2016 IEEE 55th Conference on Decision and Control (CDC)
ARIA Resort & Casino
December 12-14, 2016, Las Vegas, USA

State-Based Communication Design for Wireless Control Systems

Konstantinos Gatsis, Alejandro Ribeiro, and George J. Pappas

***Abstract*—We examine the problem of a sensor communicating over a wireless channel to an actuator in order to control a plant that is perturbed by a random disturbance. By allowing the sensor to adapt online to the stochastic system state, we develop transmission policies with guarantees on average control performance and required average communication resources. More specifically we design policies with guarantees either on a linear combination of these two objectives, or with guarantees with respect to a hard constraint on the average communication resources. Based on approximate dynamic programming we prove as well as illustrate in simulations that our policies outperform policies that do not adapt online to the stochastic system state.**

I. INTRODUCTION

Modern cyber-physical systems, e.g., smart homes or urban infrastructures, rely on wireless communication to transfer information between sensors and actuators at different physical locations. However communication resources are typically constrained, e.g., wireless sensors are battery-powered, or suffer from transmission uncertainties such as packet drops. Recent research efforts indicate that online adaptation to the physical process (plant state) is instrumental in achieving a good tradeoff between control performance and communication resources. A typical example of state-based communication policies is event-based control [1]–[6] where the system state is continuously monitored and communication between sensors and actuators is triggered only when some event occurs, e.g., the state exceeds some threshold. This framework results in average communication rates lower than standard periodic control setups and without significant losses in control performance. When allocating general wireless communication resources, e.g., transmit power, adaptation to both plant and channel conditions becomes apparent [7].

Despite recent interest in state-based communication mechanisms, providing guarantees on the amount of average communication resources required for such mechanisms or their achieved long term control performance remains a challenge [5]. Moreover, the design of optimal state-based communication mechanisms is computationally hard. As shown in [1], [7] this problem can be posed as a Markov decision process problem, but dynamic programming algorithms over the continuous-valued system state become impractical, motivating the development of suboptimal solutions [6]–[10].

In this paper we examine the problem of a sensor communicating wirelessly to an actuator in order to control a linear plant perturbed by random disturbances (Fig. 1). We are interested in designing state-aware sensor transmission policies which optimize average control performance subject to a given budget on average communication resources (Section II). As mentioned above, computing optimal sensor policies is computationally hard, hence we aim for suboptimal policies which however provide performance guarantees *by design*. In particular, we are interested in policies improving upon the performance of simple non-state-aware policies which meet the same desired communication budget, and which are used hereby as a reference. The methodology for improving upon the reference policies is based on approximate dynamic programming techniques, also employed in [6], [7].

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, aribeiro, pappasg}@seas.upenn.edu.

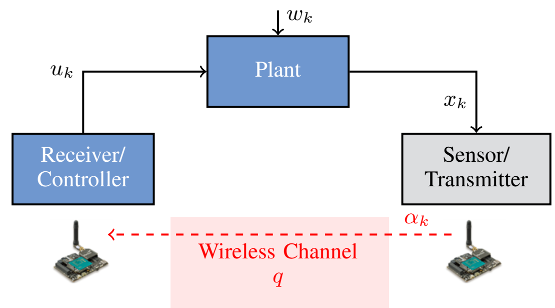

Fig. 1. Wireless Control System. A sensor measures the state of a plant perturbed by a random disturbance. The sensor decides whether to transmit the measured information over a packet-dropping wireless channel to a receiver/controller providing control inputs.

We provide a design methodology with respect to two different types of improvements upon reference policies. First in Section III we consider a Lagrangian relaxation of the original problem, leading to a linear combination of control and communication objectives. We propose a family of sensor policies which are guaranteed to improve upon the reference ones with respect to this family of linear objective combinations (Theorem 1). Under the proposed policies the sensor transmits when an explicitly constructed quadratic function of the system state exceeds a threshold. These policies are similar to the common event-based control framework [5] with the exception that they enjoy performance guarantees by design.

