2018 IEEE Conference on Decision and Control (CDC)
Miami Beach, FL, USA, Dec. 17-19, 2018

Exact minimum number of bits to stabilize a linear system

Victoria Kostina, Yuval Peres, Gireeja Ranade, Mark Sellke

***Abstract*—We consider an unstable scalar linear stochastic system, $X_{n+1} = aX_n + Z_n - U_n$, where $a \geq 1$ is the system gain, Z_n's are independent random variables with bounded α-th moments, and U_n's are the control actions that are chosen by a controller who receives a single element of a finite set $\{1, \ldots, M\}$ as its only information about system state X_i. We show that $M = \lfloor a \rfloor + 1$ is necessary and sufficient for β-moment stability, for any $\beta < \alpha$. Our achievable scheme is a uniform quantizer of the zoom-in / zoom-out type. We analyze its performance using probabilistic arguments. We prove a matching converse using information-theoretic techniques. Our results generalize to vector systems, to systems with dependent Gaussian noise, and to the scenario in which a small fraction of transmitted messages is lost.**

I. Introduction

We study the tradeoff between stabilizability of a linear stochastic system and the coarseness of the quantizer used to represent the state. The evolution of the system is described by

$$X_{n+1} = aX_n + Z_n - U_n, \tag{1}$$

where constant $a \geq 1$; X_1 and $Z_1, Z_2, \ldots$ are independent random variables with bounded α-th moments, and U_n is the control action chosen based on the history of quantized observations. More precisely, an *M-bin causal quantizer-controller* for $X_1, X_2, \ldots$ is a sequence $\{\mathsf{f}_n, \mathsf{g}_n\}_{n=1}^{\infty}$, where $\mathsf{f}_n \colon \mathbb{R}^n \mapsto [M]$ is the encoding (quantizing) function, and $\mathsf{g}_n \colon [M]^n \mapsto \mathbb{R}$ is the decoding (controlling) function, and $[M] \triangleq \{1, 2, \ldots, M\}$. At time i, the controller outputs

$$U_n = \mathsf{g}_n(\mathsf{f}_1(X_1), \mathsf{f}_2(X^2), \ldots, \mathsf{f}_n(X^n)). \tag{2}$$

The fundamental operational limit of quantized control this paper looks at is the minimum number of quantization bins to achieve β-moment stability:

$$M^\star_\beta \triangleq \min \left\{ M \colon \exists\, M\text{-bin causal quantizer-controller} \right. \\ \left. \text{such that} \quad \limsup_n \mathbb{E}\left[|X_n|^\beta\right] < \infty \right\}, \tag{3}$$

where $0 < \beta < \alpha$ is fixed.

The main result of the paper is the following theorem.

Theorem 1. *Let X_1, Z_n in (1) be independent random variables with bounded α-moments. Then for any $0 < \beta < \alpha$, the minimum number of quantization points to achieve β-moment stability is*

$$M^\star_\beta \leq \lfloor a \rfloor + 1. \tag{4}$$

The result of Theorem 1 is tight, as the following converse shows.

Theorem 2. *Let X_1, Z_n in (1) be independent random variables. Let $h(X_1) > -\infty$, where $h(X) \triangleq -\int_{\mathbb{R}} f_X(x) \log f_X(x) dx$ is the differential entropy. Then, for all $\beta > 0$,*

$$M^\star_\beta \geq \lfloor a \rfloor + 1. \tag{5}$$

In the special case of unstable scalar systems with bounded disturbances, i.e. $|Z_n| \leq B$ a.s., the results of Theorem 1 and Theorem 2 are well known from [1], [2], where it was shown that a simple uniform quantizer with the number of quantization bins in (4) stabilizes such systems. That corresponds to the special case $\alpha = \beta = \infty$.

The converse in the special case of $\beta = 2$ was proved in [3], where it was shown that it is impossible to achieve second moment stability in the system in (1) using a quantizer-controller with the number of bins $< \lfloor a \rfloor + 1$. This implies the validity of Theorem 2 for $\beta \geq 2$.

