Random-Time, State-Dependent Stochastic Drift for Markov Chains and Application to Stochastic Stabilization Over Erasure Channels

Serdar Yüksel, *Member, IEEE*, and Sean P. Meyn, *Fellow, IEEE*

Abstract—It is known that state-dependent, multi-step Lyapunov bounds lead to greatly simplified verification theorems for stability for large classes of Markov chain models. This is one component of the "fluid model" approach to stability of stochastic networks. In this paper we extend the general theory to randomized multi-step Lyapunov theory to obtain criteria for stability and steady-state performance bounds, such as finite moments. These results are applied to a remote stabilization problem, in which a controller receives measurements from an erasure channel with limited capacity. Based on the general results in the paper it is shown that stability of the closed loop system is assured provided that the channel capacity is greater than the logarithm of the unstable eigenvalue, plus an additional correction term. The existence of a finite second moment in steady-state is established under additional conditions.

Index Terms—Information theory, Markov chain Monte-Carlo (MCMC), Markov processes, networked control systems, stochastic stability.

I. INTRODUCTION

STOCHASTIC stability of Markov chains has a rich and complete theory, and forms a foundation for several other general techniques such as dynamic programming and Markov Chain Monte-Carlo (MCMC) [24]. This paper concerns extensions and application of a class of Lyapunov techniques, known as *state-dependent* drift criteria [27]. This technique is the basis of the fluid-model (or ODE) approach to stability in stochastic networks and other general models [3], [9], [10], [12], [13], [15], [22], [24].

In this paper we consider a stability criterion based on a state-dependent *random sampling* of the Markov chain of the following form: It is assumed that there is a function V on the state space taking positive values, and an increasing sequence of stopping times $\{\mathcal{T}_i : i \in \mathbb{N}_+\}$, with $\mathcal{T}_0 = 0$, such that for each i,

$$\mathsf{E}\left[V\left(x_{\mathcal{T}_{i+1}}\right) | \mathcal{F}_{\mathcal{T}_i}\right] \leq V\left(x_{\mathcal{T}_i}\right) - \delta\left(x_{\mathcal{T}_i}\right) \qquad (1)$$

where the function $\delta : \mathsf{X} \to \mathbb{R}$ is positive (bounded away from zero) outside of a "small set", and $\mathcal{F}_{\mathcal{T}_i}$ denotes the filtration of "events up to time $\mathcal{T}_i$". Under suitable conditions on the Markov chain, the drift δ, and the sequence $\{\mathcal{T}_i\}$, we establish corresponding stability and ergodicity properties of the chain. The main results extend and unify previous research on stability and criteria for finite moments obtained in [10], [15], [22], [24], [27].

Motivation for this research arose from our interest in applications to networked control, and information theory with variable length and variable delay decoding [5], [33], [34], and non-asymptotic information theory [31]. Specifically, in some network protocol, team decision and networked control applications there is only intermittent access to sensor information, control action or some common knowledge on which decisions can be made. The timing may be random, depending on availability of communication resources. One important example is the event-triggered control setup considered in [2], [32], [35], among others.

One particular example in networked control under information constraints is reported in [40], for establishing stochastic stability of adaptive quantizers for Markov sources where random stopping times are the instances when the encoder can transmit granular information to a controller. We will also consider such an application in detail in the paper. In this context, there has been a significant amount of research on stochastic stabilization of networked control systems under information constraints. For a detailed review see [36] and [23]. Stochastic stability of adaptive quantizers have been studied both in the information theory community (see [16], [18], [19]) as well as control community ([4], [21], [30], [40]). [30] provided a stability result under the assumption that a quantizer is variable-rate for systems driven by noise with unbounded support for its probability measure. [30] 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 time stages, the quantizer is not used. For such linear systems driven by unbounded noise, [40] established ergodicity, under fixed-rate constraints, through martingale methods. These papers motivated us to develop a more general theory for both random-time drift as well as the consideration of more general noise models on the channels. In a similar line of work, [17] also considered stability of the state and quantization parameters, [28] studied the problem concerning time-varying channels and provided a necessity and sufficiency result for boundedness of second moments, [39] studied the problem of control over an erasure channel in the absence of noise, and [43] considered discrete noisy channels with

noiseless feedback for systems driven by Gaussian noise using the machinery developed in the current paper. [41] obtained conditions for the existence of an invariant probability measure for noisy channels, considering deterministic, state-dependent drift, based on the criteria developed in [27].

We believe our results will provide constructive tools to address related stability issues in a large class of networked control problems.

The contributions of this paper can be summarized as follows
- Stochastic stability theory for Markov chains based on random-time, state-dependent stochastic drift criteria. A range of conditions are used to establish conditions for positive recurrence, and existence of finite moments.
- The results are applied to stochastic stabilization over an erasure network, where a linear system driven by Gaussian noise is controlled. This paper establishes that in such an application of stabilization of an unstable system driven by Gaussian noise over erasure channels, for stochastic stability, Shannon capacity is sufficient (up to an additional correction term). For the existence of finite moments, however, more stringent criteria are needed. Regarding information rate requirements, our construction is tight up to an additional symbol, in comparison with necessary conditions presented in [28].

The remainder of the paper is organized as follows. The implications of the drift criterion (1) to various forms of stochastic stability are presented in the next section. The rest of the paper focuses on an application to control over a lossy erasure network with quantized observations. The paper ends with concluding remarks in Section IV. Proofs of the stochastic stability results and other technical results are contained in the Appendix.

II. STOCHASTIC STABILITY

A. Preliminaries

We let $\boldsymbol{\phi} = \{\phi_t, t \geq 0\}$ denote a Markov chain with state space X. The basic assumptions of [25] are adopted: It is assumed that X is a complete separable metric space, that is locally compact; its Borel σ-field is denoted $\mathcal{B}(\mathsf{X})$. The transition probability is denoted by P, so that for any $\phi \in \mathsf{X}$, $A \in \mathcal{B}(\mathsf{X})$, the probability of moving in one step from the state ϕ to the set A is given by $\mathsf{P}(\phi_{t+1} \in A | \phi_t = \phi) = P(\phi, A)$. The n-step transitions are obtained via composition in the usual way, $\mathsf{P}(\phi_{t+n} \in A | \phi_t = \phi) = P^n(\phi, A)$, for any $n \geq 1$. The transition law acts on measurable functions $f : \mathsf{X} \to \mathbb{R}$ and measures μ on $\mathcal{B}(\mathsf{X})$ via,

$$Pf(\phi) := \int_{\mathsf{X}} P(\phi, dy)f(y), \quad \phi \in \mathsf{X}$$

and

$$\mu P(A) := \int_{\mathsf{X}} \mu(d\phi)P(\phi, A), \quad A \in \mathcal{B}(\mathsf{X}).$$

A probability measure π on $\mathcal{B}(\mathsf{X})$ is called invariant if $\pi P = \pi$. That is,

$$\int \pi(d\phi)P(\phi, A) = \pi(A), \quad A \in \mathcal{B}(\mathsf{X}).$$

For any initial probability measure ν on $\mathcal{B}(\mathsf{X})$ we can construct a stochastic process with transition law P, and satisfying $\phi_0 \sim \nu$. We let P_ν denote the resulting probability measure on sample space, with the usual convention for $\nu = \delta_\phi$ when the initial state is $\phi \in \mathsf{X}$. When $\nu = \pi$ then the resulting process is stationary.

There is at most one stationary solution under the following irreducibility assumption. For a set $A \in \mathcal{B}(\mathsf{X})$ we denote,

$$\tau_A := \min(t \geq 1 : \phi_t \in A) \tag{2}$$

Definition 2.1: Let φ denote a sigma-finite measure on $\mathcal{B}(\mathsf{X})$.
(i) The Markov chain is called φ-*irreducible* if for any $\phi \in \mathsf{X}$, and any $B \in \mathcal{B}(\mathsf{X})$ satisfying $\varphi(B) > 0$, we have

$$\mathsf{P}_\phi\{\tau_B < \infty\} > 0.$$

(ii) A φ-irreducible Markov chain is *aperiodic* if for any $\phi \in \mathsf{X}$, and any $B \in \mathcal{B}(\mathsf{X})$ satisfying $\varphi(B) > 0$, there exists $n_0 = n_0(\phi, B)$ such that

$$P^n(\phi, B) > 0 \qquad for\ all \quad n \geq n_0.$$

(iii) A φ-irreducible Markov chain is *Harris recurrent* if $\mathsf{P}_\phi(\tau_B < \infty) = 1$ for any $\phi \in \mathsf{X}$, and any $B \in \mathcal{B}(\mathsf{X})$ satisfying $\varphi(B) > 0$. It is *positive Harris recurrent* if in addition there is an invariant probability measure π.

Intimately tied to φ-irreducibility is the existence of a suitably rich collection of "small sets", which allows Nummelin's splitting technique to be applied leading to verification for Harris recurrence. A set $A \in \mathcal{B}(\mathsf{X})$ is small if there is an integer $n_0 \geq 1$ and a positive measure μ satisfying $\mu(\mathsf{X}) > 0$ and

$$P^{n_0}(x, B) \geq \mu(B), \quad for\ all \quad x \in A, \text{ and } B \in \mathcal{B}(\mathsf{X}).$$

Small sets are analogous to compact sets in the stability theory for φ-irreducible Markov chains. In most applications of φ-irreducible Markov chains we find that any compact set is small—In this case, $\boldsymbol{\phi}$ is called a *T-chain* [25].

To relax the φ-irreducibility assumption we can impose instead the following continuity assumption: A Markov chain is (weak) Feller if the function Pf is continuous on X, for every continuous and bounded function $f : \mathsf{X} \to \mathbb{R}$.

We next introduce criteria for positive Harris recurrence for φ-irreducible Markov chains, and criteria for the existence of a steady-state distribution π for a Markov chain satisfying the Feller property.

B. Drift Criteria for Positivity

We now consider specific formulations of the random-time drift criterion (1). Throughout the paper the sequence of stopping times $\{\mathcal{T}_i : i \in \mathbb{N}_+\}$ is assumed to be non-decreasing, with $\mathcal{T}_0 = 0$. In prior work on state-dependent criteria for stability it is assumed that the stopping times take the following form,

$$\mathcal{T}_{i+1} = \mathcal{T}_i + n(\phi(\mathcal{T}_i)), \qquad i \geq 0$$

where $n : \mathsf{X} \to \mathbb{N}$ is a deterministic function of the state. The results that follow generalize state dependent drift results in [27] to this random-time setting. We note that a similar approach has been presented recently in the literature in [14] for random-time drift (see Theorem 4), which readily generalizes the state dependent drift results in [27]. The conditions presented in [14]

are more restrictive for the stopping times than what we present here. Furthermore we present discussions for existence of finite moments, as well as extensions for non-irreducible chains.

The proofs of these results are presented in the Appendix.

Theorem 2.1 is the main general result of the paper, providing a single criterion for positive Harris recurrence, as well as finite "moments" (the steady-state mean of the function f appearing in the drift condition (3)). The drift condition (3) is a refinement of (1).

