

# **Stabilization using both noisy and noiseless feedback**

Anant Sahai  
sahai@eecs.berkeley.edu

## **Abstract**

When designing a distributed control system, the system designer has a choice in how to connect the different units through communication channels. In practice, noiseless and noisy channels may coexist. Using the standard toy example of scalar stabilization, this paper shows how a small amount of noiseless feedback can perform a “supervisory” role and thereby boost the effectiveness of noisy feedback.

## **Index Terms**

Stabilization, noisy channels, noiseless channels, anytime capacity, feedback.

## Stabilization using both noisy and noiseless feedback

### I. INTRODUCTION

Distributed control has long been understood as a challenging problem, and in such situations the control signals often have a dual purpose — to achieve the control objectives as well as to communicate necessary information between the distributed systems [1], [2], [3]. Consequently, it had been recognized that there may lie deep connections between information theory and control theory [4], but establishing such connections remained out of reach. Recently however, some deep connections have begun to emerge, most notably in the context of stabilization over a limited communication medium. The reader is directed to the recent September 2004 issue of IEEE Transactions on Automatic Control and the articles (and references) therein for a more comprehensive survey, but it is worthwhile to make note of some relevant prior work here.

The first truly information-theoretic connection was given by Tatikonda with his work on sequential rate-distortion theory. This married the causality constraint critical in control problems to the rate-distortion framework provided by Shannon. By using Massey's directed mutual information and relaxing the optimization to be over all channels with the same directed mutual information instead of the original channel, it provides a lower-bound on the achievable closed-loop performance of a control system over a channel, noisy or otherwise. Because this bound is sometimes infinite, it also implies that there is a fundamental rate of information production, namely the sum of the logs of the unstable eigenvalues of the plant, that is attached to an unstable linear discrete-time process [5], [6]. The corresponding achievability result was based on noiseless, but rate-limited, channels. Nair et al. extended the noiseless rate-limited story to the case of unbounded disturbances and observation noise under suitable conditions [7], [8].

The noisy channel side of the story is distinct and involves additional subtleties. Much work was tied to specific communication channels. We had previously showed that it is possible to stabilize persistently disturbed controlled Gauss-Markov processes over suitable power-constrained AWGN (Additive White Gaussian Noise) channels [9], [10] where it turns out that Shannon capacity is tight and linear observers and controllers are sufficient to achieve stabilization [11]. In contrast, we showed that the Shannon capacity of the binary erasure channel (BEC) is not sufficient to check stabilizability and introduced the anytime capacity as a candidate figure of merit in [12]. Following up on our treatment of the BEC case, Martins et al. have studied more general erasure-type models and have also incorporated bounded model uncertainty in the plant [13]. Elia used ideas from robust control to deal with communication uncertainty in a mixed continuous/discrete context, but restricting to linear operations [14], [15].

Our work in [16], [17] established another deep connection in the noisy channel case by showing how stabilization problems over noisy channels are equivalent to anytime communication problems where the encoder has access to feedback. These results are extended in [18] to the vector case. In stabilization problems, the sense of stability desired (mean-squared, fourth-moment, etc.) determines the target anytime reliability in conjunction with the unstable eigenvalues. In contrast, the rate requirement comes only from the unstable eigenvalues just as it does in the noiseless rate-limited channel case.

The feedback anytime capacity is lower bounded by the anytime capacity without feedback in [19], [17]. The basic result is that the block random-coding error exponent  $E_r(R)$  is a lower bound to the anytime reliability at rate  $R$ . Furthermore, implementable observers and controllers can be designed that approach this lower bound by using a nearly memoryless quantization strategy for the observer and a stack-algorithm based strategy for the controller [20], [21]. However, using purely information-theoretic arguments [22] shows that the anytime reliability with feedback can be much higher and provides a new upper bound called the "focusing bound." For symmetric channels, the focusing bound says that at rate  $R = \frac{E_0(\rho)}{\rho}$ , the anytime reliability can not exceed  $E_0(\rho)$  where  $E_0(\rho)$  is the Gallager function defined as:

$$E_0(\rho) = \max_{\vec{q}} - \ln \sum_y \left[ \sum_x q_x P_{x,y}^{\frac{1}{1+\rho}} \right]^{(1+\rho)} \quad (1)$$

Note that for symmetric channels, it suffices to use a uniform  $\vec{q}$  while optimizing (1) [23].

