Learning to Control over Unknown Wireless Channels

Konstantinos Gatsis, George J. Pappas

Abstract: Emerging control applications in the Internet-of-Things are increasingly relying on communication networks and wireless channels to close the loop. Traditional model-based approaches, i.e., assuming a known wireless channel model, are focused on analyzing stability and designing appropriate controller structures. Such modeling is challenging as wireless channels are typically unknown a priori and only available through data samples. In this work we aim to design data-based controllers using channel samples and provide high confidence guarantees on the performance of these controllers when deployed over the actual unknown channel. To achieve these results we combine statistical learning (concentration inequalities) with structural properties of our problem (monotonicity with respect to the unknown channel parameters), and also provide sample complexity analysis.

Keywords: Learning Algorithms; Networked Control Systems; Statistical Analysis; Communication channels; Controller Design

1. INTRODUCTION

Wireless communication is increasingly used to connect devices in industrial control environments, teams of robotic vehicles, and the Internet-of-Things. To guarantee safety and control performance it is customary to include a known model of the wireless channel, for example an i.i.d. or Markov link quality, alongside the model of the physical system to be controlled. In such modeled-based approaches one can characterize, for example, that it is impossible to estimate or stabilize an unstable plant if its growth rate is larger than the rate at which the link drops packets (Sinopoli et al., 2004; Schenato et al., 2007; Hespanha et al., 2007), or below a certain channel capacity (Tatikonda and Mitter, 2004; Sahai and Mitter, 2006). Models also facilitate the allocation of communication resources to optimize control performance (Quevedo et al., 2012; Gatsis et al., 2015), or event-triggered control (Heemels et al., 2012).

In practice however wireless autonomous systems operate under unpredictable channel conditions following unknown distributions, with only a finite amount of collected channel measurements available (Halperin et al., 2010; Rappaport et al., 2015). The purpose of this work is the design of controllers for networked control systems when only channel sample data are available instead of channel models. We use the data to learn directly a controller and also provide high confidence guarantees about the performance of this controller deployed over the actual unknown channel. Our paper is the first to provide such data-based performance guarantees for networked control, in contrast to the extensive literature on model-based approaches mentioned above.

Learning methods have been used in control problems most commonly within the reinforcement learning and approximate dynamic programming literature (Sutton and Barto, 1998; Bertsekas, 2012), where the goal is to learn to control an unknown dynamical system from data. One approach is to learn the system dynamics model first (Kumar and Varaiya, 2015) and recent focus has been on analyzing the sample complexity of control of unknown linear systems (Abbasi-Yadkori and Szepesvári, 2011; Dean et al., 2017). Other model-free approaches focus on learning the controller directly (Sutton and Barto, 1998). In contrast our work is focused on collecting data and learning unknown wireless channel models instead of system dynamics. In the context of networked control systems very recent works from the last two years are proposing data-based approaches including deep learning for allocating resources and scheduling (Demirel et al., 2018; Redder et al., 2019; Leong et al., 2018; Wu et al., 2018; Eisen et al., 2019) as well as for controller design (Schuurmans et al., 2019; Baumann et al., 2018). Our work is the first to provide data-based performance guarantees for networked control over unknown channels. In previous work we examined the sample complexity of stability analysis of networked control systems without controller design (Gatsis and Pappas, 2018, 2019). A related recent work appeared after the original submission of the present paper, considering the control of unknown scalar systems over unknown channels in an online setting (Singh and Kumar, 2020)

Rather than analysis, in this paper we consider the design of a state feedback controller for a linear dynamical system over a Bernoulli packet-dropping channel with an unknown success rate (Section 2). Using channel sample data, i.e., a number of packet successes and failures, we develop an algorithm to design a controller (Section 3). As our algorithm depends on random channel samples, the designed

⋆ K.G. is with the Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, United Kingdom (e-mail: konstantinos.gatsis@eng.ox.ac.uk)
G.J.P. is with the Department of Electrical and Systems Engineering, University of Pennsylvania, 200 South 33rd Street, Philadelphia, PA 19104, United States (e-mails: pappasg@seas.upenn.edu)
This research is partially supported by NSF CPS-1837253, by the Intel Science and Technology Center for Wireless Autonomous Systems (ISTC-WAS) and by ARL DCIST CRA W911NF-17-2-0181.

