

Side-information in Control and Estimation

Govind Ramnarayan, Gireeja Ranade and Anant Sahai

Wireless Foundations, EECS, UC Berkeley

gramnarayan@berkeley.edu, gireeja@eecs.berkeley.edu, sahai@eecs.berkeley.edu

Abstract—As in portfolio theory, we can think of the value of side-information in a control system as the change in the “growth rate” due to side-information. A scalar counterexample (motivated by carry-free deterministic models) shows the value of side-information for control does not exactly parallel the value of side-information for portfolios. Mutual-information does not seem to be a bound here.

The concept is further explored through a spinning vector control system that is re-oriented at each time so that the control or observation direction is partially unknown. The value of side-information can be calculated in this setup and it behaves quite differently in a control vs. estimation context. A second example considers the problem of vector control over a (scalar) erasure channel, the dual problem to the estimation problem of intermittent Kalman Filtering. The value of information here is measured through the change in the critical packet-drop probability for the system. While non-causal side-information regarding the packet arrivals does not affect the critical probability for the estimation problem, we find that it can generically be very valuable for the control problem — it seems to change the scaling behavior for the control counterpart to what would be considered the “high SNR limit” in communication problems.

I. INTRODUCTION

Parameter uncertainty has a long history in control theory — the very idea of robust control is about dealing with it. Recently, the advent of networked control systems has made stochastic uncertainty models more relevant. There is now a real need to have a theory capable of dealing with side-information in control. As just one example, control theorists are interested in knowing how networked control systems behave with or without acknowledgements of dropped packets since this is relevant for choosing among practical protocols like TCP vs. UDP [1]. Acknowledgements are a kind of side-information about control channel state, but as of now, there is no theoretical guidance for how to think about it in a principled way.

Fortunately, such things have long been studied in information theory in the context of unknown fading channels [2]. Medard in [3] examines the effect of imperfect channel knowledge on capacity, and Lapidot and Shamai quantify the degradation in performance due to channel-state estimation errors by the receiver [4]. Pradhan et al. show that the duality between source and channel coding in fact extends to the case with side-information under certain conditions [5]: this is particularly interesting given the well-known parallel between source coding and portfolio theory, which we will connect to here. Further, Kotagiri and Laneman [6] study the impact of non-causal knowledge of the state in a multiple-access setting. There are many more interesting results as well, but space precludes any serious discussion here.

Moving beyond communication, the MMSE dimension looks at the value of side-information in an estimation setting. In a system with only additive noise, Wu and Verdu show that a finite number of bits of side-information regarding the additive noise cannot generically change the high-SNR scaling behavior of the MMSE [7]. Portfolio theory also gives us an understanding of side-information. The key is the doubling rate of the system, i.e. the rate at which a gambler who chooses an optimal portfolio doubles his principal. Kelly studied this through bets placed on horse races in [8]. If each race outcome is distributed according a random variable X , then the mutual information between X and Y , $I(X; Y)$, measures the gain in the doubling rate that the side-information Y provides the gambler.

Cover showed the existence of universal portfolios [9] as well the impact of side information for these [10]. This leads to a natural question: if there exists a portfolio that can perform optimally while agnostic to the parameters of the systems, under what circumstances can we design control strategies that work universally? What is the parallel in control?

Control systems, like portfolios, have an underpinning of exponential growth. Just as the investor can choose to buy and sell at each time step to maximize growth, the controller has the choice of control strategy to minimize growth (or maximize decay). Further, causality and time are important considerations in both portfolios and control. Directed mutual information captures exactly the causal information that is shared between two random variables. This connection has been made explicit for portfolio theory in [11], [12] by showing that the directed mutual information $I(X^n \rightarrow Y^n)$ is the gain in the doubling rate for a gambler due to causal side information Y^n . Of course, directed mutual information is central to control and information theory as the measure of the capacity of a channel with feedback [13].

Here, we explore the value of both causal and non-causal side information for control systems through models that involve multiplicative parameter uncertainty, where these parameters have an i.i.d. character to them. Multiplicative models exhibit fundamentally different behavior than additive noise models do. In models with additive noise, the linearity of the system and the linearity of expectation means that estimation and control problems reduce to each other — the optimal control is a deterministic function of the optimal estimate. Multiplicative noise breaks this duality and the philosophical differences in estimation and control become evident operationally as well.

