# Ergodicity Conditions For Controlled Stochastic Non-Linear Systems Under Information Constraints

by

Nicolas Garcia

A thesis submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for
the degree of Master of Science

Queen's University
Kingston, Ontario, Canada
December 2020

# Abstract

In this thesis we study the problem of stabilizing a controlled discrete time non-linear stochastic dynamical system, subject to information constraints. Consider such a system, controlled over a (possibly noisy) communication channel. An important problem is to determine the minimum channel capacity required to render the state process stable via a causal coding and control policy. In this thesis we consider this problem for the stability notion of asymptotic ergodicity of the state process, and prove lower bounds on the channel capacity necessary to achieve it. Under mild technical assumptions, we obtain that the necessary channel capacity is lower bounded by the log-determinant of the linearization, double-averaged over the state and noise space. We prove this bound by introducing a modified version of invariance entropy and utilizing the almost sure convergence of sample paths guaranteed by the pointwise ergodic theorem. The fundamental bounds obtained generalize well-known formulas for linear systems, and are in some cases more refined than those obtained for non-linear systems via information-theoretic methods.

# Acknowledgments

I am most grateful to my family for their support in my academic pursuits. Many thanks to my supervisors Serdar Yüksel and Christoph Kawan for introducing me to the world of mathematical research. This work would not have been possible without the many hours of guidance they provided me. Lastly, I would like to thank my friends and colleagues who helped make my time at Queen's a most enjoyable experience.

# Contents

# List of Figures

# Chapter 1

# Introduction

In the mathematical field of control theory, one is interested in the behavior of dynamical systems when acted upon by inputs. Such systems can be deterministic or stochastic, and can be modeled in either discrete or continuous time. Typically, the controller generates the input for the system based on knowledge of the state at previous times in order to steer the dynamical system in a desired way.

In practical applications, the assumption that the controller has instantaneous and arbitrarily precise knowledge of the state at any given time may not hold. For example, digital systems may impose maximum bit-rates on information transfer, resulting in the controller seeing only an estimation of the state. Furthermore, physical channels may introduce errors in the communication. Realistic mathematical models of such systems must therefore take these limitations into account. This need has motivated the subfield of information-based control, in which one seeks to achieve a control task despite the presence of information constraints.

This thesis deals with the fundamental problem of characterizing the smallest amount of information required to achieve a control task. We consider stochastic discrete time systems, and thus the control task is stochastic in nature.

## 1.1 Contribution of Thesis

We study the problem of stochastic stabilization of a non-linear stochastic dynamical system controlled over a finite capacity communication channel. We will have an occasion to consider both noisy and noiseless channels. The stability criterion considered is the (asymptotic) ergodicity of the state process. As a primary contribution, we develop a stochastic volume growth technique tailored to ergodicity properties, which is in contrast with the information-theoretic methods typically used to study such problems. We establish refined and more general results on information transmission requirements for making the controlled stochastic non-linear system ergodic. In particular, compared with [39], we allow arbitrary coding and control policies and do not impose an entropy growth condition a priori. Our results generalize the linear setups considered extensively in the literature.

## 1.2 Organization of Thesis

The thesis is organized as follows. Section 2.4 provides a brief literature review and presents the contributions of this thesis. Relevant definitions and theorems are stated in Section 2.1. The main results are presented and discussed in Section 3.1, while their proofs are given in Section 3.2. Some definitions and auxiliary results are outlined in the appendix.

# Chapter 2

# Background

In this chapter we state notational conventions and provide the mathematical background necessary to state and prove the main theorems in this thesis. We give a precise formulation of the stabilization problem, and conclude with a literature review.

## 2.1 Notation and Preliminaries

Throughout this thesis, $\mathbb{R}$ denotes the real numbers, $\mathbb{Z}$ the integers and $\mathbb{N}$ the strictly positive integers. $\mathbb{Z}_+$ denotes $\mathbb{N} \cup \{0\}$, $\mathbb{R}_{>0}$ the strictly positive real numbers and $\mathbb{R}_{\geq 0}$ the non-negative real numbers. We write $[a;b]$ for a discrete interval, i.e., $[a;b] = \{a, a+1, \dots, b-1, b\}$ for any $a \leq b$ in $\mathbb{Z}$. Given a topological space $\mathcal{X}$, $\mathcal{B}(\mathcal{X})$ denotes the Borel $\sigma$-algebra of $\mathcal{X}$ and $\Sigma$ denotes the space of sequences from $\mathcal{X}$, i.e., $\Sigma = \mathcal{X}^{\mathbb{Z}_+}$. If $x \in \Sigma$, we write $x_{[0,t]} \coloneqq (x_0, x_1, \dots, x_{t-1}, x_t)$ for any $t \in \mathbb{Z}_+$. By $m$ we denote the Lebesgue measure on $\mathbb{R}^N$ where $N \in \mathbb{N}$ will be clear from context. All logarithms are taken to the base 2. Given a map $f : \mathbb{R}^n \to \mathbb{R}^m$, we denote the Jacobian by $Df$ which is just the matrix of partial derivatives. We will use $\sqcup$ to emphasize that the union in question is disjoint. We let $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ represent the ceiling and floor functions respectively. Given some sequence space $\Sigma$ we denote by $\theta$ the left shift map, so that if $x := (x_n)_{n=0}^{\infty}$ we have that $(\theta x)_n = x_{n+1}$ for every $n \in \mathbb{Z}_+$. Given a map $f$, we denote by $f^n$ the $n-$fold composition of $f$ with itself whenever the expression makes sense.

## 2.2 Stochastic Stability and Ergodicity

In control theory, it is natural to consider the problem of rendering the state process of a dynamical system 'stable'. Stability notions vary greatly depending on the context and type of system. For example, the notion of set invariance is commonly used in deterministic control, where the goal is to ensure that the state process never leaves a given (usually compact) set. In the context of stochastic control, the presence of noise makes the notions of set invariance too restrictive; for example, unbounded additive noise could result in the state process leaving any given compact set regardless of the control decision. For this reason, notions of stochastic stability are considered. In these cases, one wishes to establish desirable probabilistic properties of the state process. For example, one may consider notions of recurrence. A full treatment of stochastic stability is available in [28]. In this section, we provide only the notions necessary to define and motivate the stability notion considered in this thesis; asymptotic ergodicity of the state process. We begin with some formal definitions.

**Definition 2.2.1.** *(Discrete Time Stochastic Process) Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and* $(S, d)$ *a metric space. A discrete time stochastic process defined on* $(\Omega, \mathcal{F}, P)$ *taking values in* $(S, d)$ *is a family of random variables* $x_n : \Omega \to S$ *where* $n \in \mathbb{Z}_+$.

**Definition 2.2.2.** *A topological space* $(\mathcal{X}, T)$ *is called a Polish space if and only if* $(\mathcal{X}, T)$ *is separable, metrizable, and complete with respect to at least one metric which* *induces the topology.*

In the remainder of this section, whenever we introduce a stochastic process, we will assume that a probability space $(\Omega, \mathcal{F}, P)$ on which the process is defined is lurking. Unless otherwise stated, we will also assume that the process takes values in a Polish space $(\mathcal{X}, T)$.

Consider such a process $x := (x_n)_{n=0}^{\infty}$. Writing $\Sigma := \mathcal{X}^{\mathbb{Z}_+}$ to denote the set of sequences with entries in $\mathcal{X}$, we consider the product $\sigma$-algebra defined as the smallest $\sigma$- algebra containing cylinder sets (sets of the form $\{(x_n)_{n=0}^{\infty} \in \Sigma : x_i \in B_i$ for all $i$ satisfying $n \leq i \leq m\}$ where $m, n \in \mathbb{Z}_+$, $n \leq m$, and $B_i \in \mathcal{B}(\mathcal{X})$ for every $i \in \{n, n+1, .., m-1, m\}$). Denote this $\sigma$- algebra by $\mathcal{B}(\Sigma)$. The stochastic process $x$ induces a measure $\mu$ on the measurable space $(\Sigma, \mathcal{B}(\Sigma))$ called the process measure. To define this measure, it suffices to define it on the cylinder sets. For such a set $B := (\underbrace{\mathcal{X}, \ldots, \mathcal{X}}_{\text{n - times}}, B_n, \ldots, B_m, \mathcal{X}, \mathcal{X}, \ldots)$ we define $\mu(B) := \prod_{k=n}^{m} P(\{\omega \in \Omega : X_k(\omega) \in B_k\})$. Observe that a stochastic process can be thought of as an infinite random vector $x : (\Omega, \mathcal{F}, P) \to (\Sigma, \mathcal{B}(\Sigma))$ where the process measure $\mu$ is nothing but the push-forward of $P$ through $x$. We define the shift map $\theta : \Sigma \to \Sigma$ by $(\theta x)_t := x_{t+1}, \quad \forall x = (x_t)_{t \in \mathbb{Z}_+} \in \Sigma$.

**Definition 2.2.3.** *A measure $\nu$ on $(\Sigma, \mathcal{B}(\Sigma))$ is called* stationary *if and only if $\nu(\theta^{-1}(B)) = \nu(B)$ for all $B \in \mathcal{B}(\Sigma)$, i.e., if $(\Sigma, \mathcal{B}(\Sigma), \nu, \theta)$ is a measure-preserving system.*

*A stochastic process with process measure $\mu$ is called*

- stationary *if and only if its $\mu$ is stationary, i.e., $P(\{\omega \in \Omega : x(\omega) \in B\}) = P(\{\omega \in \Omega : (\theta x)(\omega) \in B\})$ for all $B \in \mathcal{B}(\Sigma)$.*

- asymptotically mean stationary (AMS) *if and only if there exists a measure* $\tilde{Q}$ *on* $(\Sigma, \mathcal{B}(\Sigma))$ *such that*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} P(\theta^{-k}(B)) = \tilde{Q}(B) \quad \text{for all } B \in \mathcal{B}(\Sigma).$$

It can easily be shown that the measure $\tilde{Q}$ is stationary. As such, we can obtain a measure $Q$ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ by projecting $\tilde{Q}$ down to any of its coordinates which is well defined due to stationarity. It follows that for any $B \in \mathcal{B}(\mathcal{X})$

$$\lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T-1} P(x_k \in B) = Q(B).$$

We proceed by stating some key notions from ergodic theory.

**Definition 2.2.4.** *Let* $T : \Omega \to \Omega$ *be a measurable map on a probability space* $(\Omega, \mathcal{F}, P)$*.* $T$ *is called* measure-preserving *if and only if* $P(T^{-1}(A)) = P(A)$ *for all* $A \in \mathcal{F}$*. An event* $A \in \mathcal{F}$ *is* $T$-invariant *if and only if* $A = T^{-1}(A)$ *(up to a set of measure zero). We denote by* $\mathcal{F}_{\mathrm{inv}(T)}$ *the set of all* $T$*-invariant measurable sets. It is not hard to show that* $\mathcal{F}_{\mathrm{inv}(T)}$ *is a* $\sigma$*-algebra.*

**Definition 2.2.5.** *A measure-preserving map* $T : \Omega \to \Omega$ *on a probability space* $(\Omega, \mathcal{F}, P)$ *is called* ergodic *if and only if* $P(A) \in \{0, 1\}$ *for all* $A \in \mathcal{F}_{\mathrm{inv}(T)}$*. Note that ergodicity is a property of a system* $(\Omega, \mathcal{F}, P, T)$*, but often we say "*$T$ *is ergodic", or "*$P$ *is ergodic" when the other components of the system are clear from the context.*

A fundamental result in ergodic theory is the following pointwise ergodic theorem.

**Theorem 2.2.6.** *(Pointwise Ergodic Theorem) Let* $(\Omega, \mathcal{F}, P)$ *be a probability space* *and $T : \Omega \to \Omega$ a measure-preserving map. Then for any $f \in L^1(\Omega, \mathcal{F}, P)$ we have*

$$\frac{1}{N}\sum_{k=0}^{N-1} f \circ T^k \xrightarrow[N\to\infty]{a.s} \varphi$$

*for some $\varphi \in L^1(\Omega, \mathcal{F}_{\mathrm{inv}(T)}, P|_{\mathcal{F}_{\mathrm{inv}(T)}})$ satisfying $\int \varphi \, \mathrm{d}P = \int f \, \mathrm{d}P$. If, in addition, $T$ is ergodic, then $\varphi$ is almost everywhere constant and thus*

$$\frac{1}{N}\sum_{k=0}^{N-1} f \circ T^k \xrightarrow[N\to\infty]{a.s} \int f \, \mathrm{d}P.$$

Let now $x := (x_n)_{n=0}^{\infty}$ be a stochastic process with process measure $\mu$, and suppose that the system $(\Sigma, \mathcal{B}(\Sigma), \mu, \theta)$ is ergodic. Observe that if $B \in \mathcal{B}(\mathcal{X})$, we can apply the pointwise ergodic theorem to the indicator function $\mathbb{1}_{x_0 \in B} : \Sigma \mapsto \{0, 1\}$ to obtain

$$\mu\Big(\Big\{x \in \Sigma : \lim_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} \mathbb{1}_{x_0\in B}(\theta^k(x)) = \int \mathbb{1}_{x_0 \in B}(x) \, \mathrm{d}\mu\Big\}\Big) = 1.$$

Abusing notation and letting $\mu$ also denote the projection of the stationary process measure on the space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, we can rewrite the above expression as

$$\mu\Big(\Big\{x \in \Sigma : \lim_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} \mathbb{1}_{x_k\in B}(x) = \mu(B)\Big\}\Big) = 1. \tag{2.1}$$

Thus, if the stochastic process is ergodic, then the set of sample paths which visit a Borel set $B$ with frequency $\mu(B)$ is of full measure. This is a strong notion of stability, and is a key ingredient in the proofs of the theorems in this paper.