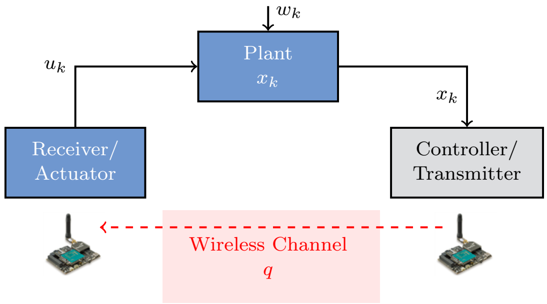

Fig. 1. Wireless Control System. A sensor measures the state of a plant perturbed by a random disturbance. The sensor transmits the measured information over a packet-dropping wireless channel to a receiver/controller providing control inputs.

controller is random and hence there is a need to provide guarantees about the performance of our controller over the true unknown wireless channel. To do this we utilize confidence bounds obtained by concentration inequalities, more specifically, Hoeffding's inequality.

Then we propose to design a controller assuming the worst channel conditions. We prove that this provides a guarantee on how large the control cost will be when such a controller is applied over the real channel. Then we also show that if we consider the best possible channel conditions, we may also provide a lower bound on the true control cost. As a result we provide performance guarantees that are data-dependent and high-confidence. The most related work is that by (Schuurmans et al., 2019) which designs controllers for general switched linear systems where the switching behavior is unknown and only available through samples and relies on distributionally robust optimization. Our work in contrast exploits the specific structure of the underlying problem (monotonicity with respect to channel parameters) and also provides high-performance guarantees about the cost achieved by the designed controller applied to the true unknown channel. We validate our theoretical analysis in numerical simulations (Section 4).

2. PROBLEM FORMULATION

We consider the control of a dynamical system over a packet dropping link. This is a standard model for control over an unreliable network or a wireless channel, for example when a controller transmits control input to be applied by a receiver – see Figure 1 and (Sinopoli et al., 2004; Schenato et al., 2007; Hespanha et al., 2007) for related examples. We suppose the system evolution is described by a linear time invariant model of the form

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

Here $x_k \in \mathbb{R}^n$ denotes the state of the system at each time k, the dynamics in open loop described by $A \in \mathbb{R}^{n\times n}$, $u_k \in \mathbb{R}^m$ is the control input, and $B \in \mathbb{R}^{n\times m}$ the input matrix. The additive terms $w_k, k \geq 0$ model an independent identically distributed (i.i.d.) noise process across time according to some known probability distribution with mean zero and positive definite covariance W.

A controller observes the state of the system, computes a desired control input, and transmits it over the channel. Then the evolution of the system depends on whether a successful transmission occurs at time k or not, indicated with variables $\gamma_k \in \{0, 1\}$. We focus on a linear controller of the form Kx_k for some gain K to be designed. At a unsuccessful transmission the control input is reset to zero, and otherwise when the transmission is successful the desired control input is applied. With such a communication and actuation model the overall system becomes

$$x_{k+1} = \begin{cases} (A + BK)\, x_k + w_k, & \text{if } \gamma_k = 1 \\ A\, x_k + w_k, & \text{if } \gamma_k = 0 \end{cases}. \tag{2}$$

The system evolution over time depends on whether the transmissions are successful or not over time as well as the chosen controller gain. In this paper we make the assumption that γ_k are independent Bernoulli random variables with a constant success probability q, and they are also independent from the system state x_k and noise w_k.

Given this model of the transmission success we are interested in the performance of the dynamical system. We will employ the usual quadratic system state and control cost, where Q, R are positive definite matrices of appropriate dimensions, with the long run average quadratic cost

$$J(K; q) = \limsup_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}[x_k^T\, Q\, x_k + u_k^T R u_k]. \tag{3}$$

The expectation at the right hand side accounts for the randomness introduced by the system disturbance and the channel. The results can also be written for the finite horizon problem. We denote the cost in (3) by $J(K, q)$ to make explicit the dependency on the controller gain K to be optimized, as well as on the the success rate of the channel q. If the channel quality q was known, then the controller design can be posed as a Linear Matrix Inequality problem using standard tools – see also Section 3.2 later.