Nair and Evans [3] showed that time-invariant fixed-rate quantizers are unable to attain bounded cost if the noise is unbounded [3], regardless of their rate. The reason is that since the noise is unbounded, over time, a large magnitude noise realization will inevitably be encountered, and the dynamic range of the quantizer will be exceeded by a large margin, not permitting recovery. This necessitates the use of adaptive quantizers of zooming type originally proposed by Brockett and Liberzon [4]. Such quantizers "zoom out" (i.e. expand their quantization intervals) when the system is far from the target and "zoom in" when the system is close to the target. They are also known to achieve input-to-state stability for linear systems with bounded disturbances [5]. Nair and Evans [3] proposed a stabilizing quantization scheme in which the number of quantization levels is finite at each step but varies with time, and showed that it suffices to use $\log_2 a$ bits on average to achieve second moment stability, as long as the system noise has bounded $2 + \epsilon$

V. Kostina (vkostina@caltech.edu) is with California Institute of Technology, Pasadena, CA. Y. Peres (peres@microsoft.com) and G. Ranade (gireeja@eecs.berkeley.edu) are with Microsoft Research, Redmond, WA. M. Sellke (msellke@gmail.com) is with University of Cambridge, Cambridge, UK. This work was supported in part by the National Science Foundation (NSF) under Grant CCF-1751356, and by the Simons Institute for the Theory of Computing.

978-1-5386-1395-5/18/$31.00 ©2018 IEEE

moment, for some $\epsilon > 0$. In this paper, we do not allow the communication rate to vary with time: our communication channel noiselessly transmits one of M messages at each time step.

The stabilizing performance of fixed-rate quantizer-controller pairs that fit the setting of this paper was studied by Yüksel [6], who proved that for Gaussian system noises,

$$M_2^\star \leq \lfloor a \rfloor + 2. \tag{6}$$

Yuksel's result leaves a gap of 1 between the upper and lower bounds. The gap might seem insignificant, especially if the gain a is large, but the gain of many realistic systems is in $[1, 2)$. The state of the art thus leaves open the question of whether such systems are stabilizable with a single-bit quantizer.

This paper resolves that question in the affirmative. We construct a controller that stabilizes linear systems with $a \in [1, 2)$ while using only 1 bit per sample to choose its control action. We show that β-moment stability is achievable as long as system noise has bounded α-moment, for some $\alpha > \beta$. The scheme and its analysis extend naturally to higher a's.

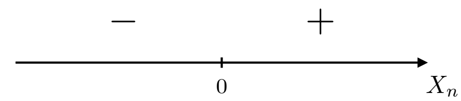
Note that both schemes [3], [6] rely on the special treatment of the overflow bins of the quantizer, which are its unbounded leftmost and rightmost bins. Once the quantizer overflows, the controllers of [3], [6] enter their zoom-out modes. Such controller strategies cannot be used with single-bit quantizers, because single-bit quantizers are always in overflow. Furthermore, as Yüksel [6] discusses, the special treatment of the overflow bin is what causes the extra 1 in (6).

In Section II, we describe our achievable scheme and present a roadmap to its technical analysis. In Section III, we give a proof of the converse in Theorem 2. Our results generalize to constant-length time delays, to control over communication channels that drop a small fraction of packets, to systems with dependent Gaussian noise, and to vector systems. These extensions are presented in Section IV. Due to space considerations most proofs are relegated to the Arxiv version [[7]][1]; wherever possible we provide the main ideas behind those proofs and point to the mathematical tools that we use.

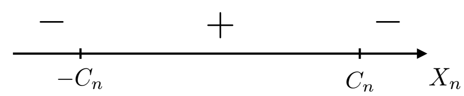
We conclude the introduction with a technical remark.

Remark 1. The assumptions in Theorem 2 that the differential entropy of X_1 is not $-\infty$ implies that X_1 must have a density. That assumption is not superficial. For example, consider $Z_i \equiv 0$ and X_1 uniformly distributed on the Cantor set, and $a = 2.9$. Clearly this system can be stabilized with 1 bit, by telling the controller at each step the undeleted third of the interval the state is at. This is lower than the result of Theorem 1, which states that $M_\beta^\star$ would be 3 if X_1 had a density. Beyond distributions with densities, we conjecture that $M_\beta^\star$ will depend on the Hausdorff dimension of the probability measure of X_1.

[1]We use double brackets [[7]] to refer to the full version of this paper.