[22] also shows that the focusing bound is asymptotically achievable for any erasure type channel or noisy channel that has a strictly positive zero-error capacity. (For non-erasure channels whose zero-error capacity is zero,

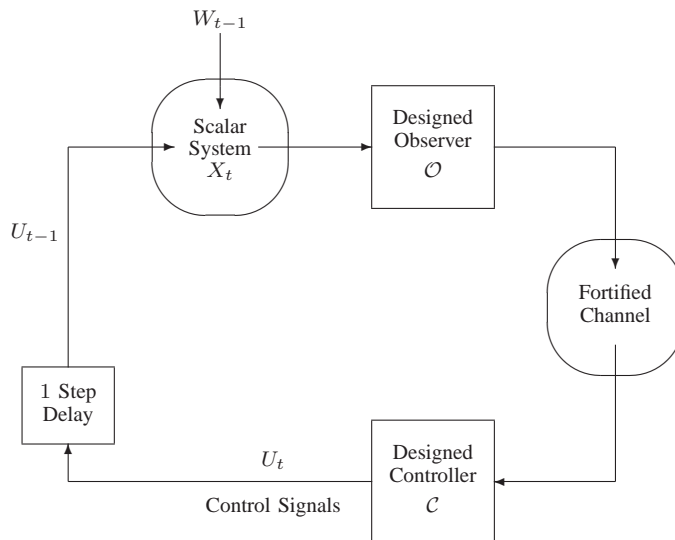


Fig. 1. Control over a fortified communication channel.

the achievability of the focusing bound is still open, although it is known that we can beat the lower-bound of  $E_r(R)$  in essentially all cases [24], [25], [26].) The achievability strategy of [22] is interesting because it has strictly bounded computation at both the encoder and decoder and furthermore, it relies on sending very low-rate “supervisory” messages using the “noiseless” aspects of the communication channel.

While it is possible to directly apply the results of [22] to the stabilization problem by means of the equivalence established in [16] and a separation architecture, it is more illuminating to give the strategy directly in the stabilization context without reference to anytime communication per-se. After setting up some notation and preliminaries in Section II, we give the stabilization scheme in Section III, before concluding in Section IV. Proof ideas are given here, but the “heavy lifting” is found in [16], [22].

## II. PRELIMINARIES

The scalar stabilization problem is illustrated in Figure 1.

$$X_{t+1} = \lambda X_t + U_t + W_t, \quad t \geq 0 \quad (2)$$

where  $\{X_t\}$  is a  $\mathbb{R}$ -valued state process.  $\{U_t\}$  is a  $\mathbb{R}$ -valued control process and  $\{W_t\}$  is a bounded noise/disturbance process s.t.  $|W_t| \leq \frac{\Omega}{2}$ . This bound is assumed to hold with certainty. For convenience, we also assume a known initial condition  $X_0 = 0$ . To make things interesting, consider  $\lambda > 1$  so the open-loop system is exponentially unstable. The observer/encoder system  $\mathcal{O}$  observes  $X_t$  and generates inputs  $a_t$  to the channel. The decoder/controller system  $\mathcal{C}$  observes channel outputs  $B_t$  and generates control signals  $U_t$ . Both  $\mathcal{O}, \mathcal{C}$  are allowed to have unbounded memory and to be nonlinear in general. The goal is stability:

*Definition 2.1:* A closed-loop dynamic system with state  $X_t$  is  $\eta$ -stable if there exists a constant  $K$  s.t.  $E[|X_t|^\eta] \leq K$  for all  $t \geq 0$ .

*Definition 2.2:* A *discrete time discrete memoryless channel (DMC)* is a probabilistic system with an input. At every time step  $t$ , it takes an input  $x_t \in \mathcal{X}$  and produces an output  $y_t \in \mathcal{Y}$  with probability  $p(y_t|x_t)$ . Both  $\mathcal{X}, \mathcal{Y}$  are finite sets. The current channel output is independent of all past random variables in the system conditioned on the current channel input.

