Stabilization with data-rate-limited feedback: 
tightest attainable bounds

Girish N. Nair *, Robin J. Evans

Centre of Expertise in Networked Decision Systems, Department of Electrical and Electronic Engineering, University of Melbourne, 
Vic. 3010, Australia

Received 26 January 2000; received in revised form 26 April 2000; accepted 27 April 2000

Abstract

This paper investigates the stabilizability of a linear, discrete-time plant with a real-valued output when the controller, 
which may be nonlinear, receives observation data at a known rate. It is first shown that, under a finite horizon cost equal to 
the mth output moment, the problem reduces to quantizing the initial output. Asymptotic quantization theory is then applied 
to directly obtain the limiting coding and control scheme as the horizon approaches infinity. This is proven to minimize a 
particular infinite horizon cost, the value of which is derived. An necessary and sufficient condition then follows for there to 
estist a coding and control scheme with the specified data rate that takes the mth output moment to zero asymptotically with 
time. If the open-loop plant is finite-dimensional and time-invariant, this condition simplifies to an inequality involving the 
data rate and the unstable plant pole with greatest magnitude. Analogous results automatically hold for the related problem 
of state estimation with a finite data rate. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Optimal control; Stabilizability; Communication; Quantization

1. Introduction

As large, digitally interconnected control systems 
become more common, it is increasingly important to 
understand how the communication and control ob- 
jectives of such systems are related. The traditional 
assumption that plant observations are available to 
controllers with infinite precision is clearly unrealistic 
when the various parts of a system are connected by 
a network that has finite information-carrying capac-
ity. More than merely introducing delay and quantiza-
tion, this limited capacity forces the question of how 
to choose the bits of information that would be most 
useful for control. In this paper we study the simplest 
type of networked system, a plant and controller linked 
by a channel with a fixed data rate. In particular, we 
seek to determine the smallest data rate needed to sta-
blize the output of a linear plant when no structural 
constraints are imposed on the coder or controller.

In recent years, a number of researchers have pro-
posed and analyzed various versions of this problem. 
In general, the focus has been on memoryless cod-
ing, in which the plant output is quantized without 
reference to its past. In [4], it was shown that if the 
output of an unstable, deterministic, discrete-time, 
linear, time-invariant (LTI) system is passed through 
a fixed, memoryless quantizer, then controllability 
in the sense of being able to make the trajectory 
approach any arbitrary point is impossible. In [9], 
the communication delays in this model were ex-
plicitly included and sufficient conditions given for
the state to eventually remain within a given bound. Wong and Brockett proved that if the initial condition of a continuous-time LTI system is within a given bounded set then memoryless coding \textit{and} control suffice to bound the state under certain conditions [16]. Investigating discrete-time, Gaussian LTI systems, Borkar and Mitter derived a separation principle and a memoryless, certainty equivalent controller, under the constraint that the coder was a memoryless quantizer acting on the innovations of a Kalman filter [1].

The only case in which schemes with memory have been dealt with in communication-limited control appears to be in [14], in which a noisy, analog channel was considered. In this paper, we assume a noiseless, digital channel and permit the coder and controller to have potentially unlimited memory. The motivation for this comes from the closely related field of communication-limited state estimation, in which recursive coding schemes have been studied extensively [3, 15, 12, 13]. Moreover, this divorces the effects of the memory constraint from those of the communication constraint proper, thereby allowing the latter to be analyzed more clearly. In the same spirit we ignore finite word-length issues, by assuming that the measurements are available to the coder with infinite precision, and focus solely on the fundamental performance limitations imposed by the finite data rate.

We consider a discrete-time, linear, time-varying and infinite-dimensional plant with no process or measurement noise and with a directly observed, real-valued output which is zero before time zero and governed by a probability density \( p \) at time zero. Our primary aim is to find the smallest data rate needed to asymptotically stabilize the output of this plant in \( m \)th moment, i.e. to take the \( m \)th moment of the output to zero with time, when no structural constraints save causality are placed on the coder or controller. With this in mind, we first consider a finite horizon cost proportional to the \( m \)th output moment and show that the problem then reduces to causally reformulating an optimal quantizer for the initial output.