We start with a simple scalar example and then define

value of side-information for control in a way that parallels information-theoretic portfolio theory. Then, we discuss two interesting vector examples, the latter of which demands a different (coarser) way of understanding the value of side-information.

II. A SCALAR EXAMPLE AND SEMI-DETERMINISTIC STORY

This section draws heavily upon our earlier Allerton paper [14] but helps make the ideas above more concrete. Consider a simple scalar control system with perfect state-observation

$$\begin{aligned} X[n+1] &= \alpha(X[n] + B[n]U[n]), \\ Y[n] &= X[n]. \end{aligned} \quad (1)$$

Suppose $B[n]$ are a series of i.i.d. random variables with mean μ_B and variance σ_B^2 , and $X[0] \sim \mathcal{N}(0, 1)$. The system is scaled by a scalar constant factor α at each time step. The aim is to choose $U[n]$, a function of $Y[n]$, so as to stabilize the system. We can show that the system (1) is mean-square stabilizable using linear strategies if $\lim_{n \rightarrow \infty} \mathbb{E}[X[n]^2] < \infty$ if $\alpha^2 < \left(\frac{\mu_B^2 + \sigma_B^2}{\sigma_B^2} \right)$.

The growth rate of the system we are considering above is related to the flow of information through the system: the randomness in the control parameter $B[n]$ impedes the controller's ability to stabilize the system. Recent works have shown that a deterministic bit-level perspective (a la [15], [16], [17]) on control systems can help elucidate the information flows in the system [18], [19].

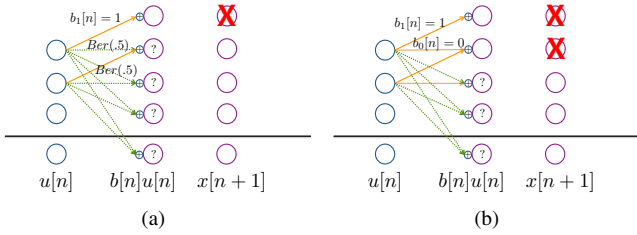


Fig. 1. This system has the highest deterministic link at level $b_{det} = 1$ and the highest unknown link at $b_{rand} = 0$. Bits $b_{-1}[n], b_{-2}[n], \dots$ are all random Bernoulli- $(\frac{1}{2})$. As a result the controller can only influence the top bits of the control going in, and can only cancel one bit of the state.

In Figure 1, we consider a simple bit-level carry-free model that illustrates this. Say the control gain $B[n]$ has one deterministic bit, so that $b_{det} = 1$, but all lower bits are random Bernoulli- $\frac{1}{2}$ bits. Then the controller can only reliably cancel $1 - 0 = 1$ bit of the state each time. The difference between the level of the deterministic bits and the level of the random bits is what determines the number of controllable bits. Clearly, if the value of b_0 was also known, then we could tolerate a growth from α of two bits at a time. We can think of this as the value of the side-information b_0 for this problem.

It is interesting to also consider the following ‘dual’ estimation problem

$$\begin{aligned} X[n+1] &= \alpha X[n], \\ Y[n] &= C[n]X[n]. \end{aligned} \quad (2)$$

where $C[n]$ are i.i.d. with a continuous density, and $X[0] \sim \mathcal{N}(0, 1)$. We know from [14] that a finite number of bits of side-information regarding the system parameter $C[n]$ do not help us estimate the system or decrease the growth rate of the error by more than a subexponential factor. Side-information is useless in this estimation problem.

III. THE VALUE OF INFORMATION

Consider a real-valued control system \mathcal{S} , with state $X[n]$, control $U[n]$ and observation $Y[n]$ at time n as below

$$\begin{aligned} X[n+1] &= \alpha \cdot f(X[n], U[n], \mathcal{T}[n]), \\ Y[n] &= g(X[n], \mathcal{T}[n]). \end{aligned} \quad (3)$$

Let $\mathcal{T}[n]$ be the set of random variables associated with the system at time n . Let $F_{\mathcal{T}[n]}$ be the set of distributions associated with them, and we assume these are known to the controller. f and g are fixed, known, deterministic functions. α is a scalar, known constant. The initial state $X[0]$ is random. The control strategy is a function $U[n] : Y_0^n \rightarrow \mathbb{R}$.