It turns out however that (2.1) holds for a larger class of processes which we characterize in the next definition.

**Definition 2.2.7.** *(AMS Ergodic) Consider a stochastic process which is AMS with asymptotic mean $Q$. If $Q$ is ergodic, we call the process AMS ergodic.*

The above notion, which we informally refer to as asymptotic ergodicity is the notion of stability that we will consider in this thesis. Let us now show that it satisfies (2.1).

**Proposition 2.2.8.** *An AMS ergodic process satisfies an equation similar to* (2.1)*. Namely, for any $B \in \mathcal{B}(\mathbb{X})$ we have that*

$$\mu\left(\left\{x \in \Sigma : \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{1}_{x_t \in B}(x) = Q(B)\right\}\right) = 1. \tag{2.2}$$

*Proof.* Let us fix a $B \in \mathcal{B}(\mathcal{X})$. By stationarity, we can project $Q$ to the space $\mathcal{X}$. By a slight abuse of notation, we also denote the projected measure by $Q$. We define

$$F := \left\{x \in \Sigma : \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{1}_{x_0 \in B}(\theta^t(x)) = Q(B)\right\}.$$

From the ergodicity assumption on $Q$, it follows that $Q(F) = 1$. Also, $F$ is invariant under $\theta$ from which we obtain that $\mu(F) = 1$ (see [17, Lem. 6.3.1 and Eq. (6.22)]). □

## 2.3 Problem Statement

We are now ready to give a precise formulation of the problem that we consider in this thesis. Consider the nonlinear stochastic dynamical system given by the equation

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

where $x_t$ and $u_t$ take values in $\mathbb{R}^N$ and $w_t$ takes values in a standard probability space. The variables $x_t, w_t$ and $u_t$ represent the state, noise, and control action at time $t$, respectively. Suppose first that the system is controlled over a noiseless channel of capacity $C$. At each time step, the encoder can transmit $M := \lfloor 2^C \rfloor$ symbols without error (the floor function is required in case that $C$ is not an integer) to the decoder. For simplicity suppose $\mathcal{M} := \{1, \ldots, M\}$ is the coding alphabet.

We will consider the case where the system is controlled over a possibly noisy communication channel. We consider the problem of determining necessary conditions on channel capacity required for the existence of coding and control policies which make the closed-loop system stochastically stable. The stability criterion considered is asymptotic ergodicity, by which we mean the existence of an asymptotically mean stationary measure which is also ergodic.

## 2.4 Literature review

The study of control under communication constraints dates back at least 50 years, as can be seen in [8] which is one of the first texts to consider the problem of control using quantized state information. The vast majority of the literature has considered linear systems, both stochastic and deterministic, and in both continuous and discrete time. In the linear case, the problem of stabilization has been worked out under a variety of setups and stability notions. Despite distinct formulations, assumptions and control objectives, a recurring theme is the characterization of the minimum data rate required for closed-loop stability as the log-sum of the unstable open-loop eigenvalues of the system dynamics matrix. Two of the earliest significant contributions are due to Wong and Brockett [36], and Baillieul [3], who showed that the state process of a scalar linear system with parameter $|a| > 1$ can be kept bounded with quantized control if and only if the data rate in bits per sample is no less than $\log_2(|a|)$. These results were some of the first instances of *data-rate theorems*, which characterize the rate of information between an encoder and controller required to accomplish a given control task. Other notable papers containing more general *data-rate theorems* are due to Delchamps [11], Tatikonda and Mitter [34], Fagnani and Zampieri [15], Delvenne [12], Matveev and Savkin [27], Nair [29], De Persis [9],Liberzon and Hespanha [26], Hespanha [19], and Savkin [33].

In the linear systems case, it is important to note that there is no distinction between global and local dynamical and control-theoretic properties. As such, the local problem of stabilization to a point, the semi-global problem of set invariance, and the global problem of stochastic stabilization can all be handled with the same tools. For nonlinear systems however, this is not so; the three aforementioned stability problems are structurally different and require distinct mathematical machinery to handle. An example of this is linearization techniques, which under regularity assumptions work well for local problems, but fail to work in general for global ones [23]. As such, the techniques for studying minimum information rates for non-linear control are fundamentally different from those used in the linear systems theory.

Evidently, the field of information-based control has achieved a certain level of maturity, and is too vast to summarize in this literature review. Many interesting and important results have been obtained in a wealth of settings. For additional data-rate theorems and control-theoretic results over finite capacity channels, the reader may wish to consult [27, 2, 16, 31]. [21] and [40] also contain relevant discussions on the subject, with the former focusing on methods from dynamical systems and the latter on systems with stochastic components.

To date, the study of stabilization of non-linear systems under communication constraints has focused primarily on deterministic systems controlled over noiseless channels. Furthermore, constructive schemes have generally been the main focus (as opposed to converse theorems). Some noteworthy results include [26] and [9]. In the former, the problem of stabilizing a continuous-time system with Lipschitz assumptions using sampled and quantized state information was considered. So long as the data rate exceeds the product of the state space dimension and the Lipschitz constant of the system dynamics function, it was shown than it is possible to achieve global asymptotic stabilization to an equilibrium point. The latter author considered feed-forward systems, and constructed an encoder and controller for a non-linear system (satisfying local Lipschitz assumptions) which achieves stabilization despite both arbitrarily large communication delays and arbitrarily small channel capacity between encoder and controller. Other constructive schemes for non-linear systems include [10] and [25].

For the task of proving more general data rate theorems for set invariance of non-linear (and linear) deterministic systems, two a priori distinct notions have proven particularly fruitful. The first is that of *topological feedback entropy*, a concept modeled after the open cover definition of topological entropy for classical dynamical systems prescribed in [1]. Presented in [30] by Nair, Evans, Mareels, and Moran, the *topological feedback entropy* measures the lowest data-rate required between a coder and controller to ensure the existence of a coding and control policy which guarantees that the system's state process remain inside a compact controlled invariant set.

The second notion is *invariance entropy*, which was introduced in [5]. The definition resembles the Bowen–Dinaburg definition of topological entropy via spanning sets (see [13]) and has analogous properties. One can even use Bowen's formula for the topological entropy to obtain the invariance entropy in the case of a linear control system [21]. To define *invariance entropy*, let $r(\tau, Q)$ denote the minimum number of control functions required to establish invariance of $Q$ on the time interval $[0, \tau]$ given an arbitrary initial state in the set. The invariance entropy of a compact subset $Q$ of the state space is then defined to be

$$h_{\text{inv}}(Q) := \lim_{\tau\to\infty} \frac{1}{\tau} \log r(\tau, Q),$$

The quantity measures how fast in time the number of open-loop control functions required to render $Q$ invariant grows. It is a measurement of the smallest average rate of information that must be transmitted to a controller to render $Q$ invariant and thus intimately linked to minimum data rates. The motivation for the above definition arises by observing that with $n$ bits of information available at the controller side, at most $2^n$ different states can be distinguished, and therefore at most $2^n$ different control inputs can be generated.

It turns out that under a strong invariance condition, the notions of *topological feedback entropy* and *invariance entropy* are equivalent, as explained in [6]. We conclude our review of dynamical notions by noting the more recent introduction of *metric invariance entropy* in [4]. This quantity provides a measure-theoretic analogue to *invariance entropy* for discrete time systems, and is based on conditionally invariant measures.

In the case of stochastic systems, the above notions can not be used directly for stabilization problems. First, asking for a compact subset of the state space to be invariant is too restrictive to be a useful notion of stability. For example, if the system is subject to unbounded noise, the state process may leave a given compact set regardless of the control policy. Secondly, If the channel is noisy, the informational content of received codewords cannot be measured by the number of distinct possible receiver outputs. As an extreme case, consider a channel where the channel inputs and outputs are independent, and hence the (information theoretic) channel capacity is zero. In this case, no reliable information can be transmitted across the channel. Due to these two observations, it is clear that distinct stability notions and methods are required for systems with stochastic components.

In stochastic control systems, many different notions of stability have been considered, such as stationarity, ergodicity, AMS, and various types of recurrence. In the context of linear systems (without information constraints), conditions for the achievement of the aforementioned stability criterion have been determined using Lyapunov methods, and we refer the reader to [40] for a full treatment of this, and other techniques. In general however, full state feedback is required to establish stabilizing policies, and thus many of the techniques in the general theory do not apply to problems with information constraints. One technique used for stochastic stabilization under communication constraints has been the use of information theoretic methods as done in [39]. These methods however do not allow one to take advantage of geometric properties of non-linear systems.

Building on work done in the deterministic non-linear system case, the authors in [23] were able to generalize the notion of invariance entropy for use in the stability analysis of discrete time stochastic systems controlled over finite-capacity channels.

The introduced quantity, called *stabilization entropy*, "is inspired by both invariance entropy and measure-theoretic entropy of dynamical systems. It is based on a characterization of measure-theoretic entropy due to Katok [20], and a generalization of said notion found in Ren et al. [32]" [23]. The authors were able to establish necessary conditions in the form of lower bounds on the channel capacity for a certain class of stochastic nonlinear systems over both noiseless and noisy channels. Such results are similar to those obtained in [39], where the stability notion considered in the first paper is AMS, and the notions considered in the second paper are AMS, ergodicity, and positive Harris recurrence. In contrast to the dynamical systems-motivated techniques used in [23], [39] relied on information-theoretic techniques and directed information (a generalization of mutual information). The distinct methods arrived at complementary results.

In this thesis, we consider the stochastic stability notion of asymptotic ergodicity for non-linear discrete time stochastic systems. Although the stability notion of ergodicity has been considered in several papers (which we summarize below), the notion of asymptotic ergodicity, to the best of our knowledge, has not been studied before. For a class of nonlinear systems controlled over noiseless channels [39], and for linear systems over Gaussian, discrete noiseless, erasure, and discrete noisy channels [40, 37, 38, 41] establish ergodicity under channel capacity constraints using information theoretic techniques.

In the thesis at hand, we provide an operationally and mathematically significant refinement, where our stability criterion is stochastic in nature, but deterministic in its sample path limits as seen in the Definition 2.2.7. Our stronger notion of stability guarantees the almost sure convergence of sample paths which asymptotically visit each subset of the state space at a frequency given by the AMS measure on the subsets. We further generalize the notion of stabilization entropy by considering a finite collection of subsets rather than one single subset of the state space, and prove stronger results using the pointwise ergodic theorem. One of our main contributions in this thesis is the development of a geometric analytical method, afforded by a stochastic volume growth approach, to study stochastic stabilization of non-linear systems under information constraints. This method is distinct from the methods relying on directed mutual information such as the ones used in [39] to study such problems. Our focus is on proving converse theorems, which aim to determine whether an unstable open loop stochastic system can be rendered stable through a coding and control policy over a (possibly noisy) finite capacity channel. To our knowledge, such theorems have only been studied in [39] and [23] for non-linear systems.

# Chapter 3

# Results

## 3.1 Bounds on channel capacity

We now state the main contributions of this thesis. Proofs can be found in the next section. Consider the system

$$x_{t+1} = f(x_t, w_t) + u_t \tag{3.1}$$

where $x_t$ and $u_t$ are $\mathbb{R}^N$-valued for some $N \in \mathbb{N}$ and $w_t$ takes values in a standard probability space $\mathbb{W}$. For a fixed $w \in \mathbb{W}$, let us denote the map $x \mapsto f(x, w)$ by $f_w$. Suppose also that the following holds:

(A1) The map $f : \mathbb{R}^N \times \mathbb{W} \to \mathbb{R}^N$ is Borel measurable.

(A2) The noise process $(w_t)_{t\in\mathbb{Z}_+}$ is i.i.d. By abuse of notation, $\nu$ denotes both the law of any individual $w_t$, as well as the process measure.

(A3) The map $f_w : \mathbb{R}^N \to \mathbb{R}^N$ is $C^1$ and injective for any $w \in \mathbb{W}$.

(A4) The initial state $x_0$ is random and independent of the noise process. We write

$\pi_0$ for the associated probability measure.

(A5) The measure $\pi_0$ is absolutely continuous with respect to the $N$-dimensional Lebesgue measure $m$, and its density (which exists by the Radon-Nikodym theorem) is bounded.

(A6) There is a constant $c > 0$ with $|\det Df_w(x)| > c$ for all $x \in \mathbb{R}^N$ and $w \in \mathbb{W}$.

We write $(\Omega, \mathcal{F}, P)$ for the probability space on which both $x_0$ and $w_t$ are modeled.

