From any-time to zero-error reliability for stabilization over Markov channels

Massimo Franceschetti, Paolo Minero

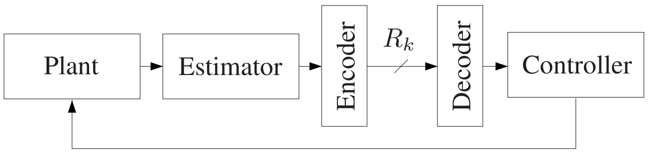

Fig. 1. Feedback loop model. The estimated state is quantized, encoded and sent to a decoder over a digital channel of state R_k that evolves in time according to a Markov process.

***Abstract*—The stochastic stability of a scalar linear system controlled over a Markov time-varying digital feedback channel with noiseless output feedback is considered. A stability threshold function of the channel's parameters and of the moment stability number m is studied. The system can be stabilized if and only if this threshold function exceeds the intrinsic entropy rate of the system, representing the growth of the state space spanned by the open loop system. The function is continuous, strictly decreasing for $m > 0$, but not generally convex. It converges to the Shannon capacity for $m \to 0$, to the zero-error capacity for $m \to \infty$, and it provides a parametric characterization of the anytime capacity of Sahai for the remaining values of m. Its operational interpretation is that of achievable communication rate, subject to a varying reliability constraint that depends on the desired stability level m.**

Applications yield a novel anytime capacity formula for the special case of the r-bit Markov erasure channel, that generalizes the one for the memoryless case obtained by Sahai in parametric, and by Xu and Sahai in explicit form. For a two-state memoryless communication channel, a converse to the achievable scheme of Yüksel and Meyn is provided, and when the rate process can take values in $\{0, \infty\}$ results for the packet erasure model are recovered. Finally, the anytime error exponent is compared to the error exponent of variable-length block-codes over Markov channels with feedback of Como, Yüksel, and Tatikonda.

The proofs rely on a novel necessary and sufficient condition for the stochastic stability of Markov jump linear systems. The sufficient condition is obtained via the idea of subsampling, while the necessary condition is based on the maximum entropy theorem and the entropy power inequality.

I. Introduction

We consider the problem of moment stabilization of a dynamical system where the estimated state is transmitted for control over a time-varying communication channel, as depicted in Fig. 4. This problem has been studied extensively in the context of networked control systems and discussed in several special issue journals dedicated to the topic [1]–[3]. Recently, it gained renewed attention due to its relevance for the design of cyberphysical systems [4]. A tutorial review of the problem with extensive references appears in [5].

The notion of Shannon capacity is in general not sufficient to characterize the trade-off between the entropy rate production of the plant, expressed by the growth of the state space spanned in open loop, and the communication rate required for its stabilization. A large Shannon capacity is useless for stabilization if it cannot be used in time for control. For the control signal to be effective, it must be appropriate to the current state of the system. Since decoding the wrong codeword implies applying a wrong signal and driving the system away from stability, applying an effective control signal depends on the history of whether previous codewords were decoded correctly or not. In essence, the stabilization problem is an example of *interactive communication*, where two-way communication occurs through the feedback loop between the plant and the controller. Error correcting codes developed independently in this context [6]–[8] have a natural tree structure representing past history and are natural candidates to be used for control. Alternative capacity notions with stronger reliability constraints than simply having a vanishing probability of error, and requiring these type of coding schemes have been proposed in the context of control, including the zero-error capacity [9], originally introduced by Shannon [10], and the anytime capacity proposed by Sahai [11], [12]–[16].

Within this general framework, we focus on the mth moment stabilization of an unstable scalar system whose state is communicated over a rate-limited channel capable of supporting R_k bits at each time step and evolving randomly in a Markovian fashion. The rate process is known casually to both encoder and decoder. Many variations of this "bit-pipe" model have been studied in the literature [17]–[39], including the case of fixed rate channel; the erasure channel where the rate process can assume value zero; and the packet loss channel, where the rate process can oscillate randomly between zero and infinity, allowing a real number with infinite precision to be transported across the channel in one time step. Connections between the rate limited and the packet loss channel have been pointed out in [32], [33], showing that results for the latter model can be recovered by appropriate limiting arguments. The additive white Gaussian channel has been considered in [40]–[44] and in this case the Shannon capacity is indeed sufficient to express the rate needed for stabilization. Extensions to the additive colored Gaussian channel [45] show that the maximum "tolerable instability" — expressed by the sum of the logarithms of the unstable eigenvalues of the system that can be stabilized by a linear controller with a given power constraint over a stationary Gaussian channel— corresponds to

M. Franceschetti is with the Dept. of Electrical and Computer Engineering University of California at San Diego, La Jolla, California 92093–0407. Email: massimo@ece.ucsd.edu

P. Minero is with the Dept. of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 . Email: pminero@nd.edu).

the Shannon feedback capacity [46], that assumes the presence of a noiseless feedback link between the output and the input of the channel and that is subject to the same power constraint. This result suggests a duality between the problems of control and communication in the presence of feedback, and indeed it has been shown that efficient feedback communication schemes can be obtained by solving a corresponding control problem [47], [48].

The major contribution of this paper is the introduction of a stability threshold function of the channel's parameters and of the moment stability number m that converges to the Shannon capacity for $m \to 0$, to the zero-error capacity for $m \to \infty$, and it provides a parametric characterization of the anytime capacity for the remaining values of m. This function yields a novel anytime capacity formula in the special case of the r-bit Markov erasure channel. To prove our results, we require some novel extensions of the theory Markov Jump Linear Systems (MJLS), that are of independent value. On the technical side, the sufficient condition for stability is obtained exploiting the idea of subsampling, while the necessary condition is based on the maximum entropy theorem and the entropy power inequality. In passim, although we do not deal with the case of vector systems directly, we point out that our results can be extended to this case exploiting usual bit-allocation techniques outlined in [32], [33], at the expense of a more technical treatment that does not add much to the engineering insight and that we wish to avoid here.

The rest of the paper is organized as follows. Some preliminary results on Markov Jump Linear Systems, necessary for our derivations are presented in Section II. Section III describes the system and channel model and introduces the stability threshold function, illustrating some of its properties. Section IV describes relationships with the anytime capacity, and provides some representative examples. Section V provides the formula for the anytime capacity of the Markov erasure channel.