(a) Sign test



(b) Magnitude test

Fig. 1. The binary quantizer uses two kinds of tests on a schedule determined by the previous $\pm$'s to produce the next $+$ or $-$.

II. Achievable scheme

A. The idea

Here we explain the idea of our achievable scheme. For readability we focus on the case $a \in [1, 2)$ and show that the system can be controlled with 1 bit. In this case we will be able to restrict to two types of tests, a *sign test* and a *magnitude test* (see Fig. 1), simplifying the description of our scheme. The straightforward extension to an arbitrary $a \geq 1$, in which the sign test is replaced by a uniform quantizer, is found in Section II-E below.

In the case of bounded noise a uniform time-invariant quantizer deterministically keeps X_n bounded [1], [2]. Indeed, when $|Z_n| \leq B$, $n = 1, 2, \ldots$ and $|X_1| \leq C_1$, if $C_1 \geq \frac{B}{1-a/2}$ one can put

$$C_2 \triangleq (a/2)C_1 + B \leq C_1, \tag{7}$$

and putting further $C_{n+1} \triangleq (a/2)C_n + B$, we obtain a monotonically decreasing to $\frac{B}{1-a/2}$ sequence numbers $\{C_n\}_{n=1}^\infty$. Setting

$$U_n = (a/2)C_n \operatorname{sgn}(X_n) \tag{8}$$

requires only 1 bit of knowledge about X_n (i.e., its sign). If $|X_n| \leq C_n$ then

$$|X_{n+1}| \leq (a/2)C_n + B = C_{n+1}, \tag{9}$$

and

$$\limsup_{n\to\infty} |X_n| \leq \frac{B}{1 - a/2}. \tag{10}$$

Actually, this is the best achievable bound on the uncertainty about the location of X_n, as a volume-division argument shows [8], [9].

When Z_n merely have bounded α-moments the above does not work because a single large value of Z_n will cause the system to explode. However we can use the idea of the bounded case with the following modification. Most of the time, in *normal*, or *zoom-in*, mode, the controller assumes the X_n are bounded by constants C_n and runs the above procedure, but occasionally, on a schedule, the controller performs a *magnitude test* and sends a bit whose sole purpose is to test whether the X_n are staying within desired bounds.

If the answer is affirmative, the controller reverts back to the normal mode, and otherwise, it enters the *emergency*, or *zoom-out*, mode, whose purpose is to look for the X_n in exponentially larger intervals until it is located, at which point it returns to the zoom-in mode while still occasionally checking for anomalies. We will show that all this can be accomplished with only 1 bit per controller action.

The intuition behind our scheme is the following. At any given time, with high probability X_n is not too large. Thus, the emergencies are rare, and when they do occur, the size of the uncertainty region tends to decrease exponentially. The zoom-in mode operates almost exactly as in the bounded case, except that we choose B large enough to diminish the probability that the noise exceeds it. We now proceed to making these intuitions precise in Section II-B.

B. The Algorithm

Here we describe the algorithm precisely and then outline the proof of why it works. Specifically, we consider the setting of Theorem 1 with $a \in [1, 2)$ and Z_n with bounded α-moments. We find U_n - a function only of the sequence of bits received from the quantizer - that achieves β-moment stability, for $0 < \beta < \alpha$.

First we prepare some constants. We fix $B \geq 1$ large enough. We set the *probing factor* $P = P(\alpha, \beta)$ - a large positive constant (how large will be explained below, but roughly P blows up as $\beta \uparrow \alpha$). We fix a small $\delta > 0$ and a large enough $k = k(a)$ so that

$$(a/2)^{k-1} a \leq 1 - 3\delta. \tag{11}$$

We proceed in "rounds" of at least $k+1$ moves, k moves in normal (zoom-in) mode and $k+1$'th move to perform a magnitude test to see whether X_n escaped the desired bounds. If that magnitude test comes back normal, the round ends; otherwise the controller enters the emergency (zoom-out) mode, whose duration is variable and which ends once the controller learns a new (larger) bound on X_n. In normal mode, we use the update rule in (8), where $C_n \geq B$ is positive. In the emergency mode, $U_n \equiv 0$ while C_n grows exponentially. A precise description of the operation of the algorithm is given below.

