Invariance and Stationarity in Non-Linear Networked Control: Deterministic and Stochastic Formulations

Christoph Kawan and Serdar Yüksel

***Abstract*— In this paper we compare two recent threads of research on the stabilization of non-linear control systems under information constraints. In a deterministic setting, characterizations of the smallest bit rate of a noiseless channel connecting a state coder to a controller have been obtained, above which local stabilization or set-invariance is possible. These characterizations use quantities similar to topological entropy in dynamical systems and lead to estimates of the corresponding bit rates in terms of Lyapunov exponents. In a stochastic setting, a decisive quantity is the channel capacity. For different notions of stochastic stability, namely stationarity, ergodicity and asymptotic mean stationarity, estimates for the smallest capacity, above which stabilization in one of these senses can be accomplished, are available. We review the results of the deterministic and the stochastic theory, highlight their analogies and compare the corresponding methods of proof.**

I. Introduction

Networked control systems challenge the paradigm of classical control theory that information can be transmitted instantaneouly, lossless and with arbitrary precision. Under the assumption that the communication between different components of a networked system is limited by the use of finite-capacity channels, it is a natural problem to determine the class of communication channels through which it is possible to stabilize the system. Already in the simplest network topology – one dynamical system and one controller connected by a digital channel – (see Fig. 1) this problem is very hard to solve for non-linear systems with currently existing techniques. This paper reviews and compares partial results that have been obtained in deterministic and in stochastic settings.

In the deterministic setup, the system is modeled by difference or differential equations of the form

$$x_{t+1} = f(x_t, u_t) \quad \text{or} \quad \dot{x}(t) = f(x(t), u(t))$$

and the channel is modeled as a 'bit-pipe', allowing a fixed number r of bits to be transmitted error-free in each time slot.

In the stochastic setup, a discrete-time model of a control system takes the form

$$x_{t+1} = f(x_t, w_t, u_t), \tag{1}$$

where w_t is a model for the noise, usually assumed to be an i.i.d. sequence of random variables. The channel in this model is also stochastic, in general, with memory and (noiseless) feedback. We present the details of this setup in Section III.

Stability Notions. On the deterministic side, we are concerned with necessary and sufficient conditions on the channel leading to invariance of a given set under some control and coding policy. On the stochastic side, we are concerned with conditions on a noisy communication channel in the presence of which there exist coding and control policies such that the controlled system is stochastically stable in the following senses: (i) $\{x_t\}$ is asymptotically stationary, or asymptotically mean stationary (AMS) and satisfies Birkhoff's sample path ergodic theorem, (ii) $\{x_t\}$ is ergodic, and in case a Markov chain can be constructed (iii) $\{x_t\}$ is a stable Markov chain.

Throughout the paper, we use the convention that $\log = \log_2$, when we are dealing with discrete-time systems and $\log = \log_e$ for continuous-time systems.

II. The Deterministic Theory

A. Entropy notions

To explain the notion of feedback entropy [9], we first recall the definition of topological entropy (cf. [1]). Let X be a compact topological space and $f : X \to X$ a continuous map. The solutions of the difference equation

$$x_{n+1} = f(x_n)$$

are given by $x_n = f^n(x_0)$, where $f^n : X \to X$ is the n-th iterate of f. If $\mathcal{U}, \mathcal{V}$ are open covers of X, we define the *pullback of* $\mathcal{U}$ *under* f by $f^{-1}\mathcal{U} := \{f^{-1}(U) : U \in \mathcal{U}\}$ and the *join of* $\mathcal{U}$ *and* $\mathcal{V}$ by $\mathcal{U} \vee \mathcal{V} := \{U \cap V : U \in \mathcal{U}, V \in \mathcal{V}\}$, which are again open covers. The last operation can be extended recursively for any finite number of open covers. We write $\mathcal{N}(\mathcal{U})$ for the minimal cardinality of a finite subcover of $\mathcal{U}$. The *entropy of* f *on the cover* $\mathcal{U}$ is given by

$$h(f;\mathcal{U}) := \lim_{n\to\infty} \frac{1}{n} \log \mathcal{N}\Big(\bigvee_{i=0}^{n-1} f^{-i}\mathcal{U}\Big),$$

where $f^{-i} = (f^i)^{-1}$ and the existence of the limit follows from subadditivity. The *topological entropy of* f is

$$h_{\text{top}}(f) := \sup_{\mathcal{U}} h(f;\mathcal{U}) \in [0,\infty) \cup \{\infty\},$$

the supremum taken over all open covers of X. Topological entropy has been constructed in strict analogy to the measure-theoretic entropy of Kolmogorov and Sinai, which we will comment on later.

