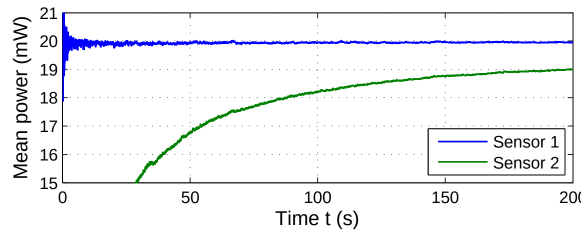

Fig. 5. Sensors' average power consumption during the online algorithm. In the limit both satisfy the power constraint $b_i = 20mW$.

IV. Numerical Simulations

We consider the frequently used benchmark example of a batch reactor [2], [8]. The continuous time plant and controller dynamics can be found in the referred works, and involve a plant with 4 states, 2 inputs and m = 2 outputs, and a PI controller with 2 states. Following Example 1 we obtain under a transmission period $T_s = 0.02s$ the discrete time switched dynamics of the form (1). Then a quadratic Lyapunov function needs to be chosen. Consider a function that would guarantee stability if each sensors transmits successfully 40% of the time, e.g., satisfying

$$\sum_{i=1}^{2} 0.4\, A_i^T P A_i + 0.2\, A_0^T P A_0 = 0.98\, P - 0.001\, I. \tag{40}$$

The value 0.98 is selected after some trials and relates to the fact that the system has an eigenvalue very close to 1 (also documented in [2]), while the term $-0.001\, I$ guarantees the left and right hand sides are almost equal.

We model the channel gains $h_{1,k}$, $h_{2,k}$ as independent over time k and also among the two sensors, both exponentially distributed with a normalized mean 1. The maximum transmit power and the power budgets are modeled as $p_{\max} = 100mW$ and $b_i = 20mW$ respectively for both sensors. The function $q(h\, p)$ is shown in Fig. 2.

We run the online algorithm of (31)-(35) in Section III, which converges to a communication protocol where sensors 1, 2 transmit with probabilities $\approx 0.54, 0.40$ respectively, slightly deviating from the values assumed in the Lyapunov construction (40). As shown in Fig. 3 the algorithm converges to a protocol that stabilizes the system according to (39), since the rate variable r_k tends to $r^* \approx 0.98$. Stability is also verified at the system state plot in Fig. 4. The resulting protocol meets the sensor's power constraints, as we see in Fig. 5 where we plot the mean power $1/N \sum_{k=1}^{N} \alpha_{i,k} p_{i,k}$ for each sensor i during the algorithm. Before convergence, sensor 2 does not transmit often enough or with enough power, explaining the large initial states in Fig. 4.

V. Concluding Remarks

In this paper we considered the problem of scheduling power-constrained sensors in wireless control systems. We developed a protocol where scheduling decisions and transmit power allocation are selected online based on the observed random wireless channel conditions and the objective is to obtain a configuration such that the control system is stable. The protocol is based on a given Lyapunov function, under which however the system might not be stabilizable. The problem of determining Lyapunov functions suitable for the scheduling algorithm requires further examination, and also relates to how the control operation is pre-designed. Future work includes as well the design of schedulers additionally adapting to plant state as in, e.g., [2], [8].

Appendix

A. Proof of Proposition 2

Using the notation introduced in the Lagrange optimizers in (27)-(29) let us define

$$s_i(\nu, \Lambda; h) = \alpha_i(\nu, \Lambda; h) p_i(\nu, \Lambda; h) - b_i, \tag{41}$$

$$S(\nu, \Lambda; h) = D_0 - \sum_{i=0}^{m} \alpha_i(\nu, \Lambda; h)\, q(h_i, p_i(\nu, \Lambda; h)) D_i - r(\nu, \Lambda) P. \tag{42}$$

This way the terms in the parentheses of steps (34), (35) can be expressed via the vector $s(\nu_k, \Lambda_k; h_k)$ and the matrix $S(\nu_k, \Lambda_k; h_k)$ respectively.