We assume that the system is controlled over a possibly noisy communication channel as depicted in Fig. 3.1. The channel has a finite input alphabet $\mathcal{M}$ and a finite output alphabet $\mathcal{M}'$. For a channel with feedback, the input $q_t$ at time $t$ is generated by a function $\gamma_t^e$ so that $q_t = \gamma_t^e(x_{[0,t]}, q'_{[0,t-1]})$ where we note that the definition is identical for a noiseless channel without feedback, since for noiseless channels feedback provides no additional information. The channel maps $q_t$ to $q'_t$ in a stochastic fashion so that $P(q'_t \in \cdot | q_t, q_{[0,t-1]}, q'_{[0,t-1]}) = P(q'_t \in \cdot | q_t)$ is a conditional probability measure on $\mathcal{M}'$ for all $t \in \mathbb{Z}_+$, for every realization $q_t, q_{[0,t-1]}, q'_{[0,t-1]}$. The controller, upon receiving the information from the channel, generates its decision at time $t$, also causally: $u_t = \gamma_t^c(q'_{[0,t]})$. Any coding and control policy of this kind is called *causal*. If the channel is noiseless, we have $\mathcal{M} = \mathcal{M}'$ and the channel capacity reduces to $C = \log |\mathcal{M}|$. If the channel is noisy and memoryless, feedback does not increase its capacity, see Section A.2.

**Theorem 3.1.1.** *Consider system* (3.1) *satisfying assumptions (A1)–(A6). Suppose the system is controlled over a discrete noiseless channel of capacity $C$ and a coding and control policy achieves that the state process is AMS ergodic with asymptotic mean*

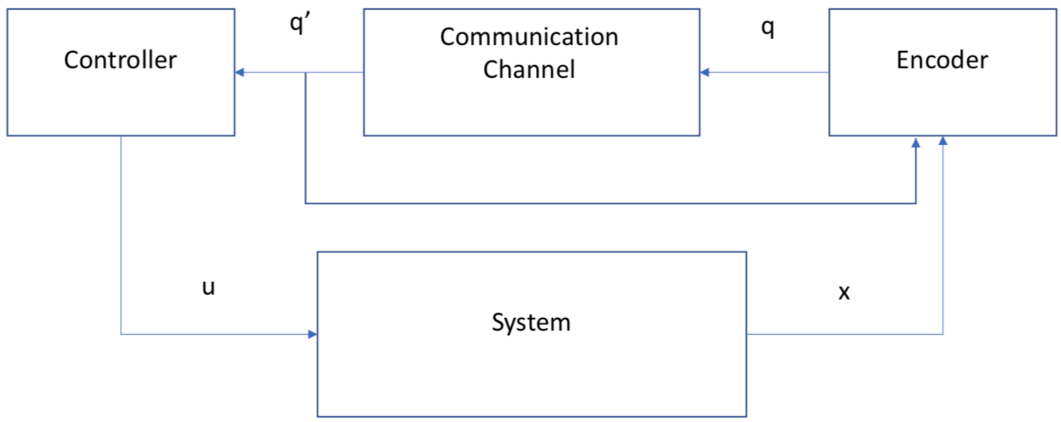


Figure 3.1: Block Diagram of Control System in Question

*$Q$. Then the capacity must satisfy*

$$\int\int \log|\det Df_w(x)|\, \mathrm{d}Q(x)\,\mathrm{d}\nu(w) \leq C.$$

Our second main theorem relaxes the condition of the channel being noiseless. On the other hand, the class of nonlinear systems considered is more restrictive.

**Theorem 3.1.2.** *Consider the scalar system*

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

*satisfying assumptions (A1)–(A5) as well as the following three assumptions:*

*(i) $|f'_w(x)| \geq 1$ for every $x \in \mathbb{R}$.*

*(ii) The support of $\pi_0$ is a compact interval $K \subseteq \mathbb{R}$.*

*(iii) The lower and upper bounds of the density of $\pi_0$ within $K$, denoted by $\rho_{\min}$ and*

*$\rho_{\max}$ respectively, satisfy $0 < \rho_{\min} \leq \rho_{\max} < \infty$.*

*Suppose that the system is controlled over a discrete memoryless channel with feedback of capacity $C$ (see Definition A.2.4) and that there exists a causal coding and control policy which results in the state process being AMS ergodic with asymptotic mean $Q$. Then the channel capacity must satisfy*

$$\int \int \log |f'_w(x)| \,\mathrm{d}Q(x)\,\mathrm{d}\nu(w) \leq C. \tag{3.2}$$

The first theorem above is a counterpart to [23, Thm. 5.1], where it was shown for systems of the form $x_{t+1} = f(x_t) + w_t + u_t$, without the ergodicity assumption on the AMS measure, that for any Borel set $B$ of finite Lebesgue measure

$$Q(B) \inf_{x \in B} \log |\det Df(x)| \leq C$$

must be satisfied. The second theorem above is a counterpart to [23, Thm. 7.1] without the ergodicity assumption on the AMS measure.

To prove Theorem 3.1.1 and Theorem 3.1.2, the stabilization entropy introduced in [23] must be generalized and a technical lemma proven. This is carried out in the next section. Before doing this, we provide a discussion of the theorems.

Observe that our lower bound on channel capacity is $\leq 0$ (and thus vacuous) if $|\det Df_w(x)| \leq 1$ for all $(x, w)$. Recall that the determinant of a square matrix represents the volume of the unit cube after it is acted on by the matrix. As such, Theorem 3.1.1 is only interesting if the system is volume-expansive on some regions of the state space. This is intuitive, since if $f$ is nowhere volume-expansive, it may be possible for the uncontrolled system to have desirable stability properties.

The results obtained here are consistent with those obtained using information-theoretic techniques in [39], but are in fact a strict refinement. A similar converse result on channel capacity was obtained in [39] under the stronger stability criterion of positive Harris recurrence of the closed-loop stochastic process. It reads as follows:

**Theorem 3.1.3.** *([39, Thm. 4.2]) Consider the system*

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

*and suppose that the following assumptions hold:*

*(i) For any fixed $w$, the function $f_w : \mathbb{R}^N \to \mathbb{R}^N$ is a $C^1$-diffeomorphism.*

*(ii) There exist $L, M \in \mathbb{R}$ such that $L \leq \log|\det Df_w(x)| \leq M$ for all $x, w \in \mathbb{R}^N$.*

*Suppose that a stationary coding and control policy (see [39] for a precise definition) is adopted so that under this policy*

*(i) the Markovian system state and encoder state is positive Harris recurrent (which implies the existence of a unique invariant measure).*

*(ii)* $\limsup_{t\to\infty} h(x_t)/t \leq 0$ *(see Definition A.2.2 for the definition of the differential entropy $h(\cdot)$).*

*Then the channel capacity must satisfy*

$$\int\int \log|\det Df_w(x)| \,\mathrm{d}Q(x)\,\mathrm{d}\nu(w) \leq C.$$

Let us now compare Theorem 3.1.1 and Theorem 3.1.2 with Theorem 3.1.3. Theorem 3.1.1 is more general in the sense that it applies to arbitrary causal coding and control policies, not just Markov ones. Moreover, it does not require the assumption of sublinear growth of the differential entropy of the state process (and the implicit assumption that the law of each state random variable $x_t$ be absolutely continuous with respect to the Lebesgue measure so that the differential entropy is well defined). Theorem 3.1.3 assumes that the state process is positive Harris recurrent which implies unique ergodicity, while Theorem 3.1.1 only assumes ergodicity of the AMS measure. On the other hand, compared with Theorem 3.1.2, Theorem 3.1.3 considers a more general class of channels (involving memory) as well as systems taking values in higher dimensions.

We conclude the discussion of the theorems with a note about achievability results. For the linear system case, it was shown in [38] that for a linear system of the form $x_{t+1} = Ax_t + w_t + u_t$ with a diagonalizable matrix $A$ (and additional technical assumptions), controlled over a DMC with channel capacity no smaller than the log-sum of the unstable eigenvalues of $A$, the AMS stochastic stability property ([23]) can be achieved. Given that AMS ergodicity implies the AMS property, we see that the lower bounds established in this thesis are tight for a class of linear systems. Achievability results have also been obtained for non-linear systems. For the system $x_{t+1} = f(x_t, u_t) + w_t$ with Gaussian i.i.d noise and technical assumptions, it was shown in [39] that ergodic stabilization is possible. We thus see that the bounds in this thesis are meaningful for non-linear systems, as the problem of ergodic stabilization is indeed feasible.

## 3.2 Proofs

In this section, we prove our two main theorems. We begin by generalizing the notion of stabilization entropy and proving a technical lemma.

### 3.2.1 Generalizing stabilization entropy

Consider system (3.1) with a fixed (open-loop) control sequence $u := (u_t)_{t\in\mathbb{Z}_+}$, a noise realization $w := (w_t)_{t\in\mathbb{Z}_+}$ and an initial state $x_0 \in \mathbb{R}^N$. For such a setup, the trajectory $x := (x_t)_{t\in\mathbb{Z}_+} \in (\mathbb{R}^N)^{\mathbb{Z}_+}$ of the state is uniquely determined. Let us denote this trajectory by $\varphi(\cdot, x_0, u, w)$ so that for any $t \in \mathbb{Z}_+$, $x_t = \varphi(t, x_0, u, w)$.

We want to find a subset of control sequences that allow to render certain subsets of the state space invariant in a probabilistic sense. This leads to the next definitions of spanning sets and stabilization entropy for finite collections of subsets of $\mathbb{R}^N$ and $\mathbb{W}$, respectively, which generalize similar notions in [23], where a single set was considered.

**Definition 3.2.1.** *Let $B \in \mathcal{B}(\mathbb{R}^N)$ and $D \in \mathcal{B}(\mathbb{W})$ be finite disjoint unions of Borel sets $B_1, \dots, B_n$ and $D_1, \dots, D_m$, respectively. Let also $R$ denote a collection of numbers $r_{k,l} \in [0,1]$ for $k \in \{1, \dots, n\}$ and $l \in \{1, \dots, m\}$ satisfying*

$$1 - r := \sum_{k=1}^{n} \sum_{l=1}^{m} (1 - r_{k,l}) \in [0,1].$$

*Fix $T \in \mathbb{N}$ and $\rho \in (0,1)$. A set of control sequences $S \subseteq (\mathbb{R}^N)^T$ is called $(T, B, D, \rho, R)$-spanning if there exists $\tilde{\Omega} \in \mathcal{F}$ such that the following conditions hold:*

- $P(\tilde{\Omega}) \geq 1 - \rho$.

- *For each $\omega \in \tilde{\Omega}$, there exists a control sequence $u \in S$ such that*

$$\frac{1}{T}|\{t \in [0; T-1] : (\varphi(t, x_0(\omega), u, w(\omega)), w_t(\omega)) \in B_k \times D_l\}| \geq 1 - r_{k,l}$$

*holds for all $k$ and $l$.*

Note that we abuse notation by calling a set $(T, B, D, \rho, R)$-spanning instead of $(T, (B_k)_{k=1}^n, (D_l)_{l=1}^m, \rho, R)$-spanning. When doing so, there is the underlying assumption that the partitions of $B$ and $D$ are fixed. No confusion should arise, since we explicitly define the partitions whenever we use the definition.

In the above definition, the fact that all random variables are modeled on a common probability space ensures that given $\omega$, the initial state and the noise sequence of length $T$ are deterministic. Intuitively speaking, a subset of control sequences of length $T$ is $(T, B, D, \rho, R)$-spanning if the probability that, for all $k, l$, we can maintain the state variable in $B_k$ and the noise variable in $D_l$ for at least $1 - r_{k,l}$ percent of the time, is at least $1 - \rho$. We want to use the size of spanning sets to quantify the difficulty of a control task, which leads to the next definition.

**Definition 3.2.2.** *For the system (3.1), we define the $(B, D, \rho, R)$-stabilization entropy by*

$$h(B, D, \rho, R) := \limsup_{T \to \infty} \frac{1}{T} \log s(T, B, D, \rho, R),$$

*where $s(T, B, D, \rho, R)$ denotes the smallest cardinality of a $(T, B, D, \rho, R)$-spanning set. We define this quantity to be $\infty$ if no or no finite spanning set exists.*

It is obvious that finite $(T, B, D, \rho, R)$-spanning sets need not exist. As we will see however, they do exist in desired scenarios.

The following lemma is instrumental to prove Theorem 3.1.1.

**Lemma 3.2.3.** *Consider system (3.1) with the assumptions of Theorem 3.1.1 (i.e., a coding and control policy exists over a noiseless channel of capacity* $C = \log|\mathcal{M}|$ *which makes the state process AMS ergodic). Let now*

- $B := \bigsqcup_{k=1}^{n} B_k \in \mathcal{B}(\mathbb{R}^N)$ *and* $D := \bigsqcup_{l=1}^{m} D_l \in \mathcal{B}(\mathbb{W})$ *be finite disjoint unions of Borel sets,*
- $\rho \in (0,1)$ *be arbitrary.*

*Next, define the sequence of numbers* $R_\epsilon := (r_{k,l})_{1 \le k \le n, 1 \le l \le m}$*, where*

$$r_{k,l} := \begin{cases} (1+\epsilon)(1 - Q(B_k)\nu(D_l)) & \text{if } Q(B_k)\nu(D_l) \in (0,1) \\ 1 & \text{if } Q(B_k)\nu(D_l) = 0 \\ \epsilon & \text{if } Q(B_k)\nu(D_l) = 1 \end{cases}$$

*and observe that for* $\epsilon > 0$ *small enough, the following conditions are satisfied:*

*(i)* $1 - r := \sum_{k=1}^{n} \sum_{l=1}^{m} (1 - r_{k,l}) \in [0,1]$.

*(ii)* $1 - (1+\epsilon)(1 - Q(B_k)\nu(D_l)) \in (0,1)$ *for all* $k, l$ *with* $Q(B_k)\nu(D_l) \in (0,1)$.