Suppose that instead of knowing the packet success rate q of the channel we have available N samples denoted by $\gamma_k, k = 0, \ldots, N-1$ drawn independently from the Bernoulli distribution with success q. In practice this data is easy to obtain, it suffices to send N packets and record whether they are received or not.

Given N channel samples, the problem we tackle is to design a controller gain K_N and provide upper and lower bounds on the cost J_N^+, J_N^-. As these values are random because they depend on the randomly taken channel sample data, we would like to have high confidence results such that

$$\mathbb{P}\Big(J_N^- \leq J(K_N; q) \leq J_N^+\Big) \geq 1 - \delta \tag{4}$$

where δ is a desired confidence level. Here the probability is with respect to the channel samples provided.

Before we proceed, we have to also ensure the problem is well defined. More specifically since we are considering an infinite horizon problem we have to ensure that the system can be stabilized. As in standard linear quadratic control problems we assume that A, B is controllable and $A, Q^{1/2}$ is observable (which in our case holds because we assumed $Q \succ 0$). It is important also to note that the above problem does not necessarily have a well-defined solution for any value of the channel quality q, but there exists a minimum channel success rate q_c such that the system cannot be stabilized if $q \leq q_c$. Hence we make the following assumption.

Assumption 1. A value $q_c \in [0, 1]$ is given such that for any channel success rate $q > q_c$ there exists a controller gain K such that the closed loop system (2) is mean square stable.

We further discuss this assumption as well as how to obtain this value in Remark 1. This value q_c will appear in the rest of the paper as follows. If not sufficient channel samples are available then it may not be possible to verify whether the actual channel quality is above the value q_c and hence we may not be able to come up with a stabilizing controller.

3. SAMPLE-BASED CONTROLLER DESIGN WITH GUARANTEES

Suppose that we have available N samples denoted by $\gamma_k, k = 0, \ldots, N-1$ drawn independently from the Bernoulli distribution with success q. Given this data the most natural approximation of the true success probability q is the sample average

$$\hat{q}_N = \frac{1}{N} \sum_{k=0}^{N-1} \gamma_k \tag{5}$$

Indeed this approximation is in some sense optimal - it maximizes the likelihood of the success rate $q \in [0, 1]$ given the data $\gamma_k, k = 0, \ldots, N-1$. In the case of unlimited data samples the sample average converges almost surely to the true underlying packet success rate by the Strong Law of Large Numbers (Durrett, 2010, Ch.2). Hence in the face of unlimited data learning a controller would be easy. However this is an asymptotic analysis. In practice only finite amount of data will be available and this motivates us to investigate a finite sample analysis.

For a finite number of samples we argue that instead of point estimates of the channel success rate, confidence intervals are more useful. We further characterize how to construct controllers and provide performance guarantees based on confidence intervals.

3.1 Confidence intervals

Our approach is based on confidence intervals constructed by the channel sample data using concentration inequalities. In particular we may employ Hoeffding's inequality (Boucheron et al., 2013, Th. 2.8). Given a desired high confidence level $1 - \delta$ where δ is a small positive value, for example of the order of 10^{-3}, and after collecting N samples, we may derive an interval where the true underlying mean lies, that is, the channel success rate in our problem, as follows.

Lemma 1. Consider a sequence $\{\gamma_k, k = 0, \ldots, N-1\}$ of i.i.d. random variables taking values in $[0, 1]$ with mean q. Let $\hat{q}_N = \frac{1}{N} \sum_{k=0}^{N-1} \gamma_k$ be the sample average. Then for any $\delta \in (0, 1)$ it holds that

$$\mathbb{P}\left(q \in \left[\hat{q}_N - \sqrt{\frac{\log(2/\delta)}{2N}}, \, \hat{q}_N + \sqrt{\frac{\log(2/\delta)}{2N}} \right] \right) \geq 1 - \delta. \tag{6}$$

where the probability is with respect to the random sequence $\{\gamma_k, k = 0, \ldots, N-1\}$.