For instance, for the system (1), $\mathcal{T}[n] = \{B[n]\}$ and $F_{\mathcal{T}[n]}$ is effectively $\{F_B\}$ at each n since $B[n]$ are i.i.d..

Parallel to the doubling rate defined in portfolio theory, we define the one-step logarithmic decay rate of a system for state $X[n]$, control strategy $U[n]$ and system randomness $F_{\mathcal{T}[n]}$ at time n as:

$$G_{\mathcal{S}}(X[n], U[n], F_{\mathcal{T}[n]}) \stackrel{\text{def}}{=} \mathbb{E} \left[\ln \frac{\|X[n]\|}{\|X[n+1]\|} \right]. \quad (4)$$

The expectation is over the randomness $F_{\mathcal{T}[n]}$. A system is logarithmically stabilizable if there exists a strategy U_0^∞ such that $\sum_{i=0}^\infty \mathbb{E} [G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})] \rightarrow \infty$. With this, we define the average decay rate of the system as

$$\bar{G}_{\mathcal{S}}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty}) \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n \mathbb{E} [G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})] \quad (5)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left[\ln \frac{\|X[0]\|}{\|X[n+1]\|} \right] \quad (6)$$

The expectation inside the sum in eq. (5) is over the random state $X[i]$. If $\bar{G}_{\mathcal{S}} > 0$ then the system is clearly logarithmically stabilizable.

The optimal control strategy determines the maximal average decay rate, $G_{\mathcal{S}}^*(F_{\mathcal{T}_0^\infty})$, with the expectation over $X[0]$.

$$G_{\mathcal{S}}^*(F_{\mathcal{T}_0^\infty}) \stackrel{\text{def}}{=} \max_{U_0^\infty} \mathbb{E} [\bar{G}_{\mathcal{S}}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty})]. \quad (7)$$

Let α^* be the maximum α such that the system is still stabilizable. Then $\ln \alpha^*$ can be thought of as the ‘tolerable growth factor’. In general, if we set $\alpha = 1$ in eq. (3), then the decay rate $G_{\mathcal{S}}^*(F_{\mathcal{T}_0^\infty})$ is equal to $\ln \alpha^*$.

Now are set to define the average value of side-information as the change in the optimal decay rate with side-information $\mathcal{Z}[n]$ (provided to the controller at time n) as

$$\mathbb{E} [G_{\mathcal{S}}^*(F_{\mathcal{T}_0^\infty | \mathcal{Z}_0^\infty}) - G_{\mathcal{S}}^*(F_{\mathcal{T}_0^\infty})]. \quad (8)$$

The expectation is taken over the random side-information vector \mathcal{Z}_0^∞ .

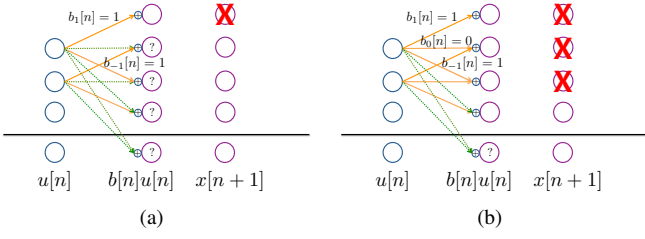


Fig. 2. Consider the following gain for the controller in (a): $b_1[n] = 1, b_{-1}[n] = 1$ are deterministically known, but all other links are Bernoulli(.5). Only a gain of $\log \alpha = 1$ can be tolerated in this case. Now, say side-information regarding the value of $b_0[n]$ is received as in (b). This suddenly buys the controller not just one, but two bits of growth.

