Markov Chains, Entropy, and Fundamental Limitations in Nonlinear Stabilization

Prashant G. Mehta   Umesh Vaidya   Andrzej Banaszuk

***Abstract*— This paper is concerned with entropy based fundamental limitation results for the nonlinear stabilization problem of a scalar dynamical system. Using methods based on stochastic dynamics, we pose the problem as control of Markov chains. It is shown that uncertainty, associated here with the unstable eigenvalue of the linearization, leads to fundamental limitations. These limitations arise as certain in-feasibility conditions for nonlinear stabilization in the presence of quantization or equivalently as positive conditional entropy of the output signal in the feedback loop. The former leads to a nonlinear stabilization result and latter to a fundamental limitation result.**

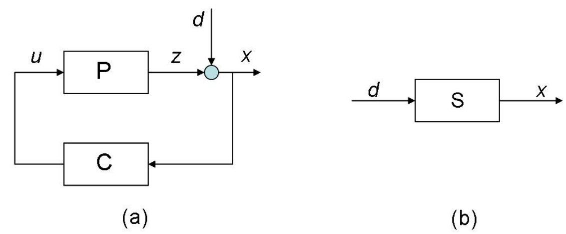

Fig. 1. (a) Feedback loop and (b) Sensitivity function.

I. Introduction

The fundamental limitations in the classical control settings address closed-loop system trade-offs and best possible performance with causal stabilizing feedback. One important result is the Bode integral formula

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\log|S(e^{i\omega})|d\omega = \sum_k \log(p_k), \tag{1}$$

where $S(e^{i\omega})$ is the transfer function of the feedback loop (in Fig. 1) from the disturbance d to output x, and p_k are unstable poles ($|p_k| > 1$) of the open loop plant; cf. Sung and Hara [1]. S is referred to as the *sensitivity function* and for an open-loop plant P and a stabilizing feedback control C, it is given by $S = \frac{1}{1+PC}$; see Fig. 1. Entropy of the signals in the feedback loop help provide another interpretation of the Bode integral formula:

$$\mathcal{H}_c(x) - \mathcal{H}_c(d) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\log|S(e^{i\omega})|d\omega. \tag{2}$$

Here, $\mathcal{H}_c(x)$ and $\mathcal{H}_c(d)$ denote the conditional entropy (see [2], [3]) of the random processes associated with the output x and disturbance d respectively. Combining Eq. (1) with Eq. (2), the open-loop unstable poles are seen to lead to a positive entropy rate. As already noted in [4], the entropy rates are well-defined even in time-domain and provide for a framework for studying fundamental limitations in nonlinear systems. In recent years, several studies have considered the entropy based extension of Bode formula to nonlinear [4], [5] and linear systems with communication constraints [6], [7]. In particular, [5] defines a Dynamical systems based notion of topological feedback entropy (TFE) and obtains explicit and tight lower bounds for this in terms of unstable eigenvalues. A closely related research effort is to quantify the impact of limited information (quantization) on the linear [8]–[10] and nonlinear stabilization [11], [12].

In this paper, we present entropy based fundamental limitation results for stabilization of a scalar nonlinear dynamical system at an unstable equilibrium. The main result appears in Section V and relates the conditional entropy (of the output x in Fig. 1) to the $\log$ value of unstable eigenvalue of the linearization taken at the equilibrium. Our original contribution lies in the framework used for obtaining these results and in bridging the statistical entropy notion of [4], [6], [7] with the Dynamical systems notion of the entropy [5]. In regards to latter, we do not consider the topological entropy but instead its statistical counterpart, the so-called measure-theoretic or metric entropy. Our framework uses in an essential way the concept of Markov chains and the entropy estimates of our paper are related to the recent work in [13]. Although, we present only the scalar results here, the journal version of this paper extends these to the multi-dimensional state feedback case. We briefly summarize the framework below.

We employ the methods of Ergodic theory to replace the deterministic *nonlinear* dynamical system by its stochastic counterpart, the so-called *linear* Perron-Frobenius operator [14]. There are three advantages to doing so. One, it is now easier to compute statistical quantities such as entropy relevant to the study of fundamental limitations. Two, the P-F operator is linear and analysis becomes much simpler and is in fact, analogous to the linear control problem. Finally, the stabilization problem with finite partition, or quantization, is easily considered using the finite-dimensional discretization of the P-F operator. This leads to controlled Markov chains.

In the stochastic context, we define the stabilization problem as shaping of the invariant measure of the controlled Markov chains. With respect to a finite partition, or quantization, the fundamental limitations arise as a limit on the

P. G. Mehta is with the Department of Mechanical & Industrial Engineering, University of Illinois at Urbana-Champaign, 1206 W. Green Street, Urbana, IL 61801 mehtapg@uiuc.edu

U. Vaidya is with the Department of Electrical & Computer Engineering, Iowa State University, Ames, IA 50014 ugvaidya@iastate.edu

A. Banaszuk is with the Systems Department, United Technologies Research Center, East Hartford, CT 06108 banasza@utrc.utc.com

maximum probability of the invariant measure or equivalently as a positive conditional entropy of the output x. The former is related to the in-feasibility of the stabilization in the presence of finite quantization and latter to the entropy based fundamental limitations results. The results demonstrate the important role of uncertainty, due to unstable open-loop dynamics here, on the feedback control problem. Markov chains provide for a useful framework to explicitly quantify this uncertainty.