The result essentially states that there is a low probability that the sample average deviates much from the true packet success rate and further provides an explicit bound on this probability. We note that the result above holds regardless of the distribution as long as it has a bounded support – see also (Gatsis and Pappas, 2019) for less conservative approaches.

3.2 Sample-based Controller Design

After we construct a high confidence interval $[q_{\min}, q_{\max}]$ for the channel quality we consider the controller design problem. Specifically if the channel quality q was known perfectly, then we can follow a standard procedure to design an optimal controller gain K that minimizes the cost $J(K, q)$. To solve this problem we can write the Bellman equation for this problem and assume a quadratic function $x^T P x$ for the optimal cost-to-go/value function. Then we get that the matrix P needs to satisfy

$$P = Q + qK^T RK + (1-q)A^T PA + q(A+BK)^T P(A+BK) \tag{7}$$

and the cost for this controller K is given by

$$J(K, q) = \text{Tr}(PW). \tag{8}$$

A completion of squares in (7) can show that the optimal controller is given by

$$K = -(R + B^T PB)^{-1} B^T PA \tag{9}$$

and then P satisfies the Riccati-like equation

$$P = Q + A^T PA - qA^T PB(R + B^T PB)^{-1} B^T PA \tag{10}$$

It is important to note here that the channel success rate q affects the matrix P and subsequently this also affects the optimal control gain K in (9).

Before we proceed, we also show how to solve equation (10) as a Linear Matrix Inequality Problem – similar to (Boyd et al., 1994, p.126-127). In particular we may pose the design of the matrix P that solves the Riccati equation as the optimal solution to the problem

$$\begin{aligned} \text{maximize} \quad & \text{Tr}(PW) \\ \text{subject to} \quad & P \preceq Q + A^T PA \\ & \qquad - qA^T PB(R + B^T PB)^{-1} B^T PA, \\ & P \succ 0 \end{aligned} \tag{11}$$

where we have converted the Riccati equation to an inequality. Applying Schur's complement to the above problem, this can be equivalently written as

$$\begin{aligned} \text{maximize} \quad & \text{Tr}(PW) \\ \text{subject to} \quad & \begin{bmatrix} Q + A^T PA - P & \sqrt{q} A^T PB \\ \sqrt{q} B^T PA & R + B^T PB \end{bmatrix} \succeq 0, \\ & P \succ 0 \end{aligned} \tag{12}$$

which a Linear Matrix Inequality problem that may be readily solved. After one finds P, substituting to (9) yields the optimal controller.

Remark 1. We can understand Assumption 1 by noting that the above Riccati equation (10) does not have necessarily have a (positive definite) solution for any value of the channel quality q. Apart from the usual requirements of A, B being controllable and $A, Q^{1/2}$ being observable, then

Algorithm 1 Sample-based Controller Design

Input: Dynamics A, B, Cost Q, R, Noise covariance W, Confidence level δ, Number of samples N, Channel samples $\gamma_0, \ldots, \gamma_{N-1} \in \{0,1\}^N$

Compute the sample average

$$\hat{q}_N = \frac{1}{N} \sum_{k=0}^{N-1} \gamma_k \tag{13}$$

Compute the high confidence lower and upper bounds

$$q_{\min} = \hat{q}_N - \sqrt{\frac{\log(2/\delta)}{2N}} \tag{14}$$

$$q_{\max} = \hat{q}_N + \sqrt{\frac{\log(2/\delta)}{2N}} \tag{15}$$

Solve the following problem

$$\begin{aligned} \text{maximize} \quad & \mathrm{Tr}(PW) \\ \text{subject to} \quad & \begin{bmatrix} Q + A^T PA - P & \sqrt{q_{\min}} A^T PB \\ \sqrt{q_{\min}} B^T PA & R + B^T PB \end{bmatrix} \succeq 0, \\ & P \succ 0 \end{aligned} \tag{16}$$

if problem (16) is feasible **then**

  Let J_N^+ be the optimal value of problem (16)