*Thus, for such a small* $\epsilon$*, the generalized stabilization entropy* $h(B, D, \rho, R_\epsilon)$ *is well-defined. (Of course,* $r$ *and the* $r_{k,l}$*'s are* $\epsilon$*-dependent, but we drop this from the notation.) Then for all* $\epsilon > 0$ *further small enough the capacity must satisfy*

$$h(B, D, \rho, R_\epsilon) \le C. \tag{3.3}$$

*Proof.* We distinguish two cases.

**Case 1:** We can remove the trivial sets with zero measure from the collections $\{B_k\}$ and $\{D_l\}$ and thus assume that $Q(B_k)\nu(D_l) > 0$ for all $(k,l)$. Indeed, if a spanning set can be found for the new collections, it is still spanning for the original ones. If $Q(B_k)\nu(D_l) = 1$ for some $(k,l)$, all the other Cartesian products have measure zero and we can remove them from the collection. Hence, this case reduces to the analysis of a single set as worked out in [23], where we considered AMS instead of AMS ergodicity as the control objective. Since AMS ergodicity implies AMS, and $h(B,D,\rho,R)$ reduces to the stabilization entropy notion used in [23] in case of a single set, the desired inequality follows.

**Case 2:** We continue by considering the case where $Q(B_k)\nu(D_l) \in (0,1)$ for all $k,l$. Let $\epsilon > 0$ be small enough such that conditions (i) and (ii) are satisfied and $\epsilon < \rho$. We will show that for any such $\epsilon$ the claim holds.

Let us denote the process measure by $\mu$, which is AMS by assumption. Let $Q$ denote the asymptotic mean, which is by assumption ergodic. As $Q$ is stationary, we can project it unambiguously to a measure on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$. By a slight abuse of notation, we denote by $Q$ both the AMS measure and its projection. Let us consider some Borel set $C \subset \mathbb{R}^N$ and let $f : (\mathbb{R}^N)^{\mathbb{Z}_+} \to \mathbb{R}$ be defined by $f((x_t)_{t\in\mathbb{Z}_+}) := \mathbb{1}_C(x_0)$. It is obvious that this function is in $L^1((\mathbb{R}^N)^{\mathbb{Z}_+})$ (with either $Q$ or $\mu$ as the measure). Recalling our ergodicity assumption, the pointwise ergodic theorem tells us that

$$\frac{1}{N}\sum_{j=0}^{N-1} f \circ \theta^j \xrightarrow[N\to\infty]{Q-a.s.} \int f \,\mathrm{d}Q = \int \mathbb{1}_C(x)\,\mathrm{d}Q(x) = Q(C).$$

Crucially however, the above convergence also happens $\mu$-almost surely (see (2.2) or [18, Lem. 7.5]). Now, for any $V \in \mathcal{B}(\mathbb{W})$, it is clear by the i.i.d. property that

$$P\Big(\Big\{\omega \in \Omega : \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{1}_V(w_t(\omega)) = \nu(V)\Big\}\Big) = 1.$$

As such, noting that $x_t$ and $w_t$ are independent at each time step $t$, it follows that

$$P\Big(\Big\{\omega \in \Omega : \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{1}_{B_k}(x_t(\omega))\mathbb{1}_{D_l}(w_t(\omega)) = Q(B_k)\nu(D_l),\ \forall k,l\Big\}\Big) = 1.$$

Let us denote the full measure set, where this convergence happens, by $\hat{\Omega}$.

We continue by defining the events

$$E_i^j := \Big\{\omega \in \Omega :\ \Big|\frac{1}{T}\sum_{t=0}^{T-1} \mathbb{1}_{B_k}(x_t(\omega))\mathbb{1}_{D_l}(w_t(\omega)) - Q(B_k)\nu(D_l)\Big| < \frac{1}{i} \qquad \forall k,l \text{ whenever } T \geq j\Big\},$$

$$E := \bigcap_{i=1}^{\infty}\bigcup_{j=1}^{\infty} E_i^j.$$

It is not hard to see that $\hat{\Omega} \subseteq E$, hence $P(E) = 1$. Furthermore, observe that $E$ is an infinite intersection of "decreasing" sets (in the containment sense). Hence,

$$P\Big(\bigcup_{j=1}^{\infty} E_i^j\Big) = 1 \quad \text{for all } i \in \mathbb{N}.$$

Let now $I_0$ be large enough such that

$$\frac{1}{I_0} \leq \epsilon(1 - Q(B_k)\nu(D_l)) \quad \text{for all } k \in \{1,\dots,n\},\ l \in \{1,\dots,m\}$$

and observe that $E^1_{I_0} \subseteq E^2_{I_0} \subseteq E^3_{I_0} \subseteq \cdots$. By continuity of probability, we have

$$\lim_{j\to\infty} P(E^j_{I_0}) = P\bigg(\bigcup_{j=1}^{\infty} E^j_{I_0}\bigg) = 1,$$

and thus there exists $J_0$ such that $P(E^j_{I_0}) \geq 1-\epsilon$ for all $j \geq J_0$. For an arbitrary $T \geq J_0$, we define the set of control sequences

$$S_T := \{u_{[0;T-1]}(\omega) : \omega \in E^T_{I_0}\}.$$

We claim that this set is $(T, B, D, \rho, R_\epsilon)$-spanning. We use the set $\tilde{\Omega}_T := E^T_{I_0} \in \mathcal{F}$ to show this, where we note that it satisfies $P(\tilde{\Omega}_T) \geq 1-\epsilon > 1-\rho$, as required. For every $\omega \in \tilde{\Omega}_T$ and all $k, l$, the control sequence $u_{[0;T-1]}(\omega)$ results in the joint state-noise process satisfying

$$\Big|\frac{1}{T}\sum_{t=0}^{T-1} \mathbb{1}_{B_k}(x_t(\omega))\mathbb{1}_{D_l}(w_t(\omega)) - Q(B_k)\nu(D_l)\Big| < \frac{1}{I_0} \leq \epsilon(1-Q(B_k)\nu(D_l)). \qquad (3.4)$$

To prove the claim, it now suffices to show that for all $\omega \in \tilde{\Omega}_T$ and $k, l$ we have

$$\frac{1}{T}|\{t \in [0;T-1] : (\varphi(t, x_0(\omega), u_{[0;T-1]}(\omega), w(\omega)), w_t(\omega)) \in B_k \times D_l\}|$$
$$\geq 1-(1+\epsilon)(1-Q(B_k)\nu(D_l)) = (1+\epsilon)Q(B_k)\nu(D_l) - \epsilon.$$

This follows directly from (3.4). Also, since the coding and control policy can generate at most $|\mathcal{M}|^T$ distinct control sequences by time $T$, it follows that $|S_T| \leq |\mathcal{M}|^T$, therefore $s(T, B, D, \rho, R_\epsilon) \leq |\mathcal{M}|^T$. Recalling that $T \geq J_0$ was arbitrary, we find that

$$\log s(T, B, D, \rho, R_\epsilon) \leq T \log |\mathcal{M}| = TC \quad \text{for all } T \geq J_0,$$

and therefore dividing by $T$ and letting $T \to \infty$ yields the desired capacity bound (3.3), which completes the proof. □

### 3.2.2 Proof of Theorem 3.1.1

*Proof.* Let $c \in (0, 1)$ be such that $c < |\det Df_w(x)|$ for all $x \in \mathbb{R}^N$ and $w \in \mathbb{W}$. Let also $\delta > 0$ and $\rho \in (0, 1)$ be arbitrary. Next, fix a partition of a Borel set $B \subset \mathbb{R}^N$ and let $D = \mathbb{W}$, respectively; let $(B_k)_{k=1}^n$ be a partition of $B$ and $(D_l)_{l=1}^m$ a partition of $D$. Suppose that $B$ has finite Lebesgue measure and

$$Q(B) > 1 - \frac{\delta}{2|\log c|},$$

where $Q$ denotes the asymptotic mean of the state process. Let $\epsilon > 0$ be small enough such that Lemma 3.2.3 holds, resulting in

$$h(B, D, \rho, R_\epsilon) \leq C,$$

where $R_\epsilon$ is the associated collection of $r_{k,l}$'s as defined in Lemma 3.2.3. Let also $1 - r := \sum(1 - r_{k,l})$. It is easy to see that $r = 1 - (1 + \epsilon)Q(B) + nm\epsilon$ (or $r = \epsilon$ if one of the $B_k \times D_l$ has full $Q \times \nu$-measure) thus we see that for every sufficiently small $\epsilon$,

$$2r < \frac{\delta}{|\log c|}. \tag{3.5}$$

Now fix a sufficiently large $T \in \mathbb{N}$ and let $S$ be a finite $(T, B, D, \rho, R_\epsilon)$-spanning set (whose existence is guaranteed by the proof of Lemma 3.2.3) with $\tilde{\Omega} \in \mathcal{F}$, $P(\tilde{\Omega}) \geq 1 - \rho$, the associated subset of $\Omega$. Also let

$$\begin{aligned} A &:= \{(w(\omega), x_0(\omega)) : \omega \in \tilde{\Omega}\}, \\ A(u) &:= \{(w, x) \in \mathbb{W}^{\mathbb{Z}_+} \times \mathbb{R}^N : \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{1}_{B_k \times D_l}(\varphi(t, x, u, w), w_t) \geq 1 - r_{k,l},\ \forall k, l\} \\ A(u, w) &:= \{x \in \mathbb{R}^N : (w, x) \in A(u)\} \end{aligned}$$

and observe that

$$A \subseteq \bigcup_{u \in S} A(u). \tag{3.6}$$

By the theorem of Fubini-Tonelli, we have

$$(\nu \times m)(A(u)) = \int m(A(u, w))\, \mathrm{d}\nu(w). \tag{3.7}$$

Let us now define a set consisting of disjoint collections of subsets of $\{0, \ldots, T-1\}$:

$$\begin{aligned} \mathbb{A} := \{\Lambda = \{\Lambda_k^l\}_{k,l} : \bigsqcup_{k=1}^{n} \bigsqcup_{l=1}^{m} \Lambda_k^l \subseteq \{0, \ldots, T-1\}, \\ |\Lambda_k^l| \geq (1 - r_{k,l})T, \forall k = 1, \ldots, n,\ l = 1, \ldots, m\} \end{aligned}$$

and note that as a consequence of the definition, $|\bigsqcup_{k=1}^{n} \bigsqcup_{l=1}^{m} \Lambda_k^l| \geq (1 - r)T$ for all $\Lambda \in \mathbb{A}$. For such a $\Lambda$, define the set

$$A(u, w, \Lambda) := \{x \in \mathbb{R}^N : (\varphi(t, x, u, w), w_t) \in B_k \times D_l \Leftrightarrow t \in \Lambda_k^l \text{ for all } k, l\}$$

and also (writing $\varphi_{t,u,w}(\cdot) := \varphi(t, \cdot, u, w)$)

$$A_t(u, w, \Lambda) := \varphi_{t,u,w}(A(u, w, \Lambda)), \quad t = 0, 1, \ldots, T-1.$$

It is not hard to see that $A(u, w) = \bigsqcup_{\Lambda \in \mathbb{A}} A(u, w, \Lambda)$ is a disjoint union, implying

$$m(A(u, w)) = \sum_{\Lambda \in \mathbb{A}} m(A(u, w, \Lambda)). \tag{3.8}$$

If $M > 0$ is an upper bound for the density of $\pi_0$, it follows that

$$1 - \rho \leq (\nu \times \pi_0)(A) \leq M \cdot (\nu \times m)(A). \tag{3.9}$$

We also have

$$A_t(u, w, \Lambda) \subseteq B_k \quad \text{whenever } t \in \Lambda_{k,l}, \ \forall \ k \in \{1, \ldots, n\}, \ l \in \{1, \ldots, m\}.$$

Next, we define the following numbers:

$$c_{k,l} := \inf_{(x,w) \in B_k \times D_l} |\det Df_w(x)|.$$

Recalling the fact that $f_w$ is injective and $C^1$, for all $(k, l)$ we have

$$m(A_{t+1}(u, w, \Lambda)) \geq c_{k,l} \cdot m(A_t(u, w, \Lambda)) \text{ whenever } t \in \Lambda_{k,l},$$
$$m(A_{t+1}(u, w, \Lambda)) \geq c \cdot m(A_t(u, w, \Lambda)) \text{ whenever } t \notin \bigsqcup \Lambda_{k,l}.$$

Letting $t^*(\Lambda_{k,l}) \coloneqq \max \Lambda_{k,l}$, $t^*(\Lambda) \coloneqq \max_{k,l} t^*(\Lambda_{k,l})$ and applying the above inequalities repeatedly, it is not hard to see that

$$m(A(u,w,\Lambda))\Big(\prod_{k=1}^{n}\prod_{l=1}^{m} c_{k,l}^{|\Lambda_{k,l}|-1}\Big)c^{rT+nm} \le m(A_{t^*(\Lambda)}(u,w,\Lambda)).$$

