The Anytime Capacity
of Markov Channels

I. INTRODUCTION

Some noteworthy references on the anytime capacity [1]–[8]. The usual references on network-theoretic approach [9]–[14]; the usual references on data rate theorem [15], [15]–[28]; some references on stabilization over AWGN channels [18], [29], [30]; some references on MJLS [31]–[34]

II. MARKOV JUMP LINEAR SYSTEMS.

Consider the scalar non-homogeneous MJLS [35] 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

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

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 $i, j \in \{1, \dots, n\}$.

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

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

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

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

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 $P^{\mathsf{T}} 2^{-mR}$, where 2^{-mR} denotes the base-2 matrix exponential of mR, i.e.,

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

The spectral radius of a matrix is the maximum among the absolute values of its eigenvalues.

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

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

then the MJLS (1) is weakly mth moment stable. In the special case $c = 0$, the inequality "$<$" in (7) is replaced by "$\leq$". Conversely, if MJLS (1) is weakly mth moment stable, then

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

In addition, for every $m \geq 2$ the inequality "$\leq$" in (8) is replaced by "$<$".

Proof: [TO WRITE] The proof of the sufficient condition is new, unpublished, and based on the idea of subsampling the original MJLS and of proving the weak stability of the subsampled system. The proof of the necessary condition with "$\leq$" sign is standard, while the proof for $m \geq 2$ is based on the maximum entropy theorem and the EPI, as presented at CDC 2013 [36]. ∎

Notice the trivial gap between the necessary and sufficient conditions in (7) and (8) in the case where $c \neq 0$ and $0 \leq m < 2$. In this regime, the EPI cannot be applied.

Lemma extends the well-known conditions for second moment stability [35] to m-th moment stability. A similar result appears in [37, Theorem 3.2] in the special case of a homogeneous MJLS driven by an i.i.d. rate process.

III. MOMENT STABILIZATION OF SCALAR SYSTEMS OVER MARKOV CHANNELS

Consider now the scalar dynamical system

$$x_{k+1} = \lambda x_k + u_k + v_k, \quad y_k = x_k + w_k, \quad k = 0, 1, \ldots, \tag{9}$$

where $|\lambda| \geq 1$. where x_k represents the state variable of the system, u_k the control input, v_k an additive disturbance independent of the initial condition x_0, y_k the sensor measurement and w_k the measurement noise, that is supposed to be independent of all other random variables.

A. Feedback 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 on (2) and has transition probability P.

The noiseless digital link can be regarded as a discrete-memoryless channel with Markov state available causally at both the encoder and the decoder. Formally, 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 $y \in \mathcal{Y}$. A channel with state is said to be 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.

The channel model studied in this paper is a discrete-memoryless channel with Markov state $(\mathcal{X} \times \mathcal{S}, p(y|x, s), \mathcal{Y})$ with $\mathcal{X} = \mathcal{Y} = \{1, \ldots, \bar{r}_n\}$, $\mathcal{S} = \{\bar{r}_1, \cdots, \bar{r}_n\}$,

$$p(y|x, s) = \begin{cases} 1 & x = y \text{ and } x \leq s \\ 0. & \text{otherwise} \end{cases} \tag{10}$$

and state transition probabilities

$$p(s_{k+1} = \bar{r}_j | s_k = \bar{r}_i) = p_{ij}. \tag{11}$$

The Shannon capacity of a general discrete-memoryless channel with Markov state was characterized in CITE . In our specific setup,

$$C_{Shannon} = \sum_{i=1}^{n} \pi_i \bar{r}_i, \tag{12}$$

where $(\pi_1, \ldots, \pi_n)$ denotes the unique stationary distribution of P. The zero-error capacity of this channel CITE is given by

$$C_{ZE} = \bar{r}_1. \tag{13}$$

B. Moment Stability

The system (9) is said to be m-th moment stable if

$$\sup_k \mathsf{E}[|X_k|^m] < \infty, \tag{14}$$

where the expectation is taken with respect to the random initial condition, the additive disturbance v_k, and the rate process R_k. We make the usual assumptions on the tail distribution of the disturbance.