1) At the start of a round at time-step m, $|X_m| \leq C_m$, the controller is silent, $U_m = 0$, and $X_{m+1} = aX_m + Z_m$. Set

$$C_{m+1} = aC_m + B, \tag{12}$$

and for each $i \in \{2, \ldots, k\}$,

$$C_{m+i} = \frac{a}{2} C_{m+i-1} + B. \tag{13}$$

In this normal mode operation, the quantizer sends a sequence of signs of X_n (see Fig. 1(a)), while the controller applies the controls (8) successively to $X_m, \ldots, X_{m+k-1}$. This normal mode operation will keep X_{m+i} bounded by C_{m+i} unless some Z_{m+i} is atypically large.

2) The quantizer applies the magnitude test to check whether $|X_{m+k}| \leq C_{m+k}$ (see Fig. 1(b)). If $|X_{m+k}| \leq C_{m+k}$, we return to step 1. If $|X_{m+k}| > C_{m+k}$, this means some Z_{m+i} was abnormally large; the system has blown up and we must do damage control. In this case we enter *emergency (zoom-out) mode* in Step 3 below.

3) In emergency mode, we repeatedly perform silent ($U_{m+k+j} \equiv 0$) magnitude tests via

$$C_{m+k+j} = P\, C_{m+k+j-1} = P^j C_{m+k} \quad j \geq 0 \tag{14}$$

until the first time τ that the magnitude test is passed, i.e.

$$\tau \triangleq \inf\left\{j \geq 0 \colon |X_{m+k+j}| \leq C_{m+k+j}\right\}. \tag{15}$$

We then set $m \leftarrow m + k + \tau$ and return to Step 1.

The controller is silent at the start of a round because it does not know the sign of X_m. Each round thus includes one silent step at the start, and $\tau \geq 0$ silent steps of the emergency mode.

C. Overview of the Analysis

We analyze the result of each round. At the start of each round m we know that X_m is contained within interval $[-C_m, C_m]$. We will show that when C_m is large, the uncertainty interval tends to decrease by a constant factor each round.

At the start of the round, $|X_m| \leq C_m$. Assume that for each $i \in \{0, 1, \ldots, k\}$, we have

$$|Z_{m+i}| \leq B. \tag{16}$$

and thus

$$|X_{m+i}| \leq C_{m+i}. \tag{17}$$

In particular, applying (11), (12) and (13), we bound the state at the end of the round as

$$\begin{aligned} |X_{m+k}| &\leq C_{m+k} && (18)\\ &\leq (1 - 3\delta)\, C_m + \frac{B}{1 - a/2}, && (19) \end{aligned}$$

which means that $C_{m+k} \leq C_m$, provided that $C_m \geq \frac{B}{3\delta(1-a/2)}$. Thus, even starting with the silent step we have successfully decreased C_m, provided that it was large enough.

What if (16) fails to hold? Because the Z_i have bounded α-moments, by the union bound and Markov's inequality, the chance (16) fails is at most

$$\mathbb{P}\left[\cup_{i=0}^{k} \{|Z_{m+i}| > B\}\right] \leq (k+1)\, \mathbb{E}\left[|Z|^\alpha\right] B^{-\alpha}. \tag{20}$$

In this case, we show that we can control the blow-up to avert a catastrophe. Recall that in emergency mode our procedure will take exponentially growing C_n (see (14)) so that we will soon observe that $|X_n| \leq C_n$. The controller then exits emergency mode and returns to the normal mode, starting a new round at time step n. Using boundedness of α-moments of Z_i, we would like to show that the chance that on step $n = m + k + j$ this *fails* is exponentially small in j. This will allow us to see that in each round starting at $X_m \in [-C_m, C_m]$, there is a high chance to shrink the magnitude of the state and a small chance to grow larger. In the next section we outline how to obtain precise moment control.

D. Technical tools and proof roadmap

Here we introduce the technical tools and give the roadmap to the proof of Theorem 1 for the case $a \in [1, 2)$. Due to space considerations full details are given in the Arxiv version [[7]].

The tools in Proposition 1 and Lemma 1 below, that will be instrumental in controlling the tails of the accumulated noise, are proved in [[7]] using elementary probability arguments.

Proposition 1. *If the random variable Z has finite α-moment, then*

$$t^{\alpha}\mathbb{P}[|Z| > t] \tag{21}$$

are bounded in t. Conversely, if (21) are bounded in t then Z has a finite β-moment for any $0 < \beta < \alpha$.