Second in Section IV we propose communication policies which are guaranteed to meet the desired average communication budget, while simultaneously improving upon control performance in comparison to reference policies. This is achieved by enforcing that *in expectation over the current plant state* and at *every time step* the sensor meets the desired communication budget. The proposed policy is again threshold-based but the threshold is now dynamically changing over time anticipating the plant state that the sensor will measure at each time step. We note that policies with average communication guarantees were also considered in [6]. However that was achieved by periodically designing deterministic transmission schedules to be followed by the sensor, i.e., adapting to the plant state only intermittently. In contrast, our policies allow the sensor to continuously adapt to and exploit online at each time step the stochastic nature of the plant state process.

Finally we describe a procedure to efficiently compute the dynamically varying thresholds of our policies. We note that the resulting policies are still computationally intense as different thresholds are required for different expected plant state values. In practice we can design a finite number of thresholds for a discretized space of expected state values. We conclude our paper with numerical simulations (Section V) and concluding remarks (Section VI).

II. Problem description

We consider a wireless control architecture where a plant is controlled over a wireless medium. A sensor measures the state of the plant and transmits it to a controller computing the plant control input – see Fig. 1. Our goal is to design the communication aspects of the problem, hence we assume the dynamics for the control system are fixed, meaning that a controller has been already designed. We assume that the evolution of the system depends on whether a transmission occurs at time k or not, indicated with variables $\gamma_k \in \{0,1\}$. We suppose the system evolution is described by a switched linear time invariant model of the form

$$x_{k+1} = \begin{cases} A_c\, x_k + w_k, & \text{if } \gamma_k = 1 \\ A_o\, x_k + w_k, & \text{if } \gamma_k = 0 \end{cases}. \tag{1}$$

Here $x_k \in \mathbb{R}^n$ denotes the state of the overall control system at each time k, which may in general include both plant and controller states – see following example or [11]. At a successful transmission the system dynamics are described by the matrix $A_c \in \mathbb{R}^{n\times n}$, where 'c' stands for closed-loop, and otherwise by $A_o \in \mathbb{R}^{n\times n}$, where 'o' stands for open-loop. We assume that A_c is asymptotically stable, implying that if system successfully transmits at each time step the state evolution $x_k, k \geq 0$ is stable. The open loop matrix A_o may be unstable.

On the other hand, the additive terms $w_k, k \geq 0$ model an independent identically distributed (i.i.d.) noise process across time according to a known probability distribution $\phi_{0,W}$ with mean zero and positive definite covariance W. We assume that the probability distribution does not contain any atoms [12, Ch. 1]. An example is Gaussian disturbance. We emphasize that knowledge of the distribution is important in our approach – see Sec. IV later. However our results hold regardless of the shape of the distribution.

We are interested in a control performance criterion that accounts for a quadratic plant state cost at each time step k as

$$x_k^T\, Q_{\gamma_k}\, x_k. \tag{2}$$

This cost is allowed to depend on whether transmission occurs ($\gamma_k = 1$), in which case it takes the form $x_k^T Q_1 x_k$, or not ($\gamma_k = 0$), in which case it becomes $x_k^T Q_0 x_k$. Both matrices Q_0, Q_1 are assumed to be positive semidefinite. Examples are presented next.

Example 1. Consider a linear plant of the form

$$x_{k+1} = Ax_k + Bu_k + w_k, \tag{3}$$

where w_k is an i.i.d. Gaussian disturbance, and a wireless sensor transmitting the state x_k to the controller. Consider first a simple control law which applies a zero input $u_k = 0$ when no information is received, and upon receiving a measurement it applies a state feedback $u_k = Kx_k$ leading to a stable closed loop mode $A + BK$. The overall networked system dynamics are expressed if the form (1) with dynamics $A_o = A$, $A_c = A + BK$. If we are interested in a usual linear quadratic regulation cost described as $x_k^T Q x_k + u_k^T R u_k$ for some given positive definite matrices Q, R. This cost is of the form (2) with $Q_o = Q$, $Q_c = Q + K^T RK$

Alternatively suppose the controller keeps a local estimate of the state according to

$$\hat{x}_k = \begin{cases} x_k, & \text{if } \gamma_k = 1 \\ A\hat{x}_{k-1} + Bu_{k-1} & \text{if } \gamma_k = 0 \end{cases}. \tag{4}$$