The following result establishes the equivalence between the m-th moment stability of (9) and the weak moment stability of (1)

Theorem 2. *There exists a control scheme that stabilizes the scalar system* (9) *in m-square sense if and only if the MJLS* (1) *is weakly mth moment stable, i.e., if and only if*

$$\log |\lambda| \lesssim -\frac{1}{m} \log \rho(P^\mathsf{T} 2^{-mR}) \triangleq R(m). \tag{15}$$

Proof: [TO WRITE] The proof is pretty much as it was sketched in the CDC 2013 paper [36]. The proof technique depends as usual on assumptions made on the noise distribution (bounded or unbounded). Since the anytime capacity result only applies to plants with bounded noise, I suggest writing the proof in this setup and mention that it can be extended to the usual case of unbounded noise. ∎
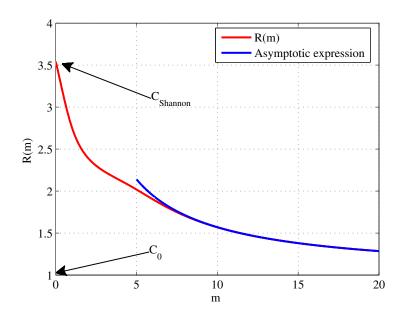
Fig. 1.   The function $R(m)$ in the special case where $P =$ and $n = 4$. The asymptotic expression is given in (17).

Proposition 3. *The following holds:*

1) $R(m)$ *is a strictly decreasing function of $m > 0$.*

2) *Convergence to the Shannon capacity:*

$$\lim_{m \to 0} R(m) = \sum_{i=1}^{n} \pi_i \bar{r}_i = C_{\text{Shannon}}, \tag{16}$$

   where C_{Shannon} denotes the Shannon capacity of the Markov channel.

3) *Convergence to the Zero Error capacity:*

$$R(m) \sim \bar{r}_1 - \frac{1}{m} \log p_{11}, \quad m \to \infty, \tag{17}$$

   hence

$$\lim_{m \to \infty} R(m) = \bar{r}_1 = C_{ZE}, \tag{18}$$

   where C_{ZE} denotes the zero-error capacity of the Markov channel.

4) *Sensitivity with respect to self-loop probabilities*

$$\frac{dR(m)}{dp_{ii}} = -\frac{2^{-m\bar{r}_{ii}}}{m\rho(P^\mathsf{T} 2^{-mR})} \frac{|D(1)|}{\sum_{i=1}^{n} |D(i)|} < 0, \tag{19}$$

where $D := \rho(P^{\mathsf{T}}2^{-mR})I - P^{\mathsf{T}}2^{-mR}$, where I denotes the $n \times n$ identity matrix, and $|D(i)|$ is the determinant of the matrix obtained by eliminating the ith row and the ith column from D. In particular, as $m \to \infty$

$$\frac{dR(m)}{dp_{11}} \sim -\frac{1}{mp_{11}\ln(2)}. \tag{20}$$

5) *The function $mR(m)$ is nonnegative, strictly increasing, and strictly concave. If $\bar{r}_1 = 0$,*

$$\lim_{m\to\infty} mR(m) = -\log p_{1,1}.$$

Proof: See Appendix A. ∎

IV. ANYTIME CAPACITY OF MARKOV CHANNELS

[We must modify this part, which is copied from the book chapter for Cuomo] Consider a system for information transmission that allows the time for processing the received codeword at the decoder to be infinite, and improves the reliability 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

$$m_0, m_1, \ldots, m_k$$

are considered for estimation, while estimates

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

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

$$\mathbb{P}((\hat{M}_{0|k}, \ldots, \hat{M}_{d|k}) \neq (M_0, \ldots, M_d)) = O(2^{-\alpha d}) \quad \text{for all } d \leq k. \tag{21}$$

The described communication system is then characterized by a rate-reliability pair (r, α). It turns out that the ability to stabilize a dynamical system depends on the ability to construct such a communication system, in terms of achievable coding and decoding schemes, with a given rate-reliability constraints.

Let the supremum of the rate r that can be achieved with reliability α be the α-anytime capacity $C_A(\alpha)$ of a given Markov channel with channel feedback. The problem of α-reliable communication and mth moment stabilization are equivalent, as shown by the following.