Note that in the carry-free model in Fig. 1, one extra bit of information about b_0 increases the tolerable growth of the system by exactly one bit. What is the potential value of R -bits of side-information? This is the answer to the optimization problem

$$V_S(R) \stackrel{\text{def}}{=} \max_{I(\mathcal{T}[n]; \mathcal{Z}[n])=R \forall n} \mathbb{E} [G_S^*(F_{T_0^\infty} | \mathcal{Z}_0^\infty) - G_S^*(F_{T_0^\infty})]. \quad (9)$$

Finally, we also define a corresponding decay rate for estimation. For the system \mathcal{S} :

$$\begin{aligned} X[n+1] &= \alpha \cdot f(X[n], \mathcal{T}[n]), \\ Y[n] &= g(X[n], \mathcal{T}[n]), \end{aligned} \quad (10)$$

we define the one-step logarithmic error decay rate as

$$G_S(\hat{X}[n], F_{\mathcal{T}[n]}) \stackrel{\text{def}}{=} \mathbb{E} \left[\ln \frac{\|X[n] - \hat{X}[n]\|}{\|X[n+1] - \hat{X}[n+1]\|} \right]. \quad (11)$$

The average logarithmic decay rate can then be defined as in the control case.

A. A control counterexample

In the portfolio theory literature, it is known that the maximum increase in doubling rate due to side-information Z for a set of stocks distributed as T is upper bounded by $I(T; Z)$. With our observation about deterministic models it is tempting to conjecture that “a bit buys a bit” and a similar bound holds for the value of information in control systems. However, we see that the following counterexample rejects this conjecture. Consider the carry-free model in Fig. 2. In Fig. 2(a) the uncertainty in $b_0[n]$ does not allow the controller to utilize the knowledge that $b_{-1}[n] = 1$. However, one bit of information $b_0[n]$ in Fig. 2(b), lets the controller buy two bits of gain in the tolerable growth rate as explained in the caption. In the case of portfolio theory, it is possible to hedge across uncertainty in the system and get “partial-credit” for uncertain quantities. This is not possible in communication¹

¹It seems that the ‘commitment’ challenge that is faced by control can also be seen in communication systems, where it is also not possible to hedge across realizations. Consider a “compound” channel made of two R -bit channels A and B but with distinct inputs, so only one can be used at a time. The message sent across one of the channels is randomly erased with probability 0.5. In this case, one bit of side-information about which channel is to be erased can buy us more than a bit: we get $\frac{R}{2}$ bits of message on average.

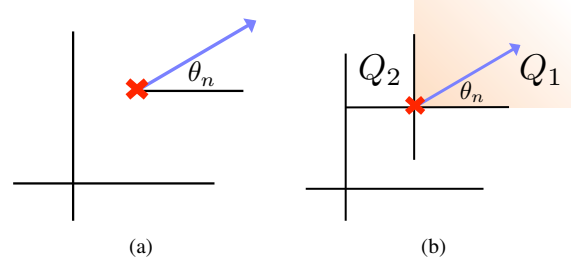


Fig. 3. A spinning control setup: the target can be moved only along the randomly chosen control direction. In version (a), the controller has perfect access to the control direction, while in version (b) only a quantized version is available.

and control systems since it is not possible to hedge a control signal in the same way one can hedge a bet.

IV. A SPINNING SYSTEM

The first example here highlights the difference in the impact of side-information for control and estimation problems.

A. A spinning controller: the control case

Consider the noiseless 2D control system \mathcal{S} in (12)

$$\begin{aligned} \begin{bmatrix} X_1[n+1] \\ X_2[n+1] \end{bmatrix} &= \alpha \left(\begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix} + U[n] \begin{bmatrix} \cos \theta_n \\ \sin \theta_n \end{bmatrix} \right), \\ Y[n] &= \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}. \end{aligned} \quad (12)$$

The controller has perfect access to the system state, but is subject to the following limitation: at each time n , the control direction is determined by a random spin, i.e. the controller may only act along the direction $[\cos \theta_n \sin \theta_n]^T$, where θ_n is drawn uniformly from $[0, 2\pi]$. Information about the control direction, θ_n , is revealed to the controller before it chooses $U[n]$. The initial state $[X_1[0] X_2[0]]^T$ is drawn randomly according to some distribution, and the goal is to drive the state to the origin. After the control acts, the system is spun again so that only the distance of the target from the origin is preserved, and the scale α is applied. This is depicted in Fig. 3(a).