The insights gained from the finite horizon analysis are then combined with asymptotic quantization theory [10, 5, 2, 8] to derive the limiting scheme as the horizon tends to infinity, \textit{without having to solve the finite horizon problem first}. We prove that this limiting coding and control scheme is optimal with respect to a certain infinite horizon cost, provided that \( p \) is continuous and satisfies certain technical conditions, and derive the optimal cost. This then leads to a necessary and sufficient condition for the plant to be asymptotically stabilizable at a data rate \( R \). If the open-loop dynamics are finite-dimensional and time-invariant, this simplifies to

\[
R > \log_2 |\lambda|, \quad (1)
\]

where \( \lambda \) is the unstable open-loop pole with largest magnitude. It is then observed that analogous results automatically hold for the problem of state estimation with a finite data rate.

2. Formulation of control problem

Consider the infinite-dimensional, time-varying, ARMA model

\[
x_{k+1} = \sum_{j=0}^{\infty} a_{k,j} x_{k-j} + b_{k,j} u_{k-j}, \quad k = 0, 1, 2, \ldots, \quad (2)
\]

where \( x_k, u_k \in \mathbb{R} \) are the output and control, respectively, at time \( k \), with \( x_k = u_k = 0 \) when \( k < 0 \), and \( a_{k,j}, b_{k,j} \in \mathbb{R}, j, k \geq 0 \), are known parameters. We assume that \( b_{0,0} \neq 0 \), \( \forall k \geq 0 \), so that the control always affects the output at the next time instant. In addition, we further assume that \( x_0 \) is a realization of a random variable \( X_0 \) on the probability space \( (\Omega, \mathcal{F}, P) \), where \( \mathcal{F} \) is the \( \sigma \)-algebra of Lebesgue sets on the real line and \( P \) is a known probability measure such that \( E[X_0]^m < \infty \) for some \( m > 0 \).

Suppose a coder observes the outputs and then sends real-time data to a distant controller over a digital channel that can carry only one symbol \( s_k \) from an alphabet \( \mathbb{Z}_M \triangleq [0, 1, \ldots, M - 1] \) during each sampling interval. The corresponding data rate is \( R \triangleq \log_2 M \) bits per interval. Neglecting the propagation delay and transmission errors, the finite data rate implies that each symbol takes one sampling interval to reach the other end of the channel. Hence, at time \( k \) the controller has \( s_0, \ldots, s_{k-1} \) available and generates

\[
u_k = v_k(\tilde{s}_{k-1}), \quad \forall k \geq 0, \quad (3)
\]

where the notation \( \tilde{y}_k \) denotes a sequence \( \{y_j\}_{j=0}^k \) and \( v_k : \mathbb{Z}_M^k \rightarrow \mathbb{R} \) is the controller function at time \( k \).

If no restrictions save causality are placed on the structure of the coder, each symbol \( s_k \) which it transmits may be a function of the sequences of past and present outputs \( \tilde{x}_k \) and past symbols \( \tilde{s}_{k-1} \). However, (2) and (3) imply that \( x_k \) in its turn is completely determined by the initial condition and \( \tilde{s}_{k-2} \), so \( s_k \) is consequently a function of \( x_0 \) and \( \tilde{s}_{k-1} \),

\[
s_k = \gamma_k(x_0, \tilde{s}_{k-1}), \quad \forall k \geq 0, \quad (4)
\]
where \( \gamma_k : \mathbb{R} \times \mathbb{Z}_M^k \rightarrow \mathbb{Z}_M \) is the coder function at time \( k \).

Note that there is no explicit communication constraint between the controller and actuator. This is obviously a reasonable assumption if they are collocated, but even otherwise the formulation above would be unchanged, since the location of the controller is purely nominal. The symbols that would be transmitted by it over an additional link to the actuator would have to be translated once again into control signals, making intermediate calculations redundant. The number \( R \) should thus be regarded as the overall rate of the complete link from sensor to actuator. In addition, we remark that there are slightly different ways in which the digital link can be defined (see [16,1] for details).

We call the pair of sequences \((\gamma, v) \triangleq (\{\gamma_k\}_{k \geq 0}, \{v_k\}_{k \geq 0})\) a coder–controller and our objective is to find one that minimizes an infinite horizon cost of the form

\[
J_m = \lim_{k \to \infty} \sup \rho_k^{-1} E[X_k]^m,
\]

where \( X_k \) is the random variable corresponding to the output \( x_k \). This compares the asymptotic behaviour of the \( m \)th output moment against some positive sequence \( \{\rho_k\}_{k \geq 0} \), which serves as a rough benchmark of the desired output behaviour. Although it does not attach a cost to the magnitudes of the controls or intermediate outputs, it does succeed in capturing the asymptotic stochastic behaviour of the closed-loop system. However, before addressing this objective, we first investigate the finite horizon cost

\[
J_{m,N} = E|X_{N+1}|^m.
\]

The insights gained then lead to an optimal solution of the infinite horizon problem and hence a necessary and sufficient condition for the system to be asymptotically stabilizable in \( m \)th moment.

