Stochastic Stability and Ergodicity of Linear Gaussian Systems Controlled over Discrete Channels

Serdar Yüksel[1]

Abstract— We present necessary and sufficient conditions for stochastic stabilizability of unstable (non-stationary) noisy linear systems over channels with memory with feedback. Stochastic stability notions include recurrence, Birkhoff sample path ergodicity and asymptotic mean stationarity, and the existence of finite second moments. We extend recent results in the literature on noiseless and erasure channels for systems driven by possibly unbounded noise. Our constructive proof uses random-time state-dependent stochastic drift criteria for stabilization of Markov chains. For asymptotic mean stationarity, it is sufficient that the capacity of a channel is (strictly) greater than the sum of the logarithms of the unstable pole magnitudes for memoryless channels and a class of channels with memory. We also present the tightness of the sufficient condition under a technical condition. We provide sufficiency conditions for the existence of finite average second moments, for such systems driven by unbounded noise, which had not been studied in the literature to our knowledge. Comparison with relevant results in the literature is presented.

I. Problem Formulation

We consider a scalar LTI discrete-time system described by

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

where x_t is the state at time t, u_t is the control input, the initial condition 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. 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 a Discrete Memoryless Channel with a finite capacity to a controller, as shown in Figure 1.

The channel source consists of state values, taking values in $\mathbb{R}$. The source is *quantized*: A quantizer Q is represented by a map $Q \colon \mathbb{R} \to \mathbb{R}$, characterized by a sequence of non-overlapping Borel-measurable bins $\{\mathcal{B}_i\}$, such that $Q(x) = q^i \in \mathbb{R}$ if and only if $x \in \mathcal{B}_i$; that is, $Q(x) = \sum_i q^i 1_{\{x \in \mathcal{B}_i\}}$. The quantizer outputs are transmitted through a channel, after being subjected to a channel encoder.

Definition 1.1: A finite-alphabet channel with memory is characterized by a sequence of finite input alphabets $\mathcal{M}^{n+1}$, finite output alphabets $\mathcal{M}'^{n+1}$, and a sequence of conditional probability measures $P_n(q'_{[0,n]}|q_{[0,n]}) : \mathcal{M}^{n+1} \times \mathcal{M}'^{n+1} \to \mathbb{R}$, with, $q'_{[0,n]} =: \{q'_0, q'_1, \ldots, q'_n\}$, $q_{[0,n]} := \{q_0, q_1, \ldots, q_n\}$.

Definition 1.2: A Discrete Memoryless Channel (DMC) is characterized by a finite input alphabet $\mathcal{M}$, a finite output alphabet $\mathcal{M}'$, and a conditional probability mass function $P(q'|q)$, from $\mathcal{M} \times \mathcal{M}'$ to $\mathbb{R}$. A DMC from $\mathcal{M}^{n+1}$ to $\mathcal{M}'^{n+1}$ satisfies the following:

$$P_{DMC}^{n+1}(q'_{[0,n]}|q_{[0,n]}) = \prod_{k=0}^{n} P_{DMC}(q'_k|q_k).$$

The quantizer outputs are transmitted through a noisy channel, hence the receiver has access to noisy versions of the quantizer/coder outputs for each time, which we denote by $q' \in \mathcal{M}'$. The quantizers and controllers are causal such that the quantizer function at time $t \geq 0$ is generated using the information vector I_t^s available at the encoder for $t > 0$:

$$I_t^s = \{I_{t-1}^s, x_t, q_{t-1}, q'_{t-1}\},$$

and $I_0^s = \{\nu_0(dx), x_0\}$, where $\nu_0 = \pi_0$ is the probability measure for the initial state. The control policies are measurable with respect to the sigma-algebra generated by I_t^c, for $t \geq 1$: $I_t^c = \{I_{t-1}^c, q'_t\}$, and are mappings to $\mathbb{R}$.
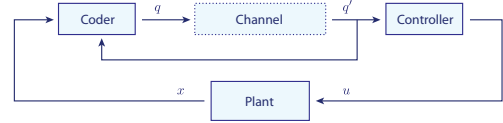

Fig. 1: Control over a discrete noisy channel with feedback.

The goal of the paper is to identify conditions on the channel under which the controlled process $\{x_t\}$ is stochastically stable in each of the following senses:

- $\{x_t\}$ is recurrent: There exists a compact set A such that $\{x_t \in A\}$ infinitely often almost surely.
- $\{x_t\}$ is asymptotically mean stationary and satisfies the sample path ergodic theorem.
- $\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} ||x_t||^2$ exists and finite almost surely.

We will make the definitions more precise, after a literature review is provided and some review of Markov chains and ergodic theory is presented.