Recall that $c \le c_{k,l}$. Now in principle, all the exponents of the $c_{k,l}$'s should be $|\Lambda_{k,l}|$, except for possibly one which should be $|\Lambda_{k,l}|-1$. We do not know which one though, so we write the weaker inequality as above. Combining this with (3.6), (3.7), (3.8) and (3.9), we obtain

$$\begin{aligned}
\frac{1}{M}(1-\rho) &\le (\nu \times m)(A) \\
&\le |S| \max_{u\in S} (\nu \times m)(A(u)) \\
&= |S| \max_{u\in S} \int m(A(u,w))\,\mathrm{d}\nu(w) \\
&= |S| \max_{u\in S} \int \sum_{\Lambda\in\mathbb{A}} m(A(u,w,\Lambda))\,\mathrm{d}\nu(w) \\
&= |S| \max_{u\in S} \sum_{\Lambda\in\mathbb{A}} \int m(A(u,w,\Lambda))\,\mathrm{d}\nu(w) \\
&\le |S| \max_{u\in S} \sum_{\Lambda\in\mathbb{A}} \int m(A_{t^*(\Lambda)}(u,w,\Lambda))c^{-(rT+nm)} \prod_{k=1}^{n}\prod_{l=1}^{m} c_{k,l}^{-(|\Lambda_{k,l}|-1)}\,\mathrm{d}\nu(w) \\
&= |S|\cdot c^{-(rT+nm)} \max_{u\in S} \sum_{t_{1,1}=(1-r_{1,1})T}^{T} \cdots \sum_{t_{n,m}=(1-r_{n,m})T}^{T} \\
&\quad \int \sum_{\Lambda\in\mathbb{A}:\ t^*(\Lambda_{k,l})=t_{k,l}\forall k,l} m(A_{t^*(\Lambda)}(u,w,\Lambda)) \prod_{k=1}^{n}\prod_{l=1}^{m} c_{k,l}^{-(|\Lambda_{k,l}|-1)}\,\mathrm{d}\nu(w) \\
&\le |S|\cdot c^{-(2rT+nm)} \max_{u\in S} \sum_{t_{1,1}=(1-r_{1,1})T}^{T} \cdots \sum_{t_{n,m}=(1-r_{n,m})T}^{T}
\end{aligned}$$

$$\int \sum_{\Lambda\in\mathbb{A}:\ t^*(\Lambda_{k,l})=t_{k,l}\forall k,l} m(A_{t^*(\Lambda)}(u,w,\Lambda)) \prod_{k=1}^{n}\prod_{l=1}^{m} c_{k,l}^{-((1-r_{k,l})T-1)}\,\mathrm{d}\nu(w).$$

In the last inequality we use that

$$\begin{aligned} c^{rT+nm}\prod_{k,l} c_{k,l}^{|\Lambda_{k,l}|-1} &= c^{rT+\sum_{k,l}|\Lambda_{k,l}|}\prod_{k,l}\Big(\frac{c_{k,l}}{c}\Big)^{|\Lambda_{k,l}|-1} \\ \geq c^{rT+\sum_{k,l}|\Lambda_{k,l}|}\prod_{k,l}\Big(\frac{c_{k,l}}{c}\Big)^{(1-r_{k,l})T-1} &= c^{rT+\sum_{k,l}|\Lambda_{k,l}|-(1-r)T+nm}\prod_{k,l} c_{k,l}^{(1-r_{k,l})T-1} \\ \geq c^{2rT+nm}\prod_{k,l} c_{k,l}^{(1-r_{k,l})T-1}. \end{aligned}$$

Observe that the sets $A_{t^*(\Lambda)}(u,w,\Lambda)$ with $\Lambda\in\mathbb{A}$, $t^*(\Lambda)$ fixed, are pairwise disjoint, since they are the images of the corresponding sets $A(u,w,\Lambda)$ under the injective map $\varphi_{t^*(\Lambda),u,w}$. Moreover, all of these sets are contained in $B$. Hence,

$$\sum_{\Lambda\in\mathbb{A}:t^*(\Lambda_{k,l})=t_{k,l}\forall k,l} m(A_{t^*(\Lambda)}(u,w,\Lambda)) \leq m(B),$$

which, together with the above chain of inequalities, implies

$$\frac{1}{M}(1-\rho)\leq |S|\cdot m(B)\cdot c^{-(2rT+nm)}\cdot\prod_{k=1}^{n}\prod_{l=1}^{m} c_{k,l}^{-((1-r_{k,l})T-1)}\prod_{k=1}^{n}\prod_{l=1}^{m}(r_{k,l}T+1).$$

Since this inequality holds for every $T$ sufficiently large, we can take logarithms on both sides, divide by $T$ and let $T\to\infty$. This results in

$$0\leq h(B,D,\rho,R_\epsilon)-2r\log c-\sum_{k=1}^{n}\sum_{l=1}^{n}(1-r_{k,l})\log c_{k,l}.$$

Recalling the definition of $r_{k,l}$, the fact that $\epsilon$ can be chosen arbitrarily small and (3.5), this leads to the estimate

$$C + \delta \geq \sum_{k=1}^{n} \sum_{l=1}^{n} Q(B_k)\nu(D_l) \inf_{(x,w)\in B_k \times D_l} \log|\det Df_w(x)|.$$

Considering the supremum of the right-hand side over all finite measurable partitions of $B$ and $\mathbb{W}$ leads to

$$C + \delta \geq \int\int \mathbb{1}_B(x) \log|\det Df_w(x)| \, \mathrm{d}Q(x) \, \mathrm{d}\nu(w),$$

where we use that the integrand is uniformly bounded below by $\log c$ (and hence, we can assume that it is non-negative). Considering now an increasing sequence of sets $B_k \subset \mathbb{R}^N$ whose union is $\mathbb{R}^N$, we can invoke the theorem of monotone convergence to obtain the desired estimate, observing that $\delta$ can be made arbitrarily small as $B_k$ becomes arbitrarily large. □

### 3.2.3 Proof of Theorem 3.1.2

*Proof.* In this proof, we assume that the controller and encoder have knowledge of the noise realization. If we can prove the bound under this setup, the bound will clearly also hold in the setup of Theorem 3.1.2. Suppose for a contradiction that a causal coding and control policy is such that the state process is AMS ergodic, but that the converse of inequality (3.2) holds. Let $r > 0$ be small enough so that

$$C < (1 - 3r) \int\int \log|f_w'(x)| \, \mathrm{d}Q(x) \, \mathrm{d}\nu(w).$$

Since we can approximate the integral by the integral over associated step functions, for any $b \in \mathbb{N}$ large enough, there exists a disjoint collection of intervals $B_1, \dots, B_{2^{b+1}}$ and a partition $D_1, \dots, D_m$ of $\mathbb{W}$ such that $B := [-b, b] = \bigsqcup_{k=1}^{2^{b+1}} B_k$, and

$$C < (1-3r) \sum_{l=1}^{m} \sum_{k=1}^{2^{b+1}} \nu(D_l) Q(B_k) \log c_{k,l},$$

where $c_{k,l} := \inf_{(x,w)\in B_k \times D_l} |f'_w(x)|$. Put $n := 2^{b+1} + 1$, and fix a $b$ (and the associated collection $(B_k)_{k=1}^{n-1}$ of intervals) further large enough such that

$$Q([-b, b])(1-r) > 1 - \frac{2.5}{2} r \tag{3.10}$$

which is possible by continuity of probability. Finally, let $B_n := \mathbb{R} \setminus \bigsqcup_{k=1}^{n} B_k$. For brevity, in the rest of the proof we write

$$m_{k,l} := Q(B_k)\nu(D_l), \quad k = 1, \dots, n, \ l = 1, \dots, m.$$

Next, we define the following sets in a slightly different manner than in the previous proof:

$$\begin{aligned} A_T(u, w) := \{ x \in \mathbb{R} : \forall k, l \text{ and } \forall N \in \{ \lceil T(1-3r) \rceil, \dots, T \}, \\ \frac{1}{N} |\{ t \in [0; N-1] : (\varphi(t, x, u, w), w_t) \in B_k \times D_l \}| \geq m_{k,l}(1-r) \}. \end{aligned}$$

It is easy to see that this set is always bounded. Later on, for appropriate parameters, we will also see that the set is nonempty. For these cases, let

$$\overline{A}_T(u, w) := [\inf A_T(u, w), \sup A_T(u, w)]$$

and let $x_0(T, u, w)$ denote the midpoint of this interval. We claim that there exists $T$ larger than some threshold $M_1 = M_1(r)$ so that for all $u, w$ and $x_1, x_2 \in A_T(u, w)$ there exists a $t^*$ with $\lceil (1-2.5r)T \rceil \leq t^* \leq T-1$ satisfying

$$\varphi(t^*, x_i, u, w) \in B \text{ for } i \in \{1, 2\}.$$

To see this, suppose otherwise. Then for at least one $i \in \{1, 2\}$ we have

$$\begin{aligned}
&|\{t \in [0; T-1] : \varphi(t, x_i, u, w) \in B\}| \leq \lceil (1-2.5r)T \rceil + \frac{1}{2}(T - \lceil (1-2.5r)T \rceil) \\
&\leq \frac{1}{2}((1-2.5r)T + 1) + \frac{1}{2}T = \frac{1}{2} + (1 + (1-2.5r))\frac{1}{2}T \\
&= \frac{1}{2} + \left(1 - \frac{2.5}{2}r\right)T < (1-r)Q(B)T,
\end{aligned}$$

where the last inequality holds for $T$ large enough from the assumption (3.10) on $Q(B)$.

This is a contradiction to $x_i \in A_T(u, w)$, which follows by recalling the definition of $A_T(u, w)$. Let now $\epsilon > 0$ and $\delta > 0$ be given. By the pointwise ergodic theorem (see the construction in the proof of Lemma 3.2.3), there exists an $M_2 := M_2(\epsilon, \delta) \in \mathbb{N}$ such that for all $T \geq M_2$

$$\begin{aligned}
P(\{\omega \in \Omega : \forall k, l, \forall N \geq (1-3r)T,& \\
\frac{1}{N}\sum_{t=0}^{N-1} \mathbb{1}_{B_k}(x_t(\omega))\mathbb{1}_{D_l}(w_t(\omega)) &\geq m_{k,l}(1-\delta)\}) > 1 - \epsilon.
\end{aligned}$$

We denote by $\tilde{\Omega}(\epsilon, \delta, M_2)$ the set of $\omega$'s for which the event within the braces of the above expression occurs. Recalling that $c_{k,l} := \inf_{(x,w) \in B_k \times D_l} |\det Df_w(x)|$ and letting

$u, w$ and $x_1, x_2 \in A_T(u,w)$ be arbitrary, we have

$$|x_1 - x_2| \leq \frac{2b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-\delta)t^*}} \leq \frac{2b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-\delta)T(1-2.5r)}} \tag{3.11}$$

which follows by noting that

$$\prod_{k,l} c_{k,l}^{(1-\delta)m_{k,l}(1-2.5r)T}|x_1 - x_2| \leq \prod_{k,l} c_{k,l}^{(1-\delta)m_{k,l}(1-2.5r)T}(|x_1| + |x_2|)$$
$$\leq \prod_{k,l} c_{k,l}^{(1-\delta)m_{k,l}t^*}|x_1| + \prod_{k,l} c_{k,l}^{(1-\delta)m_{k,l}t^*}|x_2| \leq |\varphi(t^*, x_1, u, w)| + |\varphi(t^*, x_2, u, w)| \leq 2b.$$

Given a realization $\omega \in \Omega$, we denote by $x_0(\omega)$ and $w(\omega)$ the resulting realizations of the initial state and noise sequence, respectively. Given these realizations, the control sequence is thus fully determined, and denoted by $u(\omega)$. It follows quite easily that $\omega \in \tilde{\Omega}(\epsilon, \delta, M_2)$ implies $x_0(\omega) \in A_T(u(\omega), w(\omega))$ for all $T \geq M_2$ and all $\delta < r$. Combining this with (3.11), we conclude that

$$|x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \frac{b}{\prod_{k,l} c_{k,l}^{(1-\delta)m_{k,l}(1-2.5r)T}}$$

for every $T \geq M_2(\epsilon, \delta)$ and every $\omega \in \tilde{\Omega}(\epsilon, \delta, M_2)$. Letting $\delta$ be small enough so that both $(1-3r) \leq (1-2.5r)(1-\delta)$ and $\delta < r$ hold, we conclude that

$$\liminf_{T\to\infty} P\Bigg(\Bigg\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}}\Bigg\}\Bigg) \geq 1 - \epsilon$$

and since $\epsilon > 0$ was also arbitrary, it follows that

$$\limsup_{T\to\infty} P\Big(\Big\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w(\omega))| > \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}}\Big\}\Big) = 0. \quad (3.12)$$

We will see that our initial hypothesis leads to a contradiction with the above equation. To this effect, let us choose $\alpha \in (0, 1/2)$ small enough so that for all sufficiently large $L$:

$$1 - \frac{\rho_{\min} \cdot (1-\alpha)}{2 \cdot \rho_{\max}} + \frac{\rho_{\max}^2}{2L\rho_{\min}^2} + \frac{2 \cdot \rho_{\max}}{\rho_{\min}} \frac{\alpha}{1-\alpha} < 1. \quad (3.13)$$

Let also $\tilde{\Omega} \in \mathcal{F}$ be such that $P(\tilde{\Omega}) > 1 - \alpha$, and such that for all $T$ large enough (say, larger than $C(\alpha)$),

$$|x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}}$$

for all $\omega \in \tilde{\Omega}$. The idea from here on is to treat $\tilde{\Omega}$ as "the universe", since conditioning on this set gives the above deterministic bound. We proceed by defining

$$\begin{aligned} U_T &\coloneqq \{(\gamma_0(q'_0), \ldots, \gamma_{T-1}(q'_{[0;T-1]})) \in U^T : q'_{[0;T-1]} \in (\mathcal{M}')^T\}, \\ \tilde{U}_T &\coloneqq \{(\gamma_0(q'_0(\omega)), \ldots, \gamma_{T-1}(q'_{[0;T-1]}(\omega))) \in U^T : \omega \in \tilde{\Omega}\}, \\ \tilde{R} &\coloneqq \limsup_{T\to\infty} \frac{1}{T} \log |\tilde{U}_T|. \end{aligned}$$

We now treat two distinct cases: In Case 1, we show that the condition $\tilde{R} < (1-3r)\sum_{k,l} m_{k,l} \log c_{k,l}$ cannot hold if we want to achieve the desired result. This leaves us with Case 2: the condition that $\tilde{R} \geq (1-3r)\sum_{k,l} m_{k,l} \log c_{k,l}$; however, this condition would imply $\tilde{R} > C$. We show that this cannot hold either, through a tedious argument involving a strong converse to channel coding (with feedback) and optimal transport theory. In the following, we study these two cases separately.