II. Markov Jump Linear Systems.

Consider the scalar non-homogeneous MJLS [49] with dynamics

$$z_{k+1} = \frac{\lambda}{2^{R_k}} z_k + w_k, \tag{1}$$

where $z_k \in \mathbb{R}$ with z_0 has bounded entropy, $c \geq 0$ is a constant, and $\{R_k\}_{k\geq 0}$ is a Markov rate process defined on

$$\mathcal{R} = \{\bar{r}_1, \ldots, \bar{r}_n\}, \tag{2}$$

for some integer numbers $0 \leq \bar{r}_1 < \cdots < \bar{r}_n$, and with one-step transition probability matrix P having entries

$$p_{ij} = \mathsf{P}\{R_k = \bar{r}_j | R_{k-1} = \bar{r}_i\} \tag{3}$$

for every $i, j \in \{1, \ldots, n\}$.

The system (1) is said to be weakly mth moment stable if

$$\sup_k \mathsf{E}(|Z|^m) < \infty. \tag{4}$$

Let $\mathrm{R} \in \mathbb{Z}_+^{n\times n}$ be a diagonal matrix with diagonal entries $\bar{r}_1, \ldots, \bar{r}_n$, i.e.,

$$\mathrm{R} = \mathrm{diag}(\bar{r}_1, \ldots, \bar{r}_n). \tag{5}$$

The following lemma states the necessary and sufficient condition for m-th moment stability of the system (1) in terms of the unstable mode $|\lambda|$ and the spectral radius $\rho(\cdot)$ of $\mathrm{P}^\mathsf{T} 2^{-m\mathrm{R}}$, where $2^{-m\mathrm{R}}$ denotes the base-2 matrix exponential of $m\mathrm{R}$, i.e.,

$$2^{-m\mathrm{R}} = \mathrm{diag}(2^{-m\bar{r}_1}, \cdots, 2^{-m\bar{r}_n}). \tag{6}$$

Theorem 1. *For any $m, c \in \mathbb{R}^+$, if*

$$|\lambda|^m < \frac{1}{\rho(\mathrm{P}^\mathsf{T} 2^{-m\mathrm{R}})}, \tag{7}$$

then the MJLS (1) is weakly mth moment stable. In addition, if $c = 0$, then the same statement holds with the inequality in (7) being not strict. Conversely, if the MJLS (1) is weakly mth moment stable, then

$$|\lambda|^m \leq \frac{1}{\rho(\mathrm{P}^\mathsf{T} 2^{-m\mathrm{R}})}. \tag{8}$$

In addition, if $m \geq 2$ then the inequality in (8) is strict.

The proof is given in the appendix. The sufficient condition is obtained by subsampling the original MJLS and then showing the weak stability of the subsampled system. The proof of the necessary condition with the weak inequality is standard, while the strict inequality in the case $m \geq 2$ is based on the maximum entropy theorem and the entropy power inequality (EPI). There is a gap between the necessary and sufficient conditions in (7) and (8) in the case where $c \neq 0$ and $m < 2$. In this regime the EPI cannot be applied. In the following, to express the condition for transition to instability with a single inequality, while indicating the existence of this trivial gap, we write

$$|\lambda|^m \lesssim \frac{1}{\rho(\mathrm{P}^\mathsf{T} 2^{-m\mathrm{R}})}. \tag{9}$$

Theorem 1 extends the well known conditions for second moment stability given in [49] to m-th moment stability. A similar result appears in [50, Theorem 3.2] in the special case of a homogeneous MJLS driven by an i.i.d. rate process.

III. Moment Stabilization over Markov Channels

Results on the stability of MJLS are used to characterize the stability of linear dynamical systems when the estimated state is sent to the controller over a digital communication link whose state is described by a Markov process.

A. System model

Consider the scalar dynamical system

$$\begin{aligned} x_{k+1} &= \lambda x_k + u_k + v_k, \\ y_k &= x_k + w_k, \end{aligned} \tag{10}$$

where $k \in \mathbb{N}$, and $|\lambda| \geq 1$. The variable x_k represents the state of the system, u_k the control input, v_k is an additive stochastic disturbance, y_k is the sensor measurement, and w_k

is the measurement noise. Both disturbance and noise are independent of each other and of the initial condition x_0. They are also independent of the channel state, as defined below.

B. Channel Model

The state observer is connected to the actuator through a noiseless digital communication link that at each time k allows transmission without errors of R_k bits. The rate process $\{R_k\}_{k\geq 0}$ is a homogeneous positive-recurrent Markov chain that takes values in $\mathcal{R}$ defined in (2) and that has transition probability matrix P. This noiseless digital link corresponds to a discrete-memoryless channel with Markov state available causally at both the encoder and the decoder. A channel with state is defined by a triple $(\mathcal{X} \times \mathcal{S}, p(y|x,s), \mathcal{Y})$ consisting of an input set $\mathcal{X}$, a state set $\mathcal{S}$, an output set $\mathcal{Y}$, and a transition probability matrix $p(y|x)$ for every $x \in \mathcal{X}$, $s \in \mathcal{S}$, and $y \in \mathcal{Y}$. This channel is memoryless if the output y_k at time k is conditionally independent of everything else given (x_k, s_k). The state sequence is Markov if $S_0, S_1, \ldots$ forms a Markov chain. According to these definitions, our channel model is a discrete-memoryless channel with Markov state $(\mathcal{X} \times \mathcal{S}, p(y|x,s), \mathcal{Y})$ with $\mathcal{X} = \mathcal{Y} = \{1, \ldots, \bar{r}_n\}$, $\mathcal{S} = \{\bar{r}_1, \cdots, \bar{r}_n\}$,

$$p(y|x,s) = \begin{cases} 1 & x = y \text{ and } x \leq s \\ 0. & \text{otherwise} \end{cases} \tag{11}$$

and state transition probabilities

$$p(s_{k+1} = \bar{r}_j | s_k = \bar{r}_i) = p_{ij}. \tag{12}$$

The Shannon capacity of this channel is []

$$C = \sum_{i=1}^{n} \pi_i \bar{r}_i, \tag{13}$$

where $(\pi_1, \ldots, \pi_n)$ denotes the unique stationary distribution of P.