The outline of this paper is as follows. In Section II, we present some preliminaries and notation pertaining to Perron-Frobenius operators for nonlinear dynamical system. In Section III, we formulate the nonlinear stabilization problem. In Sections IV and V, we present the fundamental limitation results in terms of maximum probability and entropy estimates respectively. Finally, in Section VI, we provide some conclusions and indicate directions of future research.

II. Preliminaries & Notation

In this paper, *discrete scalar mappings* of the form

$$x_{n+1} = T(x_n) \tag{3}$$

are considered. $T : X \to Y \subset \mathbb{R}^1$ is assumed to be continuous and $X \subset \mathbb{R}^1$ is a compact set. $\mathcal{B}(X)$ denotes the Borel σ-algebra on X and $\mathcal{M}(X)$ the vector space of bounded real valued measures on $\mathcal{B}(X)$. The stochastic Perron-Frobenius (P-F) operator for a mapping $T : X \to X$ is given by

$$\mathbb{P}[\mu](A) = \mu(T^{-1}(A)), \tag{4}$$

where $\mu \in \mathcal{M}(X)$ and $A \in \mathcal{B}(X)$. We note that $\mathbb{P}$ acts linearly on the measure space. The *invariant measures* are the fixed points of $\mathbb{P}$ that are additionally probability measures. The invariant measure is a stochastic analogue of an attractor set, such as a stable equilibrium, of the mapping; cf. [14]. For open-loop unstable systems, it is often important to consider the a mapping $T : X \to Y$ where $Y = X \cup S$. If X is compact and S is T-invariant, then Eq, (4) is well-defined on the measure space $\mathcal{M}(X)$. The resulting operator, however, is only sub-stochastic in the general case; cf., [15].

In this paper, composition maps of the form $T_1 \circ T_2 : X \to X$ will frequently be considered; $\circ$ denotes the composition of two nonlinear mappings. It is assumed that $T_1 : X \to X \cup S$ where the set S is T_1-invariant, and $T_2 : X \to X$. Suppose $\mathbb{P}_1$ and $\mathbb{P}_2$ denote the stochastic operators corresponding to the two mappings T_1 and T_2. More precisely, $\mathbb{P}_1$ need only be sub-stochastic. If follows from a simple calculation that the stochastic operator for $T_1 \circ T_2$ is given by $\mathbb{P}_1 \cdot \mathbb{P}_2$. Just as a P-F operator is a linear operator on the space of measures, a nonlinear composition of two mappings lead to linear multiplication of the corresponding operators.

The stochastic operators act on the space of measures and are thus infinite-dimensional. By taking a partition of the compact set X, denoted as $\mathcal{X}_L \doteq \{D_1, \cdots, D_L\}$ where $\cup_j D_j = X$, one can approximate these operators by a finite-dimensional matrix. On the "measure space" $\mathbb{R}^L$ associated with the partition $\mathcal{X}_L$, it is given by

$$P_{ij} = \frac{m(T^{-1}(D_j) \cap D_i)}{m(D_i)}, \tag{5}$$

m being the Lebesgue measure. This matrix is Markov (row stochastic) for stochastic $\mathbb{P}$ and sub-Markov (with row sum less than equal to 1) for sub-stochastic $\mathbb{P}$.

The above considerations also extend to a class of randomly perturbed dynamical systems

$$x_{n+1} = T(x_n, d_n), \tag{6}$$

where $d_n \in D$ models a stochastic disturbance assumed in this paper to be i.i.d with probability measure $\Omega(D)$. Random dynamical system arises because one now has a collection of mappings $\{T_d = T(\cdot, d)\}_{d \in D}$ that are randomly chosen according to the probability measure Ω. The analogue of the P-F operator for such a random dynamical system is obtained by taking an expectation

$$\mathbb{E}[\mu(T^{-1}(B, d))] = \int_D \mu(T^{-1}(B, d)) d\Omega(d) \tag{7}$$

with respect to the probability measure Ω [16]–[18]. The invariant measures are the stochastic analogue of a "random attractor" of Eq. (6).

III. Control Problem Formulation

The *stabilization problem* is defined by the closed-loop equation

$$x_{n+1} = T \circ (1 + K)(x_n) \doteq T_K(x_n), \tag{8}$$

where T is the plant and $1 + K$ is the control; 1 denotes the identity mapping and $(1 + K)(x_n) = x_n + K(x_n)$. In this paper, it is always assumed that there is a compact set $X \subset \mathbb{R}^n$ such that $T_K : X \to X$. The composition as well as the mappings T and K are assumed to be well-defined for all $x \in X$. The control is said to be *inactive* if $K = 0$.

The *disturbance rejection problem* corresponds to the random version of the stabilization problem

$$x_{n+1} = T_d \circ (1 + K_d)(x_n), \tag{9}$$

where T_d and $1 + K_d$ are random maps. Often,

$$T_d(u) = T(u) + d \tag{10}$$

where d is a random variable taking values from a given distribution. In particular, Eq. (10) leads to a disturbance rejection problem

$$x_{n+1} = T \circ [(1 + K)(x_n) + d_n)] + d_{n+1}, \tag{11}$$

where $\{d_n\}$ is disturbance, modeled as an i.i.d random process.