3. Finite horizon cost

If the initial condition were known with complete accuracy by the controller, it could easily use the system equations to generate controls that yield \( x_1 = x_2 = \cdots = 0 \). However, in the presence of the data rate constraint this is evidently impossible. One way around this is for the coder to transmit a progressively more accurate estimate of \( x_0 \) to the controller, by using the bits available at each time instant to quantize it recursively. The controller can then use these estimates to generate the controls. It is shown in this section that for linear systems of form (2), the optimal finite horizon scheme has exactly this structure. More precisely, the core of the \( J_{m,N} \)-optimal coder-controller is shown to be a causally reformulated, optimal quantizer for \( x_0 \) with \( M^N \) levels.

Observe that by using downward induction on \( k \), the output of the system can be expressed in terms of the initial condition and past controls,

\[
x_{k+1} = x_k x_0 + \sum_{j=0}^{k} \beta_k,j u_{k-j}, \quad \forall k \geq 0,
\]

where \( x_{-1} \triangleq 1 \) and \( x_k, \beta_k,j \) are given recursively by

\[
x_{k+1} \triangleq \sum_{j=0}^{k+1} a_{k+1,j} x_{k-j}, \quad \beta_{k+1,j} \triangleq b_{k+1,j} + \sum_{i=0}^{j-1} a_{k+1,i} \beta_{k-i,j-i-1}, \quad \forall k \geq 0, \ j \in [0, \ldots, k+1].
\]

Notice that the problem becomes trivial if \( x_k = 0 \) for some \( k \in [0, \ldots, N] \), since from (7) and (2) controls \( u_k, \ldots, u_N \) can then be found which yield \( x_{k+1} = \cdots = x_{N+1} = 0 \). As such, we focus on the nontrivial case when \( x_k \neq 0, \forall k \in [0, \ldots, N] \).

The next step is to change variables. Define the mappings by

\[
\eta_k(\hat{t}_{k-1}) \triangleq -x_k^{-1} \sum_{j=0}^{k} \beta_k,j v_{k-j}(\hat{t}_{k-j-1}), \quad \forall k \in [0, \ldots, N], \ \hat{t}_{k-1} \in \mathbb{Z}_M^k.
\]

so that \(-x_k \eta_k(\hat{t}_{k-1})\) represents the cumulative effect of past controls on the output at time \( k+1 \). As \( \beta_{k,0} = b_{k,0} \), which by hypothesis is nonzero, the matrix of coefficients \( \{-x_k^{-1} \beta_{k,j}\}_{k,j} \) is triangular with nonzero diagonal elements. Hence, the equation above can easily be inverted to express the coder function sequence \( \tilde{v}_N \) in terms of \( \eta_N \),

\[
v_k(\hat{t}_{k-1}) \triangleq \sum_{j=0}^{k} \tau_{k,j} \eta_{k-j}(\hat{t}_{k-j-1}), \quad \forall k \in [0, \ldots, N], \ \hat{t}_{k-1} \in \mathbb{Z}_M^k.
\]

where \( \tau_{k,j} \) is given by the recursion

\[
\tau_{k,0} \triangleq -\beta_{k,0}^{-1} x_k, \quad \tau_{k,j} \triangleq -\beta_{k,0}^{-1} \sum_{i=1}^{j} \beta_{k,i} \tau_{k-i,j-i}, \quad \forall k \in [0, \ldots, N], \ j \in [0, \ldots, k].
\]
This one-to-one correspondence between $\bar{\eta}_k$ and $\bar{v}_k$ implies that it is perfectly equivalent to optimize $J_{m,N}$ with respect to either. Substituting (7) and (9) into (6), we obtain

$$J_{m,N} = |x_N|^m E[|x_0 - \eta_N(\bar{s}_{N-1})|^m], \quad (12)$$

where $\bar{s}_{N-1}$ is the random variable corresponding to the symbol sequence $\bar{s}_{N-1}$. Now, $\eta_N(\bar{s}_{N-1})$ is a function of $x_0$ with up to $M^N$ distinct values, one for each possible symbol sequence, so the function $R_N$ defined by $R_N(x_0) \triangleq \eta_N(\bar{s}_{N-1})$ can be regarded as an $M^N$-level quantizer for $x_0$. The RHS of (12) is then simply its mean $m$th power error ($Mm$PE), scaled by $|x_N|^m$. As such, if we can find a pair $(\gamma^*, \eta^*)$ that achieves the minimum $Mm$PE for an $M^N$-level quantizer for $X_0$, it will automatically be $J_{m,N}$-optimal. We show next how to construct such a pair.