The zero-error capacity of this channel is []

$$C_0 = \bar{r}_1. \tag{14}$$

The capacities in (13) and (14) are the limiting values of a stability threshold function indicating the channel's rate-reliability constraint required to achieve a given level of stabilization. As $m \to \infty$ and the system is highly stable, then the stability threshold function tends to the zero-error capacity that has a hard reliability constraint of providing no decoding error. Conversely, as $m \to 0$ and the system's stability level decreases, then the stability threshold function tends to the Shannon capacity that has a weak reliability constraint of vanishing probability of error.

C. Stability threshold function

The system (10) is mth moment stable if

$$\sup_k \mathsf{E}[|X_k|^m] < \infty, \tag{15}$$

where the expectation is taken with respect to the random initial condition x_0, the additive disturbance v_k, and the rate process R_k. The following Theorem establishes the equivalence between the m-th moment stability of (10) and the weak moment stability of (1)

Theorem 2. *There exists a control scheme that stabilizes the scalar system* (10) *in mth momemt sense if and only if the MJLS* (1) *is weakly mth moment stable, i.e., if and only if*

$$\log|\lambda| \lessapprox -\frac{1}{m} \log \rho(\mathrm{P}^{\mathsf{T}} 2^{-m\mathrm{R}}) \triangleq R(m). \tag{16}$$

The proof is given in the appendix in the case the disturbance has bounded support. This assumotion is made for ease of presentation and to compare our results to the ones on the anytime capacity that only apply to plants with bounded disturbance [11]. The extension to unbounded disturbance can be easily obtained using standard, but more technical, adaptive encoding schemes described in [24], [32], [33].

We now mention several properties of the threshold function $R(m)$, whose proofs are given in the appendix.

Proposition 3.

1) *Monotonicity: $R(m)$ is continuous and strictly decreasing for $m > 0$.*
2) *Convergence to the Shannon capacity:*
$$\lim_{m \to 0} R(m) = \sum_{i=1}^{n} \pi_i \bar{r}_i = C. \tag{17}$$
3) *Convergence to the Zero Error capacity:*
$$R(m) \sim \bar{r}_1 - \frac{1}{m} \log p_{11}, \quad \text{as } m \to \infty, \tag{18}$$
and hence
$$\lim_{m \to \infty} R(m) = \bar{r}_1 = C_0. \tag{19}$$
4) *Sensitivity with respect to self-loop probabilities:*
$$\frac{dR(m)}{dp_{ii}} = -\frac{2^{-m\bar{r}_{ii}}}{m\rho(\mathrm{P}^{\mathsf{T}} 2^{-m\mathrm{R}})} \frac{|\mathrm{D}(1)|}{\sum_{i=1}^{n} |\mathrm{D}(i)|} < 0, \tag{20}$$
where $\mathrm{D} := \rho(\mathrm{P}^{\mathsf{T}} 2^{-m\mathrm{R}})\mathrm{I} - \mathrm{P}^{\mathsf{T}} 2^{-m\mathrm{R}}$, I denotes the $n \times n$ identity matrix, and $|\mathrm{D}(i)|$ is the determinant of the matrix obtained by eliminating the ith row and the ith column from D. We also have the asymptotic behavior
$$\frac{dR(m)}{dp_{11}} \sim -\frac{1}{mp_{11}\ln(2)} \text{ as } m \to \infty. \tag{21}$$
5) *The function $mR(m)$ is nonnegative, strictly increasing, and strictly concave. If $\bar{r}_1 = 0$, then*
$$\lim_{m \to \infty} mR(m) = -\log p_{1,1}. \tag{22}$$

IV. Anytime Capacity of Markov Channels

We now relate the stability threshold function $R(m)$ to the anytime capacity. $R(m)$ depends on both the system's stability level m and on the properties of the channel via the transition matrix P and the matrix of rate values R. We show that for the given Markov channel, it provides a parametric representation of the anytime capacity in terms of system's stability level m.

The anytime capacity is defined in the following context [13]. Consider a system for information transmission that allows the decoding time to be infinite, and improves the reliability of the estimated message as time progresses. More precisely, at each step k in the evolution of the plant a new message m_k of r bits is generated that must be sent over the channel. The coder sends a bit over the channel at each k and the decoder upon reception of the new bit updates the estimates for all messages up to time k. It follows that at time k messages

$$m_0, m_1, \dots, m_k$$

are considered for estimation, while estimates

$$\hat{m}_{0|k}, \hat{m}_{1|k}, \dots, \hat{m}_{k|k}$$

are constructed, given all the bits received up to time k. Hence, the processing operation for any message m_i continues indefinitely for all $k \geq i$. A reliability level α is achieved in the given transmission system if for all k the probability that there exists at least a message in the past whose estimate is incorrect decreases α-exponentially with the number of bits received, namely for all $d \leq k$

$$\mathsf{P}\{(\hat{M}_{0|k}, \dots, \hat{M}_{d|k}) \neq (M_0, \dots, M_d)\} = O(2^{-\alpha d}). \quad (23)$$

The described communication system is characterized by a rate-reliability pair (r, α). The work in [13] has shown that for scalar systems the ability to achieve stability depends on the ability to construct such a communication system, in terms of achievable coding and decoding schemes, with a given rate-reliability constraints.

To state this result in the context of our Markov channel, let the α-anytime capacity $C_A(\alpha)$ be the supremum of the rate r that can be achieved with reliability α. The problems of α-reliable communication and mth moment stabilization of a scalar system over a Markov channel are then equivalent in the sense of the following theorem.

Theorem 4 (Sahai, Mitter [13])**.** *The necessary and sufficient condition for mth moment stabilization of a scalar system with bounded disturbances and in the presence of channel output feedback over a Markov channel is*

$$\log|\lambda| \lesssim C_A(m \log|\lambda|). \quad (24)$$

The anytime capacity is an intermediate notion between the zero-error capacity and the Shannon capacity. The zero-error capacity requires transmission without error. The Shannon capacity requires the decoding error to tend to zero by increasing the length of the code. In the presence of disturbances, only a critical value of the zero-error capacity can guarantee the almost sure stability of the system [9]. On the other hand, for scalar systems in presence of bounded disturbances, a critical value of the anytime capacity can guarantee the ability to stabilize the system in the weaker mth moment sense.