A. Literature Review

There is a very large literature on stochastic stabilization of sources via coding, both in the information theory and control theory communities. We are unable to state the contributions of many researchers in view of space constraints. In the information theory literature, stochastic stability results are established mainly for stationary sources, which are already stable sources. In this literature, the stability of the

estimation errors as well as the encoder state processes are studied. These systems mainly involve causal and non-causal coding of stationary sources [8], [14], asymptotically mean stationary sources as well as to a limited extent on non-causal coding for non-stationary sources [12]. Real-time settings such as sigma-delta quantization schemes have also been considered in the literature. In contrast, the control theory literature has mainly considered unstable (or non-stationary) sources, in the context of networked control systems. In the following, we will provide a discussion on the contributions in the literature which are contextually close to our paper. For Gaussian channels, the fact that real-time linear schemes are rate-distortion achieving has been observed earlier in [1] for scalar systems (see also [6] on Gaussian channels regarding relevance to Shannon capacity for multi-dimensional systems). Aside from such results (which involve matching between rate-distortion achieving test channels and capacity achieving source distributions), capacity is known not to be a good measure of information reliability for channels for real-time control and estimation problems [25]. [13], [27], and [23], obtained the minimum lower bound needed for stabilization over noisy channels under a class of assumptions on the system noise and channels. [23] considered a class of quantizer policies for systems driven by noise, with unbounded support set for its probability measure, controlled over noiseless channels, and obtained necessary and sufficient conditions for the boundedness in the following sense $\limsup_{t\to\infty} E[||x_t||^2] < \infty$. [32] extended the result by replacing $\limsup$ with $\lim$. The problem of control over noisy channels has been considered later in a large number of publications including [25], [15], [16], [19], [18], [22], [30].

The particular notion of stochastic stability is very important in characterizing the conditions on the channel. [18] considered the following, when the system noise is bounded: $\limsup_{t\to\infty} |x_t| < \infty$ *a.s.*, and observed that one needs the zero-error capacity (with feedback) to be greater than a particular lower bound. A similar observation was made in [25]. [25] considered systems driven by bounded noise and considered a number of stability criteria: Almost sure stability for noise-free systems, moment stability for systems with bounded noise ($\limsup_{t\to\infty} E[|x_t|^p] < \infty$,) as well as *stability in probability* (defined in [18]) for systems with bounded noise. Stability in probability is defined as follows: For every $p > 0$, there exists a ζ such that $P(|x_t| > \zeta) < p$ for all $t \in \mathbb{N}$. [25] also offered a novel and insightful characterization for reliability for controlling unstable processes, named, any-time capacity, defined for the following criterion: $\limsup_{t\to\infty} E[|x_t|^p] < \infty$, for positive moments p. In a related context, [15], [25], [18] and [17] considered the relevance to Shannon capacity. [15] observed that when the moment coefficient goes to zero, Shannon capacity is the right measure for a channel when noise is bounded. A parallel argument is provided by [25] also for bounded noise signals. With a departure from the bounded noise assumption, [17], made the discussion in [25] more explicit and considered a more general model of multi-dimensional

systems driven by an unbounded noise process considering again *stability in probability*. [17] also showed that when the discrete noisy channel has capacity less than $\log_2(|a|)$, there exists no stabilizing scheme, and if the capacity is strictly greater than this number, there exists a stabilizing scheme.

Many network applications and networked control applications require the access of control and sensor information to be observed intermittently. Toward generating a general approach for such problems, [29] and [31] developed random-time state-dependent drift conditions leading to the existence of an invariant distribution possibly with moment constraints, extending the earlier deterministic state dependent results in [21]. Using drift arguments, [30] considered noisy (both discrete and continuous alphabet) channels, [32] considered noiseless channels and [29] considered erasure channels for positive Harris recurrence and ergodicity. We acknowledge [22], which also considered erasure channels and obtained time-varying rate conditions for control over erasure channels. Recently, [11] also considered the stochastic stability over erasure channels, parallel to the results of [29].

The question of *when does a linear system driven by unbounded noise, controlled over a channel (possibly with memory) satisfy Birkhoff's sample path ergodic theorem?*, has not been answered to our knowledge. Also, the finite moment conditions for an arbitrary discrete memoryless channel for a system driven by unbounded noise have not been investigated to our knowledge, except for the bounded noise analysis in [25]. In this paper, we will show that, the results in the literature can be strengthened to asymptotic mean stationarity and ergodicity. We will also consider conditions for finite second moment stability, which had not been studied in the literature for systems with unbounded noise.