C. Kawan is with the Fakultät für Informatik und Mathematik, Universität Passau, Innstraße 33, 94032 Passau, Germany christoph.kawan@uni-passau.de

S. Yüksel is with the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6 yuksel@mast.queensu.ca

To explain the concept of (topological) feedback entropy, let X be a topological space, U a non-empty set and

$$x_{n+1} = f(x_n, u_n) \tag{2}$$

a discrete-time control system on X, where $f : X \times U \to X$ is a map with the property that $f_u := f(\cdot, u) : X \to X$ is continuous for each $u \in U$. We merge all solutions of (2) into one map $\varphi : \mathbb{Z}_+ \times X \times U^{\mathbb{Z}_+} \to X$,

$$\varphi(n, x, \underline{u}) := f_{u_{n-1}} \circ \cdots \circ f_{u_0}(x), \; \underline{u} = (u_0, u_1, \ldots).$$

Let $Q \subset X$ be a compact set with non-empty interior satisfying the following *strong invariance condition*: For every $x \in Q$ there is $u_x \in U$ with $f(x, u_x) \in \mathrm{int}Q$.

Let $\mathcal{A}$ be an open cover of Q, $\tau \in \mathbb{N}$, and $G : \mathcal{A} \to U^\tau$ a map with components $G_0, \ldots, G_{\tau-1}$ that assign control values to all sets in $\mathcal{A}$, such that for each $A \in \mathcal{A}$ the sequence of controls $G(A)$ yields $\varphi(k, A, G(A)) \subset \mathrm{int}Q$ for $k = 1, \ldots, \tau$. Then we call the triple $(\mathcal{A}, \tau, G)$ an *invariant open cover* of Q. The existence of an invariant open cover easily follows from the strong invariance condition on Q and continuity of the maps f_u, $u \in U$, cf. [7, Sec. 2.4].

Now, for any sequence $\alpha := (A_i)_{i \geq 0}$ of sets in $\mathcal{A}$ we define the control sequence

$$\underline{u}(\alpha) := (u_0, u_1, \ldots) \text{ with } (u_l)_{l=(i-1)\tau}^{i\tau-1} \equiv G(A_{i-1}).$$

We further define for each $j \geq 1$ the open set

$$B_j(\alpha) := \{x \in X \mid \varphi(i\tau, x, \underline{u}(\alpha)) \in A_i, \; 0 \leq i < j\}.$$

Furthermore, for each $j \geq 1$, letting α run through all sequences in $\mathcal{A}$, another open cover of Q is given by

$$\mathcal{B}_j := \{B_j(\alpha) \mid \alpha \in \mathcal{A}^{\mathbb{Z}_+}\}.$$

Writing $\mathcal{N}(\mathcal{B}_j|Q)$ for the minimal cardinality of a finite subcover of $\mathcal{B}_j$, we define the *feedback entropy* by

$$h_{\mathrm{fb}}(Q) := \inf_{(\mathcal{A},\tau,G)} \lim_{j\to\infty} \frac{1}{j\tau} \log \mathcal{N}(\mathcal{B}_j|Q),$$

the infimum taken over all invariant open covers of Q. Existence of the limit follows again from subadditivity.

A conceptually simpler definition of $h_{\mathrm{fb}}(Q)$ was given in [4] under the name *invariance entropy*. Though originally defined for continuous time, invariance entropy can easily be adapted to the discrete-time setting above. For any $\tau \in \mathbb{N}$, a set $\mathcal{S} \subset U^\tau$ is called (τ, Q)-spanning if for each $x \in Q$ there is $\underline{u} \in \mathcal{S}$ with $\varphi(j, x, \underline{u}) \in \mathrm{int}Q$ for $j = 1, \ldots, \tau$. The minimal cardinality of such a set is denoted by $r_{\mathrm{inv}}(\tau, Q)$ and the *invariance entropy* of Q is defined by

$$h_{\mathrm{inv}}(Q) := \lim_{\tau\to\infty} \frac{1}{\tau} \log r_{\mathrm{inv}}(\tau, Q), \tag{3}$$

where again subadditivity guarantees the existence of the limit. In [5] it was proved that $h_{\mathrm{inv}}(Q) = h_{\mathrm{fb}}(Q)$.

