

A Random Time Stochastic Drift Result and Application to Stochastic Stabilization Over Noisy Channels

Serdar Yüksel¹

Abstract—A random time state-dependent drift result leading to various forms of stochastic stability for a Markov Chain is presented. Application to a network stabilization problem is studied. In particular, we observe that, for control over a discrete erasure channel with feedback, for recurrence or stochastic stability, it suffices to have the capacity being greater than the logarithm of the unstable eigenvalue. For the finiteness of a second moment, however, more stringent criteria are needed.

I. INTRODUCTION

Many network applications and recently popular networked control applications require the access of control and sensor information to be observed intermittently. Toward generating a solution for such problems, this paper studies random time state dependent drift conditions leading to the existence of an invariant distribution possibly with moment constraints. Earlier, Meyn and Tweedie [11] considered a number of conditions for both deterministic and random time state dependent drift conditions. In this paper, we extend some of the conditions in [11] with regard to both finite moment conditions, and the existence of an invariant distribution. We apply these results to a stabilization problem over an erasure channel.

We first present results on stochastic stability.

II. RESULTS ON STOCHASTIC STABILITY

Let us first make a number of definitions. Let $\{x_t, t \geq 0\}$ be a Markov chain with state space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$, and defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, where $\mathcal{B}(\mathbb{X})$ denotes the Borel σ -field on the complete, separable, metric space \mathbb{X} , Ω is the sample space, \mathcal{F} a sigma field of subsets of Ω , and \mathcal{P} a probability measure. Let $P(x, D) := P(x_{t+1} \in D | x_t = x)$ denote the transition probability from x to D , that is the probability of the event $\{x_{t+1} \in D\}$ given that $x_t = x$. The evolution of the probability of events is completely determined by the transition probability and the probability of the initial state, $P(dx_0)$. The probability of the event $\{x_{t+1} \in D\}$ for any t can be computed recursively by starting at $t = 0$, with $P(x_1 \in D) = \int_{\mathbb{X}} P(x_1 \in D | x_0 = x) P(dx_0)$, and iterating with a similar formula for $t = 1, 2, \dots$. See [9] for the following:

¹Mathematics and Engineering Program, Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6. Research Supported By the Natural Sciences and Engineering Research Council of Canada (NSERC). Email: yuksel@mast.queensu.ca

Definition 2.1: For a Markov chain with transition probability defined as before, a probability measure π is *invariant* on the Borel space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ if

$$\pi(D) = \int_{\mathbb{X}} P(x, D) \pi(dx), \quad \forall D \in \mathcal{B}(\mathbb{X}).$$

Definition 2.2: A Markov chain is μ -irreducible, if for any set $B \subset \mathbb{X}$, such that $\mu(B) > 0$, and $\forall x \in \mathbb{R}$, there exists some integer $n > 0$, possibly depending on B and x , such that $P^n(x, B) > 0$, where $P^n(x, B)$ is the transition probability in n stages, that is $P(x_{t+n} \in B | x_t = x)$.

Definition 2.3: A set $A \subset \mathbb{X}$ is μ -petite on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ if for some distribution \mathcal{T} on \mathbb{N} (set of natural numbers), and some measure μ ,

$$\sum_{n=0}^{\infty} P^n(x, B) \mathcal{T}(n) \geq \mu(B), \quad \forall x \in A, \text{ and } B \in \mathcal{B}(\mathbb{X}),$$

where $\mathcal{B}(\mathbb{X})$ denotes the (Borel) sigma-field on \mathbb{X} .

Theorem 2.1 ([10] Thm. 4.1): Suppose that $\{x_t\}$ is a φ -irreducible Markov chain, and suppose that there is a set $A \in \mathcal{B}(\mathbb{X})$ satisfying the following:

A is μ -petite for some μ .

A is *recurrent*: $P_x(\tau_A < \infty) = 1$ for any $x \in \mathbb{X}$.

A is *regular*: $\sup_{x \in A} E_x[\tau_A] < \infty$. Then $\{x_t\}$ is positive Harris recurrent (and thus admits a unique invariant probability measure).

Definition 2.4: A Markov chain is weak Feller, if $\int_{\mathbb{X}} P(x, dy) f(y)$ is continuous and bounded on \mathbb{X} for every continuous and bounded f on \mathbb{X} .

The following is our first result. We have omitted the proof.

Throughout this section we consider a sequence of times $\{\tau_i : i \in \mathbb{N}_+\}$ which is assumed to be non-decreasing, with $\tau_0 = 0$.

Theorem 2.2: Suppose that $\{x_t\}$ is a μ -irreducible Markov chain. Suppose moreover that there is a function $V : \mathbb{X} \rightarrow (0, \infty)$, a petite set C , and constants $\kappa \in (0, 1)$, $b \in \mathbb{R}$, such that the following hold:

$$\begin{aligned} E[V(x_{\tau_{z+1}}) | \mathcal{F}_{\tau_z}] &\leq (1 - \kappa)V(x_{\tau_z}) + b1_{\{x_{\tau_z} \in C\}} \\ E[\tau_{z+1} - \tau_z | \mathcal{F}_{\tau_z}] &\leq V(x_{\tau_z}), \end{aligned} \quad z \geq 0. \quad (1)$$

Then, the Markov chain is positive Harris recurrent.

The following provides a criterion for finite moments, which we refer to as random-time state-dependent stochastic drift.