Theorem 4. *The necessary and sufficient condition for m-moment stabilization of a scalar system with bounded disturbances and in the presence of channel output feedback over a Markov channel is*

$$\log |\lambda| \lesssim C_A(m \log |\lambda|). \tag{22}$$

Proof: [TO DO] I think this should be immediate from Sahai's paper. ∎

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*

$$C_A\big(mR(m)\big) = R(m), \tag{23}$$

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

$$m(\alpha)R\big(m(\alpha)\big) = \alpha \tag{24}$$

   and

$$C(\alpha) = R\big(m(\alpha)\big). \tag{25}$$

2) *$C_A(\alpha)$ is a nonincreasing function of $m > 0$.*

3) *Convergence to the Shannon capacity:*

$$\lim_{\alpha \to 0} C_A(\alpha) = \sum_{i=1}^{n} \pi_i r_i = C_{\text{Shannon}}, \tag{26}$$

4) *Convergence to the Zero Error capacity: If $\bar{r}_1 = 0$, then*

$$C_A(\alpha) = 0 = C_{ZE}, \quad \text{for every } \alpha \geq \log(1/p_{11}), \tag{27}$$

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

$$C_A(\alpha) \sim \bar{r}_1 \frac{\alpha}{\alpha - \log(1/p_{11})}, \quad \text{as } \alpha \to \infty, \tag{28}$$

hence

$$\lim_{\alpha \to \infty} C_A(\alpha) = \bar{r}_1 = C_{ZE}. \tag{29}$$

Proof: [Sketch] By using Theorem 7, it can shown 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

$$mR(m) = \alpha, \tag{30}$$

hence assuming equality in (15), it follows $C_A\big(mR(m)\big) = R(m)$. The remaining properties are immediate. ∎

Example IV.1. *Suppose that $n = 4, \mathcal{R} = \{1, 3, 4, 5\}$ and P is a 4x4 circulant matrix with first row equal to $\frac{1}{16}(1, 13, 1, 1)$, i.e.,*

$$P = \frac{1}{16} \begin{bmatrix} 1 & 13 & 1 & 1 \\ 13 & 1 & 1 & 1 \\ 1 & 1 & 1 & 13 \\ 1 & 1 & 13 & 1 \end{bmatrix}. \tag{31}$$

In this case $C_{Shannon} = \frac{1}{4}(1 + 3 + 4 + 5) = \frac{13}{4}$ and $C_{ZE} = 1$. Fig. 2 shows a numerical evaluation of $T(m)$ (with asymptotic approximation) and $C_\alpha(\alpha)$ in this case. Notice that both curves have the same shape and they are in fact related by an affine transformation as m grows. Notice the change of convexity for small values of m and α, as indication that $T(m)$ and $C_\alpha(\alpha)$ are not convex functions in general.
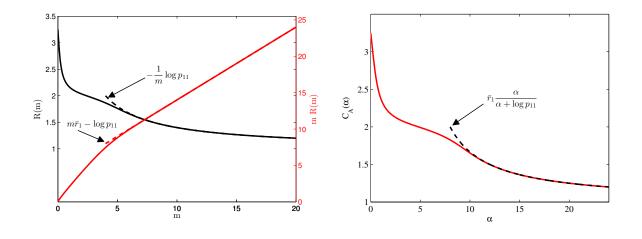
Fig. 2. $T(m)$.

Example IV.2. *Suppose now that $\mathcal{R} = \{0, 3, 4, 5\}$ and P is as in (31). Notice that the only difference with the previous example is that $\bar{r}_1$ is now 0 instead of 1. In this case $C_{Shannon} = \frac{1}{4}(0+3+4+5) = 3$ and $C_{ZE} = 0$. Fig. 3 shows a numerical evaluation of $T(m)$ (with asymptotic approximation) and $C_\alpha(\alpha)$ in this case. Notice that $C_\alpha(\alpha)$ is zero for $\alpha \geq -\log p_{11} = \log 16 = 4$.*
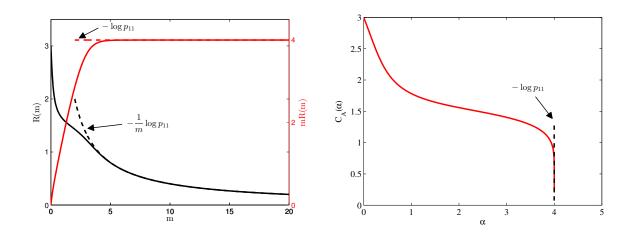

Fig. 3. $T(m)$.

A. The Markov Erasure Channel

Consider the special case of a two-state Markov process where $n = 2$ $\mathcal{R} = \{0, \bar{r}\}$, $p_{12} = q$, and $p_{21} = p$ for some $0 < p, q < 1$. In this case,

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

For this channel model we have the following result.