B. Interpretation

In spite of similarities in the construction of feedback entropy and topological entropy, which also reflect in similar properties, there is no direct relation between these two quantities. While topological entropy detects exponential complexity in the orbit structure of a system, feedback entropy measures the complexity of the control task to keep a system in a given set by applying appropriate inputs. If no escape from this set is possible, the feedback entropy is zero, no matter how complicated the orbit structure is. Hence, topological entropy is sensitive to the local behavior of the system, while feedback entropy in general is not. Interpreted in terms of information, topological entropy is a measure for the largest average rate of information about the initial state a system can generate. In contrast, feedback entropy measures the smallest rate of information about the state above which a controller is able to render the set invariant. This is the contents of the following *data-rate theorem* (cf. [9, Thm. 1]):

II.1 Theorem: *Consider system (2). Suppose that a sensor measures its states and is connected to a controller via a noiseless digital channel which carries one discrete-valued symbol per sampling interval, selected from a coding alphabet S_k of time-varying size. If the transmission data rate*

$$R = \liminf_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \log |S_j|$$

of the channel satisfies $R > h_{\mathrm{fb}}(Q)$, then a coder-controller pair exists, which renders Q invariant. If $R < h_{\mathrm{fb}}(Q)$, then no such coder-controller pair exists.

An important result in the deterministic setup is with regard to stabilization to a point: Under the assumptions that (i) f has the form (2) with $X = \mathbb{R}^N$, $U = \mathbb{R}^m$ and is continuously differentiable, (ii) there exists an equilibrium pair (x^*, u^*), i.e., $x^* = f(x^*, u^*)$, (iii) a local strong invariability condition is satisfied, namely for any $\varepsilon > 0$ there is $\rho > 0$ so that for all $\varepsilon' \in (0, \rho]$, the set $\{x : |x - x^*| \leq \varepsilon'\}$ is strongly invariant with the control value set $U = \{u : |u - u^*| \leq \varepsilon\}$, and (iv) the pair (A, B) with $A = \mathrm{D}_x f(x^*, u^*)$, $B = \mathrm{D}_u f(x^*, u^*)$ is controllable, [9] has reported that for convergence to the equilibrium an average rate $R > \sum_{|\lambda_i|>1} \log |\lambda_i|$ is sufficient, where λ_i are the eigenvalues of A.

C. Estimates in terms of Lyapunov exponents

Using a definition slightly different from (3), several estimates and formulas have been proved, relating the entropy to dynamical characteristics of the control system.

Consider now a continuous-time control system

$$\dot{x}(t) = f(x(t), u(t)), \quad u \in \mathcal{U} = L^\infty(\mathbb{R}, U), \tag{4}$$

where $U \subset \mathbb{R}^m$ is compact with $\mathrm{int}U \neq \emptyset$, M is a Riemannian manifold and $f : M \times \mathbb{R}^m \to TM$ is continuous and smooth in the first argument with $f(x, u) \in T_x M$. We let $\varphi(t, x, u)$ denote the solution of (4) for the control u and the initial state x at time $t = 0$. For simplicity, we assume that all solutions are defined for $t \geq 0$, which yields a map

$$\varphi : \mathbb{R}_{\geq 0} \times M \times \mathcal{U} \to M, \quad (t, x, u) \mapsto \varphi(t, x, u).$$

We also write $\varphi_{t,u} : M \to M$ for the diffeomorphism $\varphi(t, \cdot, u)$. A pair (K, Q) of sets with $K \subset Q \subset M$ is called *admissible* if K is compact and for each $x \in K$ there is $u \in \mathcal{U}$ with $\varphi(\mathbb{R}_{\geq 0}, x, u) \subset Q$. For $\tau > 0$, a set $\mathcal{S} \subset \mathcal{U}$ is called (τ, K, Q)*-spanning* if for each $x \in K$ there is $u \in \mathcal{S}$ with $\varphi([0, \tau], x, u) \subset Q$. Writing $r_{\rm inv}(\tau, K, Q)$ for the minimal cardinality of such a set, we define

$$h_{\rm inv}(K, Q) := \limsup_{\tau \to \infty} \frac{1}{\tau} \log r_{\rm inv}(\tau, K, Q).$$

Under additional assumptions, it is possible to obtain upper bounds for $h_{\rm inv}(K, Q)$ in terms of Lyapunov exponents. Recall that a set $D \subset M$ with non-empty interior is a *control set* if it is maximal with the property of complete approximate controllability, i.e., for each two $x, y \in D$ and every neighborhood N of y there are $u \in \mathcal{U}$ and $\tau > 0$ with $\varphi(\tau, x, u) \in N$. For a proof of the following result, see [3, Thm. 3.5 and Cor. 3]. We refer to [16, Ch. 4] for a definition and implications of the Lie algebra rank condition.

II.2 Theorem: *Let D be a control set with compact closure. Under a strong accessibility rank condition on ${\rm int}D$, for every compact set $K \subset D$, the admissible pair (K, D) satisfies*

$$h_{\rm inv}(K, D) \leq \inf_{(u,x)} \limsup_{t \to \infty} \frac{1}{t} \log^{+} \|{\rm D}\varphi_{t,u}(x)^{\wedge}\|, \qquad (5)$$

where $\log^{+} = \max\{0, \log\}$ and the infimum runs over all $(u, x) \in {\rm int}\mathcal{U} \times {\rm int}D$ such that $\varphi(t, x, u)$ is contained in a compact subset of D for all $t \geq 0$. If the system is control-affine and U is convex, the strong accessibility rank condition can be weakened to the usual Lie algebra rank condition.