Lemma 1. *Suppose $a > 1$ is fixed and Z_i are (arbitrarily coupled) random variables with uniformly bounded absolute α moments. Then the random variables*

$$\tilde{Z}_j \triangleq \sum_{i=0}^{j} a^{-i} Z_i \tag{22}$$

also have uniformly bounded absolute α-moments.

The bound in Lemma 2 below is proved in [[7]] by considering the evolution of the system over $k + 1 + \tau$ steps, where τ (15) determines the end of the round. Note that τ is a stopping time of the filtration generated by $\{X_n\}$.

Lemma 2. *Fix $B, P > 0$ and consider our algorithm described in Section II-B with these parameters. Suppose that time-step m is the start of a round, so that the round ends on time-step $m+k+\tau$. For all $1 < a < 2$ and for all $0 \le j \le \tau$, it holds that*

$$\begin{aligned} &\max\{|X_{m+1}|, \dots, |X_{m+k+j}|, C_{m+k+j}\} \qquad (23)\\ &\le Pa^{k+j}\left(2C_m + \frac{a\,B}{(2-a)(a-1)} + \sum_{\ell=0}^{k+j-1} a^{-\ell-1}|Z_{m+\ell}|\right), \end{aligned}$$

Recalling the choices of k and δ in (11), we prepare some additional constants.

- Fix $\Delta < \alpha - \beta$ an arbitrary fixed constant, e.g. $\Delta = \frac{\alpha-\beta}{3}$, so that
$$\beta = \alpha - 3\Delta. \tag{24}$$
- Fix P large enough so that
$$P/a \ge \max\left\{\left(\frac{a}{1-\delta}\right)^{\alpha-\Delta},\ 2^k,\ \frac{a^{k+1}}{2(a-1)}\right\}. \tag{25}$$

Suppose that time-step m is the start of a round, so that the round ends on time-step $m + k + \tau$, with stopping time $\tau = 0$ usually.

We define a modified sequence $\tilde{X}_n$ through, for $1 \le i \le k + \tau$,

$$\begin{aligned} \tilde{X}_{m+i} \triangleq \left(\frac{1}{1-\delta}\right)^{\tau-|i-k|_+} \qquad (26)\\ \max\{|X_{m+k}|, \dots, |X_{m+k+\tau}|, C_{m+k+\tau}\}, \end{aligned}$$

where $|\cdot|_+ \triangleq \max\{0, \cdot\}$. Clearly this definition ensures that

$$|X_{m+k+j}| \le \tilde{X}_{m+k+j} \quad 0 \le j \le \tau. \tag{27}$$

Furthermore, for all $1 \le i \le k - 1$, there exists universal constants K_1, K_2, K_3 that depend on a, k and B such that [[7]]

$$\mathbb{E}\left[|X_{m+i}|^{\beta}\right] \le K_1\,\mathbb{E}\left[\tilde{X}^{\beta}_{m+k}\right] + K_2\,\mathbb{E}\left[\tilde{X}^{\beta}_{m}\right] + K_3. \tag{28}$$

Inequalities (27) and (28) together mean that to establish $\limsup_n \mathbb{E}\left[|X_n|^{\beta}\right] < \infty$, it is sufficient to prove

$$\limsup_n \mathbb{E}[\tilde{X}^{\beta}_n] < \infty. \tag{29}$$

To establish (29), we will show that

$$\mathbb{E}[\tilde{X}^{\beta}_{m+1}] \le (1-\delta)^{\beta}\mathbb{E}[\tilde{X}^{\beta}_m] + K, \tag{30}$$

where $K = K(P, k, \delta)$ is a constant that may depend on P, k, δ (but is independent of m). Since by definition (26), $\tilde{X}_{m+i} \le \tilde{X}_{m+1} \quad i = 2, \dots, k+\tau$, (30) ensures that $\limsup_n \mathbb{E}[\tilde{X}^{\beta}_n]$ is bounded above by $\frac{K}{1-(1-\delta)^{\beta}}$.