That is, when no measurement is received the estimate is propagated according to the plant dynamics (without the unknown current disturbance). Suppose the controller applies the input $u_k = K\hat{x}_k$ at each time step. Such a control architecture is common [3], [7]. We can rewrite the overall plant-controller dynamics with respect to an augmented state $x_k, \hat{x}_{k-1}$ as a switched linear system in the form of (1) with dynamics

$$A_o = \begin{bmatrix} A & BK(A+BK) \\ 0 & (A+BK) \end{bmatrix},\; A_c = \begin{bmatrix} A+BK & 0 \\ I & 0 \end{bmatrix}. \tag{5}$$

The augmented system is also perturbed by an augmented noise vector with zero mean and covariance $\begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix}$. A quadratic regulation cost of the form $x_k^T Q x_k + u_k^T R u_k$, which is equivalent to $x_k^T Q x_k + \hat{x}_k^T K^T RK \hat{x}_k$ can again be written in the switched quadratic form of (2). □

We are interested in the design of sensor transmission policies that are efficient with respect to the control performance of such a system, as well as the utilized communication resources. More specifically, suppose at each time slot k the sensor transmits with some probability $\alpha_k \in [0,1]$. The problem we address is how to select these randomized decisions at each time step given the available information at the sensor.

When the sensor transmits, the corresponding packet might get dropped due to noise in the wireless channel [7]. We assume that the packet success probability is constant denoted by $q \in [0,1]$. The success of the transmission at time slot k becomes a Bernoulli random variable with success probability

$$\mathbb{P}(\gamma_k = 1) = q\,\alpha_k. \tag{6}$$

This packet success can be *actively controlled* by the sensor to make the control system in (1) switch in a random but controlled fashion between the two modes of operation (open and closed loop) at each time step. The transmit rates $\alpha_k, k \geq 0$ to be designed affect the performance of the control system, for which we account with the long turn average quadratic cost

$$J_{\text{control}}(\alpha_0, \alpha_1, \ldots) = \limsup_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}[x_k^T \, Q_{\gamma_k} \, x_k]. \quad (7)$$

The expectation at the right hand side accounts for the randomness introduced by the system disturbance, the channel, as well as the selected randomized sensor access policy.

Besides control performance, the sensor's transmission decisions should efficiently use available wireless communication resources, such as transmit power [7]. Assuming for simplicity that each transmission incurs a unit of normalized communication cost, we would like to account for the average communication cost

$$J_{\text{comm}}(\alpha_0, \alpha_1, \ldots) = \limsup_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}[\alpha_k]. \quad (8)$$

Our goal is to design sensor transmission *policies*, that is, transmission decisions α_k adapted to the available information at the sensor at each time step k. This information includes the current measurement x_k, as well as all previously collected local measurements $x_0, \ldots, x_{k-1}$ and the success of all previous transmissions $\gamma_0, \ldots, \gamma_{k-1}$. The latter can be made available by acknowledgments sent from the receiver/controller to the sensor.

We are interested in designing a sensor transmit policy that minimizes the control regulation cost of the system, while at the same time the average communication cost does not exceed a desired budget $\tilde{\alpha} \in [0, 1]$. This is posed as follows.

Problem 1

$$\begin{aligned} \underset{\alpha_k \in [0,1], k\geq 0}{\text{minimize}} \quad & J_{\text{control}}(\alpha_0, \alpha_1, \ldots) && (9) \\ \text{subject to} \quad & J_{\text{comm}}(\alpha_0, \alpha_1, \ldots) \leq \tilde{\alpha} && (10) \end{aligned}$$

Solving for the optimal policy is computationally hard, as it involves dynamic decisions over a continuous valued state $x_k \in \mathbb{R}^n$. Indeed this can be posed as a Markov Decision Process problem – see Remark 1. Hence we only design suboptimal policies for which, however, we provide performance guarantees. In particular we design policies which *improve upon simple non-state-aware policies with respect to the goals of Problem 1*.

To proceed with our design we make the following assumption, which is explained in the following section.

Assumption 1. The system dynamics, the packet drop rate, and the required average communication cost satisfy

$$\rho\Big(q\, \tilde{\alpha}\, A_c \otimes A_c + (1 - q\, \tilde{\alpha}) A_o \otimes A_o \Big) < 1 \quad (11)$$

where $\rho(.)$ denotes the spectral radius and $\otimes$ the Kronecker product.