By combining Theorem 2 and Theorem 4, we obtain the following result.

Theorem 5. *The following holds:*

1) *Parametric characterization of the anytime capacity: For every $m > 0$, the anytime capacity C_A satisfies*

$$C_A\big(mR(m)\big) = R(m), \quad (25)$$

i.e., for every $\alpha \geq 0$, there exists a unique $m(\alpha)$ such that

$$m(\alpha)R\big(m(\alpha)\big) = \alpha \quad (26)$$

and

$$C_A(\alpha) = R\big(m(\alpha)\big). \quad (27)$$

2) *$C_A(\alpha)$ is a nonincreasing function of $m > 0$.*
3) *Convergence to the Shannon capacity:*

$$\lim_{\alpha \to 0} C_A(\alpha) = \sum_{i=1}^{n} \pi_i r_i = C, \quad (28)$$

4) *Convergence to the Zero Error capacity: If $\bar{r}_1 = 0$, then for every $\alpha \geq \log(1/p_{11})$*

$$C_A(\alpha) = 0 = C_0. \quad (29)$$

Conversely, if $\bar{r}_1 \neq 0$, then $C_A(\alpha)$ has unbounded support and

$$C_A(\alpha) \sim \bar{r}_1 \frac{\alpha}{\alpha - \log(1/p_{11})}, \quad \text{as } \alpha \to \infty, \quad (30)$$

hence

$$\lim_{\alpha \to \infty} C_A(\alpha) = \bar{r}_1 = C_0. \quad (31)$$

The proof is given in the appendix and uses some results on the log-convexity of the spectral radius of a nonnegative matrix. These results imply that the function $\phi(m) = mR(m)$ is increasing and strictly concave, thus invertible. It follows that for every $\alpha \geq 0$, there exists a unique $m := m(\alpha)$ such that

$$mR(m) = \alpha, \quad (32)$$

hence assuming equality in (16) it follows $C_A\big(mR(m)\big) = R(m)$. The remaining properties are immediate.

We now give some representative examples of the stability threshold function, visually showing its extremal properties and its relationship with the anytime capacity.

Example IV.1. *Let $n = 4, \mathcal{R} = \{1, 3, 4, 5\}$ and P be is a 4×4 circulant matrix with first row equal to $\frac{1}{16}(1, 13, 1, 1)$, namely*

$$P = \frac{1}{16} \begin{bmatrix} 1 & 13 & 1 & 1 \\ 13 & 1 & 1 & 1 \\ 1 & 1 & 1 & 13 \\ 1 & 1 & 13 & 1 \end{bmatrix}. \quad (33)$$

In this case it is easy to compute $C = \frac{1}{4}(1 + 3 + 4 + 5) = \frac{13}{4}$ and $C_0 = 1$. Figure 2 plots the stability threshold function $R(m)$ (together with its asymptotic approximation) and the anytime capacity $C_A(\alpha)$. Both curves have the same shape and they are in fact related by an affine transformation as m grows. Furthermore, both curves have unbounded support and tend to one at infinity. There is a change of convexity for small values of m and α, as indication that $R(m)$ and $C_A(\alpha)$ are generally not convex functions. In contrast, the function $\phi(m) = mR(m)$, reported in red in the top plot of Figure 2, is strictly convex and increasing.