Theorem 2.1: Suppose that ϕ is a φ-irreducible and aperiodic Markov chain. Suppose moreover that there are functions $V : \mathsf{X} \to (0, \infty)$, $\delta : \mathsf{X} \to [1, \infty)$, $f : \mathsf{X} \to [1, \infty)$, a small set C, and a constant $b \in \mathbb{R}$, such that the following hold:

$$\mathsf{E}\left[V\left(\phi_{\mathcal{T}_{z+1}}\right) | \mathcal{F}_{\mathcal{T}_z}\right] \leq V\left(\phi_{\mathcal{T}_z}\right) - \delta\left(\phi_{\mathcal{T}_z}\right) + b1_{\{\phi_{\mathcal{T}_z} \in C\}}$$

$$\mathsf{E}\left[\sum_{k=\mathcal{T}_z}^{\mathcal{T}_{z+1}-1} f(\phi_k) | \mathcal{F}_{\mathcal{T}_z}\right] \leq \delta\left(\phi_{\mathcal{T}_z}\right), \quad z \geq 0. \tag{3}$$

Then the following hold:
  (i) ϕ is positive Harris recurrent, with unique invariant distribution π
  (ii) $\pi(f) := \int f(\phi)\pi(d\phi) < \infty$
  (iii) For any function g that is bounded by f, in the sense that $\sup_\phi |g(\phi)|/f(\phi) < \infty$, we have convergence of moments in the mean, and the Law of Large Numbers holds:
$$\lim_{t \to \infty} \mathsf{E}_\phi\left[g(\phi_t)\right] = \pi(g)$$
$$\lim_{N \to \infty} \frac{1}{N} \sum_{t=0}^{N-1} g(\phi_t) = \pi(g) \quad a.s., \ \phi \in \mathsf{X}$$
$\hfill \square$

Remark 2.1: We note that, for (ii) in Theorem 2.1, the condition that $f : \mathsf{X} \to [1, \infty)$, $\delta : \mathsf{X} \to [1, \infty)$, can be relaxed to $f : \mathsf{X} \to [0, \infty)$, $\delta : \mathsf{X} \to [0, \infty)$, provided that one can establish (i), that is the positive Harris recurrence of the chain first.

This result has a corresponding, albeit weaker, statement for a Markov chain that is Feller, but not necessarily satisfying the irreducibility assumptions:

Theorem 2.2: Suppose that ϕ is a Feller Markov chain, not necessarily φ-irreducible. If in addition (3) holds with C compact, then there exists at least one invariant probability measure. Moreover, there exists $c < \infty$ such that, under any invariant probability measure π,

$$\mathsf{E}_\pi\left[f(\phi_t)\right] = \int_{\mathsf{X}} \pi(d\phi)f(\phi) \leq c. \tag{4}$$
$\hfill \square$

We conclude by stating a simple corollary to Theorem 2.1, obtained by taking $f(\phi) = 1$ for all $\phi \in \mathsf{X}$.

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

$$\mathsf{E}\left[V\left(\phi_{\mathcal{T}_{z+1}}\right) | \mathcal{F}_{\mathcal{T}_z}\right] \leq V\left(\phi_{\mathcal{T}_z}\right) - 1 + b1_{\{\phi_{\mathcal{T}_z} \in C\}}$$
$$\sup_{z \geq 0} \mathsf{E}\left[\mathcal{T}_{z+1} - \mathcal{T}_z | \mathcal{F}_{\mathcal{T}_z}\right] < \infty. \tag{5}$$

Then ϕ is positive Harris recurrent. $\hfill \square$
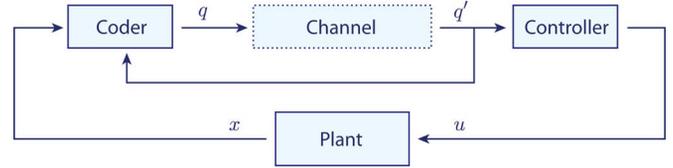
Fig. 1. Control over a discrete erasure channel with feedback. Coder represents the quantizer and the encoder.

III. APPLICATION TO STOCHASTIC STABILIZATION OVER AN ERASURE CHANNEL

The results of the previous section are now applied to a remote stabilization problem, in which the plant is open-loop unstable, and the controller has access to measurements from an erasure channel—see Fig. 1.

We begin with a scalar model; extensions to the multivariate setting are contained in Section III-E.

A. Scalar Control/Communication Model

Consider a scalar LTI discrete-time system described by

$$x_{t+1} = ax_t + bu_t + d_t, \qquad t \geq 0 \tag{6}$$

Here x_t is the state at time t, u_t is the control input, x_0 is a given initial condition, and $\{d_t\}$ is a sequence of zero-mean independent, identically distributed (i.i.d.) Gaussian random variables. It is assumed that $|a| \geq 1$ and $b \neq 0$: The system is open-loop unstable, but it is stabilizable.

This system is connected over an erasure channel with finite capacity to a controller, as shown in Fig. 1. The controller has access to the information it has received through the channel. The controller estimates the state and then applies its control. We will establish bounds on data rates which lead to various versions of stochastic stability for the closed loop system.

The details of the communication channel are specified as follows: The channel source consists of state values, taking values in $\mathbb{R}$. The source is *quantized*: The quantizer, at time $t \geq 0$, is represented by a map $Q_t : \mathbb{R} \to \mathbb{R}$, characterized by a sequence of non-overlapping intervals $\mathcal{P}_t := \{\mathcal{B}_{i,t}\}$, with $|\mathcal{P}| = K + 1$, such that $Q_t(x) = q_t^i$ if and only if $x \in \mathcal{B}_{i,t}$; that is,

$$Q_t(x) = \sum_i q_t^i \times 1_{\{x \in \mathcal{B}_{i,t}\}}.$$

The quantizer outputs are transmitted through a memoryless erasure channel, after being subjected to a bijective mapping, which is performed by the channel encoder: The channel encoder maps the quantizer output symbols to corresponding channel inputs $q \in \mathcal{M} := \{1, 2 \ldots, K + 1\}$. An encoder at time t, denoted by $\mathcal{E}_t$, maps the quantizer outputs to $\mathcal{M}$ such that $\mathcal{E}_t(Q_t(x_t)) = q_t \in \mathcal{M}$.

The controller/decoder has access to noisy versions of the coder outputs for each time, which we denote by $\{q'\} \in \mathcal{M} \cup \{e\}$, with e denoting the erasure symbol, generated according to a probability distribution for every fixed $q \in \mathcal{M}$. The channel transition probabilities are given by:

$$P(q' = i | q = i) = p, \quad P(q' = e | q = i) = 1 - p, \quad i \in \mathcal{M}.$$

For each time $t \geq 0$, the controller/decoder applies a mapping $\mathcal{D}_t : \mathcal{M} \cup \{e\} \to \mathbb{R}$, given by:

$$\mathcal{D}_t(q_t') = \mathcal{E}_t^{-1}(q_t') \times 1_{\{q_t' \neq e\}} + 0 \times 1_{\{q_t' = e\}}$$

We restrict the analysis to a class of uniform quantizers, defined by two parameters: bin size $\Delta > 0$, and an even number $K \geq 2$. The set $\mathcal{M}$ consists of $K + 1$ elements. The uniform quantizer map is defined as follows: For $k = 1, 2 \ldots, K$, as shown in the equation at bottom of the page.

We consider quantizers that are *adaptive*, so that the bin process can vary with time. The bin size Δ_t at time t is assumed to be a function of the previous value Δ_{t-1} and the past channel output q_{t-1}'.

B. Stochastic Stabilization Over an Erasure Channel

Consider the following time-invariant model. Let Υ_t denote an i.i.d. binary sequence of random variables, representing the erasure process in the channel, where the event $\Upsilon_t = 1$ indicates that the signal is transmitted with no error through the erasure channel at time t. Let $p = \mathsf{E}[\Upsilon_t]$ denote the probability of success in transmission.

The following key assumptions are imposed throughout this section: Given $K \geq 2$ introduced in the definition of the quantizer, define the *rate variables*

$$R := \log_2(K + 1) \qquad R' = \log_2(K) \qquad (7)$$

We fix positive scalars δ, α satisfying $\alpha < 1$, and

$$\alpha > |a|2^{-R'}, \qquad (8)$$

and

$$\alpha \left(|a| + \delta\right)^{p^{-1} - 1} < 1. \qquad (9)$$

We note that the (Shannon) capacity of such an erasure channel is given by $\log_2(K + 1)p$ [8]. From (7)–(9) it follows that if $\log_2(K)p > \log_2(|a|)$, then α, δ exist such that the above are satisfied.

To define the bin-size update rule we require another constant $L > 0$, chosen so that $L' =: \alpha L \geq 1$, where we take the lower bound as 1 for convenience; any positive number would suffice.

Define the mapping $H : \mathbb{R} \times \mathbb{R} \times \{0, 1\} \to \mathbb{R}$,

$$
\begin{aligned}
H(\Delta, h, p) &= |a| + \delta && \text{if} \quad |h| > 1, \quad \text{or} \quad p = 0 \\
H(\Delta, h, p) &= \alpha && \text{if} \quad 0 \leq |h| \leq 1, p = 1, \Delta \geq L \\
H(\Delta, h, p) &= 1 && \text{if} \quad 0 \leq |h| \leq 1, p = 1, \Delta < L
\end{aligned}
$$

Then with $\Delta_0 \geq L$ selected otherwise arbitrarily, define

$$
\begin{aligned}
u_t &= -\frac{a}{b}\hat{x}_t, \\
\hat{x}_t &= \mathcal{D}_t(q_t') = \Upsilon_t Q_K^{\Delta_t}(x_t), \\
\Delta_{t+1} &= \Delta_t H(\Delta_t, |h_t|, \Upsilon_t), \quad \text{where } h_t = \frac{x_t}{\Delta_t 2^{R'-1}} \quad (10)
\end{aligned}
$$

The update equations above imply that

$$\Delta_t \geq L\alpha = L' \geq 1. \qquad (11)$$

Given the channel output $q_t' \neq e$, the controller can deduce the realization of Υ_t and the event $\{|h_t| > 1\}$ simultaneously. This is due to the observation that if the channel output is not the erasure symbol, the controller knows that the signal is received with no error. If $q_t' = e$, then the controller applies 0 as its control input and enlarges the bin size of the quantizer.

Lemma 3.1: Under (10), the process (x_t, Δ_t) is a Markov chain.