Denote the points of the $Mm$PE-optimal, $M^N$-level quantizer for $X_0$ by $q_N(0), q_N(1), \ldots, q_N(M^N - 1)$, ranked from least to greatest, and set

$$\eta_N(\bar{s}_{N-1}) \triangleq q_N\left(\sum_{k=0}^{N-1} t_k M^{N-1-k}\right), \quad \forall \bar{s}_{N-1} \in \mathbb{Z}_M^N,$$

(13)

i.e. the argument of $\eta_N$ is the $M$-ary representation of that of $q_N$. By the nearest-neighbour rule [6], the optimal quantizer selects the point closest to $x_0$, so choose the symbol sequence by

$$\bar{s}_{N-1} = \arg \min_{t_{N-1} \in \mathbb{Z}_M} |x_0 - \eta_N(\bar{s}_{N-1})|^m, \quad (14)$$

breaking possible ties by selecting the greater point. As this equation can be expressed recursively, $\forall k \in [0, \ldots, N - 1]$, as

$$s_k = \gamma_{z_k}(x_0, s_{k-1}) \triangleq \arg \min_{t_k \in \mathbb{Z}_M} \left\{ \min_{t_{k+1}, \ldots, t_{N-1} \in \mathbb{Z}_M} |x_0 - \eta_N(\bar{s}_{N-1}, t_k, t_{k+1}, \ldots, t_{N-1})|^m \right\}, \quad (15)$$

it is realizable within our framework. Substituting (14) into (12), we obtain

$$|x_N|^{-m} J^{*}_{m,N} = \int \min_{\bar{s}_{N-1} \in \mathbb{Z}_M^N} |x_0 - \eta_N(\bar{s}_{N-1})|^m dP(x_0)$$

$$= \int \min_{j \in \mathbb{Z}_M^N} |x_0 - q_N(j)|^m dP(x_0). \quad (16)$$

As the last integral is simply the expression for the $Mm$PE of the optimal $M^N$-level quantizer for $X_0$ [6], the pair $(\gamma^*, \eta^*)$ is $J_{m,N}$-optimal.

We now recast the coder equation (5) in a somewhat simpler form. First extend $q_N$ over the real interval $[-1, M^N]$, so that it remains increasing and is furthermore continuous, and for convenience set $q_N(-1) = -\infty$, $q_N(M^N) = \infty$. Next, define

$$e_N(z) \triangleq (q_N(M^N z - 1) + q_N(M^N z))/2,$$

(17)

$$\forall z \in [0, 1],$$

so that $e_N(\zeta_{N-1})$ is half-way between the neighbouring quantizer points $q_N(M^N \zeta_{N-1})$ and $q_N(M^N \zeta_{N-1} + 1)$. From (14), the sequence $\bar{s}_{N-1}$ is transmitted iff the quantizer point closest to $x_0$ is $q_N(M^N \zeta_{N-1})$, equivalent to $x_0$ lying inside the interval $[e_N(\zeta_{N-1}), e_N(\zeta_{N-1} + M^N - 1)]$. As $e_N$ is increasing and continuous, $x_0$ lies in this interval iff $e_N^{-1}(x_0) \in [\zeta_{N-1}, \zeta_{N-1} + M^N - 1)$. Referring to (18), this in turn is equivalent to $\bar{s}_{N-1}$ being the first $N$ digits of the $M$-ary representation for $e_N^{-1}(x_0)$. That is, the optimal coder simply applies a transformation $e_N^{-1}$ to the initial condition and then transmits the first $N$ digits of its $M$-ary expansion. This and the previous results are encapsulated below:

**Coder 1.** First, transform the initial condition $x_0$ of system (2) to obtain $\zeta_k \triangleq c_N(x_0) \triangleq e_N^{-1}(x_0)$, where $e_N$ is given by (17). Then at time $k$, transmit the $(k + 1)$th most significant digit in the $M$-ary representation of $\zeta$ as the symbol $s_k$.