In the following section we present simple non-state-aware polices that will be used as a reference throughout the paper. Then we develop state-aware policies which are guaranteed to improve with respect to a Lagrangian relaxation of Problem 1, i.e., a linear combination of the control and communication costs. We proceed in Section IV to describe policies which by design meet the communication budget in Problem 1, while at the same time improve upon the control performance objective of Problem 1 with respect to the reference policies. This is achieved by enforcing the communication constraint to hold *in expectation over current plant state conditions*.

Remark 1. Formally our communication design problem can be seen as a Markov Decision Process (MDP) problem with state x_k and action α_k [13]. The difficulty in solving these MDP instances lies on the fact that x_k takes a continuum of values, rendering standard value or policy iteration algorithms computationally hard. See also [1], [7] for related state-aware communication design problems for control systems. Problem 1 is even more complex as it involves a constraint. □

III. Policies with Joint Control and Communication Guarantees

We begin by considering a simple non-state-aware policy that is feasible for Problem 1, and will be used as a reference policy throughout the paper. Suppose at each time step the sensor randomly decides whether to transmit or not with constant probability $\tilde{\alpha}$, that is, $\alpha_k = \tilde{\alpha}$ for all $k \geq 0$. With a slight abuse of notation we will denote this reference policy as $\tilde{\alpha}$. It is immediate that such a policy meets the communication constraint of Problem 1. In the following result we characterize the control performance of the reference policy denoted by $J_{\text{control}}(\tilde{\alpha})$. [1]

Proposition 1. *Let Assumption 1 hold. Consider the system described by (1) and the reference randomized policy* $\alpha_k = \tilde{\alpha}, k \geq 0$. *Then the system is mean square stable, i.e.,* $\limsup_{k\to\infty} \mathbb{E} x_k x_k^T < \infty$, *and the control system cost (7) equals* $J_{\text{control}}(\tilde{\alpha}) = Tr(P\,W)$ *where P is a positive semi-definite matrix satisfying the linear matrix equality*

$$P = q\, \tilde{\alpha}\, (Q_1 + A_c^T P A_c) + (1 - q\, \tilde{\alpha})(Q_0 + A_o^T P A_o). \quad (12)$$

This result is part of the random jump linear system theory [15] where it is actually shown that Assumption 1 is necessary and sufficient for stability under such a reference policy. This is the reason it is included in our work.

The interpretation of this result is that the quadratic function $x^T P x$ models the future expected control cost if the sensor is to follow this reference policy, i.e., the cost-to-go of this policy [13]. Using this function as a reference cost-to-go, we can design new policies that adapt online to the system state and have improved long term performance in comparison to the reference one. This procedure is known in the area of approximate dynamic programming as rollout policies [13, Vol. 1] – see also Remark 2 later.

We begin by considering long term performance of the system as a *Lagrangian relaxation* of the goals of Problem 1, that is,

$$J_{\text{control}}(\alpha_0, \alpha_1, \ldots) + \nu\, J_{\text{comm}}(\alpha_0, \alpha_1, \ldots) \quad (13)$$

[1] Due to space limitations the proofs are omitted but can be found in [14, Ch. 6].

for some positive weight $\nu \geq 0$. This is equivalently a linear combination of the control and communication objectives, and the weight ν can be tuned as a parameter to penalize more or less the communication resources. We explicitly describe state-aware policies which improve upon the reference ones with respect to the above relaxed objective.

Theorem 1 (Policies with Joint Guarantees)**.** *Let Assumption 1 hold. Consider the system described by (1) and for any non-negative constant $\nu \geq 0$, consider the policy $\alpha_k = \alpha^*_\nu(x_k)$ where*

$$\alpha^*_\nu(x) = \begin{cases} 1, & \text{if } x^T\, q\, M\, x \geq \nu \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

and

$$M = Q_0 + A_o^T P A_o - Q_1 - A_c^T P A_c, \tag{15}$$

and P is the positive semi-definite matrix satisfying (12). Then the average control and communication costs of this policy satisfy

$$J_{control}(\alpha^*_\nu) + \nu\, J_{comm}(\alpha^*_\nu) \leq J_{control}(\tilde{\alpha}) + \nu\, J_{comm}(\tilde{\alpha}). \tag{16}$$

A number of comments are in order. First the proposed policy (14) is not randomized like the reference policy but deterministic. It is also a threshold policy, i.e., the sensor transmits if a quadratic function of the current state exceeds some threshold. Similar threshold-based policies also appear in the event-based control framework of, e.g., [1], [3], [5]. The threshold in (14) is explicitly given by the chosen weight $\nu \geq 0$ on communication cost in the relaxed objective (13).