First we note that the vector $s(\nu, \Lambda; h)$ and the matrix $S(\nu, \Lambda; h)$ are stochastic subgradients for the dual function g in (21) at the point ν, Λ, i.e.,

$$g(\nu', \Lambda') - g(\nu, \Lambda) \le (\nu' - \nu)^T \mathbb{E}_h s(\nu, \Lambda; h) + Tr((\Lambda' - \Lambda)^T \mathbb{E}_h S(\nu, \Lambda; h) \tag{43}$$

for all $\nu' \in \mathbb{R}^m_+$ and $\Lambda' \in \mathbb{S}^n_+$. The expectations in this expression are with respect to the distribution ϕ of h. To show (43) observe that for any ν', Λ' we have by the definition of the dual function in (21) and the Lagrange minimizers in (22) that

$$g(\nu', \Lambda') \le L(\{r, \alpha, p\}(\nu, \Lambda), \nu', \Lambda') \tag{44}$$

Subtracting from each side of this inequality the term $g(\nu, \Lambda) = L(\{r, \alpha, p\}(\nu, \Lambda), \nu, \Lambda)$ and expanding the terms of the Lagrangian as in (20) we get exactly (43).

Hence steps (34), (35) of the algorithm follow random subgradient directions. Note also that subgradients are always bounded in our problem since all terms in (41), (42) are bounded. Moreover, since the primal problem (16) is always strictly feasible, it follows easily that the optimal dual variables are finite. We establish the following fact.

Fact 1. *At each k it holds that*

$$\mathbb{E}\left[\|\nu_{k+1} - \nu^*\|^2 + \|\Lambda_{k+1} - \Lambda^*\|^2 \,|\nu_k, \Lambda_k\right] \le \|\nu_k - \nu^*\|^2 + \|\Lambda_k - \Lambda^*\|^2 + 2\epsilon_k^2 B^2 - 2\epsilon_k (D^* - g(\nu_k, \Lambda_k)) \tag{45}$$

where B is a bound on the stochastic subgradients $\|s(\nu, \Lambda; h)\| \le B$ and $\|S(\nu, \Lambda; h)\| \le B$ for any ν, Λ, h.

Proof. First use the expression for ν_{k+1} in (34), with $s_k = s(\nu_k, \Lambda_k; h_k)$, to write

$$\|\nu_{k+1} - \nu^*\| = \|[\nu_k + \epsilon_k s_k]_+ - \nu^*\| \le \|\nu_k + \epsilon_k s_k - \nu^*\|, \tag{46}$$

where the last inequality holds because projecting on the positive orthant cone can only decrease the distance from a point ν^* in the orthant cone. Squaring the norms in (46), expanding the square of the right hand side, and taking expectation on both sides given ν_k, Λ_k we get

$$\mathbb{E}[\|\nu_{k+1} - \nu^*\|^2 \,|\nu_k, \Lambda_k] \le \|\nu_k - \nu^*\|^2 + \epsilon_k^2 B^2 + 2\epsilon_k (\nu_k - \nu^*)^T \mathbb{E}[s_k \,\big|\, \nu_k, \Lambda_k], \tag{47}$$

where we bounded $\|s_k\|^2 < B^2$. Similar arguments for the variable Λ_{k+1} lead to

$$\mathbb{E}[\|\Lambda_{k+1} - \Lambda^*\|^2 \,|\nu_k, \Lambda_k] \le \|\Lambda_k - \Lambda^*\|^2 + \epsilon_k^2 B^2 + 2\epsilon_k Tr((\Lambda_k - \Lambda^*)\mathbb{E}[S_k \,\big|\, \nu_k, \Lambda_k]). \tag{48}$$

The statement (45) follows from summing (47) and (48) and applying inequality (43) with the substitution $(\nu', \Lambda', \nu, \Lambda) = (\nu^*, \Lambda^*, \nu_k, \Lambda_K)$. ☐

Our goal is to use (45) to show that $\|\nu_k - \nu^*\|^2 + \|\Lambda_k - \Lambda^*\|^2 \to 0$ almost surely. The proof relies on a supermartingale convergence argument frequently used in stochastic optimization. First note that at any ν_k, Λ_k the dual function is lower than the optimal value (cf. (23)), so $D^* - g(\nu_k, \Lambda_k) \ge 0$. Hence (45) can be simplified to

$$\mathbb{E}\left[\|\nu_{k+1} - \nu^*\|^2 + \|\Lambda_{k+1} - \Lambda^*\|^2 \,|\nu_k, \Lambda_k\right] \le \|\nu_k - \nu^*\|^2 + \|\Lambda_k - \Lambda^*\|^2 + 2\epsilon_k^2 B^2. \tag{49}$$