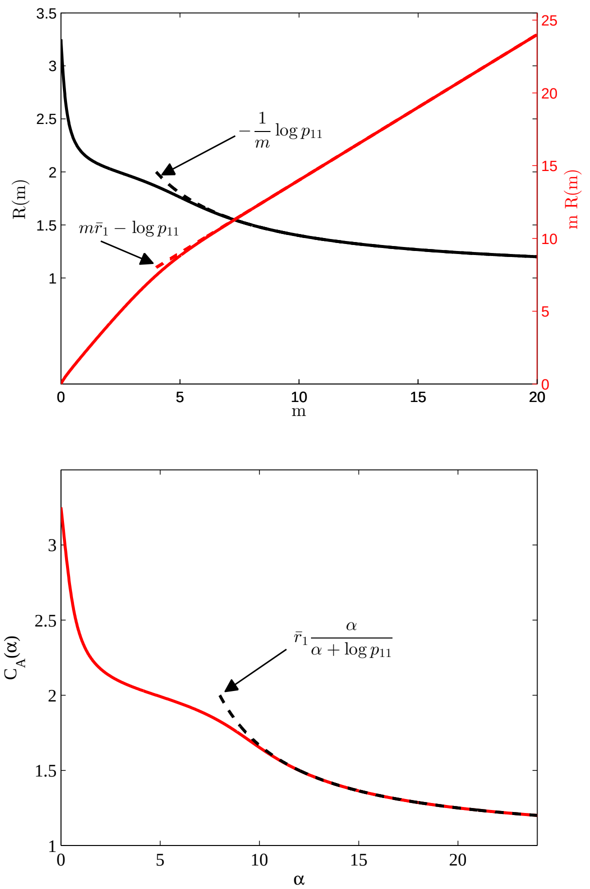

Fig. 2. Stability threshold function and anytime capacity for Example IV.1

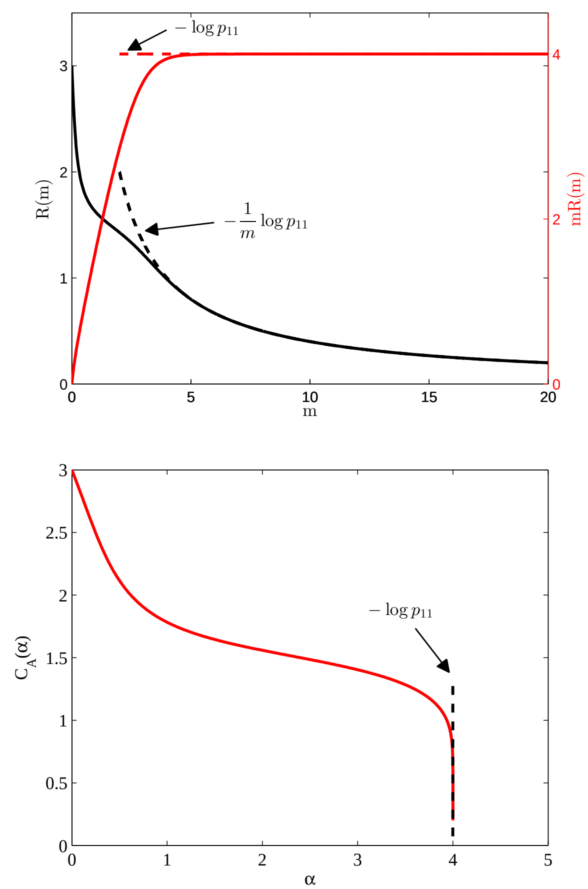

Fig. 3. Stability threshold function and anytime capacity for Example IV.2.

Example IV.2. *Let* $\mathcal{R} = \{0, 3, 4, 5\}$ *and* P *is as in* (33). *The only difference with the previous example is that* $\bar{r}_1$ *is now 0 instead than 1. In this case it is easy to compute* $C = \frac{1}{4}(0+3+4+5) = 3$ *and* $C_0 = 0$. *Figure 3 plots the stability threshold function* $R(m)$ *(together with its asymptotic approximation) and the anytime capacity* $C_A(\alpha)$. *In this case, while* $R(m)$ *has unbounded support,* $C_A(\alpha)$ *is zero for all* $\alpha \geq -\log p_{11} = \log 16 = 4$. *This occurs because the function* $\phi(m) = mR(m)$ *saturates as* $m \to \infty$*, tending to the limiting value* $-\log p_{11} = 4$.

When viewed together, the two examples above show that for some channels a communication system with an arbitrary rate-reliability pair (r, α) cannot be constructed, because the anytime capacity may have bounded support and tend abruptly to zero. However, in order to achieve mth moment stabilization it is sufficient to consider the simpler function $R(m) = C_A(mR(m))$, and construct a communication system whose reliability level depends on the desired stabilization level. It follows that we do not need to compute the whole anytime capacity if we are interested only in moment stabilization, and we may be content with determining the threshold function $R(m)$ corresponding to its parametric representation. The extremal properties of $R(m)$ determine the support of the anytime capacity corresponding to the achievable reliability level α. If $R(m) = O(1/m)$ then the anytime capacity has support bounded by the pre-constant of the asymptotic order. On the other hand, if $R(m)$ decreases at most sub-linearly to zero, or it tends to a constant zero-error capacity, then the anytime capacity has unbounded support and any reliability level α is achieved.

V. The Markov Erasure Channel

We now use the stability threshold function to compute the anytime capacity of the Markov erasure channel. Beside the memoryless erasure channel and the additive white Gaussian noise channel with input power constraint, for which the anytime capacity equals the Shannon capacity, this is the only case where an explicit anytime capacity formula can be obtained.

The Markov erasure channel corresponds to a two-state Markov process where $n = 2$ $\mathcal{R} = \{0, \bar{r}\}$, $p_{12} = q$, and $p_{21} = p$, where $0 < p, q < 1$. In this case,

$$\mathrm{P}^{\mathsf{T}} 2^{-m\mathrm{R}} = \begin{pmatrix} (1-q) & \frac{1}{2^{m\bar{r}}} p \\ q & \frac{1}{2^{m\bar{r}}}(1-p) \end{pmatrix},$$

and we have the following result, whose proof is given in the appendix.

Theorem 6. *The anytime capacity of the Markov Erasure Channel is*

$$C_A(\alpha) = \frac{\alpha \bar{r}}{\alpha + \log_2\left(\frac{1-p-2^\alpha(1-p-q)}{1-(1-q)2^{-\alpha}}\right)}, \tag{34}$$

if $0 \leq \alpha < -\log_2(1-q)$*, and 0 otherwise.*

A. Special cases

We now discuss some special cases, recovering previous results in the literature. By (34) it follows that the anytime capacity of the binary erasure channel (BEC) with Markov erasures and with noiseless channel output feedback is

$$C_A(\alpha) = \frac{\alpha}{\alpha + \log_2\left(\frac{1-p-2^\alpha(1-p-q)}{1-(1-q)2^{-\alpha}}\right)}. \tag{35}$$

By letting $q = 1 - p$, the erasure process becomes i.i.d. and we recover the anytime capacity of the memoryless BEC with erasure probability p derived by Sahai [11, page 129] (in parametric form) and by Xu [15, Theorem 1.3] (in non-parametric form)

$$C_A(\alpha) = \frac{\alpha}{\alpha + \log_2\left(\frac{1-p}{1-p2^{-\alpha}}\right)}. \tag{36}$$

By (34), letting $\alpha \to 0$, we have that

$$\lim_{\alpha \to 0} C_A(\alpha) = \frac{q}{p+q}\bar{r} = \mathsf{E}(R) = C, \tag{37}$$

where the expectation is taken with respect to the stationary distribution of P. This recovers the Shannon capacity of an $\bar{r}$-bit erasure channel with Markov erasures and with noiseless channel output feedback.

In the case $n = 2$, $\bar{r}_1 = 0$, $\bar{r}_2 = r$, and an i.i.d. rate process with $\mathsf{P}\{R_k = 0\} = p_1$ and $P\{R_k = r\} = p_2$ for all k's, then the stability condition becomes

$$|\lambda|^m \left(p_1 + p_2 2^{-mr}\right) < 1,$$

which provides a converse to the achievable scheme of Yüksel and Meyn [51, Theorem 3.3].

If we further let $r \to \infty$, then the stability condition $p_1 > 1/|\lambda|^m$ depends only on the erasure rate of the channel. In this case, our condition generalizes the packet loss model result in [34].