Theorem 2.3: Suppose that $\{x_t\}$ is a μ -irreducible Markov chain. Suppose moreover that there are functions $V : \mathbb{X} \rightarrow$

$(0, \infty)$, $\delta: \mathbb{X} \rightarrow [1, \infty)$, $f: \mathbb{X} \rightarrow [1, \infty)$, a petite set C , and a constant $b \in \mathfrak{R}$, such that the following hold:

$$\begin{aligned} E[V(x_{\tau_z+1}) | \mathcal{F}_{\tau_z}] &\leq V(x_{\tau_z}) - \delta(x_{\tau_z}) + b1_{\{x_{\tau_z} \in C\}} \\ E\left[\sum_{k=\tau_z}^{\tau_z+1-1} f(x_k) | \mathcal{F}_{\tau_z}\right] &\leq \delta(x_{\tau_z}), \quad z \geq 0. \end{aligned} \quad (2)$$

Then $\{x_t\}$ is positive Harris recurrent, and moreover $\lim_{t \rightarrow \infty} E[f(x_t)] = E_\pi[f(x)] < \infty$, with π being the invariant distribution.

By taking $f(x) = 1$ for all $x \in \mathbb{X}$, we obtain the following corollary to Theorem 2.3.

Corollary 2.1: Suppose that is a φ -irreducible Markov chain. Suppose moreover that there is a function $V: \mathbb{X} \rightarrow (0, \infty)$, a petite set C , and a constant $b \in \mathfrak{R}$, such that the following hold:

$$\begin{aligned} E[V(x_{\tau_z+1}) | \mathcal{F}_{\tau_z}] &\leq V(x_{\tau_z}) - 1 + b1_{\{x_{\tau_z} \in C\}} \\ \sup_{x \in \mathbb{X}, z \geq 0} E[\tau_{z+1} - \tau_z | x_{\tau_z} = x] &< \infty. \end{aligned} \quad (3)$$

Then $\{x_t\}$ is positive Harris recurrent.

Theorem 2.4: In the above theorems, if the irreducibility condition is replaced with the condition that the Markov Chain is weak Feller, then (i) Under Theorem 2.2, there exists an invariant distribution (ii) Under Theorem 2.3, for all invariant distributions π , $\sup_\pi \lim_{t \rightarrow \infty} E[f_t] < \infty$ exists.

Some further related results on convergence and further relaxations will be reported elsewhere.

Before ending this section, we note that the existence of an invariant distribution is useful particularly because of the sample-path individual ergodic theorem:

Theorem 2.5 (Birkhoff's Individual Ergodic Theorem): Consider a positive Harris recurrent Markov process $\{x_t\}$ taking values in \mathbb{X} , with invariant distribution $\pi(\cdot)$. Let $f: \mathbb{X} \rightarrow \mathbb{R}$ be a bounded function, measurable on $\mathcal{B}(\mathbb{X})$. Then, the following holds almost surely:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} f(x_t) = \int f(x) \pi(dx).$$

III. APPLICATION TO CONTROL OVER AN ERASURE CHANNEL

We consider a remote stabilization problem where a controller having access to measurements from a channel, and acts on a plant, which is open-loop unstable (Figure 1).

A channel coder maps the source symbols, state values, to corresponding channel inputs. This is done via quantizers. The quantizer outputs are transmitted via a noisy memoryless channel, hence the receiver has access to noisy versions of the quantizer/coder outputs for each time, which we denote by $\{q'\} \in \mathcal{M}$ generated according to a probability distribution for every fixed $q \in \{1, 2, \dots, m\} = \mathcal{M}$, which we denote by $\{P(q'|q)\}$.

Before proceeding further with the description of the system, we discuss the quantization policy investigated. A *quantizer*, Q , for a scalar continuous variable is a map $\mathbb{R} \rightarrow \mathbb{R}$,

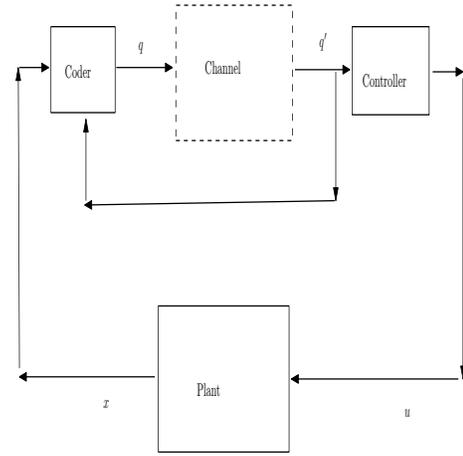


Fig. 1: Control over a noisy channel with causal feedback.

characterized by a sequence of bins $\{\mathcal{B}_i\}$ and their representation $\{q^i\}$, such that $\forall i$, $Q(x) = q^i$ if and only if $x \in \mathcal{B}_i$. Of particular interest is the class of uniform quantizers. In the following, we modify the description of a traditional uniform quantizer by assigning the same value when the state is in the overflow region² of the quantizer. As such, a uniform quantizer: $Q_K^\Delta: \mathbb{R} \rightarrow \mathbb{R}$ with step size Δ and an (odd) K number of levels satisfies the following for $k = 1, 2, \dots, K$:

$$Q_K^\Delta(x) = \begin{cases} \left(\frac{-(K+1)}{2} + k\right)\Delta, & \text{if } x \in \left[\left(\frac{-K}{2} + (k-1)\right)\Delta, \left(\frac{-K}{2} + k\right)\Delta\right) \\ 0, & \text{if } |x| > \frac{K}{2}\Delta \text{ or } x = \frac{K}{2}\Delta \end{cases}$$

A general class of quantizers are those which are adaptive. Let \mathbb{S} be a set of states for a quantizer state S . Let $F: \mathbb{S} \times \mathcal{M}' \rightarrow \mathbb{S}$ be a state update-function. An adaptive quantizer has the following state update equations:

$$S_{t+1} = F(S_t, q'_t).$$

Here, q'_t is the channel output at time t , and S_t is the *state* of the quantizer. We note that, such a quantizer is implementable since the updates can be performed both at the encoder and the decoder.