Consider the case where no information about the control direction θ_n is revealed to the controller before acting, i.e. the controller has 0 bits of information about the system randomness. Clearly, in this case the optimal control action $U[n] = 0$, and no growth α can be tolerated. The system is only stable if $\alpha \leq 1$. For the other extreme case, where the controller knows the present control direction perfectly, the following theorem characterizes the optimal strategy.

Theorem 4.1: The optimal control for the system (12) is given by the greedy strategy, i.e. $U^*[n] = -X_1[n] \cos \theta_n - X_2[n] \sin \theta_n$ for all n .

The proof follows using dynamic programming.

Corollary 4.2: The logarithmic decay rate for system (12) with perfect information at each time n about θ_n is $\ln 2$, and the tolerable growth rate is thus $\alpha^* = 2$.

Proof: The logarithmic decay rate of the system for the optimal control is given by

$$\frac{1}{\pi} \int_0^\pi \frac{1}{2} \ln \frac{|X_1[0]|^2 + |X_2[0]|^2}{|X_1[0] \cos \theta - X_2[0] \sin \theta|^2} d\theta \quad (13)$$

This integral evaluates to $\ln 2$, and hence $\alpha^* = 2$. ■

B. Partial side-information

The symmetric randomness in this example makes it easy to evaluate the impact of side-information regarding the control randomness. Instead of perfect side-information, what happens when the controller has access to only two bits of information about the control direction?

Consider the space divided into quadrants, and only the quadrant containing the direction will be revealed at time n . Say only the quadrant of the control direction θ_n , Q_1 or Q_2 , is revealed to the controller at time n (Fig. 3(b)).

Theorem 4.3: The logarithmic decay rate with two bits of side-information for the system (12) is at least .47, and the tolerable growth rate is thus at least $\alpha = 1.61$.

This also follows using dynamic programming. Similar results for k -bits of side-information are summarized in Table I. Just three bits of side-information gets the tolerable growth rate pretty close to the case of perfect side-information.

TABLE I
SYSTEM GROWTH AS A FUNCTION OF STATE-INFORMATION

Side-info	Decay rate	Tolerable growth
0 bits	0	1
1 bit	0.200	1.22
2 bits	0.477	1.61
3 bits	0.624	1.86
∞ bits	$\ln 2$	2

Note that even in the presence of noisy information about the control direction it is possible to stabilize the system for certain growth rates α . This parallels the result in [14].

C. A spinning observer: the estimation case

The behavior of the corresponding estimation problem presents a sharp contrast to the control problem. Consider the system below, where the observation directions θ_n are random

$$\begin{aligned} \begin{bmatrix} X_1[n+1] \\ X_2[n+1] \end{bmatrix} &= \alpha \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}, \\ Y[n] &= [\cos \theta_n \quad \sin \theta_n] \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}. \end{aligned} \quad (14)$$

If θ_n is perfectly known to the observer, the estimation error goes to zero after the first two observations. So the logarithmic decay rate of the error with perfect information is infinity, unlike the control case which has a small finite decay rate.

Surprisingly, the two-step observability result is quite fragile. We know from the arguments in [14] that even a slight continuous uncertainty regarding θ_n renders the estimation problem impossible. The error is not shrinking with time. Partial side-information is no more useful than no side-information at all.

V. AN INTERMITTENT CONTROLLER

This section considers the control problem that is the dual of the intermittent Kalman filtering problem [20]. In [21], Mo and Sinopoli found two interesting examples that defined corner cases for the critical erasure probability for the estimation problem. Building on this, Park and Sahai characterized the difference in the critical drop probability in the presence and absence of eigenvalue cycles [22]. The control counterpart is defined below:

$$\begin{aligned} \begin{bmatrix} X_1[n+1] \\ X_2[n+1] \end{bmatrix} &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix} + U[n] \beta[n] \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ Y[n] &= \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}. \end{aligned} \quad (15)$$

Let $\lambda_{max} = |\lambda_1| \geq |\lambda_2| \geq 1$, and $\beta[n]$ is a Bernoulli- (p) random variable. We use the terminology ‘control arrival’ in the event that $\beta[n] = 1$. [23] showed that this system can be mean-square stabilized by an LTI controller if and only if $1 - p < \frac{1}{(\lambda_1 \lambda_2)^2}$.