**Controller 1.** Upon receiving the symbol $s_{k-1}$ at time $k$, calculate the number $z_k$ using (18). Set

$$\eta_N(\bar{s}_{N-1}) \triangleq q_N(M^N \zeta_{N-1}),$$

$$\eta_N(\bar{s}_{N-1}) \triangleq c_N^{-1}(\zeta_{k-1} + 0.5/M^k),$$

(19)

$$\forall k \in [0, \ldots, N - 1],$$

where $q_N(0) < q_N(1) < \cdots < q_N(M^N - 1)$ are the points of the $Mm$PE-optimal, $M^N$-level quantizer for $X_0$, and use (10) to calculate the control $v_k(\bar{s}_{k-1})$.

We make several comments here. Firstly, the optimal coder-controller is basically a compander, i.e. it consists of a compressor $c_N$ which maps $x_0 \in \mathbb{R}$ to $\zeta \in [0, 1]$, followed by a uniform, $M^N$-level quantizer which maps this to $\zeta_{N-1}$ and then an expander $q_N(M^N \cdot)$ which transforms $\zeta_{N-1}$ into an estimate of $x_0$ [6]. Secondly, the mappings $\eta_0, \ldots, \eta_{N-1}$ are actually completely arbitrary, since they do not affect
the integrand in (16). However, the choice above is intuitively appealing, as \( \xi_k + (1/2M^{k+1}) \) is the midpoint of the interval of length \( M^{-k-1} \) which the controller knows that \( \xi \) lies in, from the sequence \( \delta_k \). Furthermore, this makes the infinite horizon analysis of the next section somewhat easier. Thirdly, although the optimal quantizer may be unique, \( \eta_N^* \) can be defined in as many different ways as there are one-to-one maps from the integers \( \mathbb{Z}_{MN} \) to the \( M \)-ary sequences \( \mathbb{Z}_M^N \). The choice of mapping taken here, as implied in Eq. (13), is one of the more tractable ones. Finally, explicit expressions for the optimal coder–controller are generally impossible to derive, since \( q_N \) can norm-

ally only be obtained numerically for a given \( p \) and \( N \) [10,6]. One of the few exceptions is when \( X_0 \) is Laplacian and \( m = 1 \), for which case a closed form solution parametrized by \( N \) and the mean and variance of \( X_0 \) has been obtained [11].

4. Infinite horizon cost

In the previous section, we observed that \( J_{m,N} \)-optimal coder–controllers are usually impossible to derive in closed form. However, we demonstrate here that when \( N \to \infty \) the limiting coder–controller can be obtained directly, without explicitly solving the finite horizon problem. We then prove that this limiting scheme is in fact optimal with respect to an infinite horizon cost of form (5), under certain mild conditions on the probability density \( p \) governing the initial output \( x_0 \).

The key is the classic result that as the number of \( \text{MmPE} \)-optimal quantizer points approaches infinity, their normalized density per unit \( x_0 \) approaches

\[
v(x_0) \triangleq \left( \int p(y) \frac{1}{(m+1)} \, dy \right)^{-1} p(x_0)^{1/(m+1)},
\]

\( \forall x_0 \in \mathbb{R}, \) \hspace{1cm} (20)

under certain technical conditions on \( p \) [5,2]. As \( q_N(M^N \xi_{N-1}) \) is the \( (M^N \xi_{N-1} + 1) \)th quantizer point by (13), the nearest-neighbour rule implies that there are \( M^N \xi_{N-1} + \mathcal{O}(1) \) points less than or equal to \( x_0 \). Observing that \( \xi_{N-1} \), being a sum of exponentially decaying terms, must converge to a number \( \xi \in [0,1) \) as \( N \to \infty \), we may define

\[
c(x_0) \triangleq \xi = \lim_{N \to \infty} \frac{M^N \xi_{N-1} + \mathcal{O}(1)}{M^N} = \int_{y \leq x_0} v(y) \, dy, \hspace{1cm} \forall x_0 \in \mathbb{R}.
\]

By analogy with Coder–Controller 1, the following scheme as the horizon \( N \) becomes unbounded is suggested:

**Coder 2.** First, transform the initial condition \( x_0 \) of system (2) to yield \( \zeta \triangleq c(x_0) \), where \( c \) is given by (21). Then at time \( k \) transmit the \((k + 1)\)th most significant digit in the \( M \)-ary representation of \( \zeta \) as the symbol \( s_k \).

**Controller 2.** Upon receiving the symbol \( s_{k-1} \) at time \( k \), calculate

\[
\eta_k(s_k) = c^{-1}(c_{k-1} + 0.5/M^k),
\]

where \( c_{k-1} \) is defined by (18), and use (10) to generate the control signal \( v_k(s_{k-1}) \).