The twist in this paper is that the noisy channel does not have to be used alone. In addition, we have a low-rate noiseless channel to help out.

*Definition 2.3:* Given a DMC  $P$  for the forward link, a  $\frac{1}{k}$ -fortified channel built around it is one in which every  $k$ -th use of  $P$  is supplemented with the ability to transmit a single noise-free bit to the receiver.

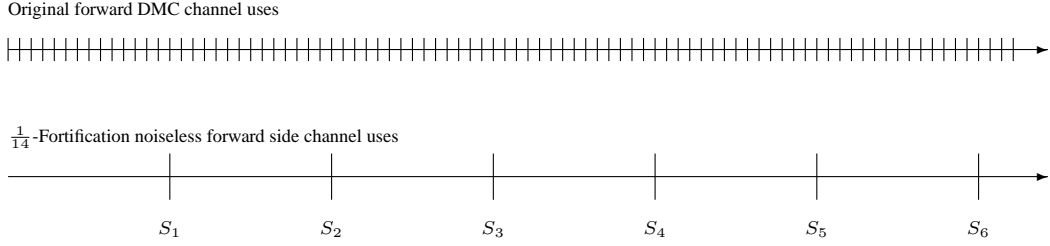


Fig. 2. Fortification illustrated: the forward noisy channel uses are supplemented with regular low-rate use of a noiseless side channel.

### III. MAIN RESULT

The main result of this paper is that the focusing bound induced fundamental limits are asymptotically achievable with even a tiny amount of fortification.

*Theorem 3.1:* For a symmetric noisy memoryless channel and any amount of fortification  $0 < k < \infty$ , it is possible to control the unstable scalar plant of (2) over that fortified channel so that the  $\eta$ -moment of  $|X_t|$  stays finite for all time if there exists any  $\epsilon > 0$  so that  $\frac{E_0(\eta+\epsilon)}{\eta+\epsilon} > \ln \lambda$  even if the observer is only allowed to observe the state  $X_t$  corrupted by bounded noise.

*Proof:* The statement of the theorem does not refer to the focusing bound, but the connection can be seen by defining  $\rho' = \eta + \epsilon$ . Then  $\frac{E_0(\rho')}{\rho'} > \ln \lambda$  is just a statement about the implicit rate  $R' = \frac{E_0(\rho')}{\rho'}$  being high enough. By multiplying both sides by  $\rho'$ , it is saying that the reliability  $E_0(\rho') > \rho' \ln \lambda > \eta \ln \lambda$  which is what is necessary for  $\eta$ -stabilization. Thus, this theorem is essentially saying that the focusing bound can be asymptotically achieved with fortification.

The control strategy itself is based on three ideas:

- Feedback of channel outputs to the observer by making the plant “dance” in an observable way following [16]. This makes the proof simpler to understand, but is later eliminated by better leveraging the next idea:
- Nonuniform sampling of the plant state at the observer for purposes of “codeword” generation for the noisy channel uses. The codewords themselves are generated using random coding. The plant state is resampled after the controller applies a “true” (as opposed to “dancing”) control.
- Supervisory use of the noiseless fortification channel by the observer to tell the controller how and when to apply a “true” control.

They are covered in greater detail in the next few subsections.

#### A. Communication through “dancing”

The approach is illustrated in Figure 3 and detailed in [16], [27]. The goal is to communicate the noisy channel outputs back to the observer so it can guide the controller. The main control signal  $U_t$  is zero except when directed by the observer as depicted in Section III-C. The actual applied control is

$$U'_t = U_t + F(b_t) - \lambda F(b_{t-1}) \quad (3)$$

where  $F(b_t) = 3\Gamma_u b_t$  and  $\Gamma_u = \Omega + (\lambda + 1)\Gamma$  where  $\Gamma$  is the bound on the noise in state observation. If  $X_o(t)$  is the boundedly noisy observation of the state  $X_t$ , the observer can compute  $X_o(t) - \lambda X_o(t-1)$  which is going to be  $U'_{t-1} + W_{t-1}$  corrupted by a bounded noise of at most  $\Gamma$  from  $X_o(t)$  and  $\lambda\Gamma$  from  $\lambda X_o(t-1)$ . The total noise in observing  $U'_{t-1}$  is thus  $\Gamma_u$  and so with this choice of  $F$ , the channel output  $b_t$  can be unambiguously determined by the observer.