One particular class of adaptive quantizers is introduced by Goodman and Gersho [3]. One such type has the following form with Q_K^Δ being a uniform quantizer with K bins and bin-size Δ and \bar{Q} determining the updates in the bin-size of the uniform quantizer as a function of the source and the current bin size:

$$q_t = Q_K^{\Delta_t}(x_t), \quad \Delta_{t+1} = \Delta_t \bar{Q}(q'_t, \Delta_t) \quad (4)$$

Here Δ_t characterizes the uniform quantizer, as it is the bin size of the quantizer at time t .

We consider an LTI discrete-time scalar system described by

$$x_{t+1} = ax_t + bu_t + d_t, \quad (5)$$

²As such, when $|x| > (K/2)\Delta$, the receiver knows that the source is in the overflow region of the quantizer.

where a is the system coefficient with $|a| \geq 1$, that is, the system is open-loop unstable. We take $b \neq 0$. Here x_t is the state at time t , u_t is the control input, x_0 is a second order random variable, and $\{d_t\}$ is a sequence of zero-mean independent, identically distributed (i.i.d.) Gaussian random variables with a finite second moment.

This system is connected over a channel with a finite capacity to an estimator (controller). The controller has access to the information it has received through the channel.

The controller in our model estimates the state and then applies its control. As such, the problem reduces to a state estimation problem since such a scalar system is controllable. Hence, the stability of the estimation error is equivalent to the stability of the state itself.

We assume the data rate (bits transmitted per channel use) to be some rate R . We will try to find a relationship between the packet loss probability and the data rate which leads to stability.

A. Literature Review with Regard to Stabilization

There has been considerable amount of research in the literature on quantizer design for such a stabilization problem, for a detailed review see [23]. Due to space constraints, we are unable to provide a detailed account here and only review some of the directly related literature to this note. Zooming type adaptive quantizers, which will be described further in the paper, have been recently introduced by Brockett and Liberzon [2], for remote stabilization of open-loop unstable, noise-free systems with arbitrary initial conditions. Nair and Evans [8] provided a stability result under the assumption that the quantizer is variable-rate and showed that for a noisy setup (with unbounded support for the noise probability measure) that on average it suffices to use more than $\log_2(|a|)$ bits to achieve a form of stability. [8] used asymptotic quantization theory to obtain a time-varying scheme, where the quantizer is used at certain intervals at a very high rate, and at other times, the quantizer is not used. In contradistinction with the result of Nair and Evans, we provide a technique which allows us to both provide a result for the case when the quantizer is fixed-rate as well as to obtain an invariance condition for a probability measure on the quantizer parameters. There is also a large body of literature on quantizer design in the communications and information theory community. One important reference is the work by Goodman and Gersho [3], where an adaptive quantizer was introduced and the adaptive quantizer's stationarity properties were investigated when the source fed to the quantizer is a second order and i.i.d. sequence. In fact, zooming type quantizers is a special class of Goodman and Gersho's adaptive quantization scheme. Kieffer and Dunham [5] have obtained conditions for the stochastic stability of a number of coding schemes when the source considered is also stable, where various forms of stability of the quantizer and the estimation error have been studied. In our case, however, the schemes in [3] and [5] are not directly applicable, as the process we consider is open-loop unstable, as well as Markovian.

References [12] and [19] showed that zero-error capacity is related for convergence in almost sure senses. Matveev and Savkin showed that, almost sure stabilization is not possible for a class of channels. In their definition, almost sure stabilization is equivalent to the following: $\limsup \|x_t\| < \infty$, almost surely. Reference [22] proposed any-time capacity as an appropriate measure of channels used in control systems. References [21], [20] and [8] studied control over noisy channels and obtained fundamental lower bounds for information transmission. Recently, [13] studied the problem concerning time-varying channels and provided a necessity and sufficiency result for boundedness of second moments.

Earlier, we had obtained conditions for the existence of an invariant distribution in [23] for noisy channels and in [17] for discrete-channels, under a fixed-rate constraint. Earlier, we had adopted the approach of using stochastic drift arguments in [23].

In this paper, our contribution here is on stochastic stabilization over erasure channels. We show that the second moment has a limit, and the individual ergodic theorem applies.

One important result in this paper is the observation that there exists a well-defined limit for the sequence of second moments $\{E[\|x_t\|_2^2], t \in \mathbb{Z}_+\}$ and this limit is finite. This result allows us to formulate an optimal quantization problem when the noise process has unbounded support and the state space for the quantizer bin edges are unbounded. This paves the way for extending the results of [1] to unstable systems for optimal quantizer design.

In view of the literature, the contributions of this note are stated as follows.

- Results on the existence of an invariant distribution is presented. To our knowledge, the first result showing that $\limsup_{t \rightarrow \infty} E[x_t^2] = \lim_{t \rightarrow \infty} E[x_t^2] < \infty$ is obtained, when the system is driven by a Gaussian disturbance. This result paves the way for solving average cost infinite horizon optimal control problems under quantization constraints.
- The obtained result uses a fixed-rate quantization scheme. This is implementable in a practical setting.

B. Results for Control over Erasure Channels

We have the following Theorem.

We first note that $P(p_t = 1) = p$, and $p_t = 1$ is the event that the signal is transmitted with no error.

Theorem 3.1: If,

$$R > \frac{\log_2(\lceil |a| \rceil + 1)}{p},$$

there exists an adaptive quantization policy such that there exists a compact set S with

$$\sup_{x \in S} E[\min(t > 0 : x_t \in S) | x_0 = x] < \infty,$$

thus S is a recurrent set.