For instance, for Laplacian \( X_0 \) with mean \( \mu \) and mean absolute deviation \( \varepsilon \), it can be shown that

\[
c(x_0) = 0.5 + \text{sign}(x_0 - \mu)(1 - e^{-|x_0 - \mu|/\varepsilon})/2,
\]

and for Gaussian \( X_0 \) with mean \( \mu \) and standard deviation \( \varepsilon \), we have

\[
c(x_0) = F \left( \frac{x_0 - \mu}{\varepsilon \sqrt{m + 1}} \right),
\]

where \( F \) is the unit normal distribution function.

We now consider whether the coder–controller above is actually optimal with respect to an infinite horizon cost of form (5). First we need to fix the weights \( \rho_k, k \geq 0 \). Observe that as \( N \to \infty \),

\[
\min_{x \in \mathbb{R}} M^{mN} \| x \|^{-m} E \| X_{N+1} \|^m
\]

\[
= \min_{x \in \mathbb{R}} M^{mN} E \| X_0 - \eta_N(S_{N-1}) \|^m
\]

\[
\to (m+1)^{-1} 2^{-m} \| p \|_{1/(m+1)}, \hspace{1cm} (23)
\]

where the limit is a well-known result of asymptotic quantization theory [2] and \( \| p \|_r = (\int p(x)^r \, dx)^{1/r} \). This is nearly what we want, except that in (5) the minimization is to be performed after the limit is taken. This suggests that an appropriate choice of weighting sequence is

\[
\rho_k = |x_{k-1}|^m/M^{m(k-1)}, \hspace{1cm} \forall k \geq 0.
\]

(24)

In order to prove that Coder–Controller 2 is \( J_{m,N} \)-optimal, we make use of the fact that it is essentially a compander and apply a result of [8]. Let \( G : \mathbb{R} \to [0,1] \) be a compressor function with a continuous and nonnegative derivative \( g \) such that \( g(x) \) decreases monotonically with \( |x| \) for sufficiently
large $|x|$. Suppose $G$ is applied to a real random variable $X$ with a probability density function $\pi$ such that $E\{g(X)^{-m}\} < \infty$ and such that, for some $\delta > 0$, both
\[
\int_{0}^{\delta} s(z)^m h(2z) \, dz = \int_{-\delta}^{1} s(z)^m h(2z - 1) \, dz < \infty,
\]
where $s = 1/gG^{-1}(\cdot)$ and $h = \pi G^{-1}(\cdot)/gG^{-1}(\cdot)$. In [8], it is shown that if $G(x)$ is quantized by $f_Q$, a $Q$-level, midpoint-based, uniform quantizer on $[0,1]$, and $G^{-1}$ subsequently applied to form an estimate of $x$, then
\[
Q^m E|X_0 - G^{-1} f_Q G(X)|^m \to (m+1)^{-1/2} - m E\{g(X)^{-m}\} \text{ as } Q \to \infty. \tag{25}
\]
We can now prove the main result of this section:

**Theorem 1.** Let the initial output $x_0$ of system (2) be governed by a continuous probability density function $p$ which decreases with $|x_0|$ for sufficiently large $|x_0|$ and satisfies $E|X_0|^{|m+n|} < \infty$, for some $m,n > 0$. Suppose further that
\[
p_c^{-1}(z) \leq Ap_c^{-1}(z), \quad \forall z \in [1 - \delta, 1],
\]
\[
p_c^{-1}(z) \leq Ap_c^{-1}(z), \quad \forall z \in [0, \delta]
\]
for some $A, \delta > 0$, where $c$ is given by (21). Then Coder–Controller 2 is $J_m$-optimal and achieves
\[
J_m = \min_{\gamma, \epsilon} \lim_{k \to \infty} \frac{1}{|M_k|} M^{m(k-1)} E|X_k|^m \tag{26}
\]
where $\alpha_k, k \geq 0$, are given by (8). Furthermore, for a given coding alphabet $M$, a coder–controller that takes $E|X_k|^m \to 0$ exists if and only if $\alpha_k/M^k \to 0$ as $k \to \infty$. \tag{27}