In (5), ${\rm D}\varphi_{t,u}(x)^{\wedge}$ denotes the linear map between the full exterior algebras of the tangent spaces T_xM and $T_{\varphi(t,x,u)}M$, induced by the derivative ${\rm D}\varphi_{t,u}(x)$. The norm of ${\rm D}\varphi_{t,u}(x)^{\wedge}$ equals the maximal value that a product of k singular values of ${\rm D}\varphi_{t,u}(x)$ can have when $1 \leq k \leq \dim M$. Using this, it is not hard to prove that the $\limsup$ on the right-hand side of (5) equals the sum of the positive Lyapunov exponents for periodic trajectories. The core argument in proving Theorem II.2 is the following: One fixes a periodic controlled trajectory $(\varphi(\cdot, x_*, u_*), u_*(\cdot))$ in ${\rm int}D$ with controllable linearization. Using complete approximate controllability on D and local controllability along $(\varphi(t, x_*, u_*), u_*(\cdot))$, every initial state in K is first steered into a neighborhood of x_* and then kept close to $\varphi(t, x_*, u_*)$ for all future times, using controls that are close to u_*. On average, the sum S of all positive Lyapunov exponents along $\varphi(\cdot, x_*, u_*)$ tells how fast one is driven away from $\varphi(t, x_*, u_*)$ without applying controls different from u_*. Hence, the growth rate of the number of different controls needed to stay close to $\varphi(t, x_*, u_*)$ is $\geq S$. This yields the upper bound (5), when the infimum runs over all regular periodic (u, x). Using the assumption on strong accessibility, a result of Coron [2] yields sufficiently many regular periodic trajectories to approximate every other trajectory that evolves in a compact subset of D.

To obtain similar lower bounds, additional dynamical structure is useful. In this case, it is convenient to assume that the system is control-affine,

$$\dot{x}(t) = f_0(x(t)) + \sum_{i=1}^{m} u_i(t) f_i(x(t)). \qquad (6)$$

If additionally U is assumed to be convex, a continuous skew-product flow $\Phi : \mathbb{R} \times (\mathcal{U} \times M) \to \mathcal{U} \times M$, called *control flow*, is induced on the extended state space by

$$\Phi_t(u, x) = (u(\cdot + t), \varphi(t, x, u)).$$

The base space $\mathcal{U}$ becomes compact metrizable with the weak*-topology of $L^{\infty}(\mathbb{R}, \mathbb{R}^m) = L^1(\mathbb{R}, \mathbb{R}^m)^*$.

We assume that Q is compact and full-time controlled invariant, i.e., for every $x \in Q$ there is $u \in \mathcal{U}$ with $\varphi(\mathbb{R}, x, u) \subset Q$. The lift of Q to $\mathcal{U} \times M$ is defined by

$$\mathcal{Q} := \{(u, x) \in \mathcal{U} \times M \ : \ \varphi(\mathbb{R}, x, u) \subset Q\}.$$

The set Q is called *uniformly hyperbolic* if for each $(u, x) \in \mathcal{Q}$ there is a direct sum decomposition

$$T_xM = E^{+}_{u,x} \oplus E^{-}_{u,x}$$

such that (i) ${\rm D}\varphi_{t,u} E^{\pm}_{u,x} = E^{\pm}_{\Phi_t(u,x)}$ for all $t \in \mathbb{R}$ and (ii) there are $c, \lambda > 0$ such that $|{\rm D}\varphi_{t,u}(x)v| \leq c^{-1}{\rm e}^{-\lambda t}|v|$ for $t \geq 0$ and $v \in E^{-}_{u,x}$, $|{\rm D}\varphi_{t,u}(x)v| \geq c{\rm e}^{\lambda t}|v|$ for $t \geq 0$ and $v \in E^{+}_{u,x}$. From the two conditions it follows that the subspaces $E^{\pm}_{u,x}$ vary continuously with (u, x). Combining [3, Thm. 5.4] and [8, Thm. 5], we obtain the following result.