**Case 1:** Let us suppose that

$$\tilde{R} < (1-3r)\sum_{k,l} m_{k,l}\log c_{k,l}. \tag{3.14}$$

Let $\epsilon > 0$ be small enough so that $\tilde{R}+2\epsilon < (1-3r)\sum_{k,l} m_{k,l}\log c_{k,l}$ and observe that for all $T$ large enough,

$$|\tilde{U}_T| \le 2^{(\tilde{R}+\epsilon)T}. \tag{3.15}$$

Recall also that $\tilde{\Omega}$ is such that for all $T$ large enough,

$$|x_0(\omega) - x_0(T, u(\omega), w(\omega))| \le \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \quad \text{for all } \omega \in \tilde{\Omega}. \tag{3.16}$$

We now fix a noise realization $w$. For all $T$ large enough so that (3.15) holds,

$$\begin{aligned} m\Big(\bigcup_{u\in\tilde{U}_T} \overline{A}_T(u,w)\Big) &\le \frac{2b\cdot 2^{(\tilde{R}+\epsilon)T}}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \le \frac{2b\cdot 2^{((1-3r)\sum_{k,l} m_{k,l}\log c_{k,l}-\epsilon)T}}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \\ &\le \frac{2b\cdot 2^{-\epsilon T}\cdot \prod_{k,l} 2^{T(1-3r)m_{k,l}\log c_{k,l}}}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} = \frac{2b}{2^{\epsilon T}}, \end{aligned}$$

where the inequalities follow by applying the union bound, and from (3.15) and (3.14). The above yields

$$\lim_{T\to\infty} m\Big(\bigcup_{u\in\tilde{U}_T} \overline{A}_T(u,w)\Big) = 0,$$

and thus by the absolute continuity and boundedness assumptions on $\pi_0$, we have

$$\lim_{T\to\infty} \pi_0\Big( \bigcup_{u\in\tilde{U}_T} \overline{A}_T(u,w)\Big)\Big) = 0.$$

On the other hand, let us define $J := \{w \in \mathbb{W}^{\mathbb{Z}_+} : P(\{\omega \in \tilde{\Omega}|w(\omega) = w\}) > 0\}$. We note that $J$ is the projection of $\tilde{\Omega}$ onto $\mathbb{R}^{\mathbb{Z}_+}$ from which the set $\{w : P(\omega \in \tilde{\Omega}|w(\omega) = w) = 0\}$ is taken out; these ensure that $J$ is a universally measurable set since the image of a Borel set under a measurable map is universally measurable [14].

We can therefore write

$$\begin{aligned}
&\limsup_{T\to\infty} P\Big(\Big\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \big|\omega \in \tilde{\Omega}\Big\}\Big) \\
&= \limsup_{T\to\infty} \Big(P\Big(\Big\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \\
&\qquad \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \big|\omega \in \tilde{\Omega}, w(\omega) \in J\Big\}\Big) \cdot P(J) \\
&+ P\Big(\Big\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \\
&\qquad \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \big|\omega \in \tilde{\Omega}, w(\omega) \in J^c\Big\}\Big) \cdot P(J^c)\Big).
\end{aligned}$$

Now, noting that $P(\tilde{\Omega}) > 1 - \alpha$ implies $\nu(J^c) \leq \alpha$, we can further write

$$\begin{aligned}
&\leq \limsup_{T\to\infty} P\Big(\Big\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \\
&\qquad \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \big|\omega \in \tilde{\Omega}, w(\omega) \in J\Big\}\Big) \cdot P(J) + \alpha.
\end{aligned}$$

Observe that for a noise realization $w \in J$, we have

$$\begin{aligned}
&\limsup_{T\to\infty} P\Big(\Big\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \\
&\qquad\qquad \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \Big| \omega \in \tilde{\Omega}, w(\omega) = w\Big\}\Big) \\
&\leq \limsup_{T\to\infty} P\Big(\Big\{\omega \in \Omega : x_0(\omega) \in \bigcup_{u \in \tilde{U}_T} \overline{A}_T(u, w) \Big| \omega \in \tilde{\Omega}, w(\omega) = w\Big\}\Big) \\
&\leq \frac{1}{P(\omega \in \tilde{\Omega} | w(\omega) = w)} \limsup_{T\to\infty} P\Big(\Big\{\omega \in \Omega : x_0(\omega) \in \bigcup_{u \in \tilde{U}_T} \overline{A}_T(u, w) | w(\omega) = w\Big\}\Big) \\
&= \frac{1}{P(\omega \in \tilde{\Omega} | w(\omega) = w)} \limsup_{T\to\infty} \pi_0\Big(\bigcup_{u \in \tilde{U}_T} \overline{A}_T(u, w)\Big) = 0,
\end{aligned}$$

where the first inequality can be justified by noting that

$$|x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \quad \Rightarrow \quad x_0(\omega) \in \overline{A}_T(u(\omega), w)$$

for all $T$ sufficiently large (see (3.16)) and the last inequality follows by independence of noise and initial state. We thus have a uniform upper bound on the limsup when conditioned on $w \in J$, hence

$$\begin{aligned}
&\limsup_{T\to\infty} P\Big(\Big\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \\
&\qquad\qquad \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \Big| \omega \in \tilde{\Omega}, w(\omega) \in J\Big\}\Big) = 0.
\end{aligned}$$

Therefore,

$$\limsup_{T\to\infty} P\Big(\Big\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \frac{b}{\prod_{k.l} c_{k,l}^{m_{k,l}(1-3r)T}} \Big| \omega \in \tilde{\Omega}\Big\}\Big) \leq \alpha,$$

which contradicts (3.12), since $\alpha < 1/2$. Hence, the proof for Case 1 is complete.

**Case 2:** Now we suppose that

$$\tilde{R} \geq (1-3r)\sum_{k,l} m_{k,l}\log c_{k,l},$$

thus by assumption we also have $\tilde{R} > C$. Recall that the proof is by contradiction. In this case, we will obtain a contradiction to a generalized version of the strong converse theorem for discrete memoryless channels with feedback (see [24] and Theorem A.2.8). Recall that by definition of $\tilde{\Omega}$, we have that for any $T$ sufficiently large, the inequality

$$|x_0(\omega) - x_0(T, u(\omega), w(\omega))| \leq \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}}$$

holds for any $\omega \in \tilde{\Omega}$. Also recall that $P(\tilde{\Omega}) > 1-\alpha$ for $\alpha$ satisfying the important assumption (3.13). As such, there must exist some noise realization $w$ such that $P(\{\omega \in \tilde{\Omega} | w(\omega) = w\}) > 1-\alpha$. This can be seen by contradiction; suppose no such realization exists. Letting $\nu$ denote the measure on the space of noise realizations, we can write

$$P(\omega \in \tilde{\Omega}) = \int P(\omega \in \tilde{\Omega} | w(\omega) = \tilde{w})\,\mathrm{d}\nu(\tilde{w}) \leq \int (1-\alpha)\,\mathrm{d}\nu(\tilde{w}) = 1-\alpha \tag{3.17}$$

which is a contradiction since $P(\tilde{\Omega}) > 1-\alpha$. The existence of such a realization $w$ yields

$$\liminf_{T\to\infty} P\Big(\Big\{\omega \in \Omega : |x_0(\omega) - x_0(T, u(\omega), w)| \leq \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}} \big| w(\omega) = w\Big\}\Big) > 1-\alpha. \tag{3.18}$$

In the remainder of the proof, we condition on the occurrence of the noise realization $w$. We follow an almost identical approach as in the proof from [23]; we will construct a sequence of codes to transmit a uniform random variable which contradicts a version of the strong converse result for DMCs. This is accomplished in four steps.

**Step 1 (Construction of bins):** For every $T \geq 1$, define $S_T \coloneqq \{x_0(T,u,w) : u \in \tilde{U}_T\}$ and enumerate the elements of this set so that

$$S_T \coloneqq \{x_1(T), \ldots, x_{n_1(T)}(T)\}. \tag{3.19}$$

We continue by defining the not necessarily disjoint collection of bins

$$B_i^T \coloneqq \Big\{x \in \mathbb{R} : |x - x_i(T)| \leq \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}}\Big\}, \quad i = 1, \ldots, n_1(T).$$

Note that for a fixed $T$, each bin has the same Lebesgue measure which we denote by $\rho_T \coloneqq (2b)/\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}$. Recalling that $P(\{\omega \in \tilde{\Omega} | w(\omega) = w\}) > 1-\alpha$, it follows that

$$1 - \alpha < \liminf_{T\to\infty} P\Big(\Big\{\omega \in \Omega : x_0(\omega) \in \bigcup_{i=1}^{n_1(T)} B_i^T \big| w(\omega) = w\Big\}\Big),$$

from which by independence of noise and initial state, we obtain

$$1 - \alpha < \liminf_{T\to\infty} \pi_0\Big(\bigcup_{i=0}^{n_1(T)} B_i^T\Big). \tag{3.20}$$

We will disregard the bins that are only partially contained in $K$. Since $\rho_T \to 0$ as $T \to \infty$ and the union of the measure of bins that are partially inside of $K$ can have at most a Lebesgue measure of $2\rho_T$, they will contribute negligible measure as $T$ gets large. Also, let us suppose without loss of generality that the ordering of the bins in

(3.19) is such that the last $n(T)$ are the ones not contained in $K$. Observing that

$$\liminf_{T\to\infty} \pi_0\Big(\bigcup_{i=0}^{n_1(T)} B_i^T\Big) = \liminf_{T\to\infty} \pi_0\Big(K \cap \bigcup_{i=0}^{n_1(T)} B_i^T\Big) = \liminf_{T\to\infty} \pi_0\Big(\bigcup_{i=0}^{n_1(T)-n(T)} B_i^T\Big)$$
$$\leq \liminf_{T\to\infty} \rho_{\max} \cdot m\Big(\bigcup_{i=0}^{n_1(T)-n(T)} B_i^T\Big) \leq \liminf_{T\to\infty}\Big(\frac{\rho_{\max} \cdot 2b \cdot (n_1(T) - n(T))}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}}\Big),$$

we obtain

$$\frac{1-\alpha}{2b \cdot \rho_{\max}} \leq \liminf_{T\to\infty} \left(\frac{(n_1(T) - n(T))}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}}\right),$$

from which we conclude that the number of bins $n_1(T) - n(T)$ which are entirely contained in $K$ must grow at an exponential rate of at least $\sum_{k,l} m_{k,l}(1-3r)\log c_{k,l}$ with $T$, just as $n_1(T)$ does. Thus, since we are concerned only with the number of bins entirely contained in $K$, we may as well assume that all are entirely in $K$ (or alternatively, relabel $n_1(T) - n(T)$ to be $n_1(T)$).

We continue by extracting a sub-collection of disjoint bins $(C_i^T)_{i=1}^{n_2(T)}$ as described in [23, App. A]. This new sub-collection has the property that

$$\frac{1}{2} m\Big(\bigcup_{i=1}^{n_1(T)} B_i^T\Big) \leq m\Big(\bigcup_{i=1}^{n_2(T)} C_i^T\Big).$$

Also, it is clear that for any given $T$, $\frac{1}{2}n_1(T) \leq n_2(T)$. Hence, we also have the exponential growth condition of

$$\limsup_{T\to\infty} \frac{1}{T} \log n_2(T) \geq (1-3r) \sum_{k,l} m_{k,l} \log c_{k,l}.$$

Analogously to [23], define the collection $(D_i^T)_{i=1}^{n_2(T)}$[1] and observe that $m(D_i^T \backslash C_i^T) \leq$

[1]These sets should not be confused with the set $D_1, \ldots, D_m \subset \mathbb{W}$.

$\rho_T$ for all $i$. Finally, for a fixed $L \in \mathbb{N}$ we join $L$ successive $D_i^T$ blocks (see [23, p. 27] for an exact formulation) to get a collection $(E_i^T)_{i=1}^{n_3(T)}$, where $n_3(T) = \lfloor \frac{n_2(T)}{L} \rfloor + 1$, possibly adding some empty sets in the last block. Again, the following holds:

$$\limsup_{T\to\infty} \frac{1}{T} \log n_3(T) \geq (1-3r) \sum_{k=1}^{n} \sum_{l=1}^{m} m_{k,l} \log c_{k,l}, \quad m(E_i^T) \geq L\rho_T.$$

We also define

$$M_T := \bigcup_{i=1}^{n_1(T)} B_i^T \qquad \overline{M}_T := \bigcup_{i=1}^{n_3(T)} E_i^T \backslash (D_{iL}^T \backslash C_{iL}^T)$$

and observe that $m(M_T) \leq 2n_2(T)\rho_T \leq 2n_3(T)L\rho_T$.