  Let P_N be the optimal solution of problem (16) and compute

$$K_N = -(R + B^T P_N B)^{-1} B^T P_N A \tag{17}$$

  Let J_N^- be the optimal value of the following problem

$$\begin{aligned} \text{maximize} \quad & \mathrm{Tr}(PW) \\ \text{subject to} \quad & \begin{bmatrix} Q + A^T PA - P & \sqrt{q_{\max}} A^T PB \\ \sqrt{q_{\max}} B^T PA & R + B^T PB \end{bmatrix} \succeq 0, \\ & P \succ 0 \end{aligned} \tag{18}$$

  return K_N, J_N^+, J_N^-.

else

  return 'Undetermined'

end if

there exists a minimum channel success rate q_c such that the system can be stabilized for all $q > q_c$ and the equation (10) indeed has a positive definite solution. One potential way to search for q_c in the range $[0, 1]$ is with bisection using the LMI formulation in (12). Picking different potential values one can seek whether problem (12) is feasible and continue either increasing or decreasing the value until q_c is localized within some interval. Alternatively q_c can be analytically computed in some special cases. In the very special case where B is full rank, one may choose the controller gain $K = -B^{-1}A$ which makes the closed loop system $A + BK = 0$. In this case it can be derived from (1) that $q_c = 1 - 1/\rho(A)^2$ where $\rho(A)$ is the spectral radius of A.

As already mentioned, the true channel parameter q is unknown in our case and instead only a high confidence interval $[q_{\min}, q_{\max}]$ for the channel quality is constructed through the data. We may in principle select some candidate channel within this interval and treat it as the real channel for the purpose of controller design. Our specific approach is shown in Algorithm 1. We propose to optimize the performance for the 'worst case' channel $q_{\min}$ and use the corresponding cost as an upper bound on the true cost of the system with this controller. We propose to also optimize the performance for the 'best case' channel $q_{\max}$ and use the result as a lower bound on the true performance of the system.

3.3 High Confidence Performance Guarantees

In this section we establish the main theoretical results of our paper regarding the guarantees of Algorithm 1. Our analysis is motivated by the following technical lemma which establishes monotonicity properties of the performance objective with respect to the channel parameters.

Proposition 1. Consider system (1) to be controlled with controller as in (2) over an i.i.d. dropping channel. For any $q > q_c$, where q_c is as in Assumption 1, let $K(q)$ denote the optimal controller that minimizes the control cost $J(K; q)$ by (3) for a channel success rate q. Let $q', q'' \in [0, 1]$ be values such that $q_c < q' \leq q''$. Then it holds that

$$J(K(q''); q'') \leq J(K(q'); q''), \text{ and} \tag{19}$$

$$J(K(q'); q'') \leq J(K(q'); q'). \tag{20}$$

The first clause of this proposition is rather obvious. It says that the cost of a controller designed for a different channel condition will be larger than the cost of a optimal controller for the true channel condition. The second statement is the most important as it characterizes a monotonicity property. It states that if we plan a controller assuming a channel condition worse than the true channel condition, and then apply this controller at the true better channel condition, then the performance that we will experience is better (the cost is no larger than what was planned). This proposition, whose proof is omitted due to space constraints but follows standard Riccati and LMI arguments, forms the basis for our main result.

Theorem 1. (Sample-based Performance Guarantees) Consider the linear system (2) to be controlled over an i.i.d. Bernoulli binary channel with an unknown success probability $q \in [0, 1]$ and assume $q > q_c$ where q_c as in Assumption 1. Consider the controller design procedure developed in Algorithm 1 using N i.i.d. channel samples drawn with success rate q. If

$$N > \frac{2\log(2/\delta)}{(q - q_c)^2}, \tag{21}$$

then the algorithm returns values K_N, J_N^+, J_N^- such that

$$\mathbb{P}\Big(J_N^- \leq J(K_N; q) \leq J_N^+\Big) \geq (1 - \delta), \tag{22}$$

where the probability is with respect to the random channel samples.