II. Stochastic Stability of Dynamical Systems and Random Processes

A. Stationary, Ergodic, and Asymptotically Mean Stationary Processes

In this subsection, we review ergodic theory, in the context of information theory (that is with the transformations being specific to the shift operation). A comprehensive discussion is available in Shields [26] and Gray [9], [10]. Let $\mathbb{X}$ be a complete, separable, metric space. Let $\mathcal{B}(\mathbb{X})$ denote the Borel sigma-field of subsets of $\mathbb{X}$. Let $\Sigma = \mathbb{X}^\infty$ denote the sequence space of all one-sided or two-sided infinite sequences drawn from $\mathbb{X}$. Thus, if $x \in \Sigma$ then $x = \{\ldots, x_{-1}, x_0, x_1, \ldots\}$ with $x_i \in \mathbb{X}$. Let $X_n : \Sigma \to \mathbb{X}$ denote the coordinate function such that $X_n(x) = x_n$. Let T denote the shift operation on Σ, that is $X_n(Tx) = x_{n+1}$. Let $\mathcal{B}(\Sigma)$ denote the smallest sigma-field containing all cylinder sets of the form $\{x : x_i \in B_i, m \leq i \leq n\}$ where $B_i \in \mathcal{B}(\mathbb{X})$, for all integers m, n. Note that $\cap_{n\geq 0} T^{-n} \mathcal{B}(\Sigma)$ is the tail σ-field: $\cap_n \sigma(x_n, x_{n+1}, \cdots)$, since $T^{-n}(A) = \{x : T^n x \in A\}$. Let μ be a stationary measure on $(\Sigma, \mathcal{B}(\Sigma))$ in the sense that $\mu(TB) = \mu(B)$ for all $B \in \mathcal{B}(\Sigma)$. The sequence of random variables $\{x_n\}$ defined on the probability space $(\Sigma, \mathcal{B}(\Sigma), \mu)$ is a stationary process.

Definition 2.1: Let P be the measure on a process. This random process is ergodic if $A = TA$ implies that $P(A) \in \{0, 1\}$.

Definition 2.2: A random process with measure P is $N-$stationary, (cyclo-stationary or periodically stationary with period N) if $P(T^N B) = \mu(B)$ for all $B \in \mathcal{B}(\Sigma)$.

Definition 2.3: A random process is $N-$ergodic if $A = T^N A$ implies that $P(A) \in \{0, 1\}$.

Definition 2.4: We say that a coordinate set $A \in \mathcal{B}(\mathbb{X})$ is coordinate-recurrent if $\sum_{k=0}^{\infty} 1_{\{X_m(T^k(x)) \in A\}} = \infty$, almost surely.

Definition 2.5: A process on a measure $(\Omega, \mathcal{F}, P)$ is asymptotically mean stationary (AMS) if there exists a probability measure $\bar{P}$ such that $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P(T^{-k}F) = \bar{P}(F)$, for all events F. Here $\bar{P}$ is called the stationary mean of P, and is a stationary process.

$\bar{P}$ is stationary since, by definition $\bar{P}(F) = \bar{P}(T^{-1}F)$, for all events F in the tail sigma field for the shift. An $N-$stationary process is AMS, see for example [2], [10] or [9] (Theorem 7.3.1).

Definition 2.6: A random process is second-moment stable almost surely if $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} X_m(T^k(x))^2 < \infty$ $a.s.$.

B. Stability of Markov Chains and Random-Time State Dependent Stochastic Drift Criteria

In this section, we review the theory of stochastic stability of Markov chains. The reader is referred to Meyn and Tweedie [20] for a detailed discussion. We let $\boldsymbol{X} = \{x_t, t \geq 0\}$ denote a Markov chain with state space $\mathbb{X}$, which is assumed to be a complete, separable, metric space, whose Borel σ-field is denoted $\mathcal{B}(\mathbb{X})$. The transition probability is denoted by P, so that for any $x \in \mathbb{X}$, $A \in \mathcal{B}(\mathbb{X})$, the probability of moving from x to A in one step is given by $P(x_{t+1} \in A \mid x_t = x) = P(x, A)$. The n-step transitions are obtained via composition in the usual way, $P(x_{t+n} \in A \mid x_t = x) = P^n(x, A)$, for any $n \geq 1$. A measure π on $\mathcal{B}(\mathbb{X})$ is called invariant if $\pi P = P$. That is, $\int \pi(dx)P(x, A) = \pi(A)$, $A \in \mathcal{B}(\mathbb{X})$. For any initial probability measure ν on $\mathcal{B}(\mathbb{X})$ we can construct a stochastic process with transition law P, and satisfying $x_0 \sim \nu$. We let P_ν denote the resulting probability measure on sample space, with the usual convention for $\nu = \delta_x$ when the initial state is $x \in \mathbb{X}$. When $\nu = \pi$ then the resulting process is stationary.

The following results and the proofs are available in [29], [31]: Let $\tau_z, z \geq 0$ be a sequence of stopping times, measurable on a filtration generated by the state process with $\tau_0 = 0$.

Theorem 2.1: [29] [31] Suppose that $\boldsymbol{X}$ is a φ-irreducible Markov chain. Suppose moreover that there is a function $V : \mathbb{X} \to (0, \infty)$, a petite set C, and constants $\kappa \in (0, 1)$, $b \in \mathbb{R}$, such that the following hold:

$$E[V(x_{\tau_{z+1}}) \mid \mathcal{F}_{\tau_z}] \leq (1-\kappa)V(x_{\tau_z}) + b1_{\{x_{\tau_z} \in C\}}$$
$$E[\tau_{z+1} - \tau_z \mid \mathcal{F}_{\tau_z}] \leq V(x_{\tau_z}), \qquad\qquad z \geq 0. \tag{2}$$

Then, the Markov chain is positive Harris recurrent. The following provides a criterion for finite moments.