Example III.1. Consider the feedback loop in Fig. 1 with linear equations of:

$$\begin{aligned} \text{state}: \quad z_{n+1} &= a(z_n + u_n), \\ \text{output}: \quad x_n &= z_n + d_n, \\ \text{control}: \quad u_n &= k(x_n). \end{aligned} \tag{12}$$

The closed-loop equation for the output is given by

$$x_{n+1} = a((1+k)x_n - d_n) + d_{n+1}. \tag{13}$$

The disturbance rejection problem in Eq. (11) is thus the nonlinear generalization of the linear disturbance rejection problem with full-state feedback. With no disturbance, one recovers the nonlinear stabilization problem in Eq. (8) as the special case of Eq. (11). ■

Denote $\mathbb{P}_T$ and $\mathbb{P}_K$ to be the infinite-dimensional P-F operators corresponding to mappings T and $1+K$, respectively. The propagation of measures (in $\mathcal{M}(X)$) for the composition in Eq. (8) is given by

$$\mu_{n+1} = \mu_n \mathbb{P}_K \cdot \mathbb{P}_T, \tag{14}$$

where the ordering reflects the fact that controller ($\mathbb{P}_K$) acts before the plant ($\mathbb{P}_T$). If control is inactive ($K = 0$) then $\mathbb{P}_K = 1$, the identity operator. A similar stochastic description also exists for the composition in Eq. (11), where the P-F operators now correspond to the random mappings.

This also leads in a natural way to a control problem expressed in terms of finite-dimensional Markov chains as follows. Consider $\mathcal{X}_L \doteq \{D_1, \cdots, D_L\}$, a finite partition of X together with the associated measure space $\mathbb{R}^L$. Let P_T and P_K denote the finite-dimensional sub-Markov or Markov matrices on this measure space. The Eq. (14) is *formally* replaced by its finite-dimensional analogue

$$\mu_{n+1} = \mu_n P_K \cdot P_T. \tag{15}$$

A. Stabilization Problem

For the infinite-dimensional case, the stabilization problem is given by the closed-loop Eq. (8) or its stochastic analogue in Eq. (14). With a finite partition, the starting point of the stabilization problem is in fact taken to be the disturbance rejection problem

$$x_{n+1} = T_d \circ (1 + K_d)(x_n). \tag{16}$$

Given a partition $\mathcal{X}_L = \{D_1, \ldots, D_L\}$, we associate a random vector $\mathcal{D} = \{\partial(1), \ldots, \partial(L)\}$, where $\partial(i)$ is uniformly distributed perturbation with support on cell D_i. Given $z \in D_i \subset X$, define a random variable

$$y = z + \partial(|z|), \tag{17}$$

where $|z| = i$. Intuitively, $\partial(|z|)$ randomizes the state z with in the cell. y and z are both located in the same cell D_i and contain the same information modulo the partition. We call such a random perturbation to be certain w.r.t the finite partition. In a suitable limit of taking finer partitions, one recovers the infinite-dimensional problem. We use $\mathcal{D}$ to define T_d and K_d for the stabilization problem w.r.t $\mathcal{X}_L$:

$$\begin{aligned} T_d(u) &\doteq T(u) + \partial(|T(u)|), \\ (1+K_d)(u) &\doteq (1+K)(u) + \partial(|(1+K)(u)|). \end{aligned} \tag{18}$$

As an example, consider a linear expanding map

$$z_{n+1} = a z_n, \tag{19}$$

where $a = 2$ is the expansion rate. The control objective is to *stabilize* the equilibrium 0. Using Eq. (16) and (18), the closed-loop equation is given by

$$x_{n+1} = a((1+k)(x_n) - d_n) + d_{n+1} \doteq a_d(1+k_d)(x_n), \tag{20}$$

where $(1+k)$ is a possibly nonlinear control mapping. For the infinite-dimensional case, the disturbance $\{d_n\}$ is assumed to be zero. For the finite partition $\mathcal{X}_L$, the disturbance is defined using $\mathcal{D}$. This is illustrated with the aid of a concrete choice of partition.

Consider two cells $D_1 = [0, \epsilon]$, $D_2 = [\epsilon, 2\epsilon]$, and denote $X = D_1 \cup D_2$ and $S = (2\epsilon, \infty)$. For Eq. (19), the plant $T : X \to X \cup S$, where S is T-invariant. As a result, the sub-stochastic operator for the restriction $T : X \to X \cup S$ is well-defined. On $\mathcal{X}_2 = \{D_1, D_2\}$, the 2×2 sub-Markov matrix is

$$P_T = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 0 \end{bmatrix}. \tag{21}$$

With respect to the partition $\mathcal{X}_2$, the control objective is to "stabilize" the *cell* D_1 containing 0. The notion of stability is stochastic and will be made precise in the following subsection. On $\mathcal{X}_2$, denote by P_K the Markov chain corresponding to control $(1+k)$. The Markov chain for the closed-loop equation (20) is formally written as

$$\mu_{n+1} = \mu_n P_K \cdot P_T. \tag{22}$$

Note that Eq. (22) provides a *linear* description of Eq. (20) irrespective of whether $1 + k$ is linear, nonlinear, or even stochastic map. As an example, consider

$$P_K = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}. \tag{23}$$

The closed-loop P-F is given by

$$P_K \cdot P_T = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}. \tag{24}$$

The interpretation here is that P_T and P_K correspond not to the *deterministic* open-loop plant (a) and control ($1 + k$) rather the closed-loop Eq. (20) with a suitable choice of disturbance $\{\partial(1), \partial(2)\}$. For $\mathcal{X}_2$, the random variable $\partial(1)$ may be constructed as follows. Let r be uniformly distributed on $[0, \epsilon]$. For $z \in D_1$,

$$\begin{aligned} \partial(1) &= r \text{ if } z + r < \epsilon, \\ &= r - \epsilon \text{ otherwise.} \end{aligned} \tag{25}$$

Any particular P_K corresponds to numerous deterministic and random control mappings $(1 + k)$. As an example, corresponding to P_K in Eq. (23), the control may be a linear gain $k = -0.5$, a random gain where k is a random variable chosen from set $[-0.5, -1)$ or a nonlinear quantization based feedback controller such that

$$\begin{aligned} (1+k)x_n &= x_n \quad \text{for } x_n \in [0, \epsilon] \\ &= x_n - \epsilon \text{ for } x_n \in (\epsilon, 2\epsilon]. \end{aligned} \tag{26}$$

Finally note that $a_d(1 + k_d) : X \to X$ (in Eq. (20)) for all these cases. The only invariant measure for $P_K \cdot P_T$ in this example is given by $[1/2, 1/2]$. Formally, this implies that asymptotically the trajectory goes to cell D_1 with probability $1/2$, i.e., the cell D_1 is not stable in the conventional sense.