Here, we investigate the impact of partial non-causal side-information about control arrivals on the critical probability. The logarithmic decay rate of the system does not serve as a good measure in this problem since the probability of the system state eventually being set to zero is 1, but the interesting question is that of the rate of decay. The change in the critical probability can be thought of as a proxy for the value of side-information and we explore how it changes.

Non-causal look-ahead regarding the sequence $\beta[n]$ does not change the critical probability for the observation problem. We are interested in understanding the effect of this side-information for the control problem.

Our first observation is to note that with infinite look-ahead on the sequence $\beta[n]$, the controller is able to plan for future arrivals and will be able to set the state to zero on the second arrival.

Theorem 5.1: The critical probability for the control problem eq. (15) is $\frac{1}{\lambda_{max}^2}$.

The proof follows from a time-reversal argument: the control problem is the same as the observation problem if time is reversed and we are only waiting on the first two arrivals. The result follows from arguments in [24]. ■

With this background, we consider the case with $\lambda_1 = 2$ and $\lambda_2 = -2$. These eigenvalues form a cycle of period two. We can further separate out the angle and the magnitude of the eigenvalues and write the gain matrix $A = 2 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ is a matrix that rotates a vector by $\frac{\pi}{2}$, and $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is the identity matrix.

Theorem 5.2: A greedy control strategy that projects the state vector in the control direction $U[n] = -[2 \quad -2] [X_1[n] X_2[n]]^T$ is the optimal strategy.

The basic idea is that controls applied at even and odd times cannot substitute for each other and one of each is essential.

These arguments for the 2D-case generalize to systems of dimension- k with an eigenvalue cycle of period- k and the same

results hold: the estimation critical probability and the control critical probability are exactly the same. Further, since the strategy is time-invariant, non-causal knowledge of the pattern of arrivals would have no impact on the rate of the decay of the system state. Here, side-information has no value.

On the other hand, consider a general aperiodic system. The optimal strategy in this case is unclear without any look-ahead, however, we can implement a finite-horizon dynamic programming solution to the problem to numerically evaluate the critical probability for a given system. This is explored in Fig. 4, where the critical probabilities for the problem are plotted against the maximum eigenvalue. The first interesting observation is that the optimal dynamic programming solution with no look-ahead decays exactly as the optimal LTI strategy: the lines for $\frac{1}{(\lambda_1 \lambda_2)^2}$ and the optimal dynamic programming strategy ($q = 0$) are on top of each other [25].

With infinite look-ahead, the critical probability decays as $\frac{1}{\lambda_2^{2m,az}}$, which is the topmost line. To understand the behavior in between, we plot the the strategy with one-step look-ahead at each time ($q = 1$), and one-step look-ahead with probability half ($q = 0.5$). The optimal dynamic programming strategies are compared to a simple align-and-kill strategy (AK) that assumes that the next control will arrive. The strategies seem to converge at high eigenvalues and this warrants further exploration. That the slopes are different suggests that the side-information has a *scaling* effect on the mapping between the eigenvalues and the critical erasure probability.

The change in the slopes of the curves shows that channel predictability gets more valuable with increasing eigenvalues. A system designer might prefer a noisier but more predictable channel, even though it has lower anytime reliability. In contrast to the observation problem, knowledge of future control directions seems to be the useful side-information for the general aperiodic control problem.

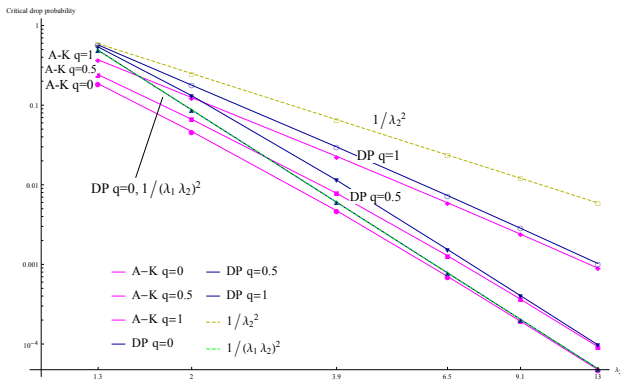


Fig. 4. Critical erasure probability vs. mag. of max eigenvalue $\frac{\lambda_1}{\lambda_2} = 1.18$.