**Step 2 (Auxiliary coding scheme):** We now construct a sequence of codes to transmit information over the channel. We will transmit a quantized version of the initial state random variable $x_0$. The quantization will be done using the bins constructed earlier. For a fixed $L$ and for each $T$, we will construct a code. Note that we are considering a channel with feedback, which can be used by the encoding function. For a given $T$, the encoding and decoding processes are specified as follows.

**Encoder**: We give to the encoder the noise realization $w$ that we have conditioned on throughout, the function $f$ corresponding to the system dynamics, and the fixed causal coding and control policy. In the classical notion of a code, the encoding function is a deterministic map and given the (system) noise sequence realization, this is the case here. The transmitted codeword is determined as follows. For an initial state realization $x_0$, the first symbol of the codeword is $q_0 = \gamma_0^e(x_0)$. Now, because the channel has feedback, the encoder can determine $u_0$ by applying the decoding function of the fixed causal coding and control policy to the output of the channel resulting from the first codeword symbol $q_0$. Thus, using the fixed and known noise realization $w$, $x_1$ can be computed. Then, the second codeword symbol $q_1 = \gamma_1^e(x_0, x_1, u_0)$ is computed again using the causal coding and control policy, and so on until $q_{T-1}$ is determined (note that the encoder makes use of the channel feedback from the channel, and thus we use the generalized version of the strong converse theorem for channel capacity to obtain a contradiction). We are essentially viewing the coding and control policy as a scheme from which the initial state can be estimated at the controller end of the channel.

**Decoder**: At time $T$, the decoder has received $T$ symbols from the channel, which are used to compute the control decisions $u_0, \ldots, u_{T-1}$ according to the fixed causal coding and control policy. The decoder also has knowledge of the noise sequence $w$ and uses it to compute the point $x_0(T, u, w)$. Our goal is to use the received channel output and control sequence to reconstruct the index $Y$ of the bin $E_Y^T$ containing $x_0$. We do this by looking at the point $x_0(T, u, w)$ for the observed control sequence $u$. Note that $w$ can be thought of as deterministic since we are conditioning on its occurrence. Recall also that $x_0(T, u, w)$ is the "midpoint" of the set $A_T(u, w)$, and *can be computed without knowledge of the initial state* $x_0$. We simply decide on our guess $\tilde{Y}$ of the index as follows.

- If $x_0(T, u, w) \in M_T$, take the index $i$ of the set $E_i^T$ containing $x_0(T, u, w)$.
- If $x_0(T, u, w) \notin M_T$, then decide randomly between $i$ and $i+1$, where $i$ is the index of the set $E_i^T$ that $x_0(T, u, w)$ belongs to.

**Analysis of probability of the error for the code**. To study the probability of error, let $Y$ be a random variable on the indices $\{1, \ldots, n_3(T)\}$, where $P(Y = i) = \pi_0(E_i^T)$. We analyze $P(\tilde{Y} \neq Y)$.

First, by construction of the bins and the estimation scheme, we have

$$P\Big(\tilde{Y} \neq Y \big| x_0 \in \overline{M}_T, |x_0 - x_0(T,u,w)| \leq \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}}\Big) = 0$$

and

$$P\Big(\tilde{Y} \neq Y \big| x_0 \in M_T \backslash \overline{M}_T, |x_0 - x_0(T,u,w)| \leq \frac{b}{\prod_{k,l} c_{k,l}^{m_{k,l}(1-3r)T}}\Big) \leq \frac{1}{2}.$$

As such, from (3.18), it is not hard to see that for every $T$ sufficiently large,

$$P(Y \neq \tilde{Y}) \leq \frac{1}{2}\pi_0(M_T \backslash \overline{M}_T) + \alpha.$$

By an analysis exactly as in [23], we have

$$\pi_0(\mathcal{M}_T \backslash \overline{M}_T) \leq \frac{1}{L}\frac{\rho_{\max}}{\rho_{\min}}\pi_0(M_T).$$

Combining the above two inequalities, we obtain

$$\sum_{i=1}^{n_3(T)} P(Y=i)P(\tilde{Y} \neq Y | Y=i) \leq \frac{1}{2L}\frac{\rho_{\max}}{\rho_{\min}}\pi_0(M_T) + \alpha.$$

**Step 3 (Introduction of an auxiliary uniform random variable):** In order to obtain a contradiction to the strong converse theorem for DMCs, we need to transmit a random variable *uniformly* distributed on the indices $1, \ldots, n_3(T)$. Let us call this random variable $W = W_T$. Of course, at any time step, $W$ must be conditionally independent from the channel output, given the channel input. To obtain the desired contradiction, we must show that $\lim_{T\to\infty} P(W \neq \tilde{Y}) < 1$. Before considering this quantity, note that by following exactly the same steps as in [23], we obtain

$$\pi_0(M_T) \leq \rho_{\max} \cdot m(M_T) \leq 2n_3(T)\rho_{\max} \cdot L\rho_T$$

and also

$$\sum_{i=1}^{n_3(T)} \frac{1}{n_3(T)} P(\tilde{Y} \neq Y | Y = i) \leq \frac{\alpha + \frac{\rho_{\max}\pi_0(M_T)}{4L\rho_{\min}}}{\frac{\rho_{\min}\pi_0(M_T)}{2\rho_{\max}}}.$$

Again as in [23] we have

$$P(W \neq \tilde{Y}) = \sum_{i=1}^{n_3(T)} P(W = i)P(\tilde{Y} \neq W | W = i) \leq P(Y \neq W) + \frac{\alpha + \frac{\rho_{\max}\pi_0(M_T)}{4L\rho_{\min}}}{\frac{\rho_{\min}\pi_0(M_T)}{2\rho_{\max}}}. \tag{3.21}$$

**Step 4 (Application of optimal transport):** Recall the independence condition mentioned above that $W$ must satisfy. To achieve this, one could adjoin $W$ to the common probability space using the product measure, thus keeping $W$ independent from all other random variables. Observe however, that the random variable $x_0$ satisfies the independence condition that we require $W$ to satisfy. As such, we are free to choose any possible coupling between $W_T$ and $x_0$ while still ensuring that $W$ will remain independent form the channel output given the channel input (in particular, $x_0$ and $W$ need not be independent). We will take advantage of this observation.

Consider (3.21) and note that if the limit as $T \to \infty$ of the right-hand side is strictly less than 1, then we will have the desired contradiction with the strong converse. As such, we proceed by finding a coupling between $W$ and $x_0$ which makes $P(Y \neq W)$ small enough so that the limit is less than 1.

We continue by letting $\mu$ denote the law of $Y$. That is, for every index $i \in 1, \ldots, n_3(T)$, $\mu(i) = \pi_0(E_i^T)$. Let also $\nu$ represent the law of $W$, i.e., a uniform measure on the set $\{1, \ldots, n_3(T)\}$. We now invoke Lemma A.1.3, which guarantees the existence of a coupling $(Y, W) : (\Omega, \mathcal{F}, P) \to \{1, \ldots, n_3(T)\}^2$ such that

$$P(Y \neq W) = \frac{1}{2} \sum_{i=1}^{n_3(T)} |\mu(i) - \nu(i)|.$$

Let now $A = \{i \in \{1, \ldots, n_3(T)\} : \mu(i) \geq \nu(i)\}$ and observe that

$$\begin{aligned} 1 - \sum_{i=1}^{n_3(T)} \min(\mu(i), \nu(i)) &= \frac{1}{2} \sum_{i=1}^{n_3(T)} \mu(i) + \frac{1}{2} \sum_{i=1}^{n_3(T)} \nu(i) - \sum_{i \in A} \nu(i) - \sum_{i \in A^c} \mu(i) \\ &= \frac{1}{2} \sum_{i \in A} \mu(i) - \frac{1}{2} \sum_{i \in A^c} \mu(i) - \frac{1}{2} \sum_{i \in A} \nu(i) + \frac{1}{2} \sum_{i \in A^c} \nu(i) \\ &= \frac{1}{2} \Big( \sum_{i \in A} \mu(i) - \nu(i) \Big) + \frac{1}{2} \Big( \sum_{i \in A^c} \nu(i) - \mu(i) \Big) = \frac{1}{2} \sum_{i=1}^{n_3(T)} |\mu(i) - \nu(i)|. \end{aligned}$$

Thus, we can write

$$P(Y \neq W) = \frac{1}{2} \sum_{i=1}^{n_3(T)} |\mu(i) - \nu(i)| = 1 - \sum_{i=1}^{n_3(T)} \min(\mu(i), \nu(i)).$$

To get an upper bound for the right-hand side, note that

$$\begin{aligned} \mu(i) = \pi_0(E_i^T) \geq \rho_{\min} \cdot m(E_i^T) &= \frac{n_3(T)}{n_3(T)} \cdot m(E_i^T) \cdot \rho_{\min} \\ &\geq \frac{n_2(T) \cdot \rho_T \cdot \rho_{\min}}{n_3(T)} \geq \frac{m(M_T) \cdot \rho_{\min}}{2 \cdot n_3(T)} \geq \frac{\pi_0(M_T) \cdot \rho_{\min}}{2 \cdot \rho_{\max} \cdot n_3(T)} \geq \frac{\rho_{\min} \cdot (1 - \alpha)}{2 \cdot \rho_{\max} \cdot n_3(T)}. \end{aligned}$$

Recalling that $\nu(i) = 1/n_3(T)$ for each $i$, we have $\min(\mu(i), \nu(i)) \geq (\rho_{\min} \cdot (1-\alpha))/(2 \cdot \rho_{\max} \cdot n_3(T))$ for all $i$, and therefore

$$P(Y \neq W) \leq 1 - \frac{\rho_{\min} \cdot (1 - \alpha)}{2 \cdot \rho_{\max}}.$$

Combining with (3.21), we obtain

$$P(W \neq \tilde{Y}) \leq 1 - \frac{\rho_{\min} \cdot (1-\alpha)}{2 \cdot \rho_{\max}} + \frac{\alpha + \frac{\rho_{\max}\pi_0(M_T)}{4L\rho_{\min}}}{\frac{\rho_{\min}\pi_0(M_T)}{2\rho_{\max}}}$$

which holds for all $T$ sufficiently large. We now evaluate the right-hand side to determine its behavior as $T$ tends to infinity. We have

$$\begin{aligned}
&\limsup_{T\to\infty}\left(1 - \frac{\rho_{\min} \cdot (1-\alpha)}{2 \cdot \rho_{\max}} + \frac{\alpha + \frac{\rho_{\max}\pi_0(M_T)}{4L\rho_{\min}}}{\frac{\rho_{\min}\pi_0(M_T)}{2\rho_{\max}}}\right) \\
&\leq 1 - \frac{\rho_{\min} \cdot (1-\alpha)}{2 \cdot \rho_{\max}} + \frac{\rho_{\max}^2}{2L\rho_{\min}^2} + \frac{2 \cdot \alpha \cdot \rho_{\max}}{\rho_{\min}} \limsup_{T\to\infty} \frac{1}{\pi_0(M_T)} \\
&\leq 1 - \frac{\rho_{\min} \cdot (1-\alpha)}{2 \cdot \rho_{\max}} + \frac{\rho_{\max}^2}{2L\rho_{\min}^2} + \frac{2 \cdot \rho_{\max}}{\rho_{\min}} \frac{\alpha}{1-\alpha},
\end{aligned}$$

where the last inequality follows from (3.20). Recall now that throughout, $L \in \mathbb{N}$ was fixed but arbitrary. Taking $L$ large enough so that (3.13) holds, and writing $T$-subscripts to emphasize $T$-dependence, we obtain $\limsup_{T\to\infty} P(W_T \neq \tilde{Y}_T) < 1$, which is a contradiction, since it negates the strong converse theorem for DMCs with feedback. Hence, the proof is complete. □

# Chapter 4

# Conclusion and Future Work

In conclusion, this thesis has focused on the problem of stabilizing a stochastic non-linear discrete time control systems subject to information constraints. The results of the thesis provide necessary conditions for stability of such systems when controlled over both noisy or noiseless channels. The results complement those obtained via information theoretic methods for such problems, and generalize well known formulas for linear systems. The techniques in this thesis build upon the earlier notion of the stabilization entropy used to study noise free systems, and rely on a stochastic volume growth approach, combined with the pointwise ergodic theorem.

There are several directions for future work. The first is to impose further conditions on the system dynamics, and use the additional structure to prove more precise channel capacity estimates. One such requirement may be to impose hyperbolicity (see [22] for a hyperbolicity definition for control systems), and focus only on the unstable system dynamics. A second possibility is the generalization of the techniques in this thesis to include continuous time systems modeled by stochastic differential equations. Lastly, it would be interesting to explore achievability schemes for non-linear systems.

# Bibliography

[1] Roy L Adler, Alan G Konheim, and M Harry McAndrew. Topological entropy. *Transactions of the American Mathematical Society*, 114(2):309–319, 1965.