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

$$C_A(\alpha) = \frac{\alpha \bar{r}}{\alpha + \log_2 \left(\frac{1 - p - 2^\alpha (1 - p - q)}{1 - (1 - q) 2^{-\alpha}} \right)}, \tag{32}$$

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

A few remarks are in order:

Remark 1. *Eq.* (33) *provides the anytime reliability of a binary erasure channel (BEC) with Markov erasures and with noiseless channel output feedback. If, in particular, we let $q = 1 - p$, then the erasure process becomes i.i.d. and* (33) *recovers the anytime capacity of the memoryless BEC with erasure probability p derived by Sahai [3, page 129] (in parametric form) and by Xu [38, Theorem 1.3] (in non-parametric form)*

$$C_A(\alpha) = \frac{\alpha}{\alpha + \log_2 \left(\frac{1 - p}{1 - p 2^{-\alpha}} \right)}, \tag{33}$$

Remark 2. *Observe that $\lim_{\alpha \to 0} C_A(\alpha) = \frac{q}{p + q} \bar{r} = \mathbb{E}(R) = C_{Shannon}$, where the expectation is taken with respect to the stationary distribution of P. This limiting value is the Shannon capacity $C_{Shannon}$ of an $\bar{r}$-bit erasure channel with Markov erasures and with noiseless channel output feedback.*

$C_A(\alpha) = 0$ *for any $\alpha >= -\log_2(1 - q)$ because the probability that d consecutive zeros occurs is $(1 - q)^d = 2^{-d(-\log_2(1 - q))}$.*

Remark 3. *If we specialize to the case $n = 2$, $\bar{r}_1 = 0$, $\bar{r}_2 = r$, and an i.i.d. rate process with*

$P\{R_k = 0\} = p_1$ *and* $P\{R_k = r\} = p_2$ *for all* k*'s, then*

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

which provides a converse to the achievable scheme in [40, Theorem 3.3], and 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 generalize the packet loss model result in [9].*
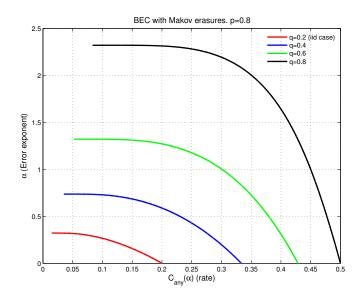

Fig. 4. Anytime reliability of a BEC with Markov erasure for different recovery probability q and $p = 0.8$.

TO DO: Similarly to what Sahai did in the case of the memoryless BEC in his CDC 2004 presentation (http://www.eecs.berkeley.edu/~sahai/Presentations/CDC2004.pdf), we can compare the the anytime error exponent to error exponents of block codes. Specifically, we can compare the curves in Fig. 4 obtained by evaluating (33) with the results by Como, Yüksel, and Tatikonda in [39, Section VI].

V. AUXILIARY RESULTS

We begin with a theorem of Friedland on the log-convexity of the spectral radius of a nonnegative matrix (superconvexity as Kingman [41] called it).

Theorem 7 (Friedland Theorem 4.2 [42])**.** *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 $\alpha \in (0,1)$,

$$\phi(\alpha D_1 + (1 - \alpha)D_2) \leq \alpha \phi(D_1) + (1 - \alpha)\phi(D_2). \tag{34}$$

Moreover, if A is irreducible and the diagonal entries of A are positive (or A is fully indecomposable) then equality holds in (34) if and only if

$$D_1 - D_2 = cI \tag{35}$$

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

Theorem 8 (Cohen Theorem 1 [43]). *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,*

$$0 < \frac{d\rho(A)}{da_{11}} = \frac{|D(1)|}{\sum_{i=1}^{n} |D(i)|} < 1 \tag{36}$$

VI. EXTENSIONS

A. More 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

$$\mathsf{E}[|Z|^m] \to c_m, \qquad \text{as } m \to \infty. \tag{37}$$

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*

$$|\lambda|^m < \frac{1}{\rho(P^{\mathsf{T}} 2^{-mR})}. \tag{38}$$

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

APPENDIX A

PROOF OF PROPOSITION 3

1) This is an immediate consequence of the log-convexity of the spectral radius of a non-negative matrix. Let $D_1 = -m \operatorname{diag}(\bar{r}_1, \cdots, \bar{r}_n) \ln 2$, $D_2 = 0_{n \times n}$ 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

$$
\begin{aligned}
\frac{n}{m} \log \rho(P^\mathsf{T} 2^{-mR}) &= \alpha \log \rho(P' e^{D_1}) + (1 - \alpha) \log \rho(P' e^{D_2}) \\
&> \log \rho(P' e^{\alpha D_1 + (1-\alpha) D_2}) \\
&= \log \rho(H^{(n)}),
\end{aligned}
\tag{39}
$$

which implies $T(n) > T(m)$.