The matrix M in the policy is explicitly defined in (15) in terms of the plant dynamics A_o, A_c, the quadratic plant state costs Q_0, Q_1, and the reference matrix P, which by (12) depends on the required communication budget $\tilde{\alpha}$ and the packet drop rate q. We interpret $x^T M x$ as the relative *value of transmitting* the current state x. Indeed (15) involves a difference between matrices corresponding to the future long term control costs of not transmitting (open loop) and that of closing the loop, with P modeling the future reference cost of the system. By (14) the sensor transmits whenever the cost of closing the loop is at least ν units better than the cost of not transmitting. We note that the matrix M is not necessarily positive semi-definite.

The most important observation about Theorem 1 is that the proposed policy comes with a performance guarantee. For any given parameter $\nu \geq 0$, the proposed policy improves upon the reference policy with respect to the linear combination of control and communication costs (cf. (16)). Geometrically, we can think of average communication and control costs as two axes (Fig. 2) where the reference policy corresponds to the point $(\tilde{\alpha}, \mathrm{Tr}(PW))$. The proposed policy (14) is guaranteed to lie to the left of the line with slope $-\nu$ going through that reference point. As we will also see in numerical simulations this results in significant performance improvements in practice.

Unfortunately there are no separate guarantees about either control or communication performance of the proposed policy (14). In particular, it is not certain that the communication budget (10) in Problem 1 is met. A heuristic would be to

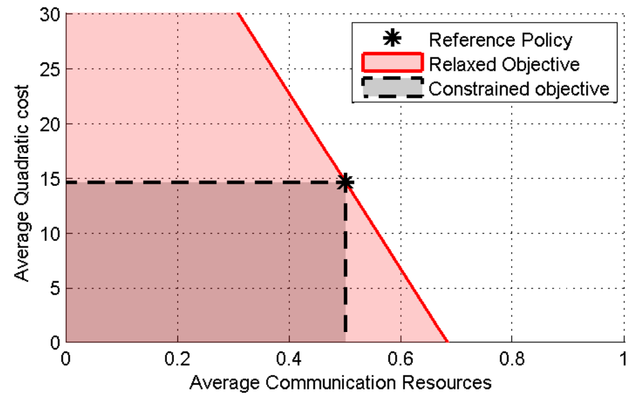

Fig. 2. Illustration of policies with respect to control and communication costs. The reference policy is the non-state-aware policy meeting the desired communication budget of Problem 1. For a fixed weight ν in the relaxed objective (13), the policy described in Theorem 1 lies at the left of the line with slope $-\nu$. On the other hand, the policy of Theorem 2 meets the desired communication guarantees by design and also improves upon the reference control cost, hence it lies below the reference point.

increase the weight $\nu >> 0$, but that deteriorates control performance as can be seen in the geometric interpretation of Fig. 2. In the following section we remedy this limitation by developing policies similar to (14) which are additionally enforced by design to meet the communication requirements.

IV. Policies with guaranteed communication costs

The policies in Theorem 1 of the previous section are stationary functions of the form $\alpha : \mathbb{R}^n \to [0, 1]$ mapping plant state conditions x_k to sensor transmission decisions $\alpha_k = \alpha(x_k)$. In this section we enhance these policies to enforce that *in expectation over the current plant state conditions* the average communication budget (10) in Problem 1 is met. That is, we design transmission functions $\alpha(.)$ such that $\mathbb{E}[\alpha(x_k)] \leq \tilde{\alpha}$ where the expectation is over the distribution of x_k. We characterize this distribution as follows.