Then consider the non-negative random variable

$$a_k = \|\nu_k - \nu^*\|^2 + \|\Lambda_k - \Lambda^*\|^2 + \sum_{\ell=k}^{\infty} 2\epsilon_l^2 B^2, \tag{50}$$

which depends on the sequence (filtration) $\mathcal{F}_k = \{\nu_{0:k}, \Lambda_{0:k}\}$. Note that a_k is bounded (hence integrable) because ν_k, Λ_k generated by (34), (35) are bounded at every k and also the stepsizes are square summable. By the relation (49) it easily follows that a_k satisfies $\mathbb{E}[a_{k+1} \,\big|\, \mathcal{F}_k] \le a_k$. Such a stochastic process is called a supermartingale [16, Ch. 5]. Moreover, a non-negative supermartingale converges almost surely to some limit random variable [16, Th. 5.2.9]. Observe that the second summand $\sum_{\ell=k}^{\infty} 2\epsilon_l^2 S^2$ of a_k in (50) is deterministic and converges to 0 because of square summability of the stepsizes. Hence the random variable

$\|\nu_k-\nu^*\|^2+\|\Lambda_k-\Lambda^*\|^2$ converges almost surely (to some random variable).

To arrive at a contradiction suppose the limit random random variable is not identically zero.Equivalently, with probability $\delta>0$ we have $\|\nu_k-\nu^*\|^2+\|\Lambda_k-\Lambda^*\|^2\geq\epsilon$ for some $\epsilon>0$ for all sufficiently large k. This implies that ν_k,Λ_k are bounded away from the optimal, hence

$$\mathbb{E}\sum_{k=0}^{\infty}2\epsilon_k(D^*-g(\nu_k,\Lambda_k))=+\infty. \tag{51}$$

Note however that taking expectation in (45) and iterating for $k=0,\dots,N-1$ we get

$$\mathbb{E}\left[\|\nu_N-\nu^*\|^2+\|\Lambda_N-\Lambda^*\|^2\right]\leq\|\nu_0-\nu^*\|^2+\|\Lambda_0-\Lambda^*\|^2 \\ +\sum_{k=0}^{N-1}2\epsilon_k^2B^2-\mathbb{E}\sum_{k=0}^{N-1}2\epsilon_k(D^*-g(\nu_k,\Lambda_k)). \tag{52}$$

The left hand side is non-negative, but (51) implies that in the limit as $N\to\infty$ the right hand side becomes negative. This is a contradiction. Therefore it must be that $\|\nu_k-\nu^*\|^2+\|\Lambda_k-\Lambda^*\|^2$ converges to zero with probability 1. By continuity of the (concave) dual function $g(\nu,\Lambda)$ we also have that $g(\nu_k,\Lambda_k)$ converges to $g(\nu^*,\Lambda^*)=D^*$ a.s.

B. Proof of Theorem 2

We will show (38) and (39) by relating them to the online optimization of problem (16) and its dual. Note that the algorithm in (31)-(33) selects r_k,p_k,α_k as functions of ν_k,Λ_k,h_k, where ν_k,Λ_k further depend on the observed history $h_{0:k-1}$. Using the notation introduced in (41), (38) can be equivalently be written as

$$\limsup_{k\to\infty}\mathbb{E}_{h_k}s(\nu_k,\Lambda_k;h_k)\leq 0, \tag{53}$$

where we suppressed the dependence on the history $h_{0:k-1}$ in (38) by the variables ν_k,Λ_k. The expectation in (53) is only with respect to h_k, not the past (random) history.

Equation (39) can be transformed in a similar way. Since (39) needs to hold for any vector x we may use the semidefinite notation, as we did in converting (15) to (18), to rewrite (39) equivalently as

$$\limsup_{k\to\infty}\{\mathbb{E}_{h_k}S(\nu_k,\Lambda_k;h_k)+(r(\nu_k,\Lambda_k)-r^*)P\}\preceq 0, \tag{54}$$

where we made use of the notation introduced in (42).

By Prop. 2 $\nu_k,\Lambda_k\to\nu^*,\Lambda^*$ a.s., hence by continuity of the function $r(\nu,\Lambda)$ in (27) we have that $r(\nu_k,\Lambda_k)\to r(\nu^*,\Lambda^*)$ a.s. We also note that $r^*=r(\nu^*,\Lambda^*)$ as follows from Prop.1(b) and the uniqueness of the Lagrangian minimizer r (cf.(26), (27). Hence $r(\nu_k,\Lambda_k)\to r^*$ a.s. The rest of the proof shows that $\mathbb{E}_{h_k}s(\nu_k,\Lambda_k;h_k)$, $\mathbb{E}_{h_k}S(\nu_k,\Lambda_k;h_k)$ become non-positive in the limit almost surely, so that (53) and (54) hold true.