The intuition behind the definition for $\tilde{X}_n$ is as follows. We want to construct a dominating sequence $\tilde{X}_n$ with the expected decrease property in (30). During emergency mode, the original sequence X_n may increase on average during rounds. The sequence $\tilde{X}_n$ in (26) takes the potential increase during each round up front, achieving the desired expected decrease property. We will see that P in (25) is chosen so that the constant-factor decrease of the system is preserved when switching between rounds.

To show (30), we define the filtration $\mathcal{F}_n$ as follows: $\mathcal{F}_n$ is the σ-algebra generated by the sequences $Z_1, Z_2, \dots, Z_{n-1}$ and $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n$. Unless n is the end of a round, knowledge of $\tilde{X}_n$ involves a peek into the future, so $\mathcal{F}_n$ encompasses slightly more information than the naive notion of "information up to time n".

Define

$$Y_n \triangleq \frac{\tilde{X}_{n+1}}{\tilde{X}_n + \frac{B}{(1-a/2)(1-3\delta)}}. \tag{31}$$

We show (30) by the means of the following two statements, where m is the transition between rounds:

(a) For sufficiently large k and P in (11) and (25), respectively, it holds that [2]

$$\mathbb{P}\left[Y_m \ge t|\mathcal{F}_m\right] = O\left(t^{-(\alpha-\Delta)}\right), \tag{32}$$

[2]Throughout this section, the implicit constants $O()$ may depend on P, k, δ (but are independent of n and $B \ge 1$).

(b) As $B \to \infty$,

$$\mathbb{P}\left[Y_m \leq 1 - 3\delta \mid \mathcal{F}_m\right] \to 1. \tag{33}$$

The statement (33) is shown as follows. By Markov's inequality (20), with probability converging to 1 as $B \to \infty$, all terms $Z_m, \ldots, Z_{m+k}$ are within $[-B, B]$, and $\tau = 0$. In such a case, applying (19) and recalling (26), we get

$$\tilde{X}_{m+1} = \max\{|X_{m+k}|, C_{m+k}\} \tag{34}$$

$$\leq (1 - 3\delta)\, \tilde{X}_m + \frac{B}{1 - a/2}, \tag{35}$$

which implies that $Y_m \leq 1 - 3\delta$, establishing (33).

The proof of (32) is lengthy [[7]] and relies on Proposition 1, Lemma 1 and Lemma 2; it is omitted here.

We use (32) and (33) to show (30) as follows. First, observe that by (32) and Proposition 1, $\{Y_m|\mathcal{F}_m\}$ has bounded $\beta + \Delta$ - moment since we assumed (24) when choosing Δ. Furthermore, since the right side of (32) is independent of $\mathcal{F}_m$, the $\beta + \Delta$ - moment of Y_m is bounded uniformly in m. Now, pick $p > 1$ so that $\beta p \leq \beta + \Delta$, and let q satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Write

$$\mathbb{E}\left[Y_m^\beta \mid \mathcal{F}_m\right]$$

$$\leq (1-3\delta)^\beta + \mathbb{E}\left[Y_m^\beta 1\{Y_m > 1 - 3\delta\} \mid \mathcal{F}_m\right] \tag{36}$$

$$\leq (1-3\delta)^\beta + \left(\mathbb{E}\left[Y_m^{\beta p} \mid \mathcal{F}_m\right]\right)^{\frac{1}{p}} \left(\mathbb{P}\left[Y_m > 1 - 3\delta \mid \mathcal{F}_m\right]\right)^{\frac{1}{q}} \tag{37}$$

$$\to (1-3\delta)^\beta, \quad B \to \infty, \tag{38}$$

where (37) is by Hölder's inequality, and the second term in (37) vanishes as $B \to \infty$ due to (33) and uniform boundedness of the $\beta + \Delta$ - moment of $\{Y_m \mid \mathcal{F}_m\}$. Note that convergence in (38) is uniform in m. It follows that for a large enough B (how large depends on the values of P, k, δ),

$$\mathbb{E}\left[Y_m^\beta \mid \mathcal{F}_m\right] \leq (1-2\delta)^\beta. \tag{39}$$

Rewriting (39) using (31) yields

$$\mathbb{E}[\tilde{X}_{m+1}^\beta \mid \mathcal{F}_m] \leq (1-2\delta)^\beta \left(\tilde{X}_m + \frac{B}{(1-a/2)(1-3\delta)}\right)^\beta \tag{40}$$

$$\leq (1-\delta)^\beta \tilde{X}_m^\beta + K, \tag{41}$$

which implies (30).