B. Objective of Fundamental Limitations

The objective of a fundamental limitations study is to obtain controller-independent performance bounds on stabilization and the disturbance rejection problem. For the infinite-dimensional problem, it is assumed that K is stabilizing, i.e.,

1) $T \circ (1 + K) : X \to X$ for some compact set $X \subset X$,
2) $T_K(0) = 0$ and 0 is asymptotically stable w.r.t. initial conditions in X.

For the finite-dimensional problem with partition $\mathcal{X}_L$, one *assumes* a stabilizing K and considers the disturbance rejection problem in Eq. (11), where T_d and K_d denote the random perturbations (using $\mathcal{D}$) of T and K respectively. The following lemma, quoted without proof, clarifies the relationship between the attractor set of Eq. (11) and the invariant measure of the discrete formulation $P_K \cdot P_T$.

Lemma III.2. *Consider a partition $\mathcal{X}_L$ with disturbance $\mathcal{D}$. Denote P_T and P_K as the Markov chains for T and $(1+K)$ respectively. Suppose, the closed-loop $T_d \circ (1+K_d) : X \to X$, and $P_K \cdot P_T$ has a unique invariant measure $\mu = [\mu_1, \ldots, \mu_L]$. Then, the output $\{x_n\}$ for the closed-loop Eq. (11) has a stationary distribution given by*

$$\textit{Prob}(x_n \in D_i) = \mu_i, \tag{27}$$

and uniform in D_i.

The proof relies on the fact that the discrete Markov chains P_T and P_K ares in fact the random perturbations of the maps T and K; cf., [18]. Here, the disturbance d is chosen so that $P_K \cdot P_T$ is the stochastic analogue of $T_d \circ (1 + K_d)$. In the limit of taking finer partitions, one recovers the stabilization problem $T \circ (1 + K)$ and its infinite-dimensional stochastic analogue $\mathbb{P}_K \cdot \mathbb{P}_T$.

In this paper, fundamental limitations are expressed within the stochastic framework, i.e., by replacing T_d and a given stabilizing K_d by P_T and P_K. For this, it is useful to extend the notion of stability in terms of invariant measures of $P_K \cdot P_T$ as follows.

Definition III.3. Consider a closed-loop system $T_d \circ (1 + K_d) : X \to X$ together with an attractor set $A \subset X$ and a corresponding unique physical invariant measure μ. A set $S \subset \mathcal{B}(X)$ is **q-stable** if

$$\mu(S) = q. \tag{28}$$

A set $S \subset \mathcal{B}(X)$ is *stable* if it is q-stable with $q = 1$. Uniqueness of physical measure and $\mu(S) = 1$ implies that for typical initial conditions $x_0 \in X$, asymptotically $x_n = (T_d \circ (1+K_d))^n(x_0) \to S$ as $n \to \infty$ with probability 1; cf., [14]. The q-stability provides a weaker notion of stability that is useful to the study of fundamental limitations in nonlinear stabilization

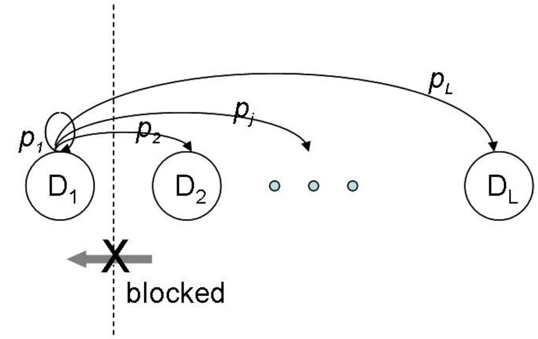

Fig. 2. Markov chain where the first row and column are given by Eqs. (29) and (30) respectively.

IV. Stabilization with Markov Chains

In this section, the stabilization problem for a finite partition is considered. Consider $\mathcal{X}_L = \{D_1, D_2, \ldots, D_L\}$ together with a matrix P_T whose first row denoted as

$$[P_T]_1 = [p_1, p_2, \ldots, p_j, \ldots, p_L] \tag{29}$$

is assumed to be given. We assume that P_T is either a Markov or a sub-Markov matrix, so $p_i \geq 0$ and $\sum_i p_i \leq 1$. Additionally, assume that the first column denoted as

$$[P_T]^1 = [p_1, 0, \ldots, 0]'. \tag{30}$$

has atmost one non-zero entry p_1. The resulting Markov chain is drawn in Fig. 2. It corresponds to certain dynamical systems with unstable equilibrium in cell D_1. For e.g., the Markov chain in Eq. (21) has

$$\begin{array}{l} [1/2, 1/2], \\ [1/2, 0]' \end{array} \tag{31}$$

as its first row and column respectively. For the finite-dimensional case, the control objective is to design P_K to q-stabilize D_1 with maximum possible value of q. For this problem:

1) Theorem IV.1 present a controller-independent upper bound on the maximum possible value of q and
2) Theorem IV.2 presents the control (Markov chain P_K) that achieves this bound.