As we argued in the proof of Prop. 2 the vector $\mathbb{E}_{h_k}s(\nu_k,\Lambda_k;h_k)$ and matrix $\mathbb{E}_{h_k}S(\nu_k,\Lambda_k;h_k)$ are subgradients of the dual function $g(\nu_k,\Lambda_k)$ with respect to ν_k and Λ_k respectively (cf.(43)). Since g is concave and $\nu_k,\Lambda_k\to\nu^*,\Lambda^*$, all limit points of the sequence of subgradients $\mathbb{E}_{h_k}s(\nu_k,\Lambda_k;h_k)$ and $\mathbb{E}_{h_k}S(\nu_k,\Lambda_k;h_k)$ are subgradients at ν^*,Λ^* [17, Prop. 4.2.3].

Thus for (53) and (54) we need to show that all subgradients of g at ν^*,Λ^* are non-positive. It can be shown that under Assumption 2 the subgradient takes a unique value. This fact is omitted due to space limitations but can be found in [11, Lemma 1]. Then by Prop.1(b) the value of the subgradient of g at ν^*,Λ^* can be computed as the constraint slack of the the optimal primal variables r^*,α^*,p^*, which is non-positive because the optimal point is primal feasible.

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] M. Donkers, W. Heemels, N. Van De Wouw, and L. Hetel, “Stability analysis of networked control systems using a switched linear systems approach,” *IEEE Transactions on Automatic Control*, vol. 56, no. 9, pp. 2101–2115, 2011.

[3] W. Zhang, M. Branicky, and S. Phillips, “Stability of networked control systems,” *IEEE Control Systems Magazine*, vol. 21, no. 1, pp. 84–99, 2001.

[4] D. Hristu-Varsakelis, “Feedback control systems as users of a shared network: Communication sequences that guarantee stability,” in *Proc. of the 40th IEEE Conference on Decision and Control, 2001*, vol. 4, 2001, pp. 3631–3636.

[5] L. Zhang and D. Hristu-Varsakelis, “Communication and control codesign for networked control systems,” *Automatica*, vol. 42, no. 6, pp. 953–958, 2006.

[6] J. Le Ny, E. Feron, and G. J. Pappas, “Resource constrained lqr control under fast sampling,” in *Proc. of the 14th international conference on Hybrid systems: computation and control*. ACM, 2011, pp. 271–280.

[7] L. Meier, J. Peschon, and R. M. Dressler, “Optimal control of measurement subsystems,” *IEEE Transactions on Automatic Control*, vol. 12, no. 5, pp. 528–536, 1967.

[8] G. C. Walsh, H. Ye, and L. G. Bushnell, “Stability analysis of networked control systems,” *IEEE Transactions on Control Systems Technology*, vol. 10, no. 3, pp. 438–446, 2002.

[9] D. Hristu-Varsakelis and P. R. Kumar, “Interrupt-based feedback control over a shared communication medium,” in *Proc. of the 41st IEEE Conference on Decision and Control, 2002*, vol. 3, 2002, pp. 3223–3228.

[10] M. H. Mamduhi, A. Molin, and S. Hirche, “Stability analysis of stochastic prioritized dynamic scheduling for resource-aware heterogeneous multi-loop control systems,” in *Proc. of the 52nd IEEE Conference on Decision and Control (CDC 2013)*, 2013, pp. 7390–7396.

[11] K. Gatsis, M. Pajic, A. Ribeiro, and G. J. Pappas, “Opportunistic scheduling of control tasks over shared wireless channels,” in *ACM/IEEE 5th International Conference on Cyber-Physical Systems (ICCPS 2014)*, 2014, to appear. Available at http://www.seas.upenn.edu/~kgatsis/papers/GatsisEtal_ICCPS14.pdf.

[12] K. Gatsis, A. Ribeiro, and G. J. Pappas, “Optimal power management in wireless control systems,” *IEEE Transactions on Automatic Control*, 2014, to appear. Available online by IEEE Early Access.

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

[14] O. L. Costa and M. D. Fragoso, “Stability results for discrete-time linear systems with markovian jumping parameters,” *Journal of mathematical analysis and applications*, vol. 179, no. 1, pp. 154–178, 1993.

[15] A. Ribeiro, “Optimal resource allocation in wireless communication and networking,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–19, 2012.

[16] R. Durrett, *Probability: theory and examples*. Cambridge University Press, 2010.

[17] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex analysis and optimization*. Athena Scientific, 2003.