Proof: The system state dynamics can be expressed $x_{t+1} = ax_t - a\hat{x}_t + d_t$, where $\hat{x}_t = \Upsilon_t Q_K^{\Delta_t}(x_t)$. It follows that the pair process (x_t, Δ_t) evolves as a nonlinear state space model,

$$
\begin{aligned}
x_{t+1} &= a\left(x_t - \Upsilon_t Q_K^{\Delta_t}(x_t)\right) + d_t \\
\Delta_{t+1} &= \Delta_t H\left(\Delta_t, \left|\frac{x_t}{2^{R'-1}\Delta_t}\right|, \Upsilon_t\right), \qquad (12)
\end{aligned}
$$

in which (d_t, Υ_t) is i.i.d.. Thus, (x_t, Δ_t) form a Markov chain (see [25, Ch. 2]). $\square$

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

Theorem 3.1: For an adaptive quantizer satisfying (7)–(9), suppose that the quantizer bin sizes are such that their base-2 logarithms are integer multiples of some scalar s, and $\log_2(H(\cdot))$ takes values in integer multiples of s. Then the process (x_t, Δ_t) forms a positive Harris recurrent Markov chain, with a unique invariant probability measure π. If the integers taken are relatively prime (that is they share no common divisors except for 1), then the invariant probability measure is independent of the value of the integer multiplying s. $\square$

Under slightly stronger conditions we obtain a finite second moment:

Theorem 3.2: Suppose that the assumptions of Theorem 3.1 hold, and in addition we have the bound

$$a^2 \left(1 - p + \frac{p}{(2^R - 1)^2}\right) < 1. \qquad (13)$$

It then follows that for each initial condition (x_0, Δ_0),

$$\lim_{t \to \infty} \mathsf{E}\left[x_t^2\right] = \mathsf{E}_\pi\left[x_0^2\right] < \infty.$$

$\square$

Remark 3.1: We note that Minero *et al.* [28], in Theorem 4.1, observed that a necessary condition for mean square stability is that the following holds:

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

Thus, our sufficiency proof almost meets this bound except for an additional transmitted symbol.

$$
Q_K^\Delta(x) = \begin{cases} \left(k - \frac{1}{2}(K + 1)\right)\Delta, & \text{if} \quad x \in \left[\left(k - 1 - \frac{1}{2}K\right)\Delta, \left(k - \frac{1}{2}K\right)\Delta\right) \\ \left(\frac{1}{2}(K - 1)\right)\Delta, & \text{if} \quad x = \frac{1}{2}K\Delta \\ 0, & \text{if} \quad x \notin \left[-\frac{1}{2}K\Delta, \frac{1}{2}K\Delta\right]. \end{cases}
$$
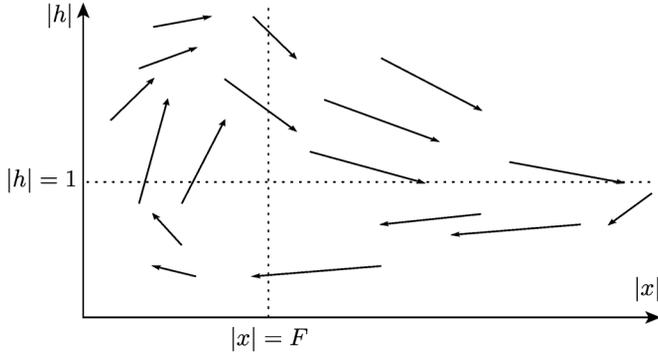
Fig. 2. Drift in the Markov Process. When under-zoomed, the error increases on average and the quantizer *zooms out*; when perfectly-zoomed, the error decreases and the quantizer *zooms in*.

We now consider the mth moment case. This moment can become useful for studying multi-dimensional systems for a sequential analysis of the modes, as we briefly discuss in Section III-E.

Theorem 3.3: Consider the scalar system in (6). Let $m \in \mathbb{N}$, suppose that the assumptions of Theorem 3.1 hold, and in addition we have the inequality,

$$|a|^m \left(1 - p + \frac{p}{(2^R - 1)^m} \right) < 1.$$

It then follows that with the adaptive quantization policy considered, $\lim_{t \to \infty} \mathsf{E}[|x_t|^m] = \mathsf{E}_\pi[|x_0|^m] < \infty$. $\square$

C. Connections With the Drift Criteria and the Proof Program

Stability of the control/communication model is established using the random-time stochastic drift criteria presented in the previous section, applied to the Markov chain $\boldsymbol{\phi} = (x_t, \Delta_t)$ (see Lemma 3.1). We provide an overview here, and the details can be found in the Appendix.

Fig. 2 provides some intuition on the construction of stopping times and the Lyapunov functions. Recall that $h_t = x_t / (2^{R'-1} \Delta_t)$ was introduced in (10). The arrows shown in the figure denote the mean one-step increments of (x_t, h_t): That is, the arrow ν with base at (x, h) is defined by,
$$\nu = \mathsf{E}\left[(x_{t+1}, h_{t+1}) - (x_t, h_t) | (x_t, h_t) = (x, h) \right]$$

With $F > 0$ fixed, and with $F' = F2^{-(R'-1)}$, two sets are used to define the small set in the drift criteria, $C_x = \{x : |x| \leq F\}$ for $F > 0$, and $C_\Delta = \{\Delta : \Delta \leq F'\}$. Denote $C_h = \{h : |h| \leq 1\}$, and assume that $F > 0$ is chosen sufficiently large so that $(x_t, \Delta_t) \in C_x \times C_\Delta$ whenever $(x_t, h_t) \in C_x \times C_h$. When x_t is outside C_x and h_t outside C_h (the *under-zoomed phase* of the quantizer), there is a drift for h_t towards C_h. When the process x_t reaches C_h (the *perfectly-zoomed phase* of the quantizer), then the process drifts towards C_x.

We next construct the sequence of stopping times required in the drift criteria of Section II-B. The controller can receive meaningful information regarding the state of the system when two events occur concurrently: the channel carries information with no error, and the source lies in the granular region of the quantizer: That is, $x_t \in [-(1/2)K\Delta_t, (1/2)K\Delta_t]$ (or $|h_t| \leq 1$) and $\Upsilon_t = 1$. The stopping times are taken to be the times at

which both of these events occur. With, $|h_0| \leq 1$, $\Upsilon_0 = 1$, we define $\mathcal{T}_0 = 0$ and

$$\mathcal{T}_{z+1} = \inf \{ k > \mathcal{T}_z : |h_k| \leq 1, \Upsilon_k = 1 \}, \quad z \in \mathbb{N}.$$

These stopping times are nearly geometric when the bin size is large. The proof of Proposition 3.1 is presented in Section V-B1.

Proposition 3.1: The discrete probability measure $\mathsf{P}(\mathcal{T}_{z+1} - \mathcal{T}_z = k | x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z})$ satisfies,

$$(1-p)^{k-1} \leq \mathsf{P}\left(\mathcal{T}_{z+1} - \mathcal{T}_z \geq k | x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z} \right) \leq (1-p)^{k-1} + o(1),$$

where $o(1) \to 0$ as $\Delta_{\mathcal{T}_z} \to \infty$ uniformly in $x_{\mathcal{T}_z}$. $\square$

The next step is to establish irreducibility structure. The proof of the following is contained in Section V-B2.

Proposition 3.2: Under the assumptions of Theorem 3.1, the chain (x_t, Δ_t) is φ-irreducible for some φ, it is aperiodic, and all compact sets are small. $\square$

We now provide sketches of the proofs of the main results. The details are collected together in the Appendix.

Sketch of Proof of Theorem 3.1: The logarithmic function $V_0(x_t, \Delta_t) = \log(\Delta^2) + B_0$ for some $B_0 > 0$ serves as the Lyapunov function in (3), with $f(x, \Delta)$ set as a constant. Note that by (11), $V_0(x_t, \Delta_t) > 0$.

Together with Propositions 3.1 and 3.2, we apply Theorem 2.1 in the special case of Corollary 2.1. Proposition 3.2 implies the existence of a unique invariant measure. Details of the proof are presented in Section V-B3. $\square$

Sketch of Proof of Theorem 3.2: A quadratic Lyapunov function $V_2(x_t, \Delta_t) = \Delta_t^2$ is used, along with $\delta(x_t, \Delta_t) = \epsilon \Delta_t^2$ for some $\epsilon > 0$, and $f(x_t, \Delta_t) = \xi x_t^2$ for some $\xi > 0$. The bound (3) is established in Section V-B4, so that the limit $\lim_{t \to \infty} \mathsf{E}[x_t^2]$ exists and is finite by Theorem 2.1. Details of the proof are presented in Section V-B4. $\square$

Sketch of Proof of Theorem 3.3: Theorem 2.1 is applied with the Lyapunov function $V_m(x_t, \Delta_t) = \Delta_t^m$, $\delta(x_t, \Delta_t) = \epsilon \Delta_t^m$ for some $\epsilon > 0$, and $f(x_t, \Delta_t) = \xi |x_t|^m$ for some $\xi > 0$. Details of the proof are presented in Section V-B4. $\square$

D. Simulation

Consider a linear system

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

with $a = 2.5$, $\{d_t\}$ is an i.i.d. $N(0, 1)$ Gaussian sequence. The erasure channel has erasure probability $1 - p = 0.1$. For stability with a finite second moment, we employ a quantizer with rate

$$\log_2 \left(\left\lceil \sqrt{\frac{p}{\frac{1}{a^2} - (1-p)}} \right\rceil + 1 \right) = \log_2(5)$$

bits. That is, a uniform quantizer with 5 bins. We have taken $L' = 1$. Figs. 3 and 4 illustrate the conclusions of the stochastic stability results presented in Theorems 3.1 and 3.2. The plots show the under-zoomed and perfectly-zoomed phases, with the peaks in the plots showing the under-zoom phases. For the plot with 5 levels, the system is positive Harris recurrent, since the update equations are such that $\alpha = 0.629$, $\delta = 0.025$ and
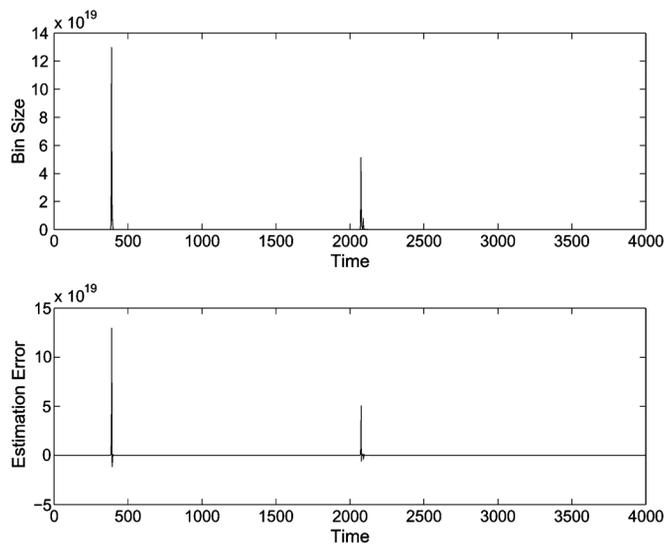
Fig. 3.   Sample path for a stochastically stable system with a 5-bin quantizer.
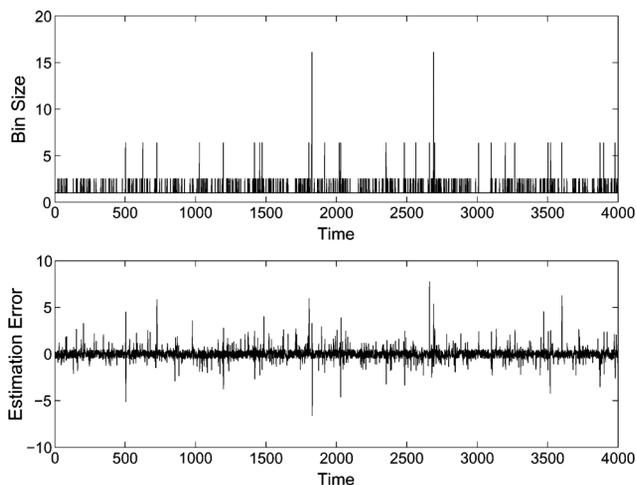
Fig. 4.   Sample path with a 17-bin quantizer; a more desirable path.