If we assume that there is a finite channel output alphabet, then because of the cancellation of the past  $F(b_{t-1})$ , this dance has no lasting impact on the state and thus can be ignored for stability considerations.

#### B. Nonuniform sampling and control choice

As in [12], the controller must keep track of a total window of uncertainty in which the state lies. The observer also tracks this window so as to communicate to the controller. Call the current uncertainty window  $\Delta_t$ . At a

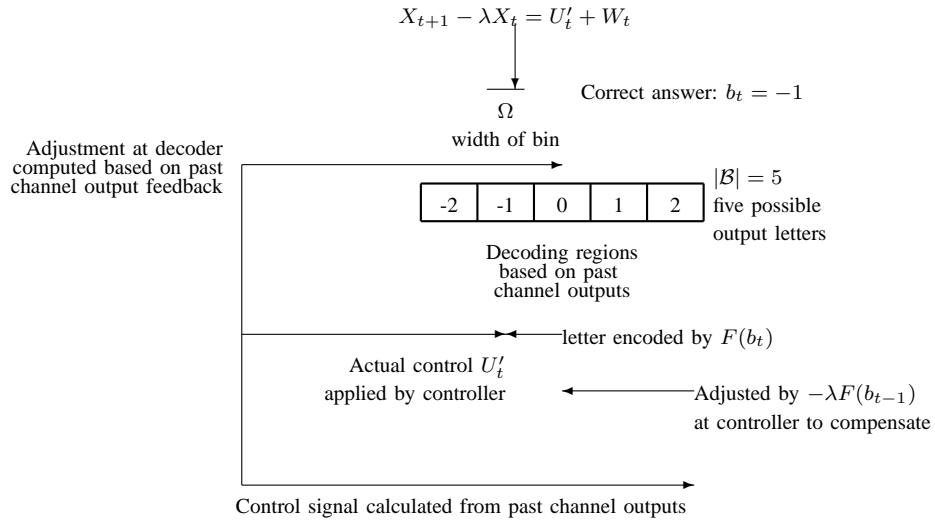


Fig. 3. How to communicate the channel outputs through the plant with state observations only. The controller restricts its main control signal to be either zero, or calculated with a delay of 1 time unit. It is adjusted by  $-\lambda F(b_{t-1})$  to eliminate the effect of the past communication. The final control signal applied is shifted slightly to encode which  $b_t$  was received. The observer uses the past  $b_0^{t-1}$  to align its decoding regions and then reads off  $b_t$  by using  $X_{t+1} - \lambda X_t$ .

sampling time, the observer divides this window into  $\exp(nR)$  (large) different regions and sees where the state observation is. Each region is assigned an infinitely long iid codeword drawn uniformly from the channel input alphabet according to the optimizing  $\vec{q}$ . Letters from this codeword are transmitted on the noisy channel until the state is resampled.

If the controller faced an uncertain window of  $\Delta_t$  at time  $t$ , then without any information, the window will grow according to:

$$\Delta_{t+1} = \lambda \Delta_t + \Omega \quad (4)$$

If the controller learns correctly which region the observation lies in, then the retrospective uncertainty shrinks to at most  $\Gamma + \Delta_t \exp(-nR)$ . If this occurs  $T$  time steps in the future, the resulting uncertainty is at most

$$\Delta_{t+T} = \lambda^T (\exp(-nR) \Delta_t + \Gamma + \frac{\Omega \lambda}{\lambda - 1}) \quad (5)$$

and knowledge of the window allows the controller to apply a “true” control that will center this uncertainty window around 0.

As in [12], it is how the randomness in  $\lambda^T$  compares to  $\exp(nR)$  that will determine whether or not the system is  $\eta$ -stable in closed loop.