Theorem 2.2: [29] [31] Suppose that $\boldsymbol{x}$ is a φ-irreducible Markov chain. Suppose moreover that there are functions $V : \mathbb{X} \to (0, \infty)$, $\delta : \mathbb{X} \to [1, \infty)$, $f : \mathbb{X} \to [1, \infty)$, a petite set C, and a constant $b \in \mathbb{R}$, such that the following hold:

$$E[V(x_{\tau_{z+1}}) \mid \mathcal{F}_{\tau_z}] \leq V(x_{\tau_z}) - \delta(x_{\tau_z}) + b1_{\{x_{\tau_z} \in C\}}$$
$$E\Big[\sum_{k=\tau_z}^{\tau_{z+1}-1} f(x_k) \mid \mathcal{F}_{\tau_z}\Big] \leq \delta(x_{\tau_z}), \qquad\qquad z \geq 0. \tag{3}$$

Then $\boldsymbol{X}$ is positive Harris recurrent, and moreover $\pi(f) < \infty$, with π being the invariant distribution.

III. STOCHASTIC STABILIZATION OVER A DMC

A. Asymptotic Mean Stationarity and $n-$ergodicity

Theorem 3.1: For a controlled linear source given in (1) over a DMC under any causal encoding and controller policy, to satisfy the AMS property under the following condition:

$$\liminf_{T\to\infty} \frac{1}{T} h(x_T) \leq 0,$$

the channel capacity C must satisfy

$$C \geq \log_2(|a|).$$

Proof: See Section V. $\diamond$

The above condition is almost sufficient as well.

Theorem 3.2: For the existence of a compact recurrent set, the following is sufficient: $C > \log_2(|a|)$.

Proof: See Section VI. $\diamond$

For the proof, we consider the following update algorithm. The algorithm and its variations have been used in source coding and networked control literature: See for example the earlier papers [8], [14], [4] (zooming algorithms); and more recent ones [23], [19], [18], [17]. Let n be a given block length. We will consider a class of uniform quantizers, defined by two parameters, with bin size $\Delta > 0$, and an even number $K(n) \geq 2$. The uniform quantizer map is defined as follows: For $k = 1, 2 \ldots, K(n)$,

$$Q_{K(n)}^\Delta(x) = \begin{cases} (k - \frac{1}{2}(K(n)+1))\Delta, \\ x \in [(k-1-\frac{1}{2}K(n))\Delta, (k-\frac{1}{2}K(n))\Delta) \\ \mathcal{Z}, \quad x \notin [-\frac{1}{2}K(n)\Delta, \frac{1}{2}K(n)\Delta), \end{cases}$$

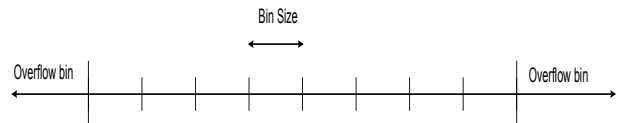where $\mathcal{Z}$ denotes the overflow symbol in the quantizer. See Figure 2.

Fig. 2: A modified uniform quantizer.

At every sampling instant $t = kn, k = 0, 1, 2, \ldots$, the source coder $\mathcal{E}_t^s$ quantizer output symbols in $\mathbb{R} \cup \{\mathcal{Z}\}$ to a set $\mathcal{M}(n) = \{1, 2, \ldots, K(n) + 1\}$. A channel encoder $\mathcal{E}_t^c$ maps the elements in $\mathcal{M}(n)$ to corresponding channel inputs $q_{[kn,(k+1)n-1]} \in \mathcal{M}^n$.

For each time $t = kn - 1, k = 1, 2, 3, \ldots$, the channel decoder applies a mapping $\mathcal{D}_{tn} : \mathcal{M}'^n \to \mathcal{M}(n)$, such that

$$c'_{(k+1)n-1} = \mathcal{D}_{kn}(q'_{[kn,(k+1)n-1]}).$$

Finally, the controller runs an estimator:

$$\hat{x}_{kn} = (\mathcal{E}^s_{kn})^{-1}(c'_{(k+1)n-1}) \times 1_{\{c'_{(k+1)n-1} \neq \mathcal{Z}\}}$$

In our setup, the bin size of the uniform quantizer Δ_t, which acts as the state of the quantizer, is assumed to be a function of the previous state Δ_{t-1} and the past n channel outputs. We assume that the encoder has access to the previous channel outputs. Thus, such a quantizer is implementable at both the encoder and the decoder. With $K(n) > \lceil |a|^n \rceil$, $R(n) = \log_2(K(n) + 1)$, let us define $R'(n) = \log_2(K(n))$ and let $R'(n) > n \log_2(\frac{|a|}{\alpha})$, for some $\alpha, 0 < \alpha < 1$ and $\delta > 0$ with

$$a^2 \geq (|a| + \delta)\alpha. \tag{4}$$

We will consider the following update rules in the controller actions and the quantizers. For $t \geq 0$ and with $\Delta_0 > L$ for some $L \in \mathbb{R}_+$, and $\hat{x}_0 \in \mathbb{R}$, consider:

$$u_t = -1_{\{t=(k+1)n-1\}} \frac{a^n}{b} \hat{x}_{kn},$$
$$\Delta_{(k+1)n} = \Delta_{kn} \bar{Q}(\Delta_{kn}, c'_{(k+1)n-1}), \tag{5}$$

where c' denotes the decoder output variable. If we use $\delta > 0$ and $L > 0$ such that,

$$\bar{Q}(\Delta, c') = (|a| + \delta)^n \quad \text{if} \quad c' = \mathcal{Z},$$
$$\bar{Q}(\Delta, c') = \alpha^n \quad \text{if} \quad c' \neq \mathcal{Z}, \Delta > L$$
$$\bar{Q}(\Delta, c') = 1 \quad \text{if} \quad c' \neq \mathcal{Z}, \Delta \leq L, \tag{6}$$

we will show that a recurrent set exists. We note that the above imply that $\Delta_t \geq L\alpha^n =: L'$ for all $t \geq 0$.

Thus, when the decoder output is the overflow signal, then the quantizer is zoomed out. Otherwise, the quantizer is zoomed in, when the current bin size is not smaller than L.

We note that the stability result for such a scheme can be verified with random-time stochastic drift conditions. This is because of the fact that, the quantizer helps reduce the uncertainty on the system state only when the state is in the *granular* region of the quantizer. The times when the state is in this region are random. Furthermore, correct information reaches the decoder when there is no error in transmission.

We now make the quantizer bin size process countable and the sampled process $\{x_{tn}, \Delta_{tn}\}$ *irreducible*. The following establishes the AMS property.

Theorem 3.3: For an adaptive quantizer satisfying Theorem 3.2, suppose that the quantizer bin sizes are such that their logarithms are integer multiples of some scalar s, and $\log_2(\bar{Q}(\cdot))$ takes values in integer multiples of s. Suppose the integers taken are relatively prime (that is they share no common divisors except for 1). Then the sampled process $\{x_{tn}, \Delta_{tn}\}$ forms a positive Harris recurrent Markov chain at sampling times on the space of admissible quantizer bins and state values. Furthermore, the process $\{x_t, \Delta_t\}$ is AMS.

B. Finite Second Moment

For a given coding scheme with block-length n and a message set $\mathcal{M}(n) = \{1, 2, \ldots, K(n) + 1\}$, and a decoding function $\gamma : \mathcal{M}'^n \to \{1, 2, \ldots, K(n) + 1\}$ define:

$$P_e(n) := \max_{c \in \mathcal{M}(n)} P(\gamma(q'_{[0,n-1]}) \neq c | c \quad is \quad sent),$$

When the block-length is clear from the context, we drop the index n. Due to space constraints, the proofs for this section are not included in the paper.

Theorem 3.4: A sufficient condition for second moment stability over a discrete memoryless channel (DMC) is that:

$$\lim_{n \to \infty} e^{H(\kappa_n)}(P_e(n))^{\kappa_n}(|a| + \delta)^{2n} < 1,$$

$$\lim_{n \to \infty} P_e(n)(|a| + \delta)^{2n} + \alpha^2 < 1,$$

$R'(n) > n \log_2(\frac{|a|}{\alpha})$, $H(\kappa) = -\kappa \log(\kappa) - (1 - \kappa) \log(1 - \kappa)$ and

$$\kappa_n = \frac{1}{\log_{(|a|+\delta)/|a|}(((|a| + \delta)/\alpha) + \frac{1}{n} \log_{\frac{(|a|+\delta)}{|a|}}(2)}.$$

Remark: For a DMC with block length n, Shannon's random coding [5] satisfies: $P_e(n) \leq e^{-nE(R)+o(n)}$, this is uniform for all codewords $c \in \{1, 2, \ldots, M(n)\}$ with c' being the decoder output. Here $\frac{o(n)}{n} \to 0$ as $n \to \infty$ and $E(R) > 0$ for $0 < R < C$. Thus, under the above conditions, the exponent under random coding should satisfy $E(R) > \frac{2 \log_2(|a|+\delta)}{\kappa}$. $\diamond$

Assumption A0 The DMC has non-zero zero error capacity. $\diamond$

Theorem 3.5: Under Assumption **A0**, a sufficient condition for second moment stability is:

$$\lim_{n \to \infty} P_e(n)(|a| + \delta)^{2n} < 1,$$

and rate $R'(n) > n \log_2(\frac{|a|}{\alpha})$.