II.3 Theorem: *Let Q be a compact uniformly hyperbolic chain control set of (6) with non-empty interior such that $\mathcal{Q}$ is an isolated invariant set. Assume that the Lie algebra rank condition holds on Q. Then Q is the closure of a control set D and for each compact $K \subset D$ of positive volume,*

$$h_{\rm inv}(K, Q) = \inf_{(u,x) \in \mathcal{Q}} \limsup_{t \to \infty} \frac{1}{t} \log\Big|\det {\rm D}\varphi_{t,u}|_{E^{+}_{u,x}}\Big|. \qquad (7)$$

The proof of this result also reveals a qualitative statement. To render Q invariant with a bit rate arbitrarily close to $h_{\rm inv}(K, Q)$, no control strategies are necessary that are more complicated than stabilization of a periodic orbit in ${\rm int}Q$. If a similar result holds without uniform hyperbolicity, is a completely open question. We also note that in the discrete-time case the inequality '$\geq$' in (7) can be proved in the same way as for continuous time. The upper estimate, however, is more problematic, because a discrete-time analogue of Coron's result for smooth systems does not seem to exist.

III. A Stochastic Formulation: Ergodicity, Stationarity and Asymptotic Mean Stationarity

The results for deterministic systems pose questions on set stability which are not sufficient to study stochastic setups which may inevitably require that the state process visits every open set in the unbounded state space, such as $\mathbb{R}^n$. Stochasticity also allows for studying control over general noisy channels and establishing connections with information theory. The measure-theoretic (also known as Kolmogorov-Sinai or metric) entropy is more relevant to information-theoretic as well as random noise-driven stochastic contexts, since in this case, one considers the *typical* distinguishable orbits of a dynamical system and not all of the sample paths the system can take.

Consider now an N-dimensional controlled non-linear system described by the discrete-time equations

$$x_{t+1} = f(x_t, u_t, w_t),$$

for a (measurable) function f, with $\{w_t\}$ being an independent and identically distributed (i.i.d.) system noise process.

This system is connected over a noisy channel with a finite capacity to a controller, as shown in Fig. 1. The controller has access to the information it has received through the channel. A source coder maps the source symbols, state values, to corresponding channel inputs. The channel inputs are transmitted through a channel; we assume that the channel is discrete with input alphabet $\mathcal{M}$ and output alphabet $\mathcal{M}'$.

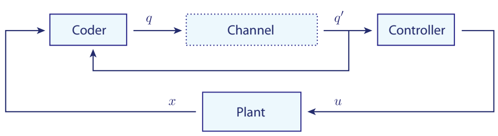

Fig. 1. Control over a noisy channel with feedback.

We refer by a *Coding Policy* Π, to a sequence of functions $\{\gamma_t^e, t \geq 0\}$ which are causal such that the channel input at time t, $q_t \in \mathcal{M}$, under Π is generated by a function of its local information, i.e.,

$$q_t = \gamma_t^e(\mathcal{I}_t^e), \quad \mathcal{I}_t^e = \{x_{[0,t]}, q'_{[0,t-1]}\},$$

where $q_t \in \mathcal{M}$, the channel input alphabet given by $\mathcal{M} = \{1, 2, \dots, M\}$. Here, we have the notation $x_{[0,t-1]} = \{x_s, 0 \leq s \leq t-1\}$ for $t \geq 1$. The channel maps q_t to q'_t in a stochastic fashion so that $P(q'_t|q_t, q_{[0,t-1]}, q'_{[0,t-1]})$ is a conditional probability measure on $\mathcal{M}'$ for all $t \in \mathbb{Z}_+$. If this expression is equal to $P(q'_t|q_t)$, the channel is said to be memoryless, i.e., the past variables do not affect the channel output q'_t given the current channel input q_t. The receiver/controller, upon receiving the information from the channel, generates its decision at time t, also causally: An admissible causal controller policy is a sequence of functions $\gamma = \{\gamma_t\}$, $\gamma_t : \mathcal{M}'^{t+1} \to \mathbb{R}^m$, so that $u_t = \gamma_t(q'_{[0,t]})$. We call such encoding and control policies *causal* or *admissible*.

This section is concerned with necessary and sufficient conditions on information channels in a networked control system for which there exist coding and control policies such that the controlled system is stochastically stable in one or more of the following senses: (i) $\{x_t\}$ is asymptotically stationary, or asymptotically mean stationary (AMS) and satisfies Birkhoff's sample path ergodic theorem, (ii) $\{x_t\}$ is ergodic, and in case a Markov chain can be constructed, (iii) $\{x_t\}$ is a stable Markov chain.

The following definition (see [17, Def. 3.1]) will be useful.

III.1 Definition: *Channels are said to be of* **Class A** *type, if*

- *they satisfy the following Markov chain condition:*

$$q'_t \leftrightarrow q_t, q_{[0,t-1]}, q'_{[0,t-1]} \leftrightarrow \{x_0, w_s, s \geq 0\}, \tag{8}$$

- *their capacity with feedback is given by*

$$C = \lim_{T\to\infty} \max_{\{P(q_t|q_{[0,t-1]}, q'_{[0,t-1]}),\ 0 \leq t \leq T-1\}} \frac{1}{T} I(q_{[0,T-1]} \to q'_{[0,T-1]}), \tag{9}$$

where the directed mutual information is defined by

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

Memoryless channels belong to this class; for such channels, feedback does not increase the capacity [12]; see [18] for examples of channels that belong to this class.