In the following section, we use these results to express fundamental limitations for the stabilization problem in Eq. (8).

Theorem IV.1. *Consider P_T defined on a partition $\mathcal{X}_L$ with the structure in Fig. 2 and the first row and column given in Eqs. (29) and (30) respectively. Let P_K denote* **any** *control Markov matrix on $\mathcal{X}_L$ without additional assumptions on its structure. If $\mu = [\mu_1, \ldots]$ is an invariant measure for the closed-loop Markov matrix $P_K \cdot P_T$ then*

$$\mu_1 \leq p_1, \tag{32}$$

i.e., the maximum possible value for q-stabilization of D_1 is given by $q = p_1$.

Proof. Let $\mu = \mu P_K \cdot P_T$. Due to the assumption on the first column of P_T (see Eq. (30)),

$$\mu_1 =< \mu, [P_K]^1 > p_1, \quad (33)$$

where $[P_K]^1$ denotes the first column of the Markov chain P_K and $<,>$ denotes the standard inner product for the two vectors. If μ is a probability measure then $\sum_i \mu_i = 1$ and

$$\mu_1 \leq \sum_i \mu_i \cdot \max_i [P_K]_{i1} \cdot p_1 \leq 1 \cdot p_1. \quad (34)$$

This gives the desired inequality. □

Theorem IV.2. *Assume the notation of the Theorem IV.1 and let $p_1 > 0$. Suppose using some control P_K, D_1 is made p_1-stable. Then*

1) $\sum_{i=1}^{L} p_i = 1$,
2) *For every i with $p_i \neq 0$, the i^{th} row of the control $[P_K]_i = [1, 0, \ldots]$.*
3) *$\mu = [p_1, \ldots, p_L]$ denotes an invariant probability measure of the closed-loop Markov chain $P_K \cdot P_T$.*

Proof. Using Eq. (33) together with the fact that $\sum_i \mu_i = 1$,

$$\mu_1 = p_1 \text{ implies } [P_K]_{i1} = 1 \text{ whenever } \mu_i \neq 0. \quad (35)$$

In particular, because $\mu_1 = p_1 > 0$, $[P_K]_{11} = 1$ and because P_K is a Markov matrix, its first row

$$[P_K]_1 = [1, 0, \ldots]. \quad (36)$$

This show (2) for the particular case of $i = 1$. By matrix multiplication,

$$[P_K \cdot P_T]_1 = [p_1, p_2, \ldots, p_L]. \quad (37)$$

Now, if $P_K \cdot P_T$ is a Markov matrix, necessarily

$$\sum_i p_i = 1. \quad (38)$$

This shows (1). Since μ is an invariant measure,

$$[\mu_1, \mu_2, \ldots, \mu_L] \begin{bmatrix} p_1 & p_2 & \cdots & p_L \\ \times & \times & & \times \\ \vdots & \vdots & & \vdots \end{bmatrix} = [\mu_1, \mu_2, \ldots, \mu_L]. \quad (39)$$

Hence,

$$\mu_i = \mu_1 p_i + \ldots \geq p_1 p_i \quad (40)$$

and thus $\mu_i > 0$ for all i such that $p_i > 0$. Using the condition in Eq. (35),

$$[P_K]_i = [1, 0, \ldots] \text{ whenever } p_i > 0. \quad (41)$$

This shows (2). Once again, by matrix multiplication

$$[P_K \cdot P_T]_i = [p_1, p_2, \ldots, p_L] \text{ whenever } p_i > 0 \quad (42)$$

Finally, because of Eqs. (38) and (42), it is easy to verify that $\mu_j = p_j$ is an invariant probability measure for the Markov chain $P_K \cdot P_T$. □

These theorems show limitations on a) maximum achievable value of q for q-stability of D_1 and b) the resulting invariant measure. Either of these are a function of *only* the properties of the open-loop Markov chain P_T. The key assumption needed for the conclusions is the structure of P_T w.r.t. D_1; expressed by Eq. (30). Where this equation fails to hold, these restrictions are no longer valid as shown by the following example.

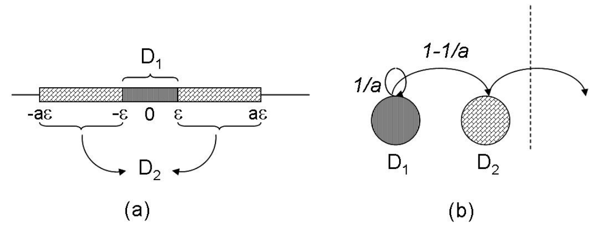

Fig. 3. The phase-space partition $\mathcal{X}_2$ and the Markov chain P_T.

Example IV.3. Consider a 2-state Markov chain on $X = \{D_1, D_2\}$

$$P_T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (43)$$

Using the notation from Theorem IV.1, $p_1 = 0$. However, a control with

$$P_K = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \quad (44)$$

q-stabilizes D_1 with $q = 1$. The resulting invariant measure for the closed-loop $P_K \cdot P_T$ is given by $[1, 0]$. ■

The following section uses these results to derive fundamental limitations for the scalar stabilization problem. These can also be expressed in terms of entropy and the entropy based formulas are the counterpart of the Bode integral formula in Eq. (1).