Remark: The non-zero zero-error capacity result is related to the notion of any-time capacity proposed by Sahai and Mitter for bounded noise systems: A channel is $\alpha-$ any-time reliable for a sequential coding scheme if: $P(m^i(t) \neq m^i(t)) \leq D2^{-\alpha(t-i)}$ for all i, t. Here m^i is a message sent at time 1, estimated at time t using received channel outputs up until time i. Sahai and Mitter also considered a block-coding setup, for the case when the noise is bounded, and were able to obtain a similar rate/reliability criterion. It is worth emphasizing that, the reliability for sending one symbol for the under-zoom phase allows an improvement in the reliability requirements drastically. $\diamond$

IV. CHANNELS WITH MEMORY

Definition 4.1: Let **Class A** be the set of channels which satisfy the Markov chain condition:

$$q'_t \leftrightarrow q_t, q_{[0,t-1]}, q'_{[0,t-1]} \leftrightarrow x_{[0,t]},$$

for all $t \geq 0$ and whose capacity with feedback is given by:

$$C = \lim_{T \to \infty} \max_{\{P(q_t | q_{[0,t-1]}, q'_{[0,t-1]}), 0 \leq t \leq T-1\}}$$
$$\frac{1}{T} I(q_{[0,T-1]} \to q'_{[0,T-1]}),$$

where $I(q_{[0,T-1]} \to q'_{[0,T-1]})$ is defined to be

$$\sum_{t=1}^{T-1} I(q_{[0,t]}; q'_t | q'_{[0,t-1]}) + I(q_0; q'_0),$$

and is called the directed mutual information (see [28]). Such a class includes discrete memoryless channels, finite state stationary Markov channels which are indecomposable, and a class of stationary ergodic symmetric channels with memory, for example see [28].

Theorem 4.1: For a linear system controlled over a noisy channel with memory with feedback in **Class A**, if the channel capacity is less than $\log_2(|a|)$ then the AMS property under the following condition $\liminf_{T\to\infty} \frac{1}{T} h(x_T) \leq 0$, cannot be satisfied under any policy.

V. SKETCH OF PROOF OF THEOREM 3.1 AND 4.1

Let us define

$$R_T = \max_{\{P(q_t|q_{[0,t-1]}, q'_{[0,t-1]}), 0 \leq t \leq T-1\}}$$
$$\frac{1}{T} \sum_{t=0}^{T-1} I(q'_t; q_{[0,t]} | q'_{[0,t-1]}). \qquad (7)$$

Observe that for $t > 0$:

$$I(q'_t; q_{[0,t]} | q'_{[0,t-1]}) = H(q'_t | q'_{[0,t-1]}) - H(q'_t | q_{[0,t]}, q'_{[0,t-1]})$$
$$= H(q'_t | q'_{[0,t-1]}) - H(q'_t | q_{[0,t]}, x_t, q'_{[0,t-1]})$$
$$\geq H(q'_t | q'_{[0,t-1]}) - H(q'_t | x_t, q'_{[0,t-1]}) = I(x_t; q'_t | q'_{[0,t-1]})$$

It follows that since R_T is assumed to have a limit,

$$\lim_{T\to\infty} R_T$$
$$\geq \limsup_{T\to\infty} \frac{1}{T} \left(\sum_{t=1}^{T-1} \left(I(x_t; q'_t | q'_{[0,t-1]}) \right) + I(x_0; q'_0) \right)$$
$$= \limsup_{T\to\infty} \frac{1}{T} \left(\sum_{t=1}^{T-1} h(x_t | q'_{[0,t-1]}) - h(x_t | q'_{[0,t]}) \right)$$
$$\geq \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T-1} \left(h(ax_{t-1} + bu_{t-1} + d_{t-1} | q'_{[0,t-1]}) \right.$$
$$\left. - h(x_t | q'_{[0,t]}) \right)$$
$$\geq \log_2(|a|) - \liminf_{T\to\infty} \left(\frac{1}{T} h(x_{T-1}) \right)$$

In the above, we have skipped a number of steps due to space constraints. The inequalities follow from the fact that conditioning reduces entropy, control actions are uniquely determined from channel outputs and and that the system noise process is independent. By the hypothesis,

$$\liminf_{t\to\infty} \frac{1}{t} h(x_t) \leq 0,$$

it must be that $\lim_{T\to\infty} R_T \geq \log_2(|a|)$. Thus, the capacity also needs to satisfy this bound. $\diamond$

VI. PROOF OF THEOREM 3.2:

For a DMC with block length n, it is known that there exist coding schemes, by Shannon's random coding [5], which satisfy:

$$P_e(n) := P(c' \neq c | c \quad is \quad transmitted) \leq e^{-nE(R)+o(n)},$$

uniformly for all codewords $c \in \{1, 2, \ldots, M(n)\}$ with c' being the decoder output. Here $\frac{o(n)}{n} \to 0$ as $n \to \infty$ and $E(R) > 0$ for $0 < R < C$.

Toward the proof, we will first obtain a supporting result.

Lemma 6.1: The sampled process (x_{tn}, Δ_{tn}) is a Markov chain.

The above follows from the observations that, the channel is memoryless, the encoding is Shannon's random-coding and the control policies use the channel outputs received in the last block.