For the proofs of the results in this section see [18].

A. Supporting results and a generalization of Bode's Integral Formula for non-linear systems

Here, instead of a general $\mathbb{R}^N$-valued non-linear state model (1) we will consider non-linear systems of the form

$$x_{n+1} = f(x_n, w_n) + Bu_n, \tag{10}$$
$$x_{n+1} = f(x_n) + Bu_n + w_n, \tag{11}$$
$$x_{n+1} = f(x_n, u_n) + w_n. \tag{12}$$

In all of the models above, x_n is the $\mathbb{R}^N$-valued state, w_n is the $\mathbb{R}^N$-valued noise variable, u_n is $\mathbb{R}^s$-valued and w_n assumed to be an independent noise process with $w_n \sim \nu$. We assume throughout that f is continuously differentiable in the state variable. For a differentiable function $f : \mathbb{R}^n \to \mathbb{R}^m$, we write $Jf(x)$ for the Jacobian matrix at $x \in \mathbb{R}^n$.

1 Assumption *In the models considered above $f(\cdot, w) : \mathbb{R}^N \to \mathbb{R}^N$ is invertible for every realization of w.*

In the following $|J(f)|$ will denote the absolute value of the determinant of the Jacobian. Furthermore, with $f_w(x) = f(x, w)$, we define $J(f(x, w)) := J(f_w(x))$.

2 Assumption *There exist $M_1, L_1 \in \mathbb{R}$ so that for all x, w*

$$L_1 \leq \log_2 |J(f(x, w))| \leq M_1$$

Let $\pi_t(B) = P(x_t \in B)$ for all Borel B, i.e., π_t is the marginal occupation probability for the state process x_t.

III.2 Theorem: *Consider the networked control problem over a Class A channel. (i) Let f have the form (10), (ii) Assumptions 1 and 2 hold, and (iii) x_0 have finite differential* *entropy. a) If, under an admissible coding and control policy,* $\liminf_{t\to\infty} h(x_t)/t \leq 0$*, it must be that*

$$C \geq \liminf_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \int \pi_t(dx)\Big(\int \nu(dw)\log|Jf_w(x)|\Big). \quad (13)$$

This result also holds if the $\liminf$*'s are replaced by* $\limsup$*'s. In either case, if* $L := \inf_{x,w}\log|Jf_w(x)|$*, then* $C \geq L$*.*

The proof of Theorem III.2 reveals an interesting connection with and a generalization of Bode's Integral Formula (and what is known as the *waterbed effect*) to non-linear systems, which we state formally in the following.

III.3 Theorem: *(i) Let* f *have the form in (10) (ii) Assumptions 1 and 2 hold, and (iii)* x_0 *have finite differential entropy. If there is an admissible coding and control policy with* $\limsup_{t\to\infty} h(x_t)/t \leq 0$ *it must be that*

$$\limsup_{T\to\infty} \tfrac{1}{T} I(q_{[0,T-1]} \to q'_{[0,T-1]}) \geq \limsup_{T\to\infty} \tfrac{1}{T}\sum_{t=0}^{T-1}\int \pi_t(dx)\big(\int \nu(dw)\log|Jf_w(x)|\big).$$

B. Asymptotic mean stationarity and ergodicity

In the following, we build on the approaches in [19], [14] to account for non-linearity of the system. Consider system (11), under a given control policy, controlled over a channel.

3 Assumption *We assume*

$$M := \sup_{x\in\mathbb{R}^N} \log|Jf(x)| < \infty,$$
$$L := \inf_{x\in\mathbb{R}^N} \log|Jf(x)| > -\infty.$$

III.4 Proposition: *Consider system (11) controlled over a Class A type noisy channel with feedback. Assume that* $h(x_0) < \infty$ *under Assumption 3. If* $C < L$*,*

$$\limsup_{T\to\infty} P(|x_T| \leq b(T)) \leq 1 - \frac{(L-C)}{M}$$

for all $b(T) > 0$ *such that* $\lim_{T\to\infty}\log(b(T))/T = 0$*.*

An implication of this result follows.

III.5 Theorem: *Consider system (11) controlled over a Class A type noisy channel with feedback, and let Assumptions 1 and 3 hold. If, under some causal encoding and controller policy, the state process is AMS, then the channel capacity* C *must satisfy* $C \geq L$*.*

We note that the corresponding results for linear systems in [19] are recovered as a special case.