With, $K = 2^R$, let us define $R' = \log_2(2^R - 1)$. We will consider the following update rules. For $t \geq 0$ and with $\Delta_0 \in$

\mathbb{R} selected arbitrarily, consider:

$$\begin{aligned} u_t &= -\frac{a}{b}\hat{x}_t, \\ \hat{x}_t &= p_t Q_K^{\Delta_t}(x_t), \\ \Delta_{t+1} &= \Delta_t \bar{Q}(|\frac{x_t}{\Delta_t 2^{R'-1}}|, p_t) \end{aligned} \quad (6)$$

If we use $\delta, \epsilon, \eta > 0$ with $\eta < \epsilon$ such that,

$$\begin{aligned} \bar{Q}(x, p) &= |a| + \delta & \text{if } |x| > 1, \text{ or } p = 0 \\ \bar{Q}(x, p) &= \frac{|a|}{2^{R'-\eta}} & \text{if } 0 \leq |x| \leq 1, p = 1, \Delta > L \\ \bar{Q}(x, p) &= 1 & \text{if } 0 \leq |x| \leq 1, p = 1, \Delta \leq L \end{aligned}$$

with $\frac{\sqrt{E[d_t^2]}}{L \frac{|a|}{2^{R'-\eta}} - 2^{R'-1}} < \delta$. Note that, the above imply that $\Delta_t \geq L \frac{|a|}{2^{R'-\eta}} =: L'$.

Our result on the existence and uniqueness of an invariant distribution is the following.

Theorem 3.2: Under the setup of Theorem 3.1, for a zooming quantizer, if the quantizer bin sizes are such that their (base-2) logarithms are integer multiples of some scalar s and $\log_2(\bar{Q}(\cdot))$ take values in integer multiples of s , then the jointly Markov process (x_t, Δ_t) forms a positive (Harris) recurrent Markov chain, and, as such, has a unique invariant distribution, given an initial condition. If the integers taken are relatively prime (that is they share no common divisors except for 1), then the invariant distribution is independent of the initial condition (the value of the integer multiplying s).

Theorem 3.3: If, under Theorems 3.1 and 3.2, we have that,

$$a^2(1 - p + \frac{p}{(2^R - 1)^2}) < 1$$

it follows that

$$\lim_{t \rightarrow \infty} E[x_t^2] < \infty,$$

and this limit is independent of the initial states of the quantizer and the system.

Remark: We note that the above results are directly applicable to multi-dimensional systems when the system matrix can be diagonalizable. \diamond

C. Proofs

1) *Proof of Theorem 3.1: Recurrence:* Toward the proof, we will first obtain a supporting result.

Lemma 3.1: Let $\mathcal{B}(\mathbb{R} \times \mathbb{R}_+)$ denote the Borel σ -field on $\mathbb{R} \times \mathbb{R}_+$. It follows that

$$\begin{aligned} P\left((x_t, \Delta_t) \in (C \times D) | (x_{t-1}, \Delta_{t-1}), \dots, (x_0, \Delta_0)\right) \\ = P\left((x_t, \Delta_t) \in (C \times D) | (x_{t-1}, \Delta_{t-1})\right), \end{aligned}$$

$\forall (C \times D) \in \mathcal{B}(\mathbb{R} \times \mathbb{R}_+)$, i.e. (x_t, Δ_t, p_t) is a Markov chain.

Let us define $h_t := \frac{x_t}{\Delta_t 2^{R'-1}}$. Consider the following sets: $C_x = \{x : |x| \leq E\}$, $C_h = \{h : |h| \leq 1\}$, with $E = 2^{R'-1}L'$. Further, let another set be $C'_x = \{x : |x| \leq F\}$, with a sufficiently large F value to be derived below. We will

study the expected number of time stages between visits of $\{(x_t, h_t)\}$ to $C'_x \times C_h$. Consider the drift of the (x_t, h_t) process in Figure 2: When x_t, h_t are in $C_x \times C_h$, the expected drift increases both $|h|$ and $|x|$. When the x_t process gets outside C'_x and h_t outside C_h (under-zoomed), there is a drift for h_t towards C_h , however, $|x_t|$ will keep increasing on average. Finally, when the process hits C_h (perfect-zoom), then the process drifts towards C'_x . We first show that the sequence

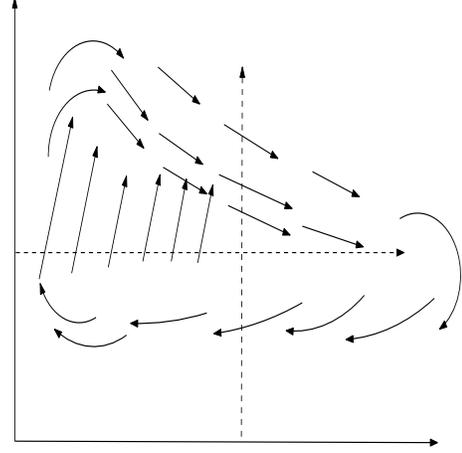


Fig. 2: Drift in the Error Process. The x-axis is $|x|$ and the y-axis is $|h|$: When under-zoomed, the error increases on average; when perfectly-zoomed, the error decreases.