$\log_2(H(\cdot)) \in \{-0.6744, 0, 1.363\}$. These values satisfy the ir-reducibility condition since $-0.6744 = -1.363/2$, and hence the communication conditions in Theorem 3.1 are satisfied. Fur-thermore, (9) is satisfied since

$$\alpha \left(|a| + \delta\right)^{p^{-1}-1} < 0.698 < 1.$$

Increasing the bit rate by only two bits in Fig. 4 leads to a much more desirable sample path, for which the severity of rare events is reduced.

E. Extension to Multi-Dimensional Systems

The control laws and analysis can be extended to multi-di-mensional models. Consider the multi-dimensional linear system

$$x_{t+1} = Ax_t + Bu_t + d_t, \tag{14}$$

where $x_t \in \mathbb{R}^n$ is the state at time t, u_t is the control input, and $\{d_t\}$ is a sequence of zero-mean independent, identically distributed (i.i.d.) $\mathbb{R}^n$-valued zero-mean Gaussian random vari-ables. Here A is the system matrix with at least one eigen-value greater than 1 in magnitude, so that the system is open-loop unstable. Without any loss of generality, we assume A

to be in Jordan form. Suppose that B is invertible for ease in presentation.

The approach for the scalar systems is applicable, however some extension is needed. Toward this goal, one can adopt two approaches.

In one scheme, one could consider a sequential stabilization of the scalar components. In particular, we can perform an anal-ysis by considering a lower to upper sequential approach, con-sidering stabilized modes at particular stopping times as noise with a finite moment. Using an inductive argument, one can first start with the lowest mode (in the matrix diagonal) of the system, and stabilize that mode so that there is a finite invariant mth mo-ment of the state. We note that the random process for the upper mode might not have Markov dynamics for its marginal, but the joint system consisting of all modes and quantizer parameters will be Markov. Such a sequential scheme ensures a successful application of the scalar analysis presented in this paper to the vector case. One technicality that arises in this case is the fact that the effective disturbance affecting the stochastic evolution of a repeated mode in a Jordan block is no longer Gaussian, but can be guaranteed to have a sufficiently light tail distribution by Theorem 3.3.

Another approach is to adopt the discussions in [44] and Sec-tion IV of [43] (see also [23] for related constructions) and apply a vector quantizer by transmitting the quantizer bits for the en-tire $\mathbb{R}^n$-valued state. In particular, by defining a vector quantizer as a product of scalar quantizers along each (possibly general-ized) eigenvector with a common under-zoom bin, letting h^i de-note the ratio of the state and the bin range of the corresponding ith scalar quantizer and defining a sequence of stopping times as follows with $\mathcal{T}_0 = 0$ and for $z \in \mathbb{N}$

$$\mathcal{T}_{z+1} = \inf \left\{ k > \mathcal{T}_z : \left| h_k^i \right| \leq 1, i = 1, 2 \ldots, n, \quad \Upsilon_k = 1 \right\},$$

the analysis can be carried over through a geometric bound on the distribution of subsequent stopping times, obtained by an application of the union bound. See [43] for details.

IV. Conclusion

This paper contains two main contributions. One is on a gen-eral drift approach for verifying stochastic stability of Markov Chains. The other is on stabilization over erasure channels. We believe that the results presented in this paper will have many applications within the context of network stability, and net-worked control systems as well as information theoretic appli-cations. Important previous research on performance bounds for variable-length decoding schemes use stopping time arguments [5], [31], and this could form the starting point of a Lyapunov analysis.

The methods of this paper can be applied to a large class of networked control systems and problems involving delay-sen-sitive information transmission. For networked control systems, the effects of randomness in delay for transmission of sensor or controller signals (see for example [6], [29]) is an application area where the research reported here is relevant. Another re-lated area is event-triggered feedback control systems [2], [32], [35].

Rates of convergence under random-time drift is one direc-tion of future research. It is apparent that the nature of the drift

as well as the distribution of stopping times used for drift will play a role in the rate of convergence. We refer the reader to [11] and [7], for results when the drift times are deterministic.

Positive Harris recurrence can be a crude measure of stability, as seen in the numerical results in this paper. Computing the sensitivity of performance to the bit rate is an important future research problem for practical applications. For example, it was observed in Fig. 4 that increasing the bit rate by only two bits leads to much more desirable sample paths, and the magnitudes of rare events are significantly reduced. A Markovian framework is valuable for sensitivity analysis, as applied in the reinforcement learning literature (see commentary in Section 17.7 of [25]).

APPENDIX I

A. Proofs of the Stochastic Stability Theorems

1) Proof of Theorem 2.1 (i): The proof is similar to the proof of the Comparison Theorem of [25]: Define the sequence $\{M_z : z \geq 0\}$ by $M_0 = V(\phi_0)$, and for $z \geq 0$,

$$M_{z+1} = V\left(\phi_{\mathcal{T}_{z+1}}\right) + \sum_{k=0}^{z} \left(\delta\left(\phi_{\mathcal{T}_k}\right) - b1_{\left\{\phi_{\mathcal{T}_k} \in C\right\}}\right)$$

Under the assumed drift condition we have,

$$\mathsf{E}\left[M_{z+1}|\mathcal{F}_{\mathcal{T}_z}\right] \leq V\left(\phi_{\mathcal{T}_z}\right) + \sum_{k=0}^{z-1} \left(\delta\left(\phi_{\mathcal{T}_k}\right) - b1_{\left\{\phi_{\mathcal{T}_k} \in C\right\}}\right),$$

which implies the super-martingale bound,

$$\mathsf{E}\left[M_{z+1}|\mathcal{F}_{\mathcal{T}_z}\right] \leq M_z$$

For a measurable subset $C \subset \mathsf{X}$ we denote the first hitting time for the sampled chain,

$$\zeta_C = \min\left\{z \geq 1 : \phi_{\mathcal{T}_z} \in C\right\} \tag{15}$$

Define $\zeta_C^n = \min(n, \zeta_C)$ for any $n \geq 1$. Then $\mathsf{E}[M_{\zeta_C^n}] \leq M_0$ for any $n \in \mathbb{Z}$, and

$$\mathsf{E}\left[\sum_{k=0}^{\zeta_C^n - 1} \delta\left(\phi_{\mathcal{T}_k}\right)|\mathcal{F}_0\right] \leq M_0 + b.$$

Applying the bound $\mathsf{E}[\sum_{k=\mathcal{T}_z}^{\mathcal{T}_{z+1}-1} f(\phi_k)|\mathcal{F}_{\mathcal{T}_z}] \leq \delta(\phi_{\mathcal{T}_z})$ and that $f(\phi) \geq 1$, the following bound is obtained from the smoothing property of the conditional expectation:

$$\mathsf{E}\left[\mathcal{T}_{\zeta_C^n}|\mathcal{F}_0\right] = \mathsf{E}\left[\sum_{i=0}^{\zeta_C^n - 1} \mathsf{E}[\mathcal{T}_{i+1} - \mathcal{T}_i]|\mathcal{F}_0]\right]$$

$$\leq \mathsf{E}\left[\sum_{i=0}^{\zeta_C^n - 1} \delta\left(\phi_{\mathcal{T}_i}\right)|\mathcal{F}_0\right] \leq M_0 + b$$

Hence by the monotone convergence theorem,

$$\mathsf{E}[\tau_C] \leq \mathsf{E}\left[\mathcal{T}_{\zeta_C}\right] = \lim_{n \to \infty} \mathsf{E}\left[\mathcal{T}_{\zeta_C^n}|\mathcal{F}_0\right] \leq M_0 + b.$$

Consequently we obtain that

$$\sup_{\phi \in C} \mathsf{E}[\tau_C] < \infty,$$

as well as recurrence of the chain, $\mathsf{P}_\phi(\tau_C < \infty) = 1$ for any $\phi \in \mathsf{X}$. Positive Harris recurrence now follows from [26] Thm. 4.1. $\qquad\square$

The following result is key to obtaining moment bounds. The inequality (17) is known as *drift condition (V3)* [25]. Define,

$$V_f^*(\phi) := \mathsf{E}_\phi\left[\sum_{t=0}^{\tau_C - 1} f(\phi_t)\right] \quad \phi \in \mathsf{X}. \tag{16}$$

Lemma 5.1: Suppose that ϕ satisfies all of the assumptions of Theorem 2.1, except that the ψ-irreducibility assumption is relaxed. Then, there is a constant b_f such that the following bounds hold

$$PV_f^* \leq V_f^* - f + b_f 1_C \tag{17}$$
$$V_f^*(\phi) \leq V(\phi) + b_f, \quad \phi \in \mathsf{X}. \tag{18}$$

Proof: The drift condition (17) is given in Theorem 14.0.1 of [25].

The proof of (18) is based on familiar super-martingale arguments: Denote $M_0 = V(\phi_0)$, and and for $z \geq 0$,

$$M_{z+1} = V\left(\phi_{\mathcal{T}_{z+1}}\right) - \sum_{k=0}^{\mathcal{T}_{z+1}-1} \left(-f(\phi_k) + b1_{\left\{\phi_{\mathcal{T}_k} \in C\right\}}\right) \tag{19}$$

The super-martingale property for $\{M_z\}$ follows from the assumed drift condition:

$$\mathsf{E}\left[M_{z+1}|\mathcal{F}_{\mathcal{T}_z}\right] = M_z + \mathsf{E}\left[V\left(\phi_{\mathcal{T}_{z+1}}\right) - V\left(\phi_{\mathcal{T}_z}\right)\right.$$
$$\left. + \sum_{k=\mathcal{T}_z}^{\mathcal{T}_{z+1}-1} \left(f(\phi_k) - b1_{\left\{\phi_{\mathcal{T}_k} \in C\right\}}\right)|\mathcal{F}_{\mathcal{T}_z}\right]$$
$$\leq M_z \tag{20}$$

As in the previous proof we bound expectations involving the stopping time ζ_C beginning with its truncation $\zeta_C^n = \min(n, \zeta_C)$.

The super-martingale property gives $\mathsf{E}[M_{\zeta_C^n}] \leq M_0$, and once again it follows again by the monotone convergence theorem that V_f^* satisfies the bound (18) as claimed. $\qquad\square$

2) Proof of Theorem 2.1 (ii) and (iii): The existence of a finite moment follows from Lemma 5.1 and the following generalization of Kac's Theorem (see [25, Theorem 10.4.9]):

$$\pi(f) := \int \pi(d\phi)f(\phi) = \int_A \pi(d\phi)\mathsf{E}_\phi\left(\sum_{t=0}^{\tau_A - 1} f(\phi_t)\right), \tag{21}$$

where A is any set satisfying $\pi(A) > 0$, and $\tau_A = \inf(t \geq 1 : \phi_t \in A)$. The super-martingale argument above ensures that the expectation under the invariant probability measure is bounded by recognizing C as a recurrent set.

(iii) now follows from the ergodic theorem for Markov chains, see [25, Theorem 17.0.1]. $\qquad\square$

3) Proof of Theorem 2.2: The existence of an invariant probability measure in (i) follows from Theorem 12.3.4 of [25] (the

solution to the drift condition (V2) can be taken to be the mean hitting time, $V(\phi) = \mathsf{E}_\phi[\tau_C]$). See also [20, Theorem 3.1].