### C. Supervision through the noiseless channel outputs

Due to the “dancing” of Section III-A, the observer knows what channel outputs have been received so far. Thus, it can monitor whether the channel outputs are good enough for the controller to successfully decode which region the state was in at the last state sampling time. When that occurs, it sends a 1 over the noiseless channel and otherwise it sends a zero. Thus, the random variable  $T$  is the time till successful decoding of an infinitely long codeword chosen out of  $\exp(nR)$  possible codewords. So far,  $nR$  was just an arbitrary constant. To interpret it, think of  $\ln \lambda < R < \frac{E_0(\eta+\epsilon)}{\eta+\epsilon}$ . Choose  $\rho$  so that  $R = \frac{E_0(\rho)}{\rho}$ . Think of  $n$  as a constant that is a large multiple of  $k$ .

At this point, the detailed analysis in [22] explains what happens. The key points are summarized here.

- The  $T$  represent the inter-sampling times of the state. This process is iid across time since the codewords are randomly chosen and the noisy channel is memoryless.
- If  $\rho \leq 1$  works, then the scheme described above works without modification and the random variable  $T$  is strictly dominated by  $n + \tilde{T}$  where the  $\tilde{T}$  is a geometric random variable with probability of failure given by  $\exp(-E_0(\rho))$ .

- c. If  $\rho > 1$  is required, then the scheme above must be modified to use list decoding among the top  $l = \lceil \rho \rceil$  entries with the next  $\lceil \log_2 l \rceil$  noiseless channel uses used to disambiguate the list. This is a small overhead that is negligible when  $nR$  is large. With that modification, everything from (b) still holds in that  $T \leq n + \lceil \log_2 l \rceil k + \tilde{T}$ .

At this point, we give the key ideas in the proof:

Using (b), we can rewrite (5) to get

$$\begin{aligned} \Delta_{t+T} &= \lambda^T (\exp(-nR)\Delta_t + \Gamma + \frac{\Omega\lambda}{\lambda-1}) \\ &\leq \lambda^{n+\tilde{T}} (\exp(-nR)\Delta_t + \Gamma + \frac{\Omega\lambda}{\lambda-1}) \\ &= \lambda^{\tilde{T}} (\lambda^n \exp(-nR)\Delta_t + \Gamma\lambda^n + \frac{\Omega\lambda^{1+n}}{\lambda-1}) \\ &< \lambda^{\tilde{T}} (\exp[-n(R - \ln \lambda)] \Delta_t + \Gamma\lambda^n + \frac{\Omega\lambda^{1+n}}{\lambda-1}) \end{aligned}$$

Call this upper bound  $\bar{\Delta}_{t+T}$  so:

$$\bar{\Delta}_{t+T} = \lambda^{T-n} (\exp[-n(R - \ln \lambda)] \bar{\Delta}_t + \Gamma\lambda^n + \frac{\Omega\lambda^{1+n}}{\lambda-1}) \quad (6)$$

where we assume that the next successful control application after time  $t$  occurs at time  $t + T$ .

It is worthwhile interpreting the random variable  $\tilde{T}$  as the inter-arrival time for a packet-erasure channel and  $\exp[n(R - \ln \lambda)]$  as the number of possible messages carried by a single packet. Since  $R > \ln \lambda$ , the size of the packet can be made as large as we want by picking a large<sup>1</sup>  $n$ . Viewed this way, the uncertainty window  $\Delta$  is very likely to shrink from one arrival to the next if it starts out large.

Asymptotically, the problem behaves like the very large packet erasure-channel case with a packet-erasure probability of  $\exp(-E_0(\rho))$ . For  $n$  large enough, the closed-loop system will have an  $\eta$ -moment as long as  $\lambda^\eta \exp(-E_0(\rho)) < 1$  or by taking logs:  $E_0(\rho) > \eta \ln \lambda$  which is true by assumption.

#### D. Event-based sampling is simpler than “dancing”

While the above three-part strategy makes for an easy proof, it invites criticism for being very unrealistic:

- The “dancing” of the plant seems to be excessive. When the channel output alphabet is large, this gives a tremendous performance hit as the channel outputs are communicated back through the plant. This also does not work if there is a continuous or other infinite channel output alphabet.
- The “observer” here essentially has to run a copy of the channel decoder in the controller as well as interpret the plant dancing. There is thus duplication of effort in the whole system.