**Proof.** Note that $M^{-1} E|X_k|^m = M^{m(k-1)} E|X_0 - \eta_{k-1}(\hat{S}_{k-2})|^m$. As $\eta_{k-1}(\hat{S}_{k-2})$ may be expressed as the compander output $c^{-1} f_{M^{1-c}}(X_0)$, with $f_{M^{1-c}}$ being the $M^{k-1}$-level, midpoint-based, uniform quantizer on $[0,1]$, our first step is to show that the compensator $c$ satisfies the conditions of [8]. Its derivative $c' = \gamma = \gamma \cdot k^{-1} p(\cdot)^{1/(m+1)}$, where $\gamma = \int p(X_0)^{1/(m+1)} \, dx_0$, so $c'$ is evidently continuous, nonnegative and monotonically decreasing for large enough $x_0$, by hypothesis on $p$. Furthermore, $E\{g(X_0)^{-m}\} = \|p\|_{1/(m+1)}$ which, as remarked in [2], is guaranteed to be bounded by Hölder’s inequality if $E|X_0|^{-m} < \infty$. Next, note that $h = k p c^{-1}(\cdot)^m(\cdot)^{1/(m+1)}$, so that $h(2z) \leq A' h(z)$, $\forall z \in [0, \delta]$, where $A' = A'M^{1-c}$. Hence,
\[
\int_0^\delta s(z)^m h(2z) \, dz \leq A' \int_0^\delta s(z)^m h(z) \, dz = A' k^m \int_{-\delta}^{\delta} p(x_0)^{1/(m+1)} \, dx_0,
\]
which is finite. The boundedness of the remaining integral can be proven in exactly the same way. Hence, all the preconditions for (25) hold, so that as $k \to \infty$,
\[
M^{m(k-1)} E|X_0 - \eta_{k-1}(\hat{S}_{k-2})|^m \to (m+1)^{-1/2} - m E\{g(X)^{-m}\} = (m+1)^{-1/2} 2^{-m} ||p||_{1/(m+1)}.
\]
Now observe that for any coder–controller $(\gamma', \eta')$,
\[
\lim_{k \to \infty} \sup_{\gamma', \eta'} M^{m(k-1)} E|X_0 - \eta_{k-1}(\hat{S}_{k-2})|^m \geq \lim_{k \to \infty} \min_{\gamma', \eta'} M^{m(k-1)} E|X_0 - \eta_{k-1}(\hat{S}_{k-2})|^m = ||p||_{1/(m+1)} (m+1)^{2m}/(m+1)^{2m},
\]
by (23). As the scheme above achieves this lower bound, it is optimal.

To prove the sufficiency of (27), suppose it holds. Eq. (26) indicates that if Coder–Controller 2 is applied then the sequence $\{x_{k-1} - \eta_{k-1}(\hat{S}_{k-2})\}$ is bounded, which forces $E|X_k|^m$ to approach zero. To prove necessity, suppose that (27) does not hold. For any coder–controller,
\[
M^{m(k-1)} E|X_k|^m \geq \frac{1}{|M_k|} M^{m(k-1)} E|X_0|^m \geq \min_{\gamma', \eta'} M^{m(k-1)} E|X_0 - \eta_{k-1}(\hat{S}_{k-2})|^m \quad \text{as } k \to \infty,
\]
where we have again made use of the limit from (23). Hence, for any coder–controller and $\epsilon > 0$, $\exists k' > 0$ such that
\[
E|X_k|^m \geq \frac{1}{M^{m(k-1)}} \left( ||p||_{1/(m+1)} - \epsilon \right), \quad \forall k \geq k'.
\]
By hypothesis the RHS cannot approach zero, so no coder–controller exists which takes $E|X_k|^m \to 0$. \qed
Consider the linear system below:

\[ \dot{x}_k = \sum_{j=1}^{d-1} a_j x_{k-j}, \]

where \( a_0 = 0 \) and \( a_2 = \cdots = a_{d} = 0 \). The solution to this is of the form

\[ x_k = \sum_{j=0}^{d-1} h_j \theta_j^k, \]

where \( \theta_0, \ldots, \theta_{d-1} \) are the poles of the system and \( h_0, \ldots, h_{d-1} \) are constants, with possibly a polynomial dependence on \( k \) if a pole is repeated. Hence, if \( \lambda \) is the pole with largest magnitude then \( x_k \sim \lambda^k \) for large \( k \), to within a polynomial factor in \( k \). Substituting this into (27), we see that this system is asymptotically stabilizable in \( m \)th moment iff \( M > |\lambda| \). As the data rate \( R = \log_2 M \), we obtain the condition \( R > \log_2 |\lambda| \). This makes precise the notion that the more unstable a system is, the higher the data rate needed to stabilize it. This result is summarized below.