The proof of (ii) is similar. Rather than work with the mean return time to C, we consider the function V_f^* defined in Lemma 5.1. We have by the Comparison Theorem of [25],

$$0 \leq P^n V_f^* \leq V_f^* + n b_f - \sum_{t=0}^{n-1} P^t f$$

Hence for any $\phi \in \mathsf{X}$,

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} P^t f(\phi) \leq b_f. \tag{22}$$

Suppose that π is any invariant probability measure. Fix $N < \infty$, let $f_N = \min(N, f)$, and apply Fatou's Lemma as follows,

$$\pi(f_N) = \limsup_{n\to\infty} \pi\left(\frac{1}{n} \sum_{t=0}^{n-1} P^t f_N\right)$$

$$\leq \pi\left(\limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} P^t f_N\right) \leq b_f.$$

Fatou's Lemma is justified to obtain the first inequality, because f_N is bounded. The second inequality holds by (22) and since $f_N \leq f$. The monotone convergence theorem then gives $\pi(f) \leq b_f$. □

B. Proofs of Stability: Stochastic Stabilization Over an Erasure Channel

1) Proof of Proposition 3.1: We obtain upper and lower bounds below.

Lemma 5.2: The discrete probability measure $\mathsf{P}(\mathcal{T}_{z+1} - \mathcal{T}_z = k | x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z})$ satisfies

$$\mathsf{P}\left(\mathcal{T}_{z+1} - \mathcal{T}_z \geq k | x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z}\right) \leq (1-p)^{k-1} + G_k\left(\Delta_{\mathcal{T}_z}\right),$$

where $G_k(\Delta_{\mathcal{T}_z}) \to 0$ as $\Delta_{\mathcal{T}_z} \to \infty$ uniformly in $x_{\mathcal{T}_z}$. □

Proof: Denote for $k \in \mathbb{N}$,

$$\Theta_k := \mathsf{P}\left(\mathcal{T}_{z+1} - \mathcal{T}_z \geq k | x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z}\right)$$
$$= \mathsf{P}_{x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z}}\left(\mathcal{T}_{z+1} - \mathcal{T}_z \geq k\right). \tag{23}$$

Without any loss, let $z = 0$, $\mathcal{T}_0 = 0$, so that $\Theta_k = \mathsf{P}_{x_0, \Delta_0}(\mathcal{T}_1 \geq k)$.

Now, at time 0, upon receiving a message successfully, the estimation error satisfies $|x_0 - \hat{x}_0| \leq \Delta_0/2$, as such we have that $|a||x_0 + (b/a)u_0| \leq |a|\Delta_0/2$.

The probability Θ_k for $k \geq 2$ is bounded as follows:

$$\Theta_k = \mathsf{P}_{x_0, \Delta_0}\left(\bigcap_{s=1}^{k-1}(\Upsilon_s = 0) \cup (|h_s| > 1)\right)$$

$$\leq \mathsf{P}_{x_0, \Delta_0}\left(\bigcap_{s=1}^{k-1}(\Upsilon_s = 0)\right.$$

$$\left.\cup \left(|x_s| \geq 2^{R'-1}(|a|+\delta)^{s-1}\alpha\Delta_0\right)\right)$$

$$= \mathsf{P}_{x_0, \Delta_0}\left(\bigcap_{s=1}^{k-1}(\Upsilon_s = 0)\right.$$

$$\cup \left(\left|a^s\left(x_0 + \sum_{i=0}^{s-1}a^{-i-1}d_i\right)\right|\right.$$

$$\left.\left.\geq (|a|+\delta)^{s-1}2^{R'-1}\alpha\Delta_0\right)\right)$$

$$\leq \mathsf{P}_{x_0, \Delta_0}\left(\bigcap_{s=1}^{k-2}(\Upsilon_s = 0) \cup (|h_s| > 1)\Big|\Upsilon_{k-1} = 0\right)(1-p)$$

$$+ \mathsf{P}_{x_0, \Delta_0}\left(\bigcap_{s=1}^{k-2}(\Upsilon_s = 0) \cup (|h_s| > 1)\right.$$

$$\left|a^{k-1}\left(x_0 + (b/a)u_0 + \sum_{i=0}^{k-2}a^{-i-1}d_i\right)\right|$$

$$\left.\geq (|a|+\delta)^{k-2}2^{R'-1}\alpha\Delta_0\right)$$

$$\times \mathsf{P}_{x_0, \Delta_0}\left(\left|a^{k-1}\left(x_0 + (b/a)u_0 + \sum_{i=0}^{k-2}a^{-i-1}d_i\right)\right|\right.$$

$$\left.\geq (|a|+\delta)^{k-2}2^{R'-1}\alpha\Delta_0\right) \tag{24}$$

$$\leq \mathsf{P}_{x_0, \Delta_0}\left(\bigcap_{s=1}^{k-2}(\Upsilon_s = 0) \cup (|h_s| > 1) |\Upsilon_{k-1} = 0\right)(1-p)$$

$$+ \mathsf{P}_{x_0, \Delta_0}\left(\left|a^{k-1}\left(x_0 + (b/a)u_0 + \sum_{i=0}^{k-2}a^{-i-1}d_i\right)\right|\right.$$

$$\left.\geq (|a|+\delta)^{k-2}2^{R'-1}\alpha\Delta_0\right)$$

$$= \mathsf{P}_{x_0, \Delta_0}(\mathcal{T}_1 \geq k-1)(1-p)$$

$$+ \mathsf{P}_{x_0, \Delta_0}\left(\left|a^{k-1}\left(x_0 + (b/a)u_0 + \sum_{i=0}^{k-2}a^{-i-1}d_i\right)\right|\right.$$

$$\left.\geq (|a|+\delta)^{k-2}2^{R'-1}\alpha\Delta_0\right). \tag{25}$$

In the above derivation, (24) follows from the following: For any three events M, C, D in a common probability space

$$\mathsf{P}(M \cap (C \cup D)) = \mathsf{P}((M \cap C) \cup (M \cap D))$$
$$\leq \mathsf{P}(M \cap C) + \mathsf{P}(M \cap D)$$

Now, observe that for $k \geq 2$,

$$\mathsf{P}_{x_0, \Delta_0}\left(\left|\left(x_0 + (b/a)u_0 + \sum_{i=0}^{k-2}a^{-i-1}d_i\right)\right|\right.$$

$$\left.\geq \left(\frac{|a|+\delta}{|a|}\right)^{k-2}2^{R'-1}\frac{\alpha}{|a|}\Delta_0\right)$$

$$\leq 2\mathsf{P}_{x_0, \Delta_0}\left(\sum_{i=0}^{k-2}a^{-i-1}d_i\right.$$

$$\left.\geq \left(2^{R'-1}\left(\frac{|a|+\delta}{|a|}\right)^{k-2}\frac{\alpha}{|a|} - \frac{1}{2}\right)\Delta_0\right)$$

$$\leq C\exp\left(-\frac{((\xi^{k-2}N - 1/2)\Delta_0)^2}{2\sigma'^2}\right), \tag{26}$$

where (26) follows from (8), for this condition ensures that the term

$$\left(2^{R'-1} \left(\frac{|a|+\delta}{|a|} \right)^{k-2} \frac{\alpha}{|a|} - \frac{1}{2} \right)$$

is positive for $k \geq 2$, and bounding the complementary error function by the following: $\int_q^\infty \mu(dx) \leq q^{-1} \int_q^\infty x\mu(dx)$, for $q > 0$. In the above derivation, the constants are:

$$\sigma'^2 = \frac{E\left[d_1^2\right]}{1 - |a|^{-2}}, \quad \xi = \frac{|a|+\delta}{|a|}, \quad N = \frac{2^{R'-1}}{(|a|)/\alpha},$$

and

$$C = 2\sigma' \frac{1}{\sqrt{2\pi}(2N-1)\Delta_0/2}.$$

Let us define:

$$\Xi_k := \frac{\left((\xi^{k-2}N - 1/2)\Delta_0\right)^2}{2\sigma'^2}$$

and

$$\widetilde{\Xi}_k := \frac{\left((\xi^k N - 1/2)\Delta_0\right)^2}{2\sigma'^2}$$

We can bound the probability Θ_k defined in (23). Since a decaying exponential decays faster than any decaying polynomial, for each $m \in \mathbb{N}_+$, there exists an $M < \infty$ such that for all $k \in \mathbb{N}$,

$$Ce^{-\Xi_k} \leq M\widetilde{\Xi}_k^{-m}. \tag{27}$$

Thus, we have that

$$\mathsf{P}_{x_0,\Delta_0}\left(x_0 + \sum_{i=0}^{k-2} a^{-i-1}d_i \geq \left(\frac{|a|+\delta}{|a|}\right)^{k-2} 2^{R'-1}\frac{\alpha}{|a|}\Delta_0 \right)$$
$$\leq M\widetilde{\Xi}_k^{-m}. \tag{28}$$

Now $\Theta_1 = 1$ by definition, and for $k > 1$,

$$\Theta_k \leq \Theta_{k-1}(1-p) + Ce^{-\Xi_k}. \tag{29}$$

We obtain,

$$\Theta_k \leq \Theta_1(1-p)^{k-1} + \sum_{s=1}^{k-1}(1-p)^{k-s-1}Ce^{-\Xi_s}$$
$$\leq \Theta_1(1-p)^{k-1} + \sum_{s=1}^{k-1}M(1-p)^{k-s-1}\widetilde{\Xi}_s^{-m}$$
$$= (1-p)^{k-1} + G_k\left(\Delta_{\mathcal{T}_0}\right), \tag{30}$$

where

$$G_k\left(\Delta_{\mathcal{T}_0}\right) := \sum_{s=1}^{k-1}M(1-p)^{k-s-1}\widetilde{\Xi}_s^{-m} \tag{31}$$

It now follows that,

$$G_k\left(\Delta_{\mathcal{T}_0}\right)$$

$$=\sum_{s=1}^{k-1}M(1-p)^{k-s-1}\widetilde{\Xi}_s^{-m}$$
$$=\Delta_0^{-2m}\sum_{s=1}^{k-1}M(1-p)^{k-s-1}\left(\frac{(\xi^s N - 1/2)^2}{(2\sigma'^2)}\right)^{-m}$$
$$=\Delta_0^{-2m}(1-p)^{k-1}\sum_{s=1}^{k-1}M(1-p)^{-s}\left(\frac{(\xi^s N - 1/2)^2}{(2\sigma'^2)}\right)^{-m}$$
$$=\Delta_0^{-2m}(1-p)^{k-1}$$
$$\times \sum_{s=1}^{k-1}M(1-p)^{-s}\xi^{-2ms}\left(N - \frac{1}{2\xi^s}\right)^{-2m}(2\sigma'^2)^m$$
$$\leq \Delta_0^{-2m}(1-p)^{k-1}\Gamma_m\sum_{s=1}^{k-1}\left((1-p)\xi^{2m}\right)^{-s}$$
$$\leq \Gamma_m\Delta_0^{-2m}(1-p)^{k-1}\frac{\left((1-p)\xi^{2m}\right)^{-k} - 1}{\left((1-p)\xi^{2m}\right)^{-1} - 1}, \tag{32}$$

with $\Gamma_m = M(N - (1/2\xi))^{-2m}(2\sigma'^2)^m < \infty$. It follows that if m is taken such that

$$(1-p)\xi^{2m} > 1, \tag{33}$$

then $\lim_{\Delta_0 \to \infty} G_k(\Delta_0) = 0$, and for all $k \in \mathbb{N}$

$$\Theta_k \leq (1-p)^{k-1}\left(1 + \Gamma_m\Delta_{\mathcal{T}_0}^{-2m}\frac{1}{1 - ((1-p)\xi^{2m})^{-1}}\right) \tag{34}$$