E. Finer Quantization

For $a \geq 2$, the controller receives an element of an $\lfloor a \rfloor + 1$-element set instead of a single bit. In this case we restrict our attention to *order-statistic* tests, meaning that we split the real line into $\lfloor a \rfloor + 1$ intervals

$$(-\infty, w_{1,n}), [w_{1,n}, w_{2,n}), \ldots, [w_{\lfloor a \rfloor, n}, \infty), \tag{42}$$

and the controller receives the index $b_n \in \{0, 1, \ldots, \lfloor a \rfloor\}$ of the interval containing X_n. The only real issue is for the quantizer and the controller to agree upon a rule for the values of w_i. However, this is easy; in the obvious generalization of our algorithm to higher a, the estimate C_n of the state magnitude will still be shared knowledge at all times; the (uniform) quantizer simply breaks up the interval $[-C_n, C_n]$ into $\lfloor a \rfloor + 1$ equal parts.

In the case $a < 1$, the controller does nothing, which by Lemma 1 achieves β-moment stability.

III. Converse

In this section, we prove the converse result in Theorem 2 using information-theoretic arguments similar to those employed in [3], [10].

Proof of Theorem 2. Conditional entropy power is defined as

$$N(X|U) \triangleq \frac{1}{2\pi e} \exp\left(2h(X|U)\right) \tag{43}$$

where $h(X|U) = -\int_{\mathbb{R}} f_{X,U}(x,u) \log f_{X|U=u}(x) dx$ is the conditional differential entropy of X.

Conditional entropy power is bounded above in terms of moments (e.g. [11, Appendix 2]):

$$N(X) \leq \kappa_\beta \mathbb{E}\left[|X|^\beta\right]^{\frac{2}{\beta}} \tag{44}$$

$$\kappa_\beta \triangleq \frac{2}{\pi e}\left(e^{\frac{1}{\beta}} \Gamma\left(1 + \frac{1}{\beta}\right) \beta^{\frac{1}{\beta}}\right)^2, \tag{45}$$

Thus,

$$\kappa_\beta \mathbb{E}\left[|X_n|^\beta\right]^{\frac{2}{\beta}} \geq N\left(X_n\right) \tag{46}$$

$$\geq N\left(X_n | U^{n-1}\right), \tag{47}$$

where (47) holds because conditioning reduces entropy. Next, we show a recursion on $N\left(X_n|U^{n-1}\right)$:

$$N\left(X_n|U^{n-1}\right) = N(\mathsf{A}X_{n-1} + Z_{n-1}|U^{n-1}) \tag{48}$$

$$\geq a^2 N(X_{n-1}|U^{n-1}) + N(Z_{n-1}) \tag{49}$$

$$\geq a^2 N(X_{n-1}|U^{n-2}) \exp\left(-2r\right) + N(Z_{n-1}), \tag{50}$$

where (49) is due to the conditional entropy power inequality:[3]

$$N(X+Y|U) \geq N(X|U) + N(Y|U), \tag{51}$$

which holds as long as X and Y are conditionally independent given U, and (50) is obtained by weakening the constraint $|U_{n-1}| \leq M$ to a mutual information constraint $I(X_{n-1}; U_{n-1}|U^{n-2}) \leq \log M = r$ and observing that

$$\min_{P_{U|X}:\, I(X;U) \leq r} h(X|U) \geq h(X) - r. \tag{52}$$

It follows from (50) that $r > \log a$ is necessary to keep $N\left(X_n|U^{n-1}\right)$ bounded. Due to (47), it is also necessary to keep β-th moment of X_n bounded. □

[3]Conditional EPI follows by convexity from the unconditional EPI [12], [13].

IV. Generalizations

In this section, we state the generalizations of our results in several directions. To show these generalizations we can still use our tools in Section II-D with gentle modifications. The proofs are omitted here and are found in [[7]].

A. Constant-Length Time Delays

Many systems have a finite delay in feedback. To model this, we can force U_n to depend on only the feedback up to round $n-\ell$, i.e.

$$U_n = \mathsf{g}_n(\mathsf{f}_1(X_1), \mathsf{f}_2(X^2), \ldots, \mathsf{f}_{n-\ell}(X^{n-\ell})), \tag{53}$$

where $\mathsf{f}_n(X^n)$ is the quantizer's output at time n, as before.