B. Error exponents

We now compare the the anytime error exponent to error exponents of block codes. Specifically, we can compare the curves in Fig. 4 obtained by evaluating (34) with the results by Como, Yüksel, and Tatikonda in [52, Section VI]. *to do*

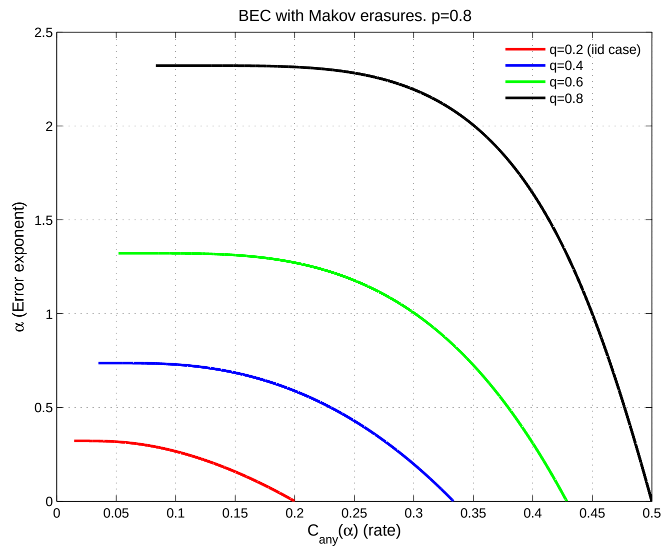

Fig. 4. Anytime reliability of a BEC with Markov erasure for different recovery probability q and $p = 0.8$.

Appendix A
Auxiliary Results

We begin with a theorem of Friedland on the log-convexity of the spectral radius of a nonnegative matrix (superconvexity as Kingman [53] calls it).

Theorem 7 (Friedland Theorem 4.2 [54])**.** *Let $\mathcal{D}_n$ be the set of $n \times n$ real-valued diagonal matrices. Let $\rho(A)$ refer to the spectral radius of a matrix A. Let A be a fixed $n \times n$ non-negative matrix having a positive spectral radius. Define $\phi : \mathcal{D}_n \to \mathbb{R}$ by $\phi(D) := \log \rho(e^D A)$. Then $\phi(D)$ is a convex functional on $\mathcal{D}_n$. Specifically: for every $D_1, D_2 \in \mathcal{D}_n$ and $\alpha \in (0,1)$,*

$$\phi(\alpha D_1 + (1-\alpha)D_2) \leq \alpha\phi(D_1) + (1-\alpha)\phi(D_2). \tag{38}$$

Moreover, if A is irreducible and the diagonal entries of A are positive (or A is fully indecomposable) then equality holds in (37) *if and only if*

$$D_1 - D_2 = cI \tag{39}$$

for some $c \in \mathbb{R}$, where I is the identity matrix.

Theorem 8 (Cohen Theorem 1 [55])**.** *Let A be a fixed $n \times n$ non-negative matrix having a positive spectral radius. Define $D := \rho(A)I - A$, where I denotes the $n \times n$ identity matrix. Then,*

$$0 < \frac{d\rho(A)}{da_{11}} = \frac{|D(1)|}{\sum_{i=1}^n |D(i)|} < 1 \tag{40}$$

Appendix B
Additional Results on Stability of MJLS

1) Strong mth Moment Stability: We say that the system (1) is strongly mth moment stable if there exists a finite $c_m \in \mathbb{R}$ such that

$$\mathsf{E}[|Z|^m] \to c_m, \qquad \text{as } m \to \infty. \tag{41}$$

Clearly, if a MJLS is strongly stable, then it is also weakly stable.

Lemma 1. *For every positive integer number $m = 1, 2, 3, \ldots$, if the MJLS* (1) *is strongly mth moment stable if and only if*

$$|\lambda|^m < \frac{1}{\rho(P^\mathsf{T} 2^{-mR})}. \tag{42}$$

Proof. The proof technique is based on the binomial expansion and was presented in the CDC paper. ◻

APPENDIX C
PROOF OF PROPOSITION 3

1) This is an immediate consequence of the log-convexity of the spectral radius of a nonnegative matrix. Let $D_1 = -m\ \mathrm{diag}(\bar{r}_1, \cdots, \bar{r}_n)\ \ln 2$, $D_2 = 0_{n\times n}$ and $\alpha = \frac{n}{m}$. Notice that $P^{\mathsf{T}} 2^{-mR} = P' e^{D_1}$, $\log \rho(P' e^{D_2}) = 0$, and $H^{(n)} = P' e^{\alpha D_1}$. Then, for every $n < m$, by Theorem 7
$$\begin{aligned}\frac{n}{m} \log \rho(P^{\mathsf{T}} 2^{-mR}) &= \alpha \log \rho(P' e^{D_1}) + (1-\alpha) \log \rho(P' e^{D_2}) \\ &> \log \rho(P' e^{\alpha D_1 + (1-\alpha) D_2}) \\ &= \log \rho(H^{(n)}),\end{aligned} \tag{43}$$
which implies $T(n) > T(m)$.

2) Next,
$$\begin{aligned}\lim_{m\to 0} R(m) &= \lim_{m\to 0} \log \rho\big((P^{\mathsf{T}} 2^{-mR})^{\frac{1}{m}}\big) \\ &= \log \rho\big(\lim_{m\to 0} (P^{\mathsf{T}} 2^{-mR})^{\frac{1}{m}}\big) \\ &= \log \rho\left(\lim_{m\to 0}\left(\begin{bmatrix} \pi_1 2^{mr_1} & \cdots & \pi_1 2^{mr_n} \\ \vdots & \vdots & \vdots \\ \pi_n 2^{mr_1} & \cdots & \pi_n 2^{mr_n}\end{bmatrix}\right)^{\frac{1}{m}}\right) \\ &= \lim_{m\to 0} \frac{1}{m} \log \rho\left(\begin{bmatrix} \pi_1 2^{mr_1} & \cdots & \pi_1 2^{mr_n} \\ \vdots & \vdots & \vdots \\ \pi_n 2^{mr_1} & \cdots & \pi_n 2^{mr_n}\end{bmatrix}\right) \\ &= \lim_{m\to 0} \frac{1}{m} \log\left(\sum_i \pi_i 2^{mr_i}\right) \\ &= \sum_i \pi_i 2^{mr_i}\end{aligned}$$