V. Fundamental Limitations

Consider first a scalar linear dynamical system,

$$\text{state}: \quad z_{n+1} = b(z_n + u_n), \quad (45)$$

with the expansion rate $a \doteq |b| > 1$. Denote

$$\begin{aligned} D_1 &= [-\epsilon, \epsilon], \\ D_2 &= a(D_1) - D_1 = [-a\epsilon, -\epsilon] \cup [\epsilon, a\epsilon], \end{aligned} \quad (46)$$

where $a(D_1) = [-a\epsilon, a\epsilon]$. Use these cells to set

$$\begin{aligned} X &= a(D_1) = D_1 \cup D_2, \\ \mathcal{X}_2 &= \{D_1, D_2\}. \end{aligned} \quad (47)$$

Associated with the discrete partition, the first row and column of the Markov chain P_T (for the state evolution $Tz = bz$) are given by

$$\begin{aligned} [P_T]_1 &= [\frac{1}{a}, 1 - \frac{1}{a}], \\ [P_T]^1 &= [\frac{1}{a}, 0]', \end{aligned} \quad (48)$$

respectively. Figure 3 depicts the partition $\mathcal{X}_2$ and the Markov chain P_T. The equations for the output and control are

$$\begin{aligned} \text{output}: \quad x_n &= z_n + d_n \quad (49) \\ \text{control}: \quad u_n &= k(x_n), \quad (50) \end{aligned}$$

where k is assumed to a linear stabilizing gain and d_n denotes some random disturbance. The closed-loop equation for the output is given by

$$x_{n+1} = b((1+k)x_n - d_n) + d_{n+1}. \quad (51)$$

With respect to the finite partition $\mathcal{X}_2$, the disturbance $\{d_n\}$ is constructed from the random vector $\{\partial(1), 0\}$. At each n, $\partial(1)$ is used to provide a random perturbation in cell D_1. We call this a stabilization problem because the disturbance is *certain* w.r.t D_1. The stochastic analogue of Eq. (51) is the now familiar

$$\mu_{n+1} = \mu_n P_K P_T, \quad (52)$$

where P_K is the Markov chain corresponding to the stabilizing control $(1+k)$. With finite partition, the stabilization problem is posed as the q-stabilization of cell D_1. The following Theorem shows the relationship between this and the original problem.

Theorem V.1. *Suppose $u_n = k(x_n)$ be* **any** *linear stabilizing control of Eq. (45) and $\mathcal{X}_2$ be a partition in Eq. (47) with arbitrary ϵ. Then for the closed-loop Eq. (51), the cell D_1 is q-stable with maximal value of $q = \frac{1}{a}$.*

Proof. The condition for closed-loop stability is

$$|a(1+k)| < 1 \quad (53)$$

which necessarily implies that $(1+k) : a(D_1) \to D_1$. The Markov chain for the control $(1+k)$ is then given by

$$P_K = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}. \quad (54)$$

Using Eq. (48), the closed-loop Markov chain is

$$P_K \cdot P_T = \begin{bmatrix} \frac{1}{a} & 1-\frac{1}{a} \\ \frac{1}{a} & 1-\frac{1}{a} \end{bmatrix} \quad (55)$$

and its closed-loop invariant measure is

$$\mu = [\frac{1}{a}, 1-\frac{1}{a}], \quad (56)$$

i.e., the cell D_1 is $\frac{1}{a}$-stable. Using Theorem IV.1, this is also the maximum value possible. □

The inability to q-stabilize the cell D_1 for arbitrary value of q (an in particular, for $q = 1$) constitutes a fundamental limitation in stabilization. This depends only upon the open-loop dynamics, expansion rate a here, and is independent of the choice of the feedback control gain k. Conversely, as the lemma shows, *any* stabilizing control achieves the upper bound $\frac{1}{a}$. Next, larger the value of a, the larger the uncertainty in the state $\{x_n\}$ of closed-loop system – it could be anywhere in cell D_2 whose length scales as a. This uncertainty is best expressed in terms of the entropy metric:

Definition V.2. The conditional entropy of a random sequence $\{x_n\}$ is given by

$$H_c(x) = \lim_{m\to\infty} H(x_n | x_{n-1}, \ldots, x_{n-m}). \quad (57)$$

The following theorem provides an explicit estimate for the control problem.

Theorem V.3. *Consider the closed-loop Eq. (51) with the expansion rate $a > 1$. For* **any** *stabilizing control gain k, the output sequence $\{x_n\}$ has entropy given by*

$$H_c(x) = \ln(a). \quad (58)$$

Proof. The invariant measure μ in Eq. (56) implies that the stationary distribution for $\{x_n\}$ is given by,

$$\begin{aligned} \text{Prob}(x_n \in D_1) &= \frac{1}{a} \\ \text{Prob}(x_n \in D_2) &= 1 - \frac{1}{a}. \end{aligned} \quad (59)$$

This correspondence is setup by the explicit choice of the disturbance $\{d_n\}$; cf., Lemma III.2. Furthermore, x_n is uniformly distributed with in each cell and its pdf is given by

$$\begin{aligned} f(x) &= \frac{1}{a\epsilon} \quad \text{for } x \in D_1, \\ &= (1-\frac{1}{a}) \cdot \frac{1}{(a-1)\epsilon} \quad \text{for } x \in D_2, \end{aligned} \quad (60)$$

i.e., $f(x) = \frac{1}{a\epsilon}$ is the uniform distribution for $x \in a(D_1)$. A simple calculation then shows that the entropy of x_n is thus given by

$$H(x_n) = \ln(a) - \ln(\epsilon), \quad (61)$$

where $\ln(\epsilon)$ is ignored by convention of defining entropy for a continuous random variable.