At the beginning of time step k and before actually measuring state x_k, the sensor knows by (1) that x_k has a distribution $\phi_{0,W}$ centered at some current mean value denoted here by $\bar{x}_k$. For compactness we denote such a distribution as $\phi_{\bar{x}_k,W}$ and the integration with respect to it as $\mathbb{E}_{\bar{x}_k,W}$. The mean $\bar{x}_k$ evolves according to

$$\bar{x}_k = \begin{cases} A_c\, x_{k-1}, & \text{if } \gamma_{k-1} = 1 \\ A_o\, x_{k-1}, & \text{if } \gamma_{k-1} = 0 \end{cases}. \tag{17}$$

That is, if a transmission occurred at the last time step the sensor knows that the plant state evolves according to the closed loop mode A_c and expects a measurement x_k with mean $\bar{x}_k = A_c x_{k-1}$. The case of no transmission is similar. The sensor can keep track of this evolving mean value since it knows all past measurements and transmission successes.

The mean value $\bar{x}_k$ summarizes all the information available at the sensor about the distribution of x_k. We propose policies which specify that at each time step k given the current mean value $\bar{x}_k$ the sensor selects a current transmission function $\alpha_{\bar{x}_k}(.)$ and upon measuring the state x_k the sensor will transmit with the prescribed probability $\alpha_k = \alpha_{\bar{x}_k}(x_k)$. For brevity let us also denote the set of all such transmission functions given some value $\bar{x}_k$ as $\mathcal{A} = \{\alpha(.) : \mathbb{R}^n \to [0,1]\}$.

Our main result is as follows.

Theorem 2 (Policies with Communication Guarantees)**.** *Let Assumption 1 hold. Consider the setup of Theorem 1. Consider a dynamic sensor policy $\alpha^*_{\bar{x}}(x)$ defined for each mean value $\bar{x} \in \mathbb{R}^n$ as the optimal solution to the following optimization problem*

$$\underset{\alpha(.)\in\mathcal{A}}{\text{minimize}} \qquad \mathbb{E}_{\bar{x},W}[-\alpha(x)x^T qMx] \tag{18}$$

$$\text{subject to} \qquad \mathbb{E}_{\bar{x},W}[\alpha(x)] \leq \tilde{\alpha}. \tag{19}$$

Then the communication cost of this policy satisfies the desired bound

$$J_{comm}(\alpha^*_{\bar{x}}(x)) \leq \tilde{\alpha} \tag{20}$$

and the control cost of this policy satisfies

$$J_{control}(\alpha^*_{\bar{x}}(x)) \leq J_{control}(\tilde{\alpha}). \tag{21}$$

At each time step the sensor selects the current transmit function as a solution to the constrained optimization problem (18)-(19) where both the objective and the constraint are expressed *in expectation over the current plant state conditions* x_k to be measured. Constraint (19) guarantees that the communication budget $\tilde{\alpha}$ of Problem 1 is met in expectation *at each time step*, hence also in the long run (cf. (20)).

The expected cost in (18) entails the matrix M interpreted after Theorem 1, as the relative value of transmitting the current state. By minimizing (18), the sensor deviates from the reference policy at each time step but without loss in control performance (cf. (21)). See Fig. 2 for a geometric interpretation. In numerical simulations this results in significant performance improvements in practice.

To find the transmission function $\alpha^*_{\bar{x}}(x)$ given any current mean value $\bar{x} \in \mathbb{R}^n$, the sensor needs to solve problem (18)-(19) over the space of transmission functions. This is an infinite-dimensional optimization problem however it enjoys simple threshold-based solutions, as shown next.

Proposition 2. *Consider the transmission function optimization problem in (18)-(19) for some fixed $\bar{x} \in \mathbb{R}^n$. Then there exists a non-negative constant $\nu(\bar{x}) \geq 0$ such that the function*

$$\alpha^*_{\bar{x}}(x) = \begin{cases} 1, & \text{if } x^T q\, M\, x \geq \nu(\bar{x}) \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

is an optimal solution.

The proof relies on the fact that for any value $\bar{x} \in \mathbb{R}^n$ the problem (18)-(19) has zero duality gap, based on the results of [16], and that $\nu(\bar{x})$ is the optimal Lagrange dual variable.

The above proposition simplifies the search for general transmission functions in Theorem 2 to search for threshold-based functions. Moreover the resulting transmission policies $\alpha^*_{\bar{x}}(.)$ in (22) are of the same quadratic threshold nature as the policies of Theorem 1 or usual event-based policies [1], [3], [5]. The threshold value however here is not a free parameter as in Theorem 1, but it depends on the current mean value $\bar{x}_k$. A computationally efficient procedure for finding this threshold is discussed at the end of this section.