C. Stationarity and positive Harris recurrence under structured (stationary) policies

In many applications, one uses a state-space formulation for coding and control policies. In the following, we will consider stationary update rules which have the form

$$q_t = \gamma^e(x_t, m_t),\ u_t = \gamma^d(m_t, q'_t),\ m_t = \eta(m_{t-1}, q'_{t-1}) \quad (14)$$

for functions γ^e, γ^d, and η. In the form above, m is an $\mathbb{S}$-valued memory or *quantizer state* variable.

In this section, instead of asymptotic mean stationarity, we consider the stronger conditions of (asymptotic) stationarity of the controlled source process, the existence of an invariant probability measure, and ergodicity. For ease in presentation we assume that m_t takes values in a countable set. In the following, we assume that the channel is memoryless. Then, under (14), the process (x_t, m_t) forms a Markov chain.

III.6 Theorem: *Suppose that the encoding, control and the memory update laws are given by (14). (i) Let* f *have the form (10), (ii) Assumptions 1 and 2 hold, (iii)* $h(x_0) < \infty$*. For the positive Harris recurrence of the process* x_t, m_t *(which implies the existence of a unique invariant measure* π *(and thus ergodicity)),* $\limsup_{t\to\infty} h(x_t)/t \leq 0$ *implies*

$$C \geq \int \pi(dx)\Big(\int \nu(dw)\log(|Jf_w(x)|)\Big). \quad (15)$$

D. Discrete noiseless channels and a policy leading to stationarity and ergodicity

As noted earlier, for noise-free systems it typically suffices to consider a sufficiently small neighborhood of an equilibrium point to obtain stabilizability conditions, which is not necessarily the case for a system driven by an additive noise process. We consider such an example in the following.

III.7 Theorem: *Consider system (12), where* $\{w_t\}$ *is a sequence of zero-mean Gaussian random vectors and there exists a control function* $\kappa(z)$ *such that* $|f(x,\kappa(z))|_\infty \leq |a||x-z|_\infty$ *for all* $x, z \in \mathbb{R}^N$*, with* $\kappa(0) = 0$*. For the stationarity and ergodicity of* $\{x_t\}$ *(and thus a unique invariant probability measure), it suffices that* $C > N\log(|a|) + 1$*.*

IV. Connections between the Deterministic and Stochastic Formulations

A. Commonality of Lyapunov exponents and the Jacobian matrix properties

As we discussed in Section II, an important result in the deterministic formulation is with regard to stabilization to a point: Under some technical conditions stated earlier, [9] has reported that for convergence to the equilibrium an average rate $R > \sum_{|\lambda_i|>1}\log|\lambda_i|$ is sufficient, where λ_i are the eigenvalues of the Jacobian at the equilibrium point.

A further related result in spirit of the stochastic analysis is in a case where there exists a controlled invariant set Q with non-empy interior: For continuous-time systems of the form $\dot{x} = f(x,u)$, $u \in \mathcal{U}$, [4] establishes the lower bound

$$\max\Big(0, \min_{(x,u)\in Q\times U} \sum_i \frac{\partial f_i}{\partial x_i}(x,u)\Big) \quad (16)$$

on the invariance entropy of Q, where f_i is the i-th coordinate function of f. A more refined bound under a uniform hyperbolicity condition is given in Theorem II.3.

These results can be viewed to be related to Theorems III.6 and III.5 in that the average entropy growth as measured by the eigenvalues of the Jacobian matrix under the invariant probability measure is lower bounded by a minimum over the elements in the support set. In the stabilization to the point example of [9] reviewed in Section II-B, the invariant measure is a δ-measure on a single point. In the invariant set example leading to (16), the set Q can be viewed as the support of some invariant measure if such a measure exists. To a smaller extent, Theorem II.2 is related to Theorem III.7. Likewise, [13] has obtained conditions for noise-free systems controlled over noiseless channels. Due to the absence of noise, one could identify an invariant compact set, and consider a bound on the Lipschitz growth parameter for the system over this invariant set to obtain sufficiency conditions.

When the system is (say, Lebesgue) irreducible, however, due to the effect of noise, local properties are not descriptive and the invariant probability measure reflects the rate conditions and entropy growth in the system. In this case, the local growth integrated under an invariant measure gives a proper bound.

A related aspect of the stochastic theory is that the stochastic analysis does not assume compactness of the state space. Furthermore, it may not be natural to ask for a stochastic stability problem, if compactness is assumed a priori.

Finally, the presence of noise in the stochastic formulation leads to more desirable regularity properties, such as ergodicity and the existence of an invariant probability measure.

B. Entropy versus volume growth arguments

The analysis for the geometric view typically builds on partitioning the state space according to invariant subspaces and the volume growth in each subspace. The volume arguments naturally require a search for the best metric to work with, as studied extensively in [7].