Next, the closed-loop Markov chain relating x_2 to x_1 is given by

$$P_K \cdot P_T = \begin{bmatrix} \frac{1}{a} & 1-\frac{1}{a} \\ \frac{1}{a} & 1-\frac{1}{a} \end{bmatrix}. \quad (62)$$

Thus the conditional probability

$$\begin{aligned} \text{Prob}(x_2 \in D_1 | x_1 \in D_i) &= \frac{1}{a} \\ \text{Prob}(x_2 \in D_2 | x_1 \in D_i) &= 1 - \frac{1}{a}. \end{aligned} \quad (63)$$

Once again, because x_n is uniformly distributed within a cell, the conditional pdf for the stationary case is given by

$$\begin{aligned} f(x_2|x_1) &= \frac{1}{a\epsilon} \quad \text{for } x_2 \in D_1, \\ &= (1-\frac{1}{a}) \cdot \frac{1}{(a-1)\epsilon} \quad \text{for } x_2 \in D_2, \end{aligned} \quad (64)$$

i.e., $f(x_2|x_1) = \frac{1}{a\epsilon}$ for $x_2, x_1 \in a(D_1)$. Now, applying the formula for relative entropy [2], [3], we have

$$\begin{aligned} H(x_2|x_1) &= -\int_{aD_1} f(x_1) \int_{aD_1} f(x_2|x_1) \ln(f(x_2|x_1)) dx_2 dx_1 \\ &= \ln(a) - \ln(\epsilon), \end{aligned} \quad (65)$$

where $\ln(\epsilon)$ is once again ignored as a result of the convention. The proof is completed by noting that x_n depends only upon x_{n-1} and not its entire history, i.e., $\{x_n\}$ is a Markov process. For a stationary Markov process, the conditional entropy

$$H(x_n|x_{n-1},\ldots,x_{n-m}) = H(x_n|x_{n-1}) = H(x_2|x_1), \tag{66}$$

where the first equality is due to the Markov assumption and the second equality is due to the stationarity [3]. This is a crucial step that makes the estimate feasible for the state feedback problem. □

Since ϵ is arbitrary, one interprets the limit as the stabilization problem. Note that the disturbance d_n goes to zero as the partition size $\epsilon \to 0$. However, the role of disturbance can not be over-stressed. At each point, one is indeed solving a disturbance rejection problem where the disturbance is assumed to be *certain* w.r.t the partition. Using Eq. (2), one also obtains the connection with Bode integral formula.

Finally, we relax the assumption that $b(\cdot)$ in Eq. (45) is a linear dynamical system. Consider now a closed-loop equation

$$x_{n+1} = b \circ (1+k)(x_n), \tag{67}$$

where $b(0) = 0$ and $a \doteq |b'(0)| > 1$. Let k be a stabilizing control such that the linearized closed-loop equation

$$y_{n+1} = b'(0) \cdot (1 + k'(0))(y_n) \tag{68}$$

is asymptotically stable, i.e., $|a \cdot (1+k'(0))| < 1$. We note that the proof of Theorem V.3 did not use linearity of either b and $(1+k)$. In fact, we did not even use these mappings, rather only the Markov chains P_K and P_T. Below, we discuss their construction for the nonlinear problem.

We assume here that both the plant and control are smooth dynamical systems. We denote $U^\epsilon \doteq [-\epsilon, \epsilon]$, an interval neighborhood of the equilibrium. Due to scalar nature of the dynamical system and the assumption of asymptotic stability, we have in a sufficiently small neighborhood U^ϵ:

$$\begin{aligned} b(U^\epsilon) &\supset U^\epsilon, \\ b(1+k)(U^\epsilon) &\subset U^\epsilon. \end{aligned} \tag{69}$$

In particular, the Grobman-Hartman theorem [19] shows that there exists a near-identity invertible co-ordinate change $h: y \to x$ with $h(0) = 0$, $h'(0) = 1$ between the closed-loop Eqs. (67)-(68). Thus, for a suitable neighborhood of 0, a change of co-ordinate,

$$x_n = h(y_n) \tag{70}$$

allows one to obtain solutions of the nonlinear Eq. (67) in terms of solution of the linear Eq. (68). Since $b(\cdot)$ is assumed to be smooth and invertible, denote $h = b \circ h_1$. We have $h_1(0) = 0$, $h_1'(0) = \frac{1}{b'(0)} \neq 0$, and $\lim_{\epsilon \to 0} h_1(U^\epsilon) = 0$, i.e., h_1 provides another invertible change of co-ordinate near 0. By the Grobman-Hartman theorem,

$$b(1+k) : b \circ h_1(U^\epsilon) \to b \circ h_1(U^\epsilon), \tag{71}$$

for all $\epsilon \leq \epsilon_0$ chosen to be sufficiently small such that $h_1(U^\epsilon) \subset b \circ h_1(U^\epsilon)$; see Eq. (69). Denote,

$$\begin{aligned} D_1 &\doteq h_1(U^\epsilon), \\ D_2 &\doteq b \circ h_1(U^\epsilon) - h_1(U^\epsilon), \end{aligned} \tag{72}$$

and $X^\epsilon = D_1 \cup D_2 = h_1(U^\epsilon)$. Now, due to Eq. (71), we necessarily have

$$(1+k) : b \circ h_1(U^\epsilon) \to D_1. \tag{73}$$

As a result, with respect to the partition $\mathcal{X}_2^\epsilon = \{D_1, D_2\}$, the Markov chain for the stabilizing control $(1+k)$ is given by Eq. (54). So, all we have to do is derive the first row of the Markov chain P_T corresponding to the plant $b(\cdot)$ in order to estimate on the entropy. For the stabilization limit as $\epsilon \to 0$, this derivation and the resulting estimate appears in the following Corollary.