The algorithm takes as inputs a random sample sequence of channel data and returns a controller as well as high confidence upper and lower bounds on the performance of this controller. Both the controller and the upper and lower performance bounds are constructed from the channel data sequence and they depend only on values that are known and not on the unknown channel parameter q. There is one caveat of Algorithm 1 which is the requirement (21) on the minimum number of channel data. If very few channel data are provided, then the algorithm cannot estimate well the current channel condition and returns the value 'Undetermined'. This minimum number of samples depends on the true channel success rate q which is a priori unknown. If the algorithm returns 'Undetermined', then the remedy would be to collect more channel data and try again. This minimum number of samples adversely depends on how close the channel success rate q is from the value q_c in Assumption 1. This can be thought of as a measure on how close the system is to being stabilizable. For less stabilizable systems more data will be needed. A similar relationship is also analyzed in our previous work (Gatsis and Pappas, 2018, 2019) which however did not consider a controller design problem.

Proof. First, we have that by Hoeffding's inequality (Lemma 1) that the following event

$$\left\{\hat{q}_N - \sqrt{\frac{\log(2/\delta)}{2N}} \leq q \leq \hat{q}_N + \sqrt{\frac{\log(2/\delta)}{2N}}\right\} \tag{23}$$

holds with probability at least $1-\delta$.

Next note that the lower bound satisfies

$$\hat{q}_N - \sqrt{\frac{\log(2/\delta)}{2N}} \geq q - 2\sqrt{\frac{\log(2/\delta)}{2N}}. \tag{24}$$

If condition (21) also holds, then we get that the lower bound satisfies

$$\hat{q}_N - \sqrt{\frac{\log(2/\delta)}{2N}} > q_c. \tag{25}$$

Combining all the above statements, the event

$$\left\{q_c < \hat{q}_N - \sqrt{\frac{\log(2/\delta)}{2N}} \leq q \leq \hat{q}_N + \sqrt{\frac{\log(2/\delta)}{2N}}\right\} \tag{26}$$

holds with probability at least $1-\delta$. Equivalently, for the variables in Algorithm 1 we establish that

$$\{q_c < q_{\min} \leq \hat{q}_N \leq q_{\max}\} \tag{27}$$

holds with probability at least $1-\delta$.

By Assumption 1 the event $q_c < q_{\min}$ implies that optimizing the controller K for the cost $J(K; q_{\min})$ is feasible, so equivalently the problem (16) in the algorithm is feasible and the algorithm returns an answer K_N, J_N^-, J_N^+.

Next we show that (22) is true. We will make use of Proposition 1. In the notation of this proposition the algorithm returns $K_N = K(q_{\min})$ and $J_N^+ = J(K(q_{\min}); q_{\min})$. We have argued that $q_c < q_{\min} \leq q$ holds with probability at least $1-\delta$. Hence applying (20) we get that

$$J(K_N, q) = J(K(q_{\min}); q) \leq J(K(q_{\min}); q_{\min}) = J_N^+ \tag{28}$$

This proves the upper bound inequality in (22).

Then note that by design the algorithm chooses $J_N^- = J(K(q_{\max}); q_{\max})$. We have argued that $q_c < q_{\min} \leq q \leq q_{\max}$ holds with probability at least $1-\delta$. Hence applying (19) we get that

$$J_N^- = J(K(q_{\max}); q_{\max}) \leq J(K(q); q_{\max}) \tag{29}$$

Then applying (20) on the right hand side we get that

$$J(K(q); q_{\max}) \leq J(K(q); q) \tag{30}$$

And applying (19) again on the right hand side we get

$$J(K(q); q) \leq J(K(q_{\min}); q) = J(K_N, q) \tag{31}$$

Combining (29)-(31) we get the lower bound inequality in (22). This concludes the proof. ■

Remark 2. We point out that in principle we may select any candidate channel quality in the high confidence interval $[q_{\min}, q_{\max}]$ for the purpose of controller design. The advantage of the specific worst case design in Algorithm 1 is that we can provide upper and lower bound guarantees on performance with high confidence. We plan to investigate other choices in future work.