Lemma 2.
$$\lim_{m\to 0} (P^{\mathsf{T}} 2^{-mR})^{\frac{1}{m}} = \lim_{m\to 0}\left(\begin{bmatrix} \pi_1 2^{mr_1} & \cdots & \pi_1 2^{mr_n} \\ \vdots & \vdots & \vdots \\ \pi_n 2^{mr_1} & \cdots & \pi_n 2^{mr_n}\end{bmatrix}\right)^{\frac{1}{m}}$$

Proof. Let $1/m = k$. By the monotonicity property, it is sufficient to prove the claim for k integer. Let
$$A := \begin{bmatrix} p_{1,1} 2^{r_1/k} & \cdots & p_{1,n} 2^{r_1/k} \\ \vdots & \vdots & \vdots \\ p_{n,1} 2^{r_n/k} & \cdots & p_{n,n} 2^{r_n/k}\end{bmatrix} = \left(H^{-\frac{1}{k}}\right)^T$$
and
$$B := \begin{bmatrix} \pi_1 2^{r_1/k} & \cdots & \pi_n 2^{r_1/k} \\ \vdots & \vdots & \vdots \\ \pi_1 2^{r_n/k} & \cdots & \pi_n 2^{r_n/k}\end{bmatrix} = \left(\begin{bmatrix} \pi_1 2^{r_1/k} & \cdots & \pi_1 2^{r_n/k} \\ \vdots & \vdots & \vdots \\ \pi_n 2^{r_1/k} & \cdots & \pi_n 2^{r_n/k}\end{bmatrix}\right)$$
We prove that
$$\lim_{k\to\infty} A^k = \lim_{k\to\infty} B^k.$$
It is enough to show that
$$\lim_{k\to\infty} [A^k]_{i,j} = \lim_{k\to\infty} [B^k]_{i,j}.$$

Note that
$$\begin{aligned}[A^k]_{i,j} &= \sum_{l_1,\ldots,l_{k-1}} \big(p_{il_1} 2^{i/k}\big) \cdots \big(p_{l_{k-1}j} 2^{l_{k-1}/k}\big) \\ &= \sum_{l_1,\ldots,l_{k-1}} \big(p_{il_1} \cdots p_{l_{k-1}j}\big) 2^{(i+\cdots+l_{k-1})/k}\end{aligned}$$
and
$$\begin{aligned}[B^k]_{i,j} &= \sum_{l_1,\ldots,l_{k-1}} \big(\pi_{l_1} 2^{i/k}\big) \cdots \big(\pi_{l_j} 2^{l_{k-1}/k}\big) \\ &= \sum_{l_1,\ldots,l_{k-1}} \big(\pi_{l_1} \cdots \pi_{l_j}\big) 2^{(i+\cdots+l_{k-1})/k}\end{aligned}$$

$$\begin{aligned}\lim_{k\to\infty} ([A^k]_{i,j} - [B^k]_{i,j}) &= \lim_{k\to\infty} \sum_{l_1,\ldots,l_{k-1}} \big(p_{il_1} \cdots p_{l_{k-1}j} - \pi_{l_1} \cdots \pi_{l_j}\big) 2^{(i+\cdot} \\ &\le \lim_{k\to\infty} \sum_{l_1,\ldots,l_{k-1}} \big(p_{il_1} \cdots p_{l_{k-1}j} - \pi_{l_1} \cdots \pi_{l_j}\big) 2^n \\ &= 2^n \left(\left(\lim_{k\to\infty} \sum_{l_1,\ldots,l_{k-1}} p_{il_1} \cdots p_{l_{k-1}j}\right) - \left(\lim_{k\to\infty}\right.\right. \\ &= 2^n \left(\lim_{k\to\infty} [P^k]_{i,j} - \lim_{k\to\infty} [Q^k]_{i,j}\right) \\ & 0.\end{aligned}$$
$$(i + \cdots + l_{k-1})/k < n.$$
◻

REFERENCES

[1] J. Baillieul and P. J. Antsaklis (editors), *IEEE Transactions on Automatic Control,* special issue on Networked Control Systems, vol. 49, no. 9, September 2004.

[2] ——, *Proceedings of the IEEE,* special issue on Technology of Networked Control Systems, vol. 95, no. 1, Jan. 2007.

[3] M. Franceschetti, T. Javidi, P. R. Kumar, S. K. Mitter, and D. Teneketzis (editors), *IEEE Journal on Selected Areas in Communications,* special issue on Control and Communications, vol. 26, no. 4, May 2008.

[4] K.-D. Kim and P. R. Kumar, "Cyber-physical systems: A perspective at the centennial," *Proceedings of the IEEE*, vol. 100, no. 13, pp. 1287–1308, May 2012.

[5] M. Franceschetti and P. Minero, *Elements of information theory for networked control systems*, ser. Lecture notes in control and information sciences, Ch. 1, G. Como Ed. Springer, 2014.

[6] D. Forney Jr., "Convolutional codes II. Maximum-likelihood decoding," *Information and Control*, vol. 25, no. 3, pp. 222–266, 1974.

[7] L. Schulman, "Coding for interactive communication," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1745 –1756, nov 1996.

[8] R. Ostrovsky, Y. Rabani, and L. Schulman, "Error-correcting codes for automatic control," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 2931 –2941, july 2009.

[9] A. S. Matveev and A. V. Savkin, "Shannon zero error capacity in the problems of state estimation and stabilization via noisy communication channels," *International Journal of Control*, vol. 80, no. 2, pp. 241–255, 2007.

[10] C. E. Shannon, "The zero-error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, pp. 8–19, 1956.

[11] A. Sahai, "Anytime information theory," Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, 2001.

[12] G. Como, F. Fagnani, and S. Zampieri, "Anytime reliable transmission of real-valued information through digital noisy channels," *SIAM J. Control Optim.*, vol. 48, no. 6, pp. 3903–3924, Apr. 2010.

[13] 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. Inf. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.

[14] T. Şimşek, R. Jain, and P. Varaiya, "Scalar estimation and control with noisy binary observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1598–1603, Sept. 2004.

[15] Q. Xu, "Anytime capacity of the AWGN+Erasure channel with feedback," Ph.D. Thesis, University of California at Berkeley, Berkeley, CA, 2005.