$\{h_t, t \geq 0\}$ visits C_h infinitely often with probability 1 and the expected length of the excursion is uniformly bounded over all possible values of $(x, h) \in C'_x \times C_h$. Let $V(h_t) = h_t^2$ serve as a Lyapunov function. Define a sequence of stopping times for the perfect-zoom case with

$$\tau_0 = 0, \quad \tau_{z+1} = \inf\{k > \tau_z : |h_k| \leq 1\}, \quad z \in \mathbb{Z}_+$$

We have that, if $|h_t| > 1$ (under-zoomed) $E[h_{t+1}^2 | h_t, x_t] \leq \frac{(a^2 + \frac{E[d_t^2]})}{(a+\delta)^2} (h_t)^2$. Since when $|h_t| > 1$, we have that $|x_t| > 2^{R'-1}L'$, it follows that

$$E[h_{t+1}^2 | h_t, x_t] \leq \left(\frac{a^2 + \frac{E[d_t^2]}{E^2}}{(a+\delta)^2}\right) (h_t)^2.$$

If $|h_t| \leq 1$, then

$$\begin{aligned} E[h_{t+1}^2] &\leq P(p_t = 1) \frac{a^2 \frac{(\Delta_t)^2}{4} + E[d_t^2]}{(\Delta_t 2^{R'-1})^2} \left(\frac{2^{R'} - \eta}{|a|}\right)^2 \\ &\quad + P(p_t = 0) \frac{a^2 \frac{(\Delta_t)^2}{4} + E[d_t^2]}{(\Delta_t 2^{R'-1})^2 (|a| + \delta)^2} \\ &\leq \frac{a^2 \frac{L'^2}{4} + E[d_t^2]}{(L' 2^{R'-1})^2} \left(\frac{2^{R'} - \eta}{|a|}\right)^2 =: K_1 \end{aligned} \quad (7)$$

where $L' = L \frac{|a|}{2^{R'-\eta}}$ (this is a lower bound on Δ_t).

Hence, it follows that

$$E[h_{t+1}^2 - h_t^2 | h_t, x_t] \leq -\rho h_t^2 + K_1 1_{(|h_t| \leq 1)}, \quad (8)$$

where $1_{(U)}$ is the indicator function for event U with $\rho = 1 - \frac{(a^2 + \frac{E[d^2]}{E^2})}{(a+\delta)^2}$. Since for $A, B > 0$, $A^2 + B^2 \leq (A+B)^2$ it follows that the hypothesis $\sqrt{\frac{E[d_t^2]}{L'^2 R'^{-1}}} < \delta$ in the theorem statement ensures $\rho > 0$. Now, let us define $K'_1 := K_1 + 1$, $M_0 := V(h_0)$, and for $t \geq 1$,

$$M_t := V(h_t) - \sum_{i=0}^{t-1} (-\rho + K'_1 1_{(h_i \in C_h)})$$

Define a stopping time: $\tau^N = \min(N, \min\{i > 0 : V(h_i) \geq N\}, \min\{i > 0 : V(h_i) \leq 1\})$. As $|h_t| > 1$ when $h_t \notin C_h$, $E[M_{t+1}|x_s, h_s, s \leq t] \leq M_t$, $\forall t \geq 0$, it follows that, $\{M_t\}$ is a Super-Martingale sequence.

The stopping time τ^N is bounded and the Super-Martingale sequence is also bounded for $t \leq \tau^N$. Hence, we have, by the Martingale optional sampling theorem: $E[M_{\tau^N}] \leq E[M_0]$. Hence, we obtain $E[\sum_{i=0}^{\tau^N-1} \rho] \leq V(h_0) + K'_1 E[\sum_{i=0}^{\tau^N-1} 1_{(h_i \in C)}]$. Thus, $\rho E[\tau^N - 1 + 1] \leq V(h_0) + K'_1$, and by the Monotone Convergence Theorem,

$$\rho \lim_{N \rightarrow \infty} E[\tau^N] = \rho E[\tau] \leq V(h_0) + K'_1 = 1 + K'_1,$$

$$E[\tau_{z+1} - \tau_z] \leq (1 + K'_1)/\rho, \quad (9)$$

uniformly for all $h_{\tau_z} \in C_h$. By the strong Markov property (x_{τ_z}, h_{τ_z}) is also a Markov chain as $\{\tau_z < n\} \in \mathcal{F}_n$, the filtration generated by the quantizer state and the quantizer output at time n . The probability that $\tau_{z+1} \neq \tau_z + 1$, is upper bounded by the probability:

$$\begin{aligned} & pP\left(\left\{a\Delta_{\tau_z}/2 + d_{\tau_z} \geq (2^{R'} - 1)\Delta_{\tau_z} \frac{|a|}{2^{R'} - \eta}\right\} \right. \\ & \quad \left. \cup \left\{-a\Delta_{\tau_z}/2 + d_{\tau_z} \leq -(2^{R'} - 1)\Delta_{\tau_z} \frac{|a|}{2^{R'} - \eta}\right\}\right) \\ & = 2pP\left(d_{\tau_z} > (|a|/2)\Delta_{\tau_z}((2^{R'} - 1)/2^{R'} - \eta) - 1\right) \end{aligned} \quad (10)$$

Let us define:

$$P_e(\Delta_{\tau_z}) := P\left(d_{\tau_z}^2 > (\Delta_{\tau_z}(|a|/2)((2^{R'} - 1)/2^{R'} - \eta) - 1)^2\right).$$

It follows that, conditioned on increment in the error, until the next stopping time, the process will increase exponentially and hence

$$x_{\tau_{z+1}} = a^{\tau_{z+1} - \tau_z} (x_{\tau_z} + \sum_{t=0}^{\tau_{z+1} - \tau_z - 1} a^{-t-1} d_{t+\tau_z}).$$

We now show that, there exist $\psi > 0$, $|G| < \infty$ such that

$$E[\log(\Delta_{\tau_{z+1}}^2)|\Delta_{\tau_z}, h_{\tau_z}] \leq \log(\Delta_{\tau_z}^2) - \psi + G1_{(|\Delta_{\tau_z}| \leq F)} \quad (11)$$

Now, it follows that,

$$\begin{aligned} & E[\log(\Delta_{\tau_{z+1}}^2)|\Delta_{\tau_z}, h_{\tau_z}] \\ & \leq (1 - P_e(\Delta_{\tau_z}))(p) \left(2 \log\left(\frac{|a|}{2^{R'} - \eta}\right) + \log(\Delta_{\tau_z}^2)\right) \\ & \quad + \sum_{k=2}^{\infty} E[\log(\Delta_{\tau_{z+1}}^2)|\tau_{z+1} - \tau_z = k] P(\tau_{z+1} - \tau_z = k) \end{aligned} \quad (12)$$

We now bound the probability that $P(\tau_{z+1} - \tau_z \geq k)$. We first have a lemma.