$\square$

Lemma 5.3: The discrete probability measure $\mathsf{P}(\mathcal{T}_{z+1} - \mathcal{T}_z = k|x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z})$ satisfies

$$\mathsf{P}\left(\mathcal{T}_{z+1} - \mathcal{T}_z \geq k|x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z}\right) \geq (1-p)^{k-1},$$

for all realizations of $x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z}$. $\square$

Proof: This follows since

$$\mathsf{P}_{x_0,\Delta_0}\left(\bigcap_{s=1}^{k-1}(\Upsilon_s = 0) \cup (|h_s| > 1)\right) \geq \mathsf{P}_{x_0,\Delta_0}\left(\bigcap_{s=1}^{k-1}(\Upsilon_s = 0)\right).$$

$\square$

As a consequence of Lemma 5.2 and Lemma 5.3, the probability $\mathsf{P}(\mathcal{T}_{z+1} - \mathcal{T}_z = k|x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z})$ tends to $(1-p)^{k-1}p$ as $\Delta_{\mathcal{T}_z} \to \infty$. $\square$

2) Proof of Proposition 3.2: Let the values taken by $\log_2(H(\cdot))/s$ be $\{-\tilde{A}, 0, \tilde{B}\}$. Let

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

Since we have by (10)

$$\Delta_{t+1} = H\left(\Delta_t, |h_t|, \Upsilon_t\right)\Delta_t,$$

it follows that,

$$\log_2(\Delta_{t+1})/s = \log_2\left(H\left(\Delta_t, |h_t|, \Upsilon_t\right)\right)/s + \log_2(\Delta_t)/s$$

is also an integer. We will establish that $\mathbb{L}_{z_0, \tilde{A}, \tilde{B}}$ forms a communication class, where $z_0 = \log_2(\Delta_0)/s$ is the initial condition of the parameter for the quantizer. Furthermore, since the source process x_t is "Lebesgue-irreducible" (for the system noise admits a probability density function that is positive everywhere) 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. In view of these, we now establish that the Markov chain is irreducible.

Given $l, k \in \mathbb{L}_{z_0, \tilde{A}, \tilde{B}}$, there exist $N_{\tilde{A}}, N_{\tilde{B}} \in \mathbb{N}$ such that $l - k = -N_{\tilde{A}} \tilde{A} + N_{\tilde{B}} \tilde{B}$. In particular, if at time 0, the quantizer is perfectly zoomed and $\Delta_0 = 2^{sk}$, then there exists a sequence of events consisting of $N_{\tilde{B}}$ erasure events (simultaneously satisfying $|h_t| \leq 1$) and consequently $N_{\tilde{A}}$ zoom-in events taking place with probability at least:

$$
\left(p\mathsf{P}\left(d_t \in \left[-\left(\alpha 2^{R'} - |a| \right) L'/2, \left(\alpha 2^{R'} - |a| \right) L'/2 \right] \right) \right)^{N_{\tilde{A}}}
$$
$$
\times \left((1-p)\mathsf{P}\left(|d_t| \leq \delta 2^{R'-1} L' \right) \right)^{N_{\tilde{B}}} > 0, \quad (35)
$$

so that $\mathsf{P}(\Delta_{N_{\tilde{A}} + N_{\tilde{B}}} = 2^{sl} | \Delta_0 = 2^{ks}, x_0) > 0$, uniformly in x_0. In the following we will consider this sequence of events.

Now, for some distribution $\mathcal{K}$ on positive integers, $E \subset \mathbb{R}$ and Δ an admissible bin size,

$$
\sum_{n \in \mathbb{N}_+} \mathcal{K}(n)\mathsf{P}((x_n, \Delta_n) \in (E \times \{\Delta\}) | x_0, \Delta_0) \geq K_{\Delta_0, \Delta} \psi(E, \Delta)
$$

Here $K_{\Delta_0, \Delta}$, denoting a lower bound on the probability of visiting Δ from Δ_0 in some finite time, is non-zero by (35) and ψ is positive as the following argument shows: Let $t > 0$ be the time stage for which $\Delta_t = \Delta$ and thus by the construction in (35), with $|h_{t-1}| \leq 1$: $|ax_{t-1} + bu_{t-1}| \leq |a| \Delta_{t-1}/2 = (|a|/\alpha)\Delta/2$. Thus, it follows that, for $A_1, B_1 \in \mathbb{R}, A_1 < B_1$,

$$
\mathsf{P}\left(x_t \in [A_1, B_1] \Big| |ax_{t-1} + bu_{t-1}| \leq |a|\Delta_{t-1}/2, \Delta_{t-1} \right)
$$
$$
= \mathsf{P}\left(ax_{t-1} + bu_{t-1} + d_{t-1} \in [A_1, B_1] \right.
$$
$$
\left. \Big| |ax_{t-1} + bu_{t-1}| \leq |a|\Delta_{t-1}/2, \Delta_{t-1} \right)
$$
$$
\geq \min_{|z| \leq \frac{\Delta}{2}\left(\frac{|a|}{\alpha}\right)} \left(\mathsf{P}(d_{t-1} \in [A_1 - z, B_1 - z]) > 0 \right) \quad (36)
$$

Now, define the finite set $C'_\Delta := \{\Delta : L' \leq |\Delta| \leq F', (\log_2(\Delta)/s) \in \mathbb{N}\}$. The chain satisfies the recurrence property that $\mathsf{P}_{(x, \Delta)}(\tau_{C_x \times C'_\Delta} < \infty) = 1$ for any admissible (x, Δ). This follows, as before in Section V-B1, from the construction of

$$
\Theta_k(\Delta, x) := \mathsf{P}(\mathcal{T}_1 \geq k | x, \Delta),
$$

where

$$
\mathcal{T}_1 = \inf\left(k > 0 : |x_k| \leq 2^{R'-1}\Delta_k, x_0 = x, \Delta_0 = \Delta \right)
$$

and observing that $\Theta_k(\Delta, x)$ is majorized by a geometric measure with similar steps as in Section V-B1. Once a state which is perfectly zoomed, that is which satisfies $|x_t| \leq 2^{R'-1}\Delta_t$, is visited, the stopping time analysis can be used to verify that from any initial condition the recurrent set is visited in finite time with probability 1.

In view of (35), we have that the chain is irreducible.

We can now show that the set $C_x \times C'_\Delta$ is small. We will show first that this set is *petite*: A set $D \in \mathcal{B}(\mathsf{X})$ is petite if there is a probability measure $\mathcal{K}$ on the non-negative integers $\mathbb{N}$, and a positive measure μ_p satisfying $\mu(\mathsf{X}) > 0$ and

$$
\sum_{n=0}^{\infty} P^n(x, E)\mathcal{K}(n) \geq \mu_p(E), \, for \, all \, x \in D, \text{ and } E \in \mathcal{B}(\mathsf{X}).
$$

By Theorem 5.5.7 of [25], under aperiodicity and irreducibility, every petite set is small. To this end, we will establish aperiodicity at the end of the proof.

To establish the petite set property, we will follow an approach taken by Tweedie [37], [38] which considers the following test, which only depends on the one-stage transition kernel of a Markov chain: If a set S is such that, the following *uniform countable additivity* condition

$$
\lim_{n \to \infty} \sup_{x \in S} P(x, B_n) = 0, \quad (37)
$$

is satisfied for every sequence $B_n \downarrow \emptyset$, and if the Markov chain is irreducible, then S is petite (see Lemma 4 of Tweedie [38] and Proposition 5.5.5 (iii) of Meyn-Tweedie [25]).

Now, the set $C_x \times C'_\Delta$ satisfies (37), since for any given bin size Δ' in the countable space constructed above, we have that

$$
\lim_{n \to \infty} \sup_{(x, \Delta) \in C_x \times C'_\Delta} \mathsf{P}\left((x_{t+1}, \Delta_{t+1}) \in (B_n \times \Delta') \right.
$$
$$
\left. | x_t = x, \Delta_t = \Delta \right)
$$
$$
= \lim_{n \to \infty} \sup_{(x, \Delta) \in C_x \times C'_\Delta} \mathsf{P}\left((ax + bu_t + d_t, \Delta_{t+1}) \right.
$$
$$
\left. \in (B_n \times \Delta') | x_t = x, \Delta_t = \Delta \right)
$$
$$
= \lim_{n \to \infty} \sup_{(x, \Delta) \in C_x \times C'_\Delta} \mathsf{P}\left((d_t, \Delta_{t+1}) \right.
$$
$$
\in ((B_n - (ax + bu_t)) \times \Delta')
$$
$$
\left. | x_t = x, \Delta_t = \Delta \right)
$$
$$
= 0.
$$

This follows from the fact that the Gaussian random variable d_1 satisfies

$$
\lim_{n \to \infty} \sup_{d_1 \in C_0} \mathsf{P}(d_1 \in A_n) = 0,
$$

uniformly over a compact set C_0, for any sequence $A_n \downarrow \emptyset$, since a Gaussian measure admits a uniformly bounded density function.

Therefore, $C_x \times C'_\Delta$ is petite.

If the integers $\tilde{A}, \tilde{B}$ are relatively prime, then by Bézout's Lemma (see [1]), the communication class will include the bin sizes whose logarithms are integer multiples of a constant except those leading to $\Delta < L'$.