4. NUMERICAL RESULTS

We consider the control of a system with known dynamics given by

$$A = \begin{bmatrix} 2 & -1 \\ 0 & 0.9 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{32}$$

and noise covariance $W = I$. The minimum required success rate is $q_c \approx 0.75$ as in Remark 1. We also consider $Q = I$ and $R = 10I$ in (3). We suppose the true underlying channel has a packet success rate $q = 0.9$. We select a confidence level $\delta = 10^{-5}$ and we collect N i.i.d. samples (packet successes and failures) from this channel that we feed to Algorithm 1. For each value of N the algorithm provides a data-driven controller as well as upper and lower bounds on the true performance of this controller over the true channel, plotted in Fig 2 along with the true performance of the controller. For clarity of exposition we normalize these costs by dividing with the true optimal cost of the system under this channel condition. We see that it takes 300 samples before the algorithm can provide a result (a stabilizing controller). Then as the number of samples increases the upper and lower bound tend close to each other. We also observe that for all times when the algorithm returns an answer, the provided upper and lower bounds indeed contain the true performance of the system (cf. Theorem 1). As the number of data increases the algorithm converges to the true optimal control performance.

5. CONCLUSION

Motivated by the deployment of connected autonomous systems in smart infrastructures and the Internet-of-Things in this paper we consider the problem of learning to control over unknown channels. In particular we consider the problem of designing a state feedback controller for a linear system to be controlled over a packet dropping link with unknown success probability. Using a sequence of collected channel sample data we design a controller and provide high confidence guarantees about the performance of the sample-based controller on the actual channel. To get this result we exploit the structural properties of this problem and combine them with concentration inequality results.

Future work includes extending the approach to more general channel models as well as comparisons with model-free methods such as reinforcement learning.

REFERENCES

Abbasi-Yadkori, Y. and Szepesvári, C. (2011). Regret bounds for the adaptive control of linear quadratic systems. In *Proceedings of the 24th Annual Conference on Learning Theory*, 1–26.

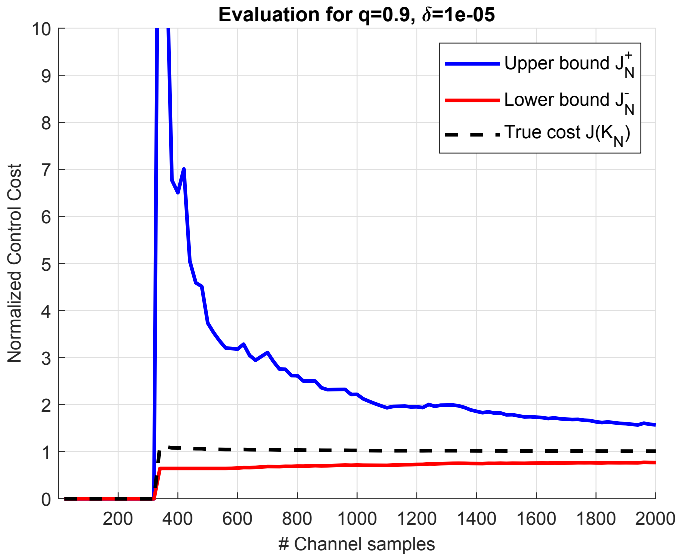

Fig. 2. We implement the sample-based controller design algorithm. After a number a channel samples the procedure is able to produce a stabilizing controller. We observe the upper and lower bounds on control cost contain the true control cost and converge to it as the number of samples increases.

Baumann, D., Zhu, J.J., Martiu, G., and Trimpe, S. (2018). Deep reinforcement learning for event-triggered control. In *2018 IEEE 57rd Annual Conference on Decision and Control (CDC)*.

Bertsekas, D.P. (2012). *Dynamic programming and optimal control*, volume 2. 4th Edition, Athena scientific Belmont, MA.

Boucheron, S., Lugosi, G., and Massart, P. (2013). *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press.