The proposed policy $\alpha^*_{\bar{x}}(x)$ is not stationary with respect to plant state, i.e., the transmission decision α_k is not just a time-invariant function of x_k but it also depends on its expected value $\bar{x}_k$ via the threshold. This dynamic anticipatory feature differentiates our policy from the usual time-invariant event-based control polices. The advantage of this dynamic feature is that it provides communication guarantees (cf. (20)).

We note that to implement the proposed policy a transmission rule solving problem (18)-(19) is needed for any possible mean value $\bar{x}_k \in \mathbb{R}^n$. This requires the solution of an infinite number of such optimization problems. In practice, as well as in the numerical simulations that follow, the space of mean values $\bar{x}_k \in \mathbb{R}^n$ can be discretized and solve instead a large number of optimization problems at the discrete points.

Remark 2. The policies in this paper are designed by modeling the future cost-to-go control performance according to the cost-to-go of the reference non-state-aware policy. The advantage is that the reference policy has a simple explicit quadratic cost-to-go function $x^T Px$ where P is given in (12). Policies improving upon a reference cost-to-go are known as rollout policies in approximate dynamic programming [13]. Rollout policies for sensor transmit power were designed in our previous work [7] but without any guarantees as in the current paper. Rollout sensor transmit policies have also been used in [6] however based on deterministic periodic schedules as references. As a result the policies in that work are updated once every period, unlike our policies in Theorems 1 and 2 which allow the sensor to continuously exploit online at each time step the stochastic plant state process. □

A. Computing Transmission Functions

As noted after Prop. 2 the threshold $\nu(\bar{x})$ corresponds to the optimal Lagrange dual variable of problem (18)-(19). We briefly describe a dual subgradient algorithm to find this optimal dual point (cf. [17, Ch.8]). A subgradient direction for the dual problem is typically given by the constraint slack of the primal problem evaluated at a primal Lagrangian minimizer. More precisely, given some dual variable $\nu_t \geq 0$ at iteration t, a corresponding primal solution is given by substituting ν_t in place of $\nu(\bar{x})$ in (22). The constraint slack of this solution with respect to (19) is given by the difference $\mathbb{P}_{\bar{x},W}[x^T q\, Mx \geq \nu_t] - \tilde{\alpha}$. Hence a dual subgradient ascent algorithm to compute the point ν_{t+1} for the next iteration is given by

$$\nu_{t+1} = \max\Big\{0,\, \nu_t + \varepsilon_t \, \Big(\mathbb{P}_{\bar{x},W}[x^T q\, Mx \geq \nu_t] - \tilde{\alpha}\Big)\Big\} \tag{23}$$

where $\varepsilon_t \geq 0$ is a stepsize, and the maximum is taken as dual points are non-negative. Iterating (23), the dual variable ν_t converges to the optimal $\nu(\bar{x})$ – see, e.g., [17, Ch. 8.2].

V. Numerical simulations

We consider a scalar control system as in the example of Section II with $A = 1.5$, $A + BK = 0.5$, Gaussian disturbance

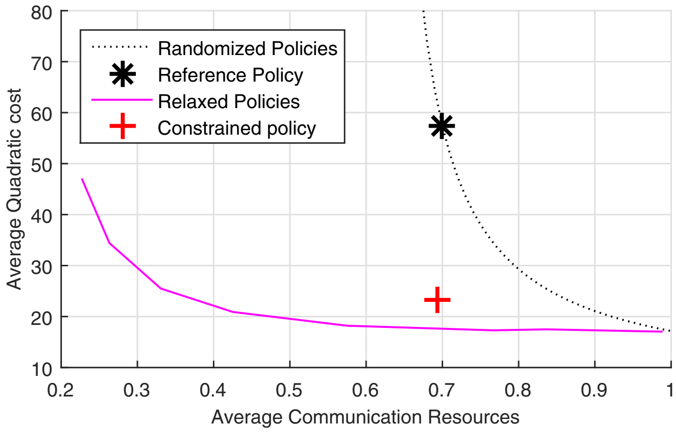

Fig. 3. Comparison of policies. The reference non-state-aware policy with the desired average communication constraint is shown. The policies of Theorem 1 based on the relaxed objective outperform the reference. The policy of Theorem 2 by design meets the communication constraint. For comparison, the set of all non-state-aware randomized policies are also plotted.

with zero mean and variance $W = 1$, and an LQR control stage cost with $Q = 1$, $R = 10$. Even though the plant is scalar, the overall control system with the plant and the estimator/controller as shown in (5) is two dimensional. We assume the packet success rate of the wireless channel is $q = 0.9$, and we are interested in solving Problem 1 with a communication budget $\tilde{\alpha} = 0.7$.