Corollary V.4. *Consider the closed-loop Eq. (51) where b and $(1+k)$ are nonlinear dynamical systems and the expansion rate $a \doteq |b'(0)| > 1$. In the limit of vanishing disturbance ($\epsilon \to 0$), the output sequence $\{x_n\}$ has entropy given by*

$$H_c(x) = \ln(a) \tag{74}$$

for **any** *stabilizing control k.*

Proof. Since P_K is given by Eq. (54), we need estimate only the first row of the Markov matrix P_T, where $Tz \doteq b(z)$ denotes the state evolution. In fact, for the two cell partition, we need estimate only the $[P_T]_{11}$ entry, denoted here by q^ϵ. The superscript is used to make the dependence upon ϵ explicit. The closed-loop Markov matrix for $\mathcal{X}_2^\epsilon$ is then given by

$$P_K \cdot P_T = \begin{bmatrix} q^\epsilon & 1-q^\epsilon \\ q^\epsilon & 1-q^\epsilon \end{bmatrix}. \tag{75}$$

Using the proof of the Theorem V.3, one obtains an entropy estimate for the disturbance rejection problem corresponding to $\mathcal{X}_2^\epsilon$ as

$$H_c(x) = -\ln(q^\epsilon). \tag{76}$$

To prove this corollary, we show that

$$\lim_{\epsilon \to 0} q^\epsilon = \frac{1}{a}. \tag{77}$$

Indeed, the $\epsilon = 0$ limit is

$$\begin{aligned} q^0 &= \lim_{\epsilon \to 0} \frac{m(T^{-1}D_1 \cap D_1)}{m(D_1)} \\ &= \lim_{\epsilon \to 0} \frac{m(T^{-1}D_1)}{m(D_1)} \\ &= \left|\frac{dT^{-1}}{dx}(0)\right| = \frac{1}{a}, \end{aligned} \tag{78}$$

where $a = |b'(0)|$ is the expansion rate. Eq. (77) and the result follows. □

VI. Conclusions

In this paper, we presented a dynamical systems based framework and fundamental limitation results for the nonlinear stabilization problem. The stochastic framework and associated discrete Markov chains allow one to both interpret and compute the entropy estimates in a rather straightforward manner. The entropy estimates correspond to the so-called measure-theoretic entropy in the Dynamical systems literature. There are two extensions of this work. One is the problem of multi-state stabilization, results for which will appear in the journal version of this paper. The other is the nonlinear disturbance rejection with state feedback which is a subject of continuing research. The Markov chain based framework presented here is expected to be particularly applicable for this problem.

REFERENCES

[1] H. K. Sung and S. Hara, "Properties of sensitivity and complementary sensitivity functions in single-input single-output digital control systems," *Int. J. Control*, vol. 48, no. 6, pp. 2429–2439, 1988.

[2] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecommunications. New York: Wiley Interscience, 1991.

[4] G. Zang and P. A. Iglesias, "Nonlinear extension of Bode's integral based on an information theoretic interpretation," *Systems and Control Letters*, vol. 50, pp. 11–29, 2003.

[5] G. N. Nair, R. J. Evans, and I. M. Y. Mareels, "Topological feedback entropy and nonlinear stabilization," *IEEE Transaction on Automatic Control*, vol. 49, no. 9, pp. 1585–1597, 2004.

[6] N. Elia, "When Bode meets Shannon: Control-oriented feedback communication schemes," *IEEE Transaction on Automatic Control*, vol. 49, no. 9, 2004.

[7] N. C. Martins and M. A. Dahleh, "Feedback control in the presence of noisy channels: "Bode-like" fundamental limitations of performance," 2004, preprint.

[8] N. Elia, "Stabilization of linear systems with limited information," *IEEE Transaction on Automatic Control*, vol. 46, no. 9, pp. 1384–99, 2001.

[9] D. Liberzon, "On stabilization of linear systems with limited information," *IEEE Transaction on Automatic Control*, vol. 48, no. 2, pp. 304–307, 2003.

[10] S. Tatikonda and S. K. Mitter, "Control under communication constraints," *IEEE Transaction on Automatic Control*, vol. 49, no. 7, pp. 1056–68, 2004.

[11] F. Fagnani and S. Zampieri, "Stability analysis and synthesis for scalar linear systems with quantized feedback," *IEEE Transaction on Automatic Control*, vol. 48, no. 9, 2003.

[12] D. Liberzon, "Stabilization of nonlinear systems with limited information feedback," *IEEE Transaction on Automatic Control*, vol. 50, no. 6, pp. 910–915, 2005.

[13] G. Froyland, "Using Ulam's method to calculate entropy and other dynamical invariants," *Nonlinearity*, vol. 12, pp. 79–101, 1999.

[14] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. New York: Springer-Verlag, 1994.

[15] S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*. Berlin: Springer-Verlag, 1993.

[16] L. Arnold, *Random Dynamical Systems*. Springer-Verlag, 1998.

[17] Y. Kifer, *Ergodic Theory of Random Transformations*, ser. Progress of Probability and Statistics. Boston: Birkhauser, 1986, vol. 10.

[18] G. Froyland, "Extracting dynamical behaviour via markov models," *Submitted to Nonlinear dynamics and statistics*, manuscript available at citeseer.ist.psu.edu/196746.html. [Online]. Available: citeseer.ist.psu.edu/196746.html

[19] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, ser. Applied Mathematical Sciences. New York, NY: Springer-Verlag, 1983, no. 42.