Boyd, S., El Ghaoui, L., Feron, E., and Balakrishnan, V. (1994). *Linear matrix inequalities in system and control theory*, volume 15. Siam.

Dean, S., Mania, H., Matni, N., Recht, B., and Tu, S. (2017). On the sample complexity of the linear quadratic regulator. *arXiv preprint arXiv:1710.01688*.

Demirel, B., Ramaswamy, A., Quevedo, D.E., and Karl, H. (2018). Deepcas: A deep reinforcement learning algorithm for control-aware scheduling. In *2018 IEEE 57rd Annual Conference on Decision and Control (CDC)*.

Durrett, R. (2010). *Probability: theory and examples*. Cambridge University Press.

Eisen, M., Rashid, M.M., Cavalcanti, D., and Ribeiro, A. (2019). Control-aware scheduling for low latency wireless systems with deep learning. *arXiv preprint arXiv:1906.06225*.

Gatsis, K., Pajic, M., Ribeiro, A., and Pappas, G.J. (2015). Opportunistic control over shared wireless channels. *IEEE Transactions on Automatic Control*, 60(12), 3140–3155.

Gatsis, K. and Pappas, G.J. (2018). Sample complexity of networked control systems over unknown channels. In *57th IEEE Conference on Decision and Control (CDC 2018)*.

Gatsis, K. and Pappas, G.J. (2019). Statistical learning for analysis of networked control systems over unknown channels. Submitted. Available on Arxiv.

Halperin, D., Hu, W., Sheth, A., and Wetherall, D. (2010). Predictable 802.11 packet delivery from wireless channel measurements. In *ACM SIGCOMM Computer Communication Review*, volume 40, 159–170. ACM.

Heemels, W., Johansson, K.H., and Tabuada, P. (2012). An introduction to event-triggered and self-triggered control. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, 3270–3285. IEEE.

Hespanha, J.P., Naghshtabrizi, P., and Xu, Y. (2007). A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1), 138–162.

Kumar, P.R. and Varaiya, P. (2015). *Stochastic systems: Estimation, identification, and adaptive control*, volume 75. SIAM.

Leong, A.S., Ramaswamy, A., Quevedo, D.E., Karl, H., and Shi, L. (2018). Deep reinforcement learning for wireless sensor scheduling in cyber-physical systems. *arXiv preprint arXiv:1809.05149*.

Quevedo, D.E., Ahlén, A., Leong, A.S., and Dey, S. (2012). On Kalman filtering over fading wireless channels with controlled transmission powers. *Automatica*, 48(7), 1306–1316.

Rappaport, T.S., MacCartney, G.R., Samimi, M.K., and Sun, S. (2015). Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design. *IEEE Transactions on Communications*, 63(9), 3029–3056.

Redder, A., Ramaswamy, A., and Quevedo, D.E. (2019). Deep reinforcement learning for scheduling in large-scale networked control systems. *arXiv preprint arXiv:1905.05992*.

Sahai, A. and Mitter, S. (2006). The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link&# 8212; part i: Scalar systems. *IEEE Transactions on Information Theory*, 52(8), 3369–3395.

Schenato, L., Sinopoli, B., Franceschetti, M., Poolla, K., and Sastry, S. (2007). Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, 95(1), 163–187.

Schuurmans, M., Sopasakis, P., and Patrinos, P. (2019). Safe learning-based control of stochastic jump linear systems: a distributionally robust approach. *arXiv preprint arXiv:1903.10040*.

Singh, R. and Kumar, P. (2020). Learning in networked control systems. *arXiv preprint arXiv:2003.09596*.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M., and Sastry, S. (2004). Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 49(9), 1453–1464.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement learning: An introduction*. MIT press Cambridge.

Tatikonda, S. and Mitter, S. (2004). Control under communication constraints. *IEEE Transactions on Automatic Control*, 49(7), 1056–1068.

Wu, S., Ren, X., Jia, Q.S., Johansson, K.H., and Shi, L. (2018). Learning optimal scheduling policy for remote state estimation under uncertain channel condition. *arXiv preprint arXiv:1810.09820*.