ACKNOWLEDGMENTS

The authors would like to thank the NSF for a GRF and CNS-0932410, CNS-1321155, and ECCS-1343398.

REFERENCES

- [1] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, 2007.
- [2] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *Information Theory, IEEE Transactions on*, vol. 44, no. 6, pp. 2148–2177, 1998.
- [3] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *Information Theory, IEEE Transactions on*, vol. 46, no. 3, pp. 933–946, 2000.
- [4] A. Lapidoth and S. Shamai, "Fading channels: how perfect need perfect side information be?" *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1118–1134, 2002.
- [5] S. S. Pradhan, J. Chou, and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *Information Theory, IEEE Transactions on*, vol. 49, no. 5, pp. 1181–1203, 2003.
- [6] S. Kotagiri and J. Laneman, "Multiaccess channels with state known to some encoders and independent messages," *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, pp. 1–14, 2008.
- [7] Y. Wu and S. Verdú, "MMSE dimension," *Information Theory, IEEE Transactions on*, vol. 57, no. 8, pp. 4857–4879, 2011.
- [8] J. Kelly Jr, "A new interpretation of information rate," *Information Theory, IRE Transactions on*, vol. 2, no. 3, pp. 185–189, 1956.
- [9] T. M. Cover, "Universal portfolios," *Mathematical finance*, vol. 1, no. 1, pp. 1–29, 1991.
- [10] T. M. Cover and E. Ordentlich, "Universal portfolios with side information," *Information Theory, IEEE Transactions on*, vol. 42, no. 2, pp. 348–363, 1996.
- [11] H. H. Permuter, Y.-H. Kim, and T. Weissman, "On directed information and gambling," in *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*. IEEE, 2008, pp. 1403–1407.
- [12] —, "Interpretations of directed information in portfolio theory, data compression, and hypothesis testing," *Information Theory, IEEE Transactions on*, vol. 57, no. 6, pp. 3248–3259, 2011.
- [13] S. Tatikonda, "Control under communication constraints," 2000.
- [14] G. Ranade and A. Sahai, "Non-coherence in estimation and control," in *51th Annual Allerton Conference on Communication, Control, and Computing*, 2013.
- [15] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *Information Theory, IEEE Transactions on*, vol. 57, no. 4, pp. 1872–1905, 2011.
- [16] U. Niesen and M. Maddah-Ali, "Interference alignment: From degrees-of-freedom to constant-gap capacity approximations," in *Information Theory Proceedings (ISIT), IEEE International Symposium on*, 2012, pp. 2077–2081.
- [17] S. Park, G. Ranade, and A. Sahai, "Carry-free models and beyond," in *IEEE International Symposium on Information Theory*, 2012, pp. 1927–1931.
- [18] P. Grover and A. Sahai, "Implicit and explicit communication in decentralized control," in *48th Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois*, 2010.
- [19] S. Y. Park and A. Sahai, "It may be easier to approximate decentralized infinite-horizon LQG problems," in *Decision and Control (CDC), IEEE 51st Annual Conference on*, 2012, pp. 2250–2255.
- [20] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1453–1464, 2004.
- [21] Y. Mo and B. Sinopoli, "A characterization of the critical value for Kalman filtering with intermittent observations," in *Decision and Control, CDC. 47th IEEE Conference on*, 2008, pp. 2692–2697.
- [22] S. Park and A. Sahai, "Intermittent Kalman filtering: Eigenvalue cycles and nonuniform sampling," in *American Control Conference (ACC), 2011*, 2011, pp. 3692–3697.
- [23] N. Elia, "Remote stabilization over fading channels," *Systems & Control Letters*, vol. 54, no. 3, pp. 237–249, 2005.
- [24] S. Y. Park, "Information flow in linear systems," Ph.D. dissertation, University of California, Berkeley, December 2013.
- [25] O. C. Imer, S. Yüksel, and T. Başar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429–1439, 2006.