The entropy analysis, however, avoids such a search and is almost universal. Differential entropy is a useful measure for how much a stochastic system generates uncertainty. On the other hand, the lower bounds obtained (in Theorems III.6 and III.5) through entropy are oblivious to the stable and unstable components of a non-linear dynamical system. That is, in the analysis that we obtained, the lower bounds consider both stable and unstable components. For systems where the stable and unstable components can be decoupled in the sense that with $x = (x^s, x^u)$ where x^s represents the stable and x^u the expansive mode, suppose that the dynamics is such that x^u is decoupled from x^s in the sense that

$$x^u_{n+1} = f_1(x^u_n, u_n, w_n).$$

but x^s may depend on x^u. In this case, the analysis in [18] can be generalized so as to include only the unstable components. However, this is a rare setup (though one which includes the important case of linear systems).

C. Agreement for the linear case

The stability criteria outlined earlier have been studied for linear systems of the form

$$x_{t+1} = Ax_t + Bu_t + Gw_t, \tag{17}$$

where $x_t \in \mathbb{R}^N$ is the state at time t, $u_t \in \mathbb{R}^m$ is the control input, and $\{w_t\}$ is a sequence of zero-mean i.i.d. $\mathbb{R}^d$-valued Gaussian random vectors. Here, (A, B) and (A, G) are controllable pairs. Assume that all eigenvalues λ_i of A are unstable. In this case, under the stochastic stability criteria studied in this paper, it has been observed that $C > \sum_{|\lambda_i|>1} \log|\lambda_i|$ is necessary and sufficient, see [18]. In particular, the stochastic and deterministic formulations are in agreement when the system is linear. For linear systems, (i) one can in a global fashion split the components into stable and unstable subspaces, and (ii) the Lyapunov exponents are global and not just local properties.

D. Convergence of the approaches

Current work focuses on merging the two approaches in the context of state estimation problems by viewing a stochastic system as a random dynamical system [10].

REFERENCES

[1] R. Adler, A. Konheim, and M. McAndrew. Topological entropy. Trans. Amer. Math. Soc., 114 (1965), 61–85.

[2] J.–M. Coron. Linearized control systems and applications to smooth stabilization. SIAM J. Control Optim. 32 (1994), 2, 358–386.

[3] A. Da Silva, C. Kawan. Invariance entropy of hyperbolic control sets. Discrete Contin. Dyn. Syst. 36, no. 1, 97–136 (2016).

[4] F. Colonius, C. Kawan. Invariance entropy for control systems. SIAM J. Control Optim. 48 (2009), 1701–1721.

[5] F. Colonius, C. Kawan and G. N. Nair. A note on topological feedback entropy and invariance entropy. Systems & Control Letters 62, 377–381 (2013).

[6] M. Franceschetti, P. Minero. Elements of information theory for networked control systems. In Information and Control in Networks (pp. 3–37). Springer International Publishing, 2014.

[7] C. Kawan. *Invariance Entropy for Deterministic Control Systems. An Introduction*. Lecture Notes in Mathematics, 2089. Springer, 2013.

[8] C. Kawan. On the structure of uniformly hyperbolic chain control sets. Systems & Control Letters 90, 71–75 (2016).

[9] G. N. Nair, R. J. Evans, I. M. Y. Mareels, W. Moran. Topological feedback entropy and nonlinear stabilization. IEEE Trans. Automat. Control 49 (2004), 9, 1585–1597.

[10] C. Kawan and S. Yüksel. State estimation under information constraints for stochastic non-linear systems. Preprint, 2016.

[11] F. Colonius. Metric invariance entropy and conditionally invariant measures. To appear in Ergodic Theory and Dynamical Systems.

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

[13] D. Liberzon. Stabilization by quantized state or output feedback: A hybrid control approach. In *Proc. IFAC 15th Triennial World Congress*, 2002.

[14] A. S. Matveev. State estimation via limited capacity noisy communication channels. *Mathematics of Control, Signals, and Systems*, 20:1–35, 2008.

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

[16] E. D. Sontag. *Mathematical control theory. Deterministic finite-dimensional systems*. Second edition. Springer, New York, 1998.

[17] S. Yüksel. Characterization of information channels for asymptotic mean stationarity and stochastic stability of non-stationary/unstable linear systems. *IEEE Transactions on Information Theory*, 58:6332–6354, October 2012.

[18] S. Yüksel. Stationary and ergodic properties of stochastic non-linear systems controlled over communication channels. *SIAM Journal on Control and Optimization*, 2016.

[19] S. Yüksel and T. Başar. *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Birkhäuser, New York, NY, 2013.