To overcome this, we just have to notice that essentially all the controller has to be told is when to stop decoding this block and when it is time to move on to the next block. We can view successful decoding as an “event” that defines when the observer will next “sample” the state process for the purpose of generating channel inputs. The above strategy uses the noiseless feedback of channel outputs through the plant to decide when successful decoding as occurred — however there is a simpler way to see this. *Whenever the controller is successful in decoding the channel outputs, it will be able to apply a control that drives the state to the neighborhood of zero!*

This gives rise to modifications to the control strategy that do not rely on the finite channel output alphabet assumption:

- The controller no longer encodes the channel outputs into the control signals by making the plant dance.
- Instead of waiting for a “go ahead” signal from the observer through the fortification channel, the controller applies the control based on its tentative best decision so far. If it gets negative feedback through the fortification channel, it backs-out the tentative control<sup>2</sup> it had applied and crosses the corresponding message off the list of possibilities. It then moves on to the new top-scoring message and applies the corresponding control until it receives positive feedback.

<sup>1</sup>Of course, picking a large  $n$  also makes the effect of the disturbance and observation noise greater, but it remains bounded no matter how large  $n$  is and that is all that matters when it comes to the existence of  $\eta$ -moments.

<sup>2</sup>By adding  $-\lambda U_{t-1}$  to its new control when it wants to countermand its previous control.

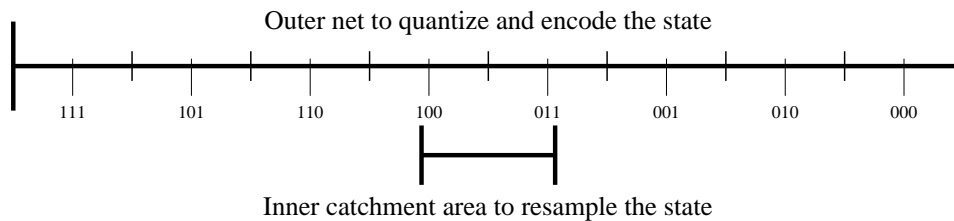


Fig. 4. The observer maintains two “nets” that it uses to catch the state. The outer “quantization net” represents everywhere the controller believes the state could possibly be based on the information it is known to have received so far. At sampling time, it is partitioned into  $e^{nR}$  uniformly sized boxes and the box corresponding to the true state is used to pick a message to send over the noisy channel. The inner net (of size  $\bar{\Delta}_{t+T}$  at time  $T$  when the last sampling event occurred at time  $t$ ) represents where the controller could drive the state given that it was able to successfully decode the current message block. Both nets grow steadily exponentially at rate  $\approx \lambda$  per unit time, and shrink suddenly whenever the controller is able to successfully decode and have this success confirmed through a fortified bit.

- Instead of using the fortification channel to tell the controller when to stop decoding and apply a control, the observer uses the fortification channel to simply tell the decoder when it has gotten the right answer. This can be observed at the observer by seeing the plant state move into the vicinity<sup>3</sup> of the origin.
- Otherwise, the observer remains essentially as before when it comes to sampling the controlled state and determining what infinitely-long random codeword to start sending incrementally.

In terms of analysis, there are only two remaining concerns:

*a) False control performance hit:* Before a true control action is confirmed, the controller might try some different false controls. However, a false control can at worst move the state to twice the outer limits of the current uncertainty window. This happens when the true control should be dealing with one extreme endpoint and the false applied control corresponds to the other endpoint. However, this factor of 2 just changes the bound on the  $\eta$ -th moment by a factor of  $2^\eta$ . It will not make a moment infinite if it were finite before.

*b) List decoding:* In the original scheme above that was made to leverage the purely channel coding derivation in [22], list-decoding was required to disambiguate some remaining uncertainty. This caused this small constant  $(1 + \log_2 l)k$  term to show up along with the original  $n$ . But since  $l$  is a constant that does not depend on  $n$ , the  $\log_2 l$  is not necessary for the result to hold. Everything will work just as well asymptotically even if we used a “unary encoding” and used  $(l + 1)k$  time-slots for the list-disambiguation. This is essentially what the new observer does since it just keeps ruling out the seemingly most likely codewords until the true one is applied.