We argue here that this makes no difference in terms of the minimum number of bits required for stability. We state the modified result next.

Theorem 3. *Let X_1, Z_n in (1) be independent random variables with bounded α-moments. Assume that $h(X_1) > -\infty$. The minimum number of quantization points to achieve β-moment stability, for any $0 < \beta < \alpha$ and with any constant delay ℓ is given by $\lfloor a \rfloor + 1$.*

B. Packet drops

Suppose that the encoder cannot send information to the controller at all time-steps. Instead, the encoder can only send information at a deterministic set $\mathcal{T} \subseteq \mathbb{N}$ of times. Formally,

$$U_n = \mathsf{g}_n(\{\mathsf{f}_n(X^n) \colon n \in \mathcal{T}\}). \tag{54}$$

As long as the density of $\mathcal{T}$ is high enough on all large, constant-sized scales, the same results go through.

Definition 1. *A set $\mathcal{T} \subseteq \mathbb{N}$ is* strongly p-dense *if there exists N such that for all n we have*

$$\frac{|n+i \colon n+i \in \mathcal{T},\ i = 0, \ldots, N-1|}{N} > p. \tag{55}$$

Note that the constant delay scenario in Section IV-A amounts to control on a strongly p-dense set, with $p \in [0,1)$ as close to 1 as desired.

Theorem 4. *Let X_1, Z_n in (1) be independent random variables with bounded α-moments. Assume that $h(X_1) > -\infty$. The minimum number of quantization points to achieve β-moment stability is $\lfloor a \rfloor + 1$, for any $0 < \beta < \alpha$ and on any strongly p-dense set with some $p \in [0,1]$ large enough so that*

$$(\lfloor a \rfloor + 1)^p > a. \tag{56}$$

C. Dependent Noise

Here we address a modification in which the noise is correlated rather than independent.

Theorem 5. *The results in Theorems 1, 3, 4 extend to the case when $\{Z_n\}$ is correlated Gaussian noise whose covariance matrix has bounded spectrum.*

D. Vector systems

The results generalize to higher dimensional systems

$$X_{n+1} = \mathsf{A}X_n + Z_n - \mathsf{B}U_n, \tag{57}$$

where A is a $d \times d$ matrix and Z_n, U_n are vectors. The dimensionality of control signals U_n can be less than d, in which case B is a tall matrix.

The idea behind our generalization to the vector case, previously explored in e.g. [3], is that we can decompose $\mathbb{R}^d$ into eigenspaces of A and rotate attention between these parts.

Theorem 6. *Consider the stochastic vector linear system in (57) with (A, B) stabilizable. Let X_1, Z_n be independent random $\mathbb{R}^d$-valued random vectors with bounded α-moments. Assume that $h(X_1) > -\infty$. Let $(\lambda_1, \ldots, \lambda_d)$ be the eigenvalues of A, and set*

$$a \triangleq \prod_{j=1}^{d} \max(1, |\lambda_j|). \tag{58}$$

Then for any $0 < \beta < \alpha$, the minimum number of quantization points to achieve β-moment stability is

$$M^\star_\beta = \lfloor a \rfloor + 1. \tag{59}$$

References

[1] J. Baillieul, "Feedback designs for controlling device arrays with communication channel bandwidth constraints," in *ARO Workshop on Smart Structures, Pennsylvania State Univ*, 1999, pp. 16–18.

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

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

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

[5] Y. Sharon and D. Liberzon, "Input to state stabilizing controller for systems with coarse quantization," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 830–844, 2012.

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

[7] V. Kostina, Y. Peres, G. Ranade, and M. Sellke, "Exact minimum number of bits to stabilize a linear system," *ArXiv preprint arXiv:1807.07686*, July 2018.

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

[9] B. G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[10] V. Kostina and B. Hassibi, "Rate-cost tradeoffs in control," *ArXiv preprint*, Oct. 2016.

[11] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 428–436, Mar. 1992.

[12] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, July and October 1948.

[13] A. J. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Information and Control*, vol. 2, no. 2, pp. 101–112, 1959.