Lemma 3.2: Let A, B, C, D be events in a probability space (Ω, \mathcal{F}, P) . It follows that,

$$\begin{aligned} & P\left((A \cup B) \cap (C \cup D)\right) \\ & \leq P((A \cup B)|C)P(C) + P((A \cup B)|D)P(D) \end{aligned} \quad (13)$$

We now apply this Lemma to obtain the following:

$$\begin{aligned} & P(\tau_{z+1} - \tau_z \geq k) \\ & = P\left(\bigcap_{s=1}^{k-1} (p_s = 0) \cup (|h_s| > 1)\right) \\ & = P\left(\bigcap_{s=1}^{k-1} (p_s = 0) \cup (|x_s| \geq |2^{R'-1}(|a| + \delta)^{s-1} \frac{|a|}{2^{R'} - \eta} \Delta_0|\right) \\ & \leq P\left(\bigcap_{s=1}^{k-1} (p_s = 0) \cup (|h_s| > 1) | p_{k-1} = 0\right) (1 - p) \\ & \quad + \left(\frac{E[\zeta^2]}{\left(\left(\frac{|a| + \delta}{|a|\right)^{k-1} |a| \left(\frac{2^{R'-1}}{2^{R'} - \eta} - 1/2\right) \Delta_0\right)^2}\right) \\ & = P(\tau_{z+1} - \tau_z \geq k - 1) (1 - p) \\ & \quad + \left(\frac{E[\zeta^2]}{\left(\left(\frac{|a| + \delta}{|a|\right)^{k-1} |a| \left(\frac{2^{R'-1}}{2^{R'} - \eta} - 1/2\right) \Delta_0\right)^2}\right) \end{aligned} \quad (14)$$

As such, we can bound the probability that $P(\tau_{z+1} - \tau_z = k)$. Here ζ is a random variable with variance

$$\sigma^2 = \frac{E[d^2]}{1 - |a|^{-2}}.$$

Let for $k \in \mathbb{Z}_+$

$$\Theta_k := P(\tau_{z+1} - \tau_z \geq k).$$

Then, we have that

$$\Theta_k \leq \Theta_{k-1} (1 - p) + \frac{E[\zeta^2]}{\kappa \Delta_0 \Lambda^{k-1}},$$

where $\Lambda > \frac{|a| + \delta}{|a|}$, with $\kappa > 0$. As such, we can obtain, recursively

$$\Theta_k \leq \Theta_1 (1 - p)^{k-1} + \frac{E[\zeta^2]}{\kappa \Delta_0^2} \left(\sum_{l=1}^{k-1} (1 - p)^{k-l-1} \Lambda^{-l}\right)$$

$$\begin{aligned}
&= (1-p)^{k-1} \left(\Theta_1 + \frac{E[\zeta^2]}{\kappa \Delta_0^2} \sum_{l=1}^{k-1} ((1-p)\Lambda)^{-l} \right) \\
&\leq (1-p)^{k-1} \left(\Theta_1 + \frac{E[\zeta^2]}{\kappa \Delta_0^2} \frac{(1-p)\Lambda^{-1}}{1 - ((1-p)\Lambda)^{-1}} \right),
\end{aligned}$$

with $\Theta_1 = 1$. We now show that

$$\Theta_k \geq (1-p)^{k-1}.$$

This follows since

$$P\left(\bigcap_{s=1}^{k-1} (p_s = 0) \cup (|h_s| > 1)\right) \geq P\left(\bigcap_{s=1}^{k-1} (p_s = 0)\right).$$

As such,

$$\begin{aligned}
&P(\tau_{z+1} - \tau_z = k) \\
&\leq (1-p)^{k-1} \left(\Theta_1 + \frac{E[\zeta^2]}{\kappa \Delta_0^2} \frac{(1-p)\Lambda^{-1}}{1 - ((1-p)\Lambda)^{-1}} \right) \\
&\quad - (1-p)^k
\end{aligned}$$

As $\Delta_0 \rightarrow \infty$, it follows that (since $\Theta_1 = 1$)

$$P(\tau_{z+1} - \tau_z = k) \rightarrow (1-p)^{k-1}(p)$$

As such, the probability distribution is asymptotically dominated by a geometrically distributed process, for large Δ_0 values.

Let us now use $V(\Delta_t) = \log(\Delta_t^2)$. We will then invoke Theorem 2.2. Now, it follows that,

$$\begin{aligned}
&E[\log(\Delta_{\tau_{z+1}}^2) | \Delta_{\tau_z}, h_{\tau_z}] \\
&\leq (1 - P_e(\Delta_{\tau_z}))(p) \left(2 \log\left(\frac{|a|}{2^{R'} - \eta}\right) + \log(\Delta_{\tau_z}^2) \right) \\
&\quad + \sum_{k=2}^{\infty} E[\log(\Delta_{\tau_{z+1}}^2)] P(\tau_{z+1} - \tau_z = k) \\
&= (1 - P_e(\Delta_{\tau_z}))(p) \left(2 \log\left(\frac{|a|}{2^{R'} - \eta}\right) + \log(\Delta_0^2) \right) \\
&\quad + (1-p) E[\log((|a| + \delta)^{2\tau_z}) + \log(\Delta_0^2/2^{2R'})] \\
&= (1 - P_e(\Delta_{\tau_z}))(p) \left(2 \log\left(\frac{|a|}{2^{R'} - \eta}\right) + \log(\Delta_0^2/2^{2R'}) \right) \\
&\quad + E[(\tau_{z+1} - \tau_z) \log((|a| + \delta)^2) + \log(\Delta_0^2/2^{2R'})]
\end{aligned} \tag{15}$$