[16] R. Sukhavasi and B. Hassibi, "Error correcting codes for distributed control," *Available on-line at http://arxiv.org: arXiv:1112.4236v2 [cs.IT]*, Dec. 2001.

[17] J. Baillieul, "Feedback designs for controlling device arrays with communication channel bandwidth constraints," in *ARO Workshop on Smart Structures*, Penn. State U., USA, Aug. 1999.

[18] ——, "Feedback designs in information-based control," in *B. Pasik-Duncan (ed.), Proceedings of the Workshop on Stochastic Theory and Control*, Lawrence, Kansas, Springer, 2001.

[19] W. S. Wong and R. Brockett, "Systems with finite communication bandwidth constraints. II. Stabilization with limited information feedback," *IEEE Transactions on Automatic Control*, vol. 44, no. 5, pp. 1049–1053, May 1999.

[20] R. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 7, pp. 1279–1289, July 2000.

[21] D. Liberzon, "On stabilization of linear systems with limited information," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 304–307, Feb. 2003.

[22] D. Delchamps, "Stabilizing a linear system with quantized state feedback," *IEEE Transactions on Automatic Control*, vol. 35, no. 8, pp. 916–924, Aug. 1990.

[23] N. Martins, M. Dahleh, and N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 438–447, Mar. 2006.

[24] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM J. Control Optim.*, vol. 43, no. 2, pp. 413–436, Feb. 2004.

[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.

[26] S. Tatikonda and S. K. Mitter, "Control under communication constraints," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, July 2004.

[27] ——, "Control over noisy channels," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1196–1201, July 2004.

[28] V. Gupta and N. Martins, "On stability in the presence of analog erasure channel between the controller and the actuator," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 175–179, Jan. 2010.

[29] S. Yüksel, "Stochastic stabilization of noisy linear systems with fixed-rate limited feedback," *IEEE Transactions on Automatic Control*, vol. 55, no. 12, pp. 2847–2853, Dec. 2010.

[30] S. Yüksel and T. Başar, "Control over noisy forward and reverse channels," *IEEE Transactions on Automatic Control*, vol. 56, no. 5, pp. 1014–1029, May 2011.

[31] V. Borkar and S. Mitter, *LQG Control with Communication Constraints. Communications,Computation, Control and Signal Processing: A Tribute to Thomas Kailath.* Kluver, 1997.

[32] P. Minero, M. Franceschetti, S. Dey, and G. Nair, "Data rate theorem for stabilization over time-varying feedback channels," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 243–255, Feb. 2009.

[33] L. Coviello, P. Minero, and M. Franceschetti, "Stabilization over Markov feedback channels: The general case." *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 349–362, 2013.

[34] V. Gupta, D. Spanos, B. Hassibi, and R. M. Murray, "Optimal LQG control across packet-dropping links," *Systems and Control Letters*, vol. 56, no. 6, pp. 439–446, 2007.

[35] V. Gupta, N. Martins, and J. Baras, "Optimal output feedback control using two remote sensors over erasure channels," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1463–1476, july 2009.

[36] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, Jan. 2007.

[37] M. Huang and S. Dey, "Stability of kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, 2007.

[38] N. Elia, "Remote stabilization over fading channels," *Systems and Control Letters*, vol. 54, no. 3, pp. 237–249, 2005.

[39] N. Elia and S. K. Mitter, "Stabilization of linear systems with limited information," *IEEE Transactions on Automatic Control*, vol. 46, no. 9, pp. 1384–1400, 2001.

[40] S. Tatikonda, A. Sahai, and S. K. Mitter, "Stochastic linear control over a communication channel," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1549–1561, Sept. 2004.

[41] J. Braslavsky, R. Middleton, and J. Freudenberg, "Feedback stabilization over signal-to-noise ratio constrained channels," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1391–1403, Aug. 2007.

[42] R. Middleton, A. Rojas, J. Freudenberg, and J. Braslavsky, "Feedback stabilization over a first order moving average Gaussian noise channel," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 163–167, 2009.

[43] J. Freudenberg, R. H. Middleton, and V. Solo, "Stabilization and disturbance attenuation over a Gaussian communication channel," *IEEE Transactions on Automatic Control*, vol. 55, no. 3, pp. 795–799, 2010.

[44] E. Silva, G. Goodwin, and D. Quevedo, "Control system design subject to snr constraints," *Automatica*, vol. 46, no. 2, Dec. 2010.

[45] E. Ardestanizadeh and M. Franceschetti, "Control-theoretic approach to communication with feedback," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, October 2012.

[46] Y.-H. Kim, "Feedback capacity of stationary Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 57–85, 2010.

[47] N. Elia, "When Bode meets Shannon: control-oriented feedback communication schemes," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1477–1488, Sept. 2004.

[48] E. Ardestanizadeh, P. Minero, and M. Franceschetti, "LQG control approach to Gaussian broadcast channels with feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5267–5278, Aug. 2012.

[49] O. Costa, D. Fragoso, and R. Marques, *Discrete-Time Markov Jump Linear Systems*, ser. Probability and its Applications. Springer, 2004.

[50] Y. Fang, K. A. Loparo, and X. Fend, "Almost sure and δ-moment stability of jump linear systems," *International Journal of Control*, vol. 59, no. 5, pp. 1281–1307, 1994.

[51] S. Yuksel and S. Meyn, "Random-time, state-dependent stochastic drift for markov chains and application to stochastic stabilization over erasure channels," *IEEE Transactions on Automatic Control*, vol. 58, no. 1, pp. 47–59, Jan. 2012.

[52] G. Como, S. Yüksel, and S. Tatikonda, "The error exponent of variable-length codes over Markov channels with feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2139–2160, 2009.

[53] J. K. Kingman, "A convexity property of positive matrices," *The Quarterly Journal of Mathematics*, vol. 12, no. 1, pp. 283–284, 1961.

[54] S. Friedland, "Convex spectral functions," *Linear and Multilinear Algebra*, vol. 9, no. 4, pp. 299–316, 1981.

[55] J. E. Cohen, "Derivatives of the spectral radius as a function of non-negative matrix elements," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 83, pp. 183–190, 3 1978.