[2] Boris Rostislavovich Andrievsky, Aleksei Serafimovich Matveev, and Aleksandr L'vovich Fradkov. Control and estimation under information constraints: Toward a unified theory of control, computation and communications. *Automation and Remote Control*, 71(4):572–633, 2010.

[3] J Baillieul. Feedback designs for controlling device arrays with communication channel bandwidth constraints. In *ARO workshop on smart structures*, pages 16–18. University Park PA, 1999.

[4] Fritz Colonius. Metric invariance entropy and conditionally invariant measures. *Ergodic Theory and Dynamical Systems*, 38(3):921–939, 2018.

[5] Fritz Colonius and Christoph Kawan. Invariance entropy for control systems. *SIAM Journal on Control and Optimization*, 48(3):1701–1721, 2009.

[6] Fritz Colonius, Christoph Kawan, and Girish Nair. A note on topological feedback entropy and invariance entropy. *Systems & Control Letters*, 62(5):377–381, 2013.

[7] Thomas M Cover. *Elements of information theory.* John Wiley & Sons, 1999.

[8] Renwick E Curry. *Estimation and control with quantized measurements.* MIT press Cambridge, MA, 1970.

[9] Claudio De Persis. n-bit stabilization of n-dimensional nonlinear systems in feedforward form. *IEEE Transactions on Automatic Control*, 50(3):299–311, 2005.

[10] Claudio De Persis and Alberto Isidori. Stabilizability by state feedback implies stabilizability by encoded state feedback. *Systems & Control Letters*, 53(3–4):249–258, 2004.

[11] David F Delchamps. Stabilizing a linear system with quantized state feedback. *IEEE transactions on automatic control*, 35(8):916–924, 1990.

[12] Jean-Charles Delvenne. An optimal quantized feedback strategy for scalar linear systems. *IEEE Transactions on Automatic Control*, 51(2):298–303, 2006.

[13] Efim I Dinaburg. On the relations among various entropy characteristics of dynamical systems. *Mathematics of the USSR-Izvestiya*, 5(2):337, 1971.

[14] Eugene B Dynkin and Alexander A Yushkevich. *Controlled Markov processes*, volume 235. Springer, 1979.

[15] Fabio Fagnani and Sandro Zampieri. A symbolic approach to performance analysis of quantized feedback systems: The scalar case. *SIAM journal on control and optimization*, 44(3):816–866, 2005.

[16] Massimo Franceschetti and Paolo Minero. Elements of information theory for networked control systems. In *Information and Control in Networks*, pages 3–37. Springer, 2014.

[17] Robert M Gray. *Probability, random processes, and ergodic properties*. Springer, 1988.

[18] Robert M Gray. *Probability, random processes, and ergodic properties*. Springer, 2009.

[19] Joao Hespanha, Antonio Ortega, and Lavanya Vasudevan. Towards the control of linear systems with minimum bit-rate. In *Proc. of the Int. Symp. on the Mathematical Theory of Networks and Syst*, page 1. Citeseer, 2002.

[20] Anatole Katok. Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Publications Mathématiques de l'IHÉS*, 51:137–173, 1980.

[21] Christoph Kawan. Invariance entropy for deterministic control systems. *Lecture notes in mathematics*, 2089, 2013.

[22] Christoph Kawan. Control of chaos with minimal information transfer. *arXiv: 2003.06935*, 2020.

[23] Christoph Kawan and Serdar Yüksel. Invariance properties of nonlinear stochastic dynamical systems under information constraints. *IEEE Transactions on Automatic Control, to appear (arXiv: 1901.02825)*, 2020.

[24] Johannes HB Kemperman. Strong converses for a general memoryless channel with feedback. In *Proceedings 6th Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes*, pages 375–409, 1971.

[25] Keyong Li and John Baillieul. Data-rate requirements for nonlinear feedback control. *IFAC Proceedings Volumes*, 37(13):997–1002, 2004.

[26] Daniel Liberzon and João P Hespanha. Stabilization of nonlinear systems with limited information feedback. *IEEE Transactions on Automatic Control*, 50(6):910–915, 2005.

[27] Alexey S Matveev and Andrey V Savkin. *Estimation and control over communication networks*. Springer Science & Business Media, 2009.

[28] Sean P Meyn and Richard L Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.

[29] Girish N Nair and Robin J Evans. Stabilizability of stochastic linear systems with finite feedback data rates. *SIAM Journal on Control and Optimization*, 43(2):413–436, 2004.

[30] Girish N Nair, Robin J Evans, Iven MY Mareels, and William Moran. Topological feedback entropy and nonlinear stabilization. *IEEE Transactions on Automatic Control*, 49(9):1585–1597, 2004.

[31] Girish N Nair, Fabio Fagnani, Sandro Zampieri, and Robin J Evans. Feedback control under data rate constraints: An overview. *Proceedings of the IEEE*, 95(1):108–137, 2007.

[32] YunLi Ren, LianFa He, JinFeng Lü, and GuoPing Zheng. Topological r-entropy and measure-theoretic r-entropy of a continuous map. *Science China Mathematics*, 54(6):1197–1205, 2011.

[33] Andrey V Savkin. Analysis and synthesis of networked control systems: Topological entropy, observability, robustness and optimal control. *Automatica*, 42(1):51–62, 2006.

[34] Sekhar Tatikonda and Sanjoy Mitter. Control under communication constraints. *IEEE Transactions on Automatic Control*, 49(7):1056–1068, 2004.

[35] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[36] Wing Shing Wong and Roger W Brockett. Systems with finite communication bandwidth constraints. ii. stabilization with limited information feedback. *IEEE Transactions on Automatic Control*, 44(5):1049–1053, 1999.

[37] Serdar Yüksel. Stochastic stabilization of noisy linear systems with fixed-rate limited feedback. *IEEE Transactions on Automatic Control*, 55:2847–2853, December 2010.

[38] Serdar 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.

[39] Serdar Yüksel. Stationary and ergodic properties of stochastic nonlinear systems controlled over communication channels. *SIAM Journal on Control and Optimization*, 54(5):2844–2871, 2016.

[40] Serdar Yüksel and Tamer Başar. *Stochastic networked control systems: Stabilization and optimization under information constraints*. Springer Science & Business Media, 2013.

[41] Serdar Yüksel and Sean P Meyn. Random-time, state-dependent stochastic drift for Markov chains and application to stochastic stabilization over erasure channels. *IEEE Transactions on Automatic Control*, 58:47–59, January 2013.

# Appendix A

# Optimal Transport and the Strong Converse from Information Theory

In this section, we state a few results required in the thesis.

## A.1 A result from optimal transport

In the proof of Theorem 3.1.2, a basic result from optimal transport is used, which we state here.

**Definition A.1.1.** *Let $\mu$ and $\nu$ be Borel probability measures on a metric space $(S, d)$. A coupling of $\mu$ and $\nu$ is a pair of random variables $X, Y$ defined on some probability space $(\Omega, \mathcal{F}, P)$ such that the law of the random variable $(X, Y)$ on $S^2$ admits $\mu$ and $\nu$ as its marginals.*

The notion of coupling can easily be generalized for the case where the measures $\mu$ and $\nu$ are on distinct spaces, however we do not require that level of generality. The total variation distance between probability measures on the same measurable space serves as a measure for how distinct they are. The definition reads as follows.

**Definition A.1.2.** *Let* $\mu$ *and* $\nu$ *be probability measures on a measurable space* $(\Omega, \mathcal{F})$*. We define the total variation distance as*

$$\|\mu - \nu\|_{TV} := 2 \sup_{A \in \mathcal{F}} |\mu(A) - \nu(A)|.$$

**Lemma A.1.3.** *Let* $(X, Y) : (\Omega, \mathcal{F}, P) \to S^2$ *be a coupling of the probability measures* $\mu$ *and* $\nu$ *on the metric space* $(S, d)$*. Then*

$$\|\mu - \nu\|_{\mathrm{TV}} \leq 2 \cdot P(\{\omega \in \Omega : X(\omega) \neq Y(\omega)\}).$$

If in addition, $S$ is a finite set, then a coupling $(X, Y)$ exists which achieves the above bound.

*Proof.* See Equation (6.11) in [35]. □

Note also that if $S$ is finite in the above setup, then a simple calculation results in

$$\|\mu - \nu\|_{\mathrm{TV}} = \sum_{x \in S} |\nu(x) - \mu(x)|.$$

Indeed, for finite $S$ let $A \coloneqq \{x \in S : \mu(x) \geq \nu(x)\}$. The result follows by noting that $\|\mu - \nu\|_{\mathrm{TV}} = |\mu(A) - \nu(A)| + |\mu(A^c) - \nu(A^c)|$. As such, a coupling $(X, Y)$ of the laws exists which satisfies

$$P(X \neq Y) = \frac{1}{2} \sum_{x \in S} |\nu(x) - \mu(x)|.$$

We make use of this identity in case 2 of the proof for the noisy channel case.

## A.2 Channel coding theorem

When considering a system controlled over a noisy channel, we make use of the strong converse of the noisy channel coding theorem. We state the necessary definitions and theorems here without proof. A detailed overview of these concepts can be found in [7].

**Definition A.2.1.** *(Entropy) Let $X$ and $Y$ be finite alphabet random variables taking values in alphabets $\mathcal{X}$ and $\mathcal{Y}$ respectively. Let $p_X$ denote the probability mass function of $X$ and $p_Y$ the probability mass function of $Y$.*

- *The entropy (in bits) of $X$ is defined to be $H(X) := -\sum_{x\in\mathcal{X}} p_X(x)\cdot\log_2(p_X(x))$.*
- *The conditional entropy of $X$ given the event $\{Y = y\}$ (in bits) is defined to be $H(X|Y = y) := -\sum_{x\in\mathcal{X}} p_{X|Y}(x,y)\log_2(p_{X|Y}(x,y))$ where $p_{X|Y}$ is the conditional probability mass function of $X$ given $Y$.*
- *The conditional entropy of $X$ given $Y$ (in bits) is defined to be $H(X|Y) := \sum_{y\in\mathcal{Y}} H(X|Y = y)p_Y(y)$.*

**Definition A.2.2.** *Let $X$ be a random variable taking values in $\mathbb{R}^n$, admitting a density $f$. We define the differential entropy of $X$ as*

$$h(X) := \int -f(x)\log_2(f(x))dm(x)$$

*where $m$ denotes the $n$-dimensional Lebesgue measure.*

**Definition A.2.3.** *(Mutual Information) Let $X$ and $Y$ be as in the previous definition. the mutual information of $X$ and $Y$ is defined to be $I(X;Y) := H(X) -$*

$H(X|Y) = H(Y) - H(Y|X)$. *It can be interpreted as the reduction in uncertainty of* $X$ *given the random variable* $Y$ *(or vice versa due to symmetry).*

**Definition A.2.4.** *(Discrete Memoryless Channel) Consider a memoryless finite alphabet channel with input alphabet* $\mathcal{X}$*, output alphabet* $\mathcal{Y}$ *and a given transition probability measure. The capacity of the channel is defined by* $C := \sup_{p(x)} I(\mathcal{X}, \mathcal{Y})$*, where the sup is taken over all possible probability measures on the input alphabet* $\mathcal{X}$*. We call such a channel a Discrete Memoryless Channel (DMC). A DMC with feedback is as above, but with the additional property that the encoder has knowledge of the channel output. The definition of capacity is identical, and it is well-known that feedback does not increase channel capacity.*

Next, we provide the definition of a code. We provide the definitions for channels without feedback, however the feedback case is very similar, the only difference being that at a given time, the encoder can use the channel output for previous inputs in generating the next codeword symbol.

**Definition A.2.5.** *For* $M, n \in \mathbb{N}$*, an* $(M, n)$*-code consists of an encoding function* $x^n : \{1, \ldots, M\} \to \mathcal{X}^n$ *and a decoding function* $g : \mathcal{Y}^n \to \{1, \ldots, M\}$*. We define the rate of an* $(M, n)$*-code by* $R := (\log M)/n$*.*

For a code as above, we call $x^n(1), x^n(2), \ldots, x^n(M)$ the codewords. Because the channel distorts the codewords, we must consider the probability that we can decode correctly. This leads to the following definition.

**Definition A.2.6.** *The maximal error of an* $(M, n)$*-code is given by*

$$\lambda^{(n)} := \max_{i=1,\ldots,M} P(g(Y^n) \neq i | X^n = x^n(i)).$$

**Definition A.2.7.** *A rate $R$ is called achievable if there exists a sequence of $(\lceil 2^{nR} \rceil, n)$-codes with the property that $\lambda^{(n)} \to 0$ as $n \to \infty$.*

The following is the strong converse of the noisy channel coding theorem in information theory.

**Theorem A.2.8.** *Consider a DMC $(\mathcal{X}, p(\cdot|\cdot), \mathcal{Y})$ of capacity $C$. Let $R > C$ and consider an arbitrary sequence of $(\lceil 2^{nR} \rceil, n)$-codes, used to transmit the uniform random variables $W_n$, uniformly distributed on the set $\{1, \ldots, 2^{nR}\}$, respectively. Then $P(W_n \neq g_n(Y^n)) \to 1$ as $n \to \infty$.*

The above theorem also holds for DMCs with feedback (see [24] for a proof). In the proof of Theorem 3.1.2, the encoding functions require that the channel has feedback, hence the need for this assumption in the theorem statement.