Since

$$\sum_{k=1}^{\infty} (1-p)^{k-1} (k)p = 1/p,$$

for large Δ_{τ_z} values

$$\begin{aligned}
&E[\log(\Delta_{\tau_{z+1}}^2) | \Delta_{\tau_z}, h_{\tau_z}] \\
&\leq \frac{1}{p} \left(2 \log_2(|a| + \delta) - 2 \log(2^{R'} - \eta) + \log(\Delta_{\tau_z}^2) \right)
\end{aligned}$$

As such, for some arbitrarily small $\eta > 0$, we require:

$$\log(2^{R'} - \eta) > \frac{1}{p} \log_2(|a| + \delta),$$

for stochastic stability. \diamond

2) Proof of Theorem 3.2: Uniqueness and Irreducibility:

In our setting, (x_t, Δ_t) form the Markov chain, as was proved in Lemma 3.1. Let $\mathbb{I}_{z_0, A, B}$ be defined as

$$\{n \in \mathbb{Z}_+, n \geq \log_2(L')/s : \exists N_A, N_B, n = -N_A A + N_B B + z_0\}.$$

Under the conditions of the paper, $\mathbb{I}_{z_0, A, B}$ forms a communication class, where $z_0 = \log_2(\Delta_0)/s$ is the initial condition of the parameter for the quantizer: Since we have $\Delta_{t+1} = \bar{Q}(|\frac{x_t}{\Delta_t 2^{R'-1}}|, p_t) \Delta_t$, it follows that

$$\log_2(\Delta_{t+1})/s = \log_2(\bar{Q}(|\frac{x_t}{\Delta_t 2^{R'-1}}|))/s + \log_2(\Delta_t)/s,$$

is also an integer. Furthermore, since the source process x_t is Lebesgue-irreducible (as the system noise admits a probability density function with positive mass on every open set), and there is a uniform lower bound L' on bin-sizes, the error process takes values in any of the admissible quantizer bins with non-zero probability. Let the values taken by $\log_2(\bar{Q}(|\frac{x_t}{\Delta_t 2^{R'-1}}|))/s$ be $\{-A, B\}$. Consider two integers $k, l \geq \frac{\log_2(L')}{s}$. For all $l, k \in \mathbb{I}_{z_0, A, B}$, that is there exist $N_A, N_B \in \mathbb{Z}_+$ such that $l - k = -N_A A + N_B B$. Consider first the case where $k > \frac{\log_2(L')}{s} + N_A A$. We show that the probability of N_A occurrences of perfect zoom, and N_B occurrences of under-zoom phases is bounded away from zero. This set of occurrences includes the event that in the first N_A time stages perfect-zoom occurs and later, successively, N_B times under-zoom phase occurs. The probability of this event is lower bounded by

$$\left(P(d_t \in [-2^{sl} - L', -2^{sl} + L']) \right)^{N_A} (1-p)^{N_B} > 0.$$

A similar analysis can be performed when $k < \frac{\log_2(L')}{s} + N_A A$, by considering the opposite order of events, where in the first N_B times, under-zoom occurs, and in the successive N_A time stages, perfect-zoom occurs. As such, for any two integers k, l and for some $p > 0$, $P(\log_2(\Delta_{t+p}) = ls | \log_2(\Delta_t) = ks) > 0$. In our setting, the recurrent set consists of both Δ_t and e_t . Hence, in this setting, the product space is the product of a discrete-space and a real space. For x_t , the Borel sets are the open intervals, which are visited infinitely often, and for Δ_t , the individual *atoms* are visited infinitely often with probability 1.

Now, we can connect the results of the previous section with Theorems 3.1 and 2.2. The recurrent set $C'_x \times C_h$ is ν -petite, for some probability measure ν as any Borel set in the state space is visited starting from $C'_x \times C_h$, and the chain is irreducible. These two imply that the chain is positive Harris recurrent.

If the integers are relatively prime, the invariant set for the quantizer parameter \mathbb{I} will be the entire set, and invariant of the initial condition. \diamond

D. Simulation

As a simulation study, we consider a linear system with the following dynamics:

$$x_{t+1} = ax_t + u_t + d_t,$$

with $a = 2.5$, $E[d_t] = 0$, $E[d_t^2] = 1$, and $\{d_t\}$ are i.i.d. Gaussian variables.

The erasure channel has packet loss probability $1 - p = 0.1$.

For stochastic stability, we use the zooming quantizer with rate $\log_2(\lceil |a|^{1/p} \rceil + 1) = 4$. We have taken $L' = 1$.

For finite second moment stability, we use the zooming quantizer with rate $\log_2(\lceil \sqrt{\frac{p}{a^2 - (1-p)}} \rceil + 1) = 5$. We have taken $L' = 1$.

The plots below (Figures 3 and 4) corroborate the stochastic stability result, by explicitly showing the under-zoomed and perfectly-zoomed phases, with the peaks in the plots showing the under-zoom phases.

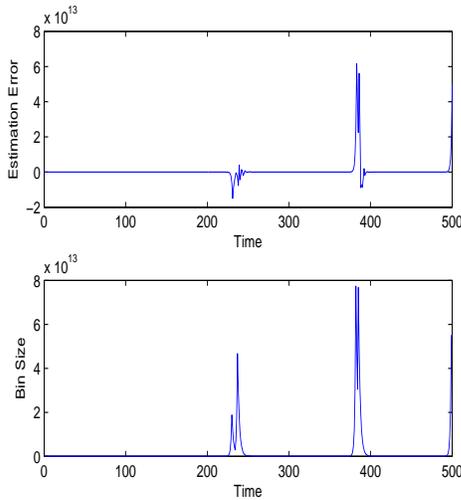


Fig. 3: Sample path for stochastically stable system.