In Fig. 3 we plot the control and communication costs of non-state-aware randomized policies of the form $\alpha_k = \hat{\alpha}$ for different $\hat{\alpha} \in [0, 1]$. For low transmission rates the system becomes unstable and the control cost diverges. We pick the reference policy $\alpha_k = 0.7$ that meets our communication budget. For a range of parameters $\nu \geq 0$, we simulate the policies $\alpha_\nu(x)$ of Theorem 1 and plot their empirical cost. As expected from Theorem 1 they outperform the reference policy with respect to linear combinations of the two objectives. Even though the theory guarantees just an improvement, in simulations we observe a significant improvement of $\approx 50\%$ in both control and communication costs.

Finally, we simulate the dynamic policy of Theorem 2 which meets the required communication budget, as also shown in simulation. Moreover, it results in a significant improvement of $\approx 60\%$ in control cost in comparison with the reference. The dynamic policy performs slightly worse that some of the stationary policies of Theorem 1. However, to the best of our knowledge, there is no efficient procedure other than by simulation in order to find those better stationary policies.

VI. Concluding Remarks

We design dynamic sensor transmission policies for wireless control systems adapting online to the physical plant measurements. Based on approximate dynamic programming ideas the proposed policies come with improved control and communication performance. Moreover, we show how our policies can be modified to meet hard average resource constraints, such as a fixed communication resource budget. We prove, as well as illustrate in simulations, that our online policies outperform policies that do not adapt to the physical system state.

Our technique can be adapted for the reverse of Problem 1, that is, optimizing communication cost while guaranteeing a desired level of control performance. Moreover, our approach can be naturally extended to allocate communication resources among multiple control systems, e.g., for scheduling over a shared wireless channel [18], guaranteeing desired communication or control performance separately for each system. Future work also includes communication policies jointly adapted to plant and varying channel conditions [7]. Finally, the design requires further exploration when packet drops are not i.i.d. but due to packet collisions from other transmitting closed loop systems [19], [20].

References

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

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

[3] 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.

[4] M. Rabi and K. H. Johansson, "Scheduling packets for event-triggered control," in *European Control Conference (ECC)*, pp. 3779–3784, 2009.

[5] W. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in *51st Annual Conference on Decision and Control*, pp. 3270–3285, 2012.

[6] D. Antunes and W. Heemels, "Rollout event-triggered control: beyond periodic control performance," *IEEE Transactions on Automatic Control*, vol. 59, no. 12, pp. 3296–3311, 2014.

[7] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Optimal power management in wireless control systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1495–1510, 2014.

[8] R. Cogill, "Event-based control using quadratic approximate value functions.," in *Proc. of the 48th IEEE Conference on Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference*, pp. 5883–5888, 2009.

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

[10] L. Li, M. Lemmon, and X. Wang, "Event-triggered state estimation in vector linear processes," in *American Control Conference (ACC)*, pp. 2138–2143, 2010.

[11] 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.

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

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

[14] K. Gatsis, *Resource-aware Design of Wireless Control Systems*. PhD thesis, University of Pennsylvania, 2016.

[15] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-time Markov jump linear systems*. Springer Science & Business Media, 2006.

[16] 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.

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

[18] K. Gatsis, M. Pajic, A. Ribeiro, and G. J. Pappas, "Opportunistic control over shared wireless channels," *IEEE Trans. on Automatic Control*, vol. 60, no. 12, pp. 3140–3155, 2015.

[19] R. Blind and F. Allgöwer, "Analysis of networked event-based control with a shared communication medium: Part I-Pure ALOHA," in *IFAC World Congress*, 2011.

[20] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Control with random access wireless sensors," in *54th IEEE Conference on Decision and Control*, pp. 318–323, 2015.