We finally show that the Markov chain is aperiodic. This follows from the fact that the smallest admissible state for the quantizer, $\Delta^* = L'$, can be visited in subsequent time stages with non-zero probability, since

$$
\left(\min_{|x| \leq \frac{\Delta^*}{2}} P\left(d_t \in \left[-2^{R'-1}\Delta^* - ax, 2^{R'-1}\Delta^* - ax \right] \right) \right) p > 0.
$$

3) Proof of Theorem 3.1: With the Lyapunov function $V_0(x_t, \Delta_t) = \log(\Delta_t^2) + B_0$, for $\Delta_{\mathcal{T}_z} > L$, we have that

$$\mathsf{E}\left[V_0\left(x_{\mathcal{T}_{z+1}},\Delta_{\mathcal{T}_{z+1}}\right)|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right]$$
$$= B_0 + \mathsf{P}\left(\mathcal{T}_{z+1}-\mathcal{T}_z=1|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right)$$
$$\times\left(2\log(\alpha)+\log\left(\Delta_{\mathcal{T}_z}^2\right)\right)$$
$$+\sum_{k=2}^{\infty}\log\left(\Delta_{\mathcal{T}_z+k}^2\right)\mathsf{P}\left(\mathcal{T}_{z+1}-\mathcal{T}_z=k|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right)$$

Thus, the drift satisfies:
$$\mathsf{E}\left[V_0\left(x_{\mathcal{T}_{z+1}},\Delta_{\mathcal{T}_{z+1}}\right)|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right]-V_0\left(x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right)$$
$$=\sum_{k=1}^{\infty}2\log\left((|a|+\delta)^{(k-1)}\alpha\right)\mathsf{P}\left(\mathcal{T}_{z+1}-\mathcal{T}_z=k|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right)$$
$$=2\sum_{k=1}^{\infty}(k-1)\log(|a|+\delta)\,\mathsf{P}\left(\mathcal{T}_{z+1}-\mathcal{T}_z=k|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right)$$
$$+2\log(\alpha). \tag{38}$$

By (31), the summability of $\sum_{k=1}^{\infty}G_k(\Delta_{\mathcal{T}_z})$, and the dominated convergence theorem,

$$\lim_{\Delta_{\mathcal{T}_z}\to\infty}\sum_{k=1}^{\infty}(k-1)\left((1-p)^{k-1}+G_k\left(\Delta_{\mathcal{T}_z}\right)-(1-p)^k\right)$$
$$=\sum_{k=1}^{\infty}\lim_{\Delta_{\mathcal{T}_z}\to\infty}(k-1)\left((1-p)^{k-1}+G_k\left(\Delta_{\mathcal{T}_z}\right)-(1-p)^k\right)$$
$$=\sum_{k=1}^{\infty}p(1-p)^{k-1}(k-1)=p^{-1}-1 \tag{39}$$

Provided (9) holds, it follows from Lemma 5.2 and Lemma 5.3 that for some $b_0>0$,

$$\lim_{\Delta_{\mathcal{T}_z}\to\infty}\left\{\mathsf{E}\left[V_0\left(x_{\mathcal{T}_{z+1}},\Delta_{\mathcal{T}_{z+1}}\right)|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right]-V_0\left(x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right)\right\}$$
$$=2\log(\alpha)+2\lim_{\Delta_{\mathcal{T}_z}\to\infty}\left\{\sum_{k=1}^{\infty}(k-1)\log(|a|+\delta)\right.$$
$$\left.\times\mathsf{P}\left(\mathcal{T}_{z+1}-\mathcal{T}_z=k|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right)\right\}$$
$$\leq -b_0. \tag{40}$$

For $\Delta_{\mathcal{T}_z}$ in a compact set and lower bounded by L' defined by (11), $\mathsf{E}[\log(\Delta_{\mathcal{T}_{z+1}}^2)|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}]$ is uniformly bounded. This follows from the representation of the drift given in (38). Finally, since,

$$G_k\left(\Delta_{\mathcal{T}_0}\right)\leq(1-p)^{k-1}\Gamma_m\Delta_{\mathcal{T}_0}^{-2m}\frac{1}{1-((1-p)\xi^{2m})^{-1}},$$

it follows that $\sum_{k=1}^{\infty}G_k(\Delta_{\mathcal{T}_0})k<\infty$ and as a result

$$\sup_{x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}}\mathsf{E}_{x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}}[\mathcal{T}_{z+1}-\mathcal{T}_z]<\infty. \tag{41}$$

Consequently, under the bound (9), there exist $b_0>0$, $b_1<\infty$, $F'>0$ such that,

$$\mathsf{E}\left[V_0\left(x_{\mathcal{T}_{z+1}},\Delta_{\mathcal{T}_{z+1}}\right)|x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right]$$
$$\leq V_0\left(x_{\mathcal{T}_z},\Delta_{\mathcal{T}_z}\right)-b_0+b_1 1_{\{|\Delta_{\mathcal{T}_z}|\leq F'\}}. \tag{42}$$

Thus combined with Proposition 3.2, (41), and (42), Corollary 2.1 leads to positive Harris recurrence. □

4) Proof of Theorem 3.2: First, let us note that by (30) and (34), for every $\kappa>0$, we can find Δ_0 sufficiently large such that

$$\lim_{t\to\infty}\frac{\mathsf{P}(\mathcal{T}_1\geq t|x_0,\Delta_0)}{(1-p+\kappa)^{t-1}}=0.$$

Under the assumed bound $(1-p)|a|^2<1$, we can fix $\kappa>0$ such that $(1-p+\kappa)|a+\delta|^2<1$.

Next, observe that for all initial conditions for which $|h_0|\leq 1$,

$$\lim_{\Delta_0\to\infty}\frac{\mathsf{E}[V_2\left(x_{\mathcal{T}_1},\Delta_{\mathcal{T}_1}\right)|x_0,\Delta_0]}{V_2(x_0,\Delta_0)}$$
$$=\lim_{\Delta_0\to\infty}\frac{\mathsf{E}\left[\Delta_{\mathcal{T}_1}^2|x_0,\Delta_0\right]}{\Delta_0^2}$$
$$=\lim_{\Delta_0\to\infty}\frac{1}{\Delta_0^2}\sum_{k=1}^{\infty}\mathsf{P}(\mathcal{T}_1=k)\mathsf{E}\left[\Delta_k^2|\mathcal{T}_1=k,x_0,\Delta_0\right]$$
$$=\lim_{\Delta_0\to\infty}\alpha^2\sum_{k=1}^{\infty}\mathsf{P}(\mathcal{T}_1=k)\left(|a|+\delta\right)^{2(k-1)}$$
$$=p\alpha^2\frac{1}{1-(1-p)\left(|a|+\delta\right)^2}, \tag{43}$$

where the last equality follows from Lemma 5.2 and the dominated convergence theorem.

Now, if (13) holds, we can find α such that $R'>\log_2(|a|/\alpha)$, and

$$\frac{p\alpha^2}{1-(1-p)\left(|a|+\delta\right)^2}<1, \tag{44}$$

and simultaneously (9) is satisfied. We note that (44) implies (9) since by Jensen's inequality:

$$\log\left(p\alpha^2+(1-p)\left(|a|+\delta\right)^2\right)$$
$$>p\log(\alpha^2)+(1-p)\log\left((|a|+\delta)^2\right),$$

and (9) is equivalent to the term on the right hand side being negative.

To establish the required drift equation, we first establish the following bound for all $z\geq0$:

$$\kappa\mathsf{E}\left[\sum_{m=\mathcal{T}_z}^{\mathcal{T}_{z+1}-1}x_m^2|x_0,\Delta_0\right]\leq\Delta_{\mathcal{T}_z}^2 2^{2(R'-1)}, \tag{45}$$

for some $\kappa>0$.

Without loss of generality take $z=0$ so that $\mathcal{T}_z=0$. Observe that for any $\chi>0$, by Hölder's inequality,

$$\mathsf{E}\left[\sum_{t=0}^{\mathcal{T}_1-1}x_t^2|x_0,\Delta_0\right]=\mathsf{E}\left[\sum_{t=0}^{\infty}1_{\{t<\mathcal{T}_1\}}x_t^2|x_0,\Delta_0\right]$$
$$\leq\sum_{t=0}^{\infty}\left(\mathsf{E}\left[(1_{\{t<\mathcal{T}_1\}})^{1+\chi}|x_0,\Delta_0\right]\right)^{\frac{1}{1+\chi}}$$
$$\times\left(\mathsf{E}\left[x_t^{2(\frac{1+\chi}{\chi})}|x_0,\Delta_0\right]\right)^{\frac{\chi}{1+\chi}}. \tag{46}$$

Moreover, for some $B_2<\infty$,

$$\mathsf{E}\left[x_t^{2(\frac{1+\chi}{\chi})}|x_0,\Delta_0\right]$$

$$
= \mathsf{E}\left[a^{2t\left(\frac{1+\chi}{\chi}\right)} \left(x_0 + \sum_{i=0}^{t-1} a^{-i-1} d_i \right)^{2\left(\frac{1+\chi}{\chi}\right)} \bigg| x_0, \Delta_0 \right]
$$

$$
= |a|^{2t\left(\frac{1+\chi}{\chi}\right)} \mathsf{E}\left[\left(x_0 + \sum_{i=0}^{t-1} a^{-i-1} d_i \right)^{2\frac{1+\chi}{\chi}} \bigg| x_0, \Delta_0 \right]
$$

$$
\leq |a|^{2t\left(\frac{1+\chi}{\chi}\right)} \mathsf{E}\left[\left(x_0 + \sum_{i=0}^{\infty} a^{-i-1} d_i \right)^{2\frac{1+\chi}{\chi}} \bigg| x_0, \Delta_0 \right]
$$

$$
= |a|^{2t\left(\frac{1+\chi}{\chi}\right)} \left(2^{R'-1} \Delta_0 \right)^{2\frac{1+\chi}{\chi}}
$$
$$
\times \mathsf{E}\left[\left(\frac{x_0 + \sum_{i=0}^{\infty} a^{-i-1} d_i}{2^{R'-1} \Delta_0} \right)^{2\frac{1+\chi}{\chi}} \bigg| x_0, \Delta_0 \right]
$$

$$
= |a|^{2t\left(\frac{1+\chi}{\chi}\right)} \left(2^{R'-1} \Delta_0 \right)^{2\frac{1+\chi}{\chi}}
$$
$$
\times \mathsf{E}\left[\left(h_0 + \frac{\sum_{i=0}^{\infty} a^{-i-1} d_i}{2^{R'-1} \Delta_0} \right)^{2\frac{1+\chi}{\chi}} \bigg| x_0, \Delta_0 \right]
$$

$$
< B_2 \left(2^{R'-1} \Delta_0 \right)^{2\frac{1+\chi}{\chi}} |a|^{2t\left(\frac{1+\chi}{\chi}\right)}, \tag{47}
$$

where the last inequality follows since for every fixed $|h_0| \leq 1$, the random variable $h_0 + (\sum_{i=0}^{\infty} a^{-i-1} d_i)/(2^{R'-1} \Delta_0)$ has a Gaussian distribution with finite moments, uniform on $\Delta_0 \geq L'$.

Thus,

$$
\mathsf{E}\left[\sum_{t=0}^{\mathcal{T}_1 - 1} x_t^2 | x_0, \Delta_0 \right]
$$
$$
\leq \sum_{t=0}^{\infty} \left(\mathsf{E}\left[(1_{\{t < \mathcal{T}_1\}})^{1+\chi} | x_0, \Delta_0 \right] \right)^{\frac{1}{1+\chi}}
$$
$$
\times \left(B_2 \left(2^{R'-1} \Delta_0 \right)^{2\frac{1+\chi}{\chi}} |a|^{2t\left(\frac{1+\chi}{\chi}\right)} \right)^{\frac{\chi}{1+\chi}}
$$
$$
= \left(2^{R'-1} \Delta_0 \right)^2 \sum_{t=0}^{\infty} (\mathsf{P}(\mathcal{T}_1 \geq t+1 | x_0, \Delta_0))^{\frac{1}{1+\chi}}
$$
$$
\times \left(B_2 |a|^{2t\left(\frac{1+\chi}{\chi}\right)} \right)^{\frac{\chi}{1+\chi}}
$$
$$
< \zeta_{B_2} \left(2^{R'-1} \Delta_0 \right)^2
$$

for some $\zeta_{B_2} < \infty$.

The last inequality is due to the fact there exists $\kappa > 0$ such that

$$
\lim_{t \to \infty} \frac{\mathsf{P}(\mathcal{T}_1 \geq t | x_0, \Delta_0)}{(1 - p + \kappa)^{t-1}} = 0,
$$

and we can pick $\chi > 0$ with $(1 - p + \kappa)|a|^{2(1+\chi)} < 1$. Such χ and κ exist by the hypothesis that $(1 - p)|a|^2 < 1$.