2) Next,

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

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

$$B := \begin{bmatrix} \pi_1 2^{r_1/k} & \cdots & \pi_n 2^{r_1/k} \\ \vdots & \vdots & \vdots \\ \pi_1 2^{r_n/k} & \cdots & \pi_n 2^{r_n/k} \end{bmatrix} = \left(\begin{bmatrix} \pi_1 2^{r_1/k} & \cdots & \pi_1 2^{r_n/k} \\ \vdots & \vdots & \vdots \\ \pi_n 2^{r_1/k} & \cdots & \pi_n 2^{r_n/k} \end{bmatrix} \right)^T$$

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

$$[A^k]_{i,j} = \sum_{l_1,\ldots,l_{k-1}} \left(p_{i l_1} 2^{i/k} \right) \cdots \left(p_{l_{k-1} j} 2^{l_{k-1}/k} \right)$$

$$= \sum_{l_1,\ldots,l_{k-1}} \left(p_{i l_1} \cdots p_{l_{k-1} j} \right) 2^{(i + \cdots + l_{k-1})/k}$$

and

$$[B^k]_{i,j} = \sum_{l_1,\ldots,l_{k-1}} \left(\pi_{l_1} 2^{i/k} \right) \cdots \left(\pi_{l_j} 2^{l_{k-1}/k} \right)$$

$$= \sum_{l_1,\ldots,l_{k-1}} \left(\pi_{l_1} \cdots \pi_{l_j} \right) 2^{(i + \cdots + l_{k-1})/k}$$

$$\lim_{k\to\infty} \left([A^k]_{i,j} - [B^k]_{i,j}\right) = \lim_{k\to\infty} \sum_{l_1,\ldots,l_{k-1}} \left(p_{il_1}\cdots p_{l_{k-1}j} - \pi_{l_1}\cdots\pi_{l_j}\right)2^{(i+\cdots+l_{k-1})/k}$$

$$\leq \lim_{k\to\infty} \sum_{l_1,\ldots,l_{k-1}} \left(p_{il_1}\cdots p_{l_{k-1}j} - \pi_{l_1}\cdots\pi_{l_j}\right)2^n$$

$$= 2^n \left(\left(\lim_{k\to\infty} \sum_{l_1,\ldots,l_{k-1}} p_{il_1}\cdots p_{l_{k-1}j}\right) - \left(\lim_{k\to\infty} \sum_{l_1,\ldots,l_{k-1}} \pi_{l_1}\cdots\pi_{l_j}\right)\right)$$

$$= 2^n \left(\lim_{k\to\infty}[P^k]_{i,j} - \lim_{k\to\infty}[Q^k]_{i,j}\right)$$

$$0.$$

$(i + \cdots + l_{k-1})/k < n.$ ∎

REFERENCES

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

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

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

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, Sept. 2004.

[13] 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.

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

[15] 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.

[16] 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.

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

[18] 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.

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

[20] 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.

[21] 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.

[22] 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.

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

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

[25] K. You and L. Xie, "Minimum data rate for mean square stabilizability of linear systems with Markovian packet losses," *IEEE Transactions on Automatic Control*, vol. 56, no. 4, pp. 772–785, Apr. 2011.

[26] ——, "Minimum data rate for mean square stabilization of discrete LTI systems over lossy channels," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2373–2378, Oct. 2010.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] P. Seiler and R. Sengupta, "Analysis of communication losses in vehicle control problems," in *Proceedings of IEEE American Control Conference*, vol. 2, 2001, pp. 1491–1496.

[32] P. Kawka and A. Alleyne, "Stability and performance of packet-based feedback control over a Markov channel," in *Proceedings of IEEE American Control Conference*, June 2006, pp. 2807–2812.

[33] J. Liu, N. Elia, and S. Tatikonda, "Capacity-achieving feedback scheme for Markov channels with channel state information," in *Proc. IEEE International Symposium on Information Theory*, July 2004.

[34] A. Sahai, S. Avestimehr, and P. Minero, "Anytime communication over the gilbert-eliot channel with noiseless feedback," in *Proc. IEEE International Symposium on Information Theory*, Oct. 2006.

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

[36] P. Minero, M. Franceschetti, and T. Lim, "Moment stabilization over markov channels," in *2013 IEEE 52nd Annual Conference on Decision and Control (CDC 2013),*, Dec 2013, pp. 6091–6096.

[37] 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.

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

[39] 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.

[40] 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.

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

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

[43] 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.