**Corollary 1.** Consider the linear system below, with time-invariant, \( d \)-dimensional open-loop dynamics,

\[ x_{k+1} = \sum_{j=0}^{d-1} a_j x_{k-j} + \sum_{j=0}^{\infty} b_{k,j} u_{k-j}, \quad \forall k \geq 0, \]

where \( x_k, u_k \in \mathbb{R} \) are the output and control, respectively, at time \( k \), \( a_j, b_{k,j} \in \mathbb{R} \) with \( b_{k,0} \neq 0 \), \( \forall j, k \geq 0 \), and \( x_k = u_k = 0 \) when \( k < 0 \). If the probability density function governing \( x_0 \) satisfies the conditions in Theorem 1, then a coder-controller with data rate \( R \) that takes \( h_k X_k |m \to 0 \) exists if and only \( R > \log_2 |\lambda| \),

where \( \lambda \) is the unstable system pole with largest magnitude.

The technical conditions on \( p \) in Theorem 1 effectively limit the speed of decay of \( pc^{-1}(z) \) as \( z \) approaches \( 0 \) and \( 1 \). They can be shown to be satisfied by any \( p \) such that \( p(y) \sim |y|^\alpha \exp(-B|y|^{\alpha}) \) for large \( |y| \) and parameters \( B, w > 0, v \in \mathbb{R} \), which includes densities such as the Gaussian and Laplacian. We conjecture that the infimum of \( J_m \) assuming only that \( p \) is Lebesgue-integrable is also given by (23), since it should be possible to construct compressors \( c_i, i \geq 0 \), which satisfy the conditions of [8] and approach \( c \) in an appropriate integral sense as \( i \to \infty \). However, it is much more difficult to prove that Coder–Controller 2 actually achieves this lower bound for any such general \( p \), despite being the limiting form of the optimal finite horizon scheme.

We remark that the results above automatically apply to the problem of output estimation under a data rate constraint [15,12]. The only differences in the problem formulation are that the controls in the system equation (2) are set to zero, the controller is replaced by an estimator

\[ \hat{x}_k = \delta_k(\hat{x}_{k-1}), \quad \forall k \geq 0 \]

and the objective is to find a coder-estimator \( \langle \gamma, \delta \rangle \triangleq \{\gamma_k \geq 0, \delta_k \geq 0 \} \) that minimizes the distortion

\[ D_m \triangleq \limsup_{k \to \infty} \rho_k^{-1} E|X_k - \hat{X}_k|^m. \quad (29) \]

The optimal coder-estimator is the same as Coder–Controller 2, except that \( \eta_k(\hat{x}_{k-1}) \) is used to generate an estimate \( \delta_k(\hat{x}_{k-1}) = x_{k-1}\eta_k(\hat{x}_{k-1}) \) rather than a control.

Finally, note that the results of this section indicate that one-bit mean coding schemes [15,7] are suboptimal with respect to the infinite horizon, mean-square-error cost \( J_2 \). Such schemes have intuitive appeal, as they proceed by simply partitioning a coding interval \( I \) containing \( x_0 \) at the conditional mean \( E[X_0|x_0 \in I] \) to form two new candidate intervals. However, it is easy to show from the discussion at the beginning of this section that the intervals formed by Coder 2 always contain equal proportions of optimal, infinite-level quantizer points. For a one-bit scheme, this means that each coding interval \([a, b]\) should be divided at the point \( u \) such that

\[ \int_a^u v(x_0) \, dx_0 = \int_u^b v(x_0) \, dx_0, \]

which in general does not coincide with \( E[X_0|x_0 \in [a, b]] \). Hence, although the conditional mean is the mean-square-error-optimal reconstruction point, it is not the \( J_2 \)-optimal partition point.
5. Conclusion

In this paper, the asymptotic stabilizability of a linear, discrete-time system with a communication constraint was investigated. Finite and infinite horizon control objectives were formulated and it was shown that the optimal finite horizon coder–controller is essentially an optimal quantizer for the initial output. Asymptotic quantization theory was then used to directly obtain the limiting scheme as the horizon approaches infinity. Under certain technical conditions on the probability density $p$ governing the initial output, this scheme was shown to be optimal in the infinite horizon sense and an expression for the optimal cost was derived. This led to a necessary and sufficient condition for the system to be asymptotically stabilizable in $m$th moment at a given data rate. Further work is currently being undertaken on relaxing the conditions on $p$ and extending the results presented here to stochastic and nonlinear systems.

References