IV. CONCLUDING REMARKS

We presented a stochastic stability result which characterizes the stability of a sampled Markov chain, where the sample times are state-dependent and random. We think that this result will have many applications in networked control systems.

We applied this result to the problem of control over an erasure channel and obtained new results for the case when the system is driven by noise with unbounded support for its probability measure.

One related question is the following: Is Shannon Capacity sufficient for stochastic stability for an arbitrary DMC with feedback?. Earlier, it was observed by Sahai and Mitter that Shannon capacity was not adequate. We had observed in [23] that, for there to be an invariant distribution, the capacity should be at least greater than the logarithm of the unstable eigenvalue. This paper showed that for erasure channels, for stochastic stability, Shannon capacity is sufficient; however, for

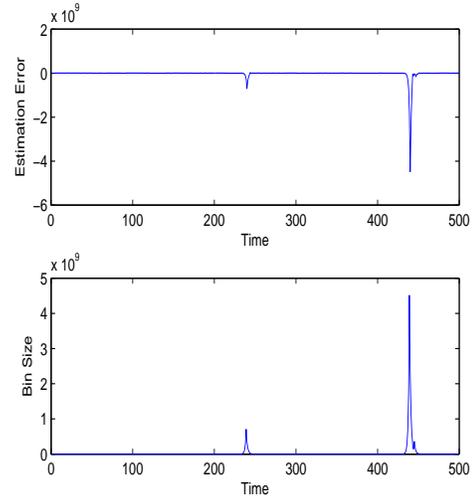


Fig. 4: Finite second moment requires more information.

finite moment conditions more stringent criteria are needed. It is not easy to extend this result to arbitrary channels, since one loses Markoviannes. Nonetheless, recurrence seems to be a property that one can verify. \diamond

REFERENCES

- [1] V. S. Borkar, S. K. Mitter, and S. Tatikonda, "Optimal sequential vector quantization of Markov sources," *SIAM J. Control and Optimization*, vol. 40, pp. 135-148, 2001.
- [2] R. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems", *IEEE Trans. on Automatic Control*, vol. 45, pp. 1279-1289, July 2000.
- [3] D. J. Goodman and A. Gersho, "Theory of an adaptive quantizer", *IEEE Trans. Commun.*, vol. 22, pp. 1037- 1045, Aug. 1974.
- [4] J. C. Kieffer, "Stochastic stability for feedback quantization schemes", *IEEE Trans. Inform. Theory*, vol. 28, pp. 248-254, March 1982.
- [5] J. C. Kieffer and J. G. Dunham, "On a type of stochastic stability for a class of encoding schemes", *IEEE Trans. Inform. Theory*, vol. 29, pp. 793-797, November 1983.
- [6] D. Liberzon, "On stabilization of linear systems with limited information", *IEEE Trans. on Automatic Control*, vol. 48, pp. 304-307, Feb 2003.
- [7] D. Liberzon and D. Nesić, "Input-to-state stabilization of linear systems with quantized state measurements", *IEEE Trans. on Automatic Control*, vol. 52, no. 5, pp. 767-781, May 2007.
- [8] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates", *SIAM Journal on Control and Optimization*, vol. 43, pp. 413 - 436, July 2004.
- [9] S. P. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, Springer Verlag, London (1993).
- [10] S. P. Meyn and R. Tweedie, "Stability of Markovian processes I: Criteria for discrete-time chains", *Adv. Appl. Probability*, vol. 24, pp. 542-574, 1992.
- [11] S. P. Meyn and R. Tweedie, "State-dependent criteria for convergence of Markov chains," *Ann. Appl. Probab.* 4:149-168, 1994.
- [12] A. S. Matveev and A. V. Savkin, *Estimation and Control over Communication Networks*, Birkhauser Boston, 2008.
- [13] P. Minero, M. Franceschetti, S. Dey and G. Nair, "Data rate theorem for stabilization over time-varying feedback channels", *IEEE Trans. on Automatic Control*, vol. 54, pp. 243-255, February 2009.
- [14] R. L. Tweedie, "Invariant measures for Markov chains with no irreducibility assumptions", in *A Celebration of Applied Probability*, *J. Appl. Prob.*, Spec. Vol. 25A pp. 275-285, 1988.
- [15] E. Nummelin, *General irreducible Markov chains and non-negative operators*.

- [16] J. B. Lasserre, "Invariant probabilities for Markov chains on a metric space", *Statistics and Probability Letters*, vol. 34, pp. 259-265, 1997.
- [17] S. Yüksel, "Stochastic stability of adaptive quantizers for Markov sources," *Proc. International Symposium on Information Theory*, June 2009.
- [18] O. C. Imer, S. Yüksel, and T. Başar, "Optimal Control of LTI Systems over Communication Networks", *Automatica*, vol. 42, pp. 1429–1440, Sept. 2006.
- [19] N. C. Martins, M. A. Dahleh and N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link", *IEEE Transactions on Automatic Control*, volume 51, pp. 438–447, March 2006.
- [20] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Aut. Control*, vol. 49, pp. 1056-1068, July 2004.
- [21] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints - part II: Stabilization with limited information feedback, *IEEE Trans. Aut. Control*, vol. 42, pp.1294–1299, Sept. 1997.
- [22] A. Sahai and S. 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, pp. 3369–3395, Aug. 2006.
- [23] S. Yüksel and T. Başar, "Control over noisy forward and reverse channels", *IEEE Trans. Automatic Control*, under review.