Hence, with $0 < \epsilon < 1 - p\alpha^2/[1 - (1 - p)(|a| + \delta)^2]$

$$
\delta(x, \Delta) = \epsilon \Delta^2, \quad f(x, \Delta) = \frac{\epsilon}{\zeta_{B_2} 2^{2(R'-1)}} x^2,
$$

C a compact set, and $V_2(x, \Delta) = \Delta^2$, Theorem 2.1 applies and $\lim_{t \to \infty} E[x_t^2] < \infty$. □

5) Proof of Theorem 3.3: The proof follows closely that of Theorem 3.2.

Again by Hölder's inequality, for any $\chi > 0$,

$$
\mathsf{E}\left[\sum_{t=0}^{\mathcal{T}_1 - 1} |x_t|^m \bigg| x_0, \Delta_0 \right] = \mathsf{E}\left[\sum_{t=0}^{\infty} 1_{\{t < \mathcal{T}_1\}} |x_t|^m \bigg| x_0, \Delta_0 \right]
$$
$$
\leq \sum_{t=0}^{\infty} \left(\mathsf{E}[(1_{\{t < \mathcal{T}_1\}})^{1+\chi} | x_0, \Delta_0] \right)^{\frac{1}{1+\chi}}
$$
$$
\times \left(\mathsf{E}\left[|x_t|^{m\left(\frac{1+\chi}{\chi}\right)} | x_0, \Delta_0 \right] \right)^{\frac{\chi}{1+\chi}}. \tag{48}
$$

As in (47), for some $B_m < \infty$,

$$
\mathsf{E}\left[|x_t|^{m\left(\frac{1+\chi}{\chi}\right)} | x_0, \Delta_0 \right] \leq B_m \left(\Delta_0^m 2^{m(R'-1)} \right)^{\frac{1+\chi}{\chi}} |a|^{mt\left(\frac{1+\chi}{\chi}\right)}
$$

and consequently,

$$
\mathsf{E}\left[\sum_{t=0}^{\tau_1 - 1} |x_t|^m | x_0, \Delta_0 \right]
$$
$$
\leq \Delta_0^m 2^{m(R'-1)}
$$
$$
\times \sum_{t=0}^{\infty} (\mathsf{P}(\mathcal{T}_1 \geq t+1 | x_0, \Delta_0))^{\frac{1}{1+\chi}} \left(B_m^{\frac{\chi}{1+\chi}} |a|^{mt} \right)
$$
$$
< \zeta_{B_m} \left(2^{R'-1} \Delta_0 \right)^m \tag{49}
$$

where, once again, the last inequality is due to the fact that there exists a $\kappa > 0$ such that

$$
\lim_{t \to \infty} \frac{\mathsf{P}(\mathcal{T}_1 \geq t | x_0, \Delta_0)}{(1 - p + \kappa)^{t-1}} = 0,
$$

and we can pick $\chi > 0$ such that $(1 - p + \kappa)|a|^{m(1+\chi)} < 1$; such χ and κ exist by the property $(1 - p)|a|^m < 1$.

Hence, with $0 < \epsilon < 1 - p\alpha^m/[1 - (1 - p)(|a| + \delta)^m]$,

$$
\delta(x, \Delta) = \epsilon \Delta^m, \quad f(x, \Delta) = \frac{\epsilon}{\zeta_{B_m} 2^{2(R'-1)}} |x|^m,
$$

C a compact set, and $V_m(x, \Delta) = \Delta^m$, Theorem 2.1 applies, establishing in the desired conclusions, and in particular that $\lim_{t \to \infty} E[|x_t|^m]$ exists and is finite. □

REFERENCES

[1] G. E. Andrews, *Number Theory*. New York: Dover, 1995.
[2] K. J. Aström and B. Bernhardsson, "Comparison of riemann and lebesgue sampling for first order stochastic systems," in *Proc. IEEE Conf. Decision Control*, Dec. 2002, pp. 2011–2016.
[3] V. S. Borkar and S. P. Meyn, "The ODE method for convergence of stochastic approximation and reinforcement learning," *SIAM J. Control Optim.*, pp. 447–469, Dec. 2000.
[4] R. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Trans. Autom. Control*, vol. 45, no. 7, pp. 1279–1289, Jul. 2000.
[5] M. V. Burnashev, "Data transmission over a discrete channel with feedback. Random transmission time," *Probl. Pered. Inf.*, vol. 12, pp. 10–30, 1976.

[6] M. B. G. Cloosterman, N. van de Wouw, W. P. M. H. Heemels, and H. Nijmeijer, "Stability of networked control systems with uncertain time-varying delays," *IEEE Trans. Autom. Control*, vol. 54, no. 7, pp. 1575–1580, Jul. 2009.

[7] S. B. Connor and G. Fort, "State-dependent foster-lyapunov criteria for subgeometric convergence of Markov chains," *Stoch. Processes Appl.*, vol. 119, no. 12, pp. 4176–4193, Dec. 2009.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[9] J. G. Dai, "On positive Harris recurrence of multiclass queueing networks: A unified approach via fluid limit models," *Annals Appl. Probability*, vol. 5, pp. 49–77, 1995.

[10] J. G. Dai and S. P. Meyn, "Stability and convergence of moments for multiclass queueing networks via fluid limit models," *IEEE Trans. Autom. Control*, vol. 40, no. 11, pp. 1889–1904, Nov. 1995.

[11] R. Douc, G. Fort, E. Moulines, and P. Soulier, "Practical drift conditions for subgeometric rates of convergence," *Ann. Appl. Probab.*, vol. 14, pp. 1353–1377, 2004.

[12] G. Fayolle, V. A. Malyshev, and M. V. Meńshikov, *Topics in the Constructive Theory of Countable Markov Chains*. Cambridge, MA: Cambridge Univ. Press, 1995.

[13] G. Fort, S. P. Meyn, E. Moulines, and P. Priouret, "ODE methods for Markov chain stability with applications to MCMC," in *Proc. 1st Int. Conf. Perform. Eval. Methodology Tools*, New York, NY, 2006, ACM.

[14] B. H. Fralix, "Foster-type criteria for Markov chains on general spaces," *J. Appl. Probability*, vol. 43, pp. 1194–1200, Dec. 2006.

[15] G. Fort, S. P. Meyn, E. Moulines, and P. Priouret, "ODE methods for skip-free Markov chain stability with applications to MCMC," *Annals Appl. Probability*, pp. 664–707, 2008.

[16] D. J. Goodman and A. Gersho, "Theory of an adaptive quantizer," *IEEE Trans. Commun.*, vol. 22, pp. 1037–1045, Aug. 1974.

[17] A. Gurt and G. Nair, "Internal stability of dynamic quantised control for stochastic linear plantsstar," *Automatica*, vol. 45, pp. 1387–1396, Jun. 2009.

[18] J. C. Kieffer, "Stochastic stability for feedback quantization schemes," *IEEE Trans. Inform. Theory*, vol. 28, pp. 248–254, Mar. 1982.

[19] J. C. Kieffer and J. G. Dunham, "On a type of stochastic stability for a class of encoding schemes," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 793–797, Nov. 1983.

[20] J. B. Lasserre, "Invariant probabilities for Markov chains on a metric space," *Stat. Probability Lett.*, vol. 34, pp. 259–265, 1997.

[21] D. Liberzon and D. Nesic, "Input-to-state stabilization of linear systems with quantized state measurements," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 767–781, May 2007.

[22] V. A. Malyšev and M. V. Meńšikov, "Ergodicity, continuity and analyticity of countable Markov chains," *Trans. Moscow Math. Soc.*, vol. 39, pp. 1–48, 1981.

[23] A. S. Matveev and A. V. Savkin, *Estimation and Control Over Communication Networks*. Boston, MA: Birkhauser, 2008.

[24] S. P. Meyn, *Control Techniques for Complex Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[25] S. P. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*. London, U.K.: Springer Verlag, 1993.

[26] 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.

[27] S. P. Meyn and R. Tweedie, "State-dependent criteria for convergence of Markov chains," *Ann. Appl. Probab.*, vol. 4, pp. 149–168, 1994.

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

[29] L. A. Montestruque and P. Antsaklis, "Stability of model-based networked control systems with time-varying transmission times," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1562–1572, Sep. 2004.

[30] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM J. Control Optim.*, vol. 43, pp. 413–436, Jul. 2004.

[31] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Variable-length coding with feedback in the non-asymptotic regime," in *Proc. IEEE Int. Symp. Inform. Theory*, Austin, TX, Jun. 2010, pp. 231–235.

[32] M. Rabi, K. H. Johansson, and M. Johansson, "Optimal stopping for event-triggered sensing and actuation," in *Proc. IEEE Conf. Decision Control*, Cancun, Mexico, Dec. 2008, pp. 3607–3612.

[33] A. Sahai, "Why block length and delay behave differently for channel coding with feedback," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 1860–1886, May 2008.

[34] 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, no. 8, pp. 3369–3395, Aug. 2006.

[35] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1680–1685, Sep. 2007.

[36] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.

[37] R. L. Tweedie, "Invariant measures for Markov chains with no irreducibility assumptions," *Celebration Appl. Probability, J. Appl. Prob., Spec.*, vol. 25A, pp. 275–285, 1988.

[38] R. L. Tweedie, "Drift conditions and invariant measures for Markov chains," *Stoch. Processes Appl.*, vol. 92, pp. 345–354, 2001.

[39] K. You and L. Xie, "Minimum data rate for mean square stabilization of discrete LTI systems over lossy channels," *IEEE Trans. Autom. Control*, vol. 55, no. 10, pp. 2373–2378, Oct. 2010.

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

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

[42] S. Yüksel, "A random time stochastic drift result and application to stochastic stabilization over noisy channels," in *Proc. Annu. Allerton Conf. Commun., Control, Comp.*, Sep. 2009, pp. 628–635.

[43] S. Yüksel, "Characterization of information channels for asymptotic mean stationarity and stochastic stability of non-stationary/unstable linear systems," *IEEE Trans. Inform. Theory*, to be published.

[44] S. Yüksel, "A tutorial on quantizer design for networked control systems: Stabilization and optimization," *Appl. Computat. Math.*, vol. 11, pp. 365–403, 2011.

Serdar Yüksel (M'02) was born in Lice, Turkey, in 1979. He received the B.Sc. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2001, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana-Champaign, in 2003 and 2006, respectively.

He was a Post-Doctoral Researcher at Yale University, Storrs, CT, before joining the Department of Mathematics and Statistics, Queen's University, Charlotte, NC, at in 2007. His research interests are on stochastic control, decentralized control, information theory, source coding theory, multi-terminal control and communication systems and stochastic processes.

Sean P. Meyn (F'02) received the B.A. degree in mathematics from the University of California at Los Angeles (UCLA) in 1982, and the Ph.D. degree in electrical engineering from McGill University, Montreal, QC, Canada, in 1987.

After 22 years as a Professor at the University of Illinois, he is now Robert C. Pittman Eminent Scholar Chair in the Department of Electrical and Computer Engineering, University of Florida, Gainesville, and Director of the new Laboratory for Cognition and Control. His research interests include stochastic processes, optimization, complex networks, information theory, and power and energy systems.