Consequently, the event-based sampling approach that uses the fortification to give simple positive and negative supervision to the controller will achieve the focusing-bound performance.  $\square$

#### IV. CONCLUSION

The main result of this paper is that a little bit of noiseless feedback used for supervision can be used to leverage the effectiveness of the noisy feedback. The difference in what can be achieved is illustrated in Figure 5. With even a tiny bit of noiseless feedback, we can achieve performance corresponding to the upper curves.

The theorem as stated is asymptotic in nature and it might appear troubling to have to use large  $n$  and thereby significantly sacrifice performance just in order to stabilize larger moments. However, by applying the “supervisory” idea again, no such sacrifice is needed. The control strategy of [20] can be applied and most of the time, it will hold the plant very close to the origin. Rarely, the state will wander out of a particular target ball. At that point, the observer can use the noiseless channel to declare an “emergency” and switch to the control strategy given in this paper. Once the state has returned to a moderate sized neighborhood of the origin, the state of emergency can be lifted and the control strategy can return to that of [20]. By making the target ball large, emergencies will only occur rarely and thus will not impact average performance. However, the emergency mode will ensure that as many higher moments of the state will exist as is possible for the particular channel in question.

<sup>3</sup>By vicinity, we simply mean that the observed state lands within a box of size  $\bar{\Delta}_{t+T}$  of the origin when considering time  $t + T$ .

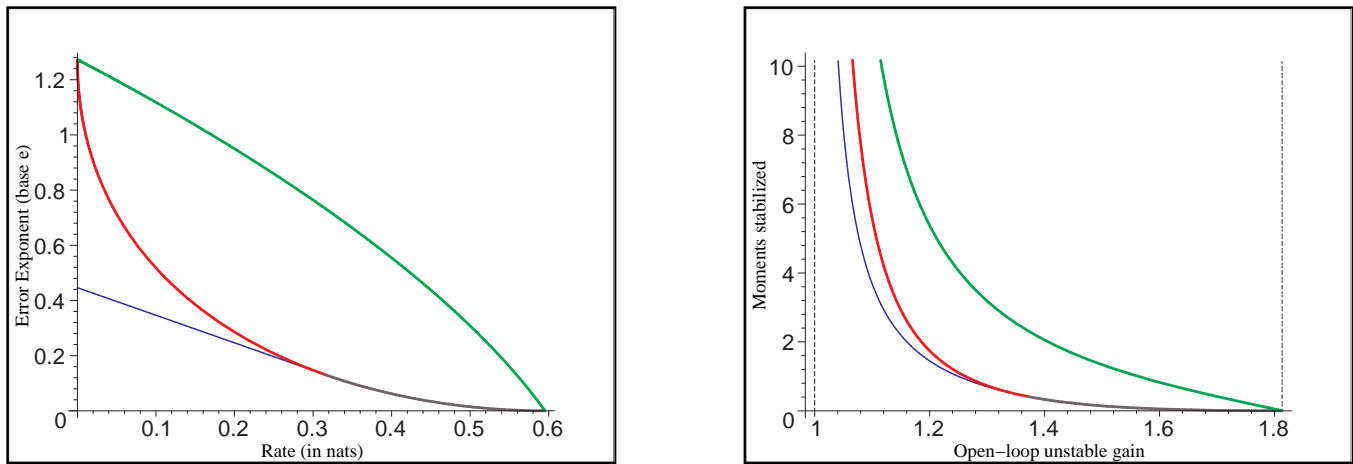


Fig. 5. The sphere-packing and focusing bounds for the binary symmetric channel with crossover probability 0.02. On the left, reliability is plotted relative to rate, while on the right, the the stabilized moment is plotted relative to the unstable gain  $\lambda$ . The focusing bound is the highest and bounds what is possible with feedback, the sphere-packing bound is below and bounds the reliability for communication without feedback. The random coding error exponent matches that bound at high-rates, but falls below at low rates. Since stabilization is equivalent to communication with feedback, the upper bound is the target.