Let us define $h_t := \frac{x_t}{\Delta_t 2^{R'(n)-1}}$. We will say that the quantizer is perfectly zoomed when $|h_1| \leq 1$, and under-zoomed otherwise. Consider the following sets: $C_x = \{x : |x| \leq F\}$, $C_h = \{h : |h| \leq 1\}$, with a sufficiently large F value to be studied below. We will study the recurrence properties for such a set. Define a sequence of stopping times for the perfect-zoom case with (where the initial state is perfectly zoomed) $z \in \mathbb{Z}_+$:

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

Denote for $k \in \mathbb{N}$, $\Theta_k := P(\tau_{z+1} - \tau_z \geq kn | \Delta_{\tau_z}, h_{\tau_z})$. Without any loss, let $z = 0$ and $\tau_0 = 0$, so that $\Theta_k = P(\tau_1 \geq k | \Delta_{\tau_0}, h_{\tau_0})$. We have the following key result, the proof of which is omitted.

Lemma 6.2: The discrete probability distribution

$$P(\tau_{z+1} - \tau_z | \Delta_{\tau_z}, h_{\tau_z})$$

is asymptotically, in the limit of large Δ_{τ_z}, dominated (majorized) by a geometrically distributed measure.

$$P(\tau_{z+1} - \tau_z \geq kn | \Delta_{\tau_z}, h_{\tau_z})$$
$$\leq \Xi(\Delta_{\tau_z}, n) e^{H(\kappa_n)(k-1)} e^{-(nE(R)-o(n))\kappa_n(k-1)},$$

where $\Xi(\Delta_{\tau_z}, n) < \infty$ and $\Xi(\Delta_{\tau_z}, n) \to 0$ as $\Delta_{\tau_z} \to \infty$ for every fixed n and

$$\kappa_n = \frac{1}{\log_{(|a|+\delta)/|a|}((|a|+\delta)/\alpha) + \frac{1}{n}\log_{\frac{(|a|+\delta)}{|a|}}(2)}.$$

Thus,

$$E[\tau_1 | \Delta_0] = \sum_{k=1}^{\infty} P(\tau_1 \geq k) \leq \sum_{t=1}^{\infty} P(\lceil \tau/n \rceil n \geq kn)$$
$$\leq n \Xi(\Delta_0, n) \frac{1}{1 - (e^{H(\kappa)} e^{-n\kappa(E(R) - \frac{o(n)}{n})})} =: K' < \infty.$$

By the strong Markov property (x_{τ_z}, h_{τ_z}) is also a Markov chain. We now apply random-time drift from Theorem 2.1

below. The probability that $\tau_{z+1} \neq \tau_z + n$, is upper bounded by the probability:

$$P_e + (1 - P_e)2P\left(\bar{d} \geq (2^{R'(n)}(\frac{\alpha}{|a|})^n - 1)\Delta_0/2\right)$$

$$\leq P_e + 2P\left(\bar{d} \geq (2^{R'(n)}(\frac{\alpha}{|a|})^n - 1)\Delta_0/2\right) =: \Upsilon(\Delta_{\tau_z})$$

If $\tau_{z+1} \neq \tau_z + n$, then this means that the error is increasing on average and the system is once-again under-zoomed at time $t = \tau_z + n$: $x_{\tau_{z+1}} = ax_{\tau_z} + d_{\tau_z}$ with $\Delta_{\tau_{z+1}} \geq \alpha^n \Delta_{\tau_z}$ (when $\Delta_{\tau_z} \geq L$). The inequality is due to the observation that, if there is an error in the transmission, the receiver will enlarge the quantizer bin. We now show that, there exist $\psi > 0$, $|G| < \infty$ such that

$$E[\log(\Delta_{\tau_{z+1}}^2)|\Delta_{\tau_z}, h_{\tau_z}] \leq \log(\Delta_{\tau_z}^2) - \psi + G1_{\{|\Delta_{\tau_z}| \leq F\}}. \quad (8)$$

It follows that,

$$E[\log(\Delta_{\tau_{z+1}}^2)|\Delta_{\tau_z}, h_{\tau_z}]$$
$$\leq (1 - \Upsilon(\Delta_{\tau_z}))(1 - P_e)\left(2n\log(\alpha) + \log(\Delta_{\tau_z}^2)\right)$$
$$+(\Upsilon(\Delta_{\tau_z}) + P_e - P_e\Upsilon(\Delta_{\tau_z}))$$
$$E\left[2(\tau_{z+1} - \tau_z)\log(|a| + \delta) + 2\log(\Delta_{\tau_z})\right]$$

It now follows from (8) that for every fixed n, if $\frac{R'(n)}{n} > \log_2(|a|/\alpha)$, $\lim_{\Delta \to \infty} \Upsilon(\Delta_{\tau_z}) = P_e(n)$. For (8) to hold, it suffices that the following equation is satisfied for large enough Δ_{τ_z} and n values, for some $\psi > 0$:

$$(\Upsilon(\Delta_{\tau_z}) + P_e(\Delta_{\tau_z}))\left\{(2K')\log(|a| + \delta)\right\}$$
$$+(1 - \Upsilon(\Delta_{\tau_z}))(1 - P_e)2n\log(\alpha) \leq -\psi < 0. \quad (9)$$

Thus, for a set off a compact set, say $\{\Delta_0 : \Delta_0 \geq F\} > 0$, for some large enough F, by $\lim_{n \to \infty} P_e(n) = 0$, and for some sufficiently large finite n, there exists a drift towards a compact set. This implies recurrence by Theorem 2.1.