## REFERENCES

- [1] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, Jan. 1968.
- [2] —, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.
- [3] S. K. Mitter and A. Sahai, "Information and control: Witsenhausen revisited," in *Learning, Control and Hybrid Systems: Lecture Notes in Control and Information Sciences 241*, Y. Yamamoto and S. Hara, Eds. New York, NY: Springer, 1999, pp. 281–293.
- [4] Y. C. Ho, M. P. Kastner, and E. Wong, "Teams, signaling, and information theory," *IEEE Trans. Automat. Contr.*, vol. 23, no. 2, pp. 305–312, Apr. 1978.
- [5] S. Tatikonda, "Control under communication constraints," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2000.
- [6] S. Tatikonda and S. K. Mitter, "Control under communication constraints," *IEEE Trans. Automat. Contr.*, vol. 49, no. 7, pp. 1056–1068, July 2004.
- [7] G. N. Nair and R. J. Evans, "Communication-limited stabilization of linear systems," in *Proceedings of the 39th IEEE Conference on Decision and control*, Sydney, Australia, Dec. 2000, pp. 1005–1010.
- [8] —, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM Journal on Control and Optimization*, vol. 43, no. 2, pp. 413–436, July 2004.
- [9] A. Sahai, S. Tatikonda, and S. K. Mitter, "Control of LQG systems under communication constraints," in *Proceedings of the 1999 American Control Conference*, San Diego, CA, June 1999, pp. 2778–2782.
- [10] S. Tatikonda, A. Sahai, and S. K. Mitter, "Stochastic linear control over a communication channel," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1549–1561, Sept. 2004.
- [11] R. Bansal and T. Basar, "Simultaneous design of measurement and control strategies for stochastic systems with feedback," *Automatica*, vol. 45, no. 5, pp. 679–694, Sept. 1999.
- [12] A. Sahai, "Evaluating channels for control: Capacity reconsidered," in *Proceedings of the 2000 American Control Conference*, Chicago, CA, June 2000, pp. 2358–2362.
- [13] N. C. Martins, N. Elia, and M. A. Dahleh, "Stabilization of uncertain systems in the presence of a stochastic digital link," *IEEE Trans. Automat. Contr.*, submitted for publication.
- [14] N. Elia, "When Bode meets Shannon: control-oriented feedback communication schemes," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1477–1488, Sept. 2004.
- [15] —, "Remote stabilization over fading channels," *Systems and Control Letters*, vol. 54, no. 3, pp. 239–249, Mar. 2005.
- [16] A. Sahai and S. K. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link. part I: scalar systems," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.
- [17] A. Sahai, "Any-time information theory," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2001.
- [18] A. Sahai and S. K. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link. part II: vector systems," *IEEE Trans. Inform. Theory*, submitted for publication. [Online]. Available: <http://www.eecs.berkeley.edu/~sahai/Papers/control-part-II.pdf>
- [19] —, "Source coding and channel requirements for unstable processes," *IEEE Trans. Inform. Theory*, Submitted, 2006. [Online]. Available: <http://www.eecs.berkeley.edu/~sahai/Papers/anytime.pdf>

- [20] A. Sahai and H. Palaiyanur, "A simple encoding and decoding strategy for stabilization over discrete memoryless channels," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2005. [Online]. Available: <http://www.eecs.berkeley.edu/~sahai/Papers/allerton05-sequential-stabilization.pdf>
- [21] H. Palaiyanur, "Sequential decoding for anytime communication," Master's thesis, University of California, Berkeley, 2006.
- [22] A. Sahai, "Why block length and delay are not the same thing," *IEEE Trans. Inform. Theory*, submitted. [Online]. Available: <http://www.eecs.berkeley.edu/~sahai/Papers/FocusingBound.pdf>
- [23] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY: John Wiley, 1971.
- [24] A. Sahai, "Why delay and block length are not the same thing for channel coding with feedback," in *Proceedings of the Information Theory and Applications Workshop*, Feb. 2006.
- [25] H. T. Simsek, "Anytime channel coding with feedback," Ph.D. dissertation, University of California, Berkeley, 2004.
- [26] T. Simsek, R. Jain, and P. Varaiya, "Scalar estimation and control with noisy binary observations," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1598–1603, Sept. 2004.
- [27] A. Sahai, "The necessity and sufficiency of anytime capacity for control over a noisy communication link," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, Paradise Island, Bahamas, Dec. 2004, pp. 1896 – 1901.