REFERENCES

[1] R. Bansal and T. Başar, "Simultaneous design of measurement and control strategies for stochastic systems with feedback," *Automatica*, 25(9):679–694, 1989.

[2] R. A. Boyles and W.A. Gardner, "Cycloergodic properties of discrete-parameter non-stationary stochastic processes", *IEEE IT*, Vol. IT-29, 1983, pp. 105-114.

[3] V. S. Borkar, "Convex Analytic Methods in Markov Decision Processes," *Handbook of Markov Decision Processes: Methods and Applications*, Kluwer, Boston, MA.

[4] R. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems", *IEEE TAC*, vol. 45, pp. 1279-1289, July 2000.

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

[6] J. Freudenberg, R. Middleton, and V. Solo, "Stabilization and disturbance attenuation over a Gaussian communication channel, *IEEE Trans. Automat. Control*, vol. 55, no. 3, pp. 795 799, 2010.

[7] J. Flum and M. Grohe, *Parameterized Complexity Theory*, Springer, 2006.

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

[9] R. M. Gray, *Probability, Random Processes, and Ergodic Properties*, New York: Springer-Verlag, 1988.

[10] R. M. Gray and J. C. Kieffer, "Asymptotically mean stationary measures," *Ann. Prob.*, vol. 8, pp. 962-973, 1980.

[11] A. Gurt and G. Nair, "Internal stability of dynamic quantised control for stochastic linear plants", *Automatica*, vol. 45, pp. 1387-1396, June 2009.

[12] T. Hashimoto and S. Arimoto, "On the rate-distortion function for the nonstationary Gaussian autoregressive process," *IEEE IT*, vol. 26, pp. 478-480, 1980.

[13] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints - part II: Stabilization with limited information feedback, *IEEE TAC*, 42(9):1294–1299, 1997.

[14] J. C. Kieffer, "Stochastic stability for feedback quantization schemes", *IEEE IT*, vol. 28, pp. 248-254, March 1982.

[15] N. C. Martins, M. A. Dahleh and N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link", *IEEE TAC*, volume 51, pp. 438 - 447, March 2006.

[16] N. C. Martins and M. A. Dahleh, "Feedback control in the presence of noisy channels: 'Bode-Like' fundamental limitations of performance," *IEEE TAC*, vol. 53, pp. 1604 - 1615, Aug. 2008.

[17] A. S. Matveev, "State estimation via limited capacity noisy communication channels, *Mathematics of Control, Signals, and Systems*, vol. 20, pp. 135, 2008.

[18] A. S. Matveev and A. V. Savkin, *Estimation and Control over Communication Networks*, Birkhauser Boston, 2008.

[19] A. S. Matveev and A. Savkin, "An analogue of Shannon information theory for detection and stabilization via noisy discrete communication channels," *SICON*, vol. 46, pp. 13231367, 2007.

[20] S. P. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, Springer Verlag, London (1993).

[21] S. P. Meyn and R. Tweedie, "Stability of Markovian processes I: Criteria for discrete-time chains", *Adv. Appl. Prob.*, vol. 24, pp. 542-574, 1992.

[22] P. Minero, M. Franceschetti, S. Dey and G. Nair, "Data rate theorem for stabilization over time-varying feedback channels", *IEEE TAC*, 54(2):243-255, 2009.

[23] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates", *SICON*, vol. 43, pp. 413 - 436, July 2004.

[24] A. Sahai, "Why delay and block length are not the same thing for channel coding with feedback," in *Proc. Information Theory and Applications*, January 2006, San Diego, CA.

[25] 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 IT*, vol. 52, pp. 3369–3395, Aug. 2006.

[26] P. C. Shields, *The ergodic theory of discrete sample paths*, AMS, Providence RI, 1996.

[27] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE TAC*, 49(7):1056-1068, 2004.

[28] S. Tatikonda and S. Mitter, "The capacity of channels with feedback, *IEEE IT*, vol. 55, no. 1, pp. 323 349, January 2009.

[29] S. Yüksel and S. P. Meyn, "Random-time, state-dependent stochastic drift for Markov Chains and application to stochastic stabilization over erasure channels", *IEEE TAC*, under review.

[30] S. Yüksel and T. Başar, "Control over noisy forward and reverse channels," *IEEE TAC*, to appear, 2011.

[31] S. Yüksel, "A random time stochastic drift result and application to stochastic stabilization over noisy channels", in *Allerton Conf.*, IL, October 2009.

[32] S. Yüksel, "Stochastic stabilization of noisy linear systems with fixed-rate limited feedback", *IEEE TAC*, vol. 55, pp. 2847-2853, December 2010.