Optimality of Walrand-Varaiya Type Policies and Approximation Results for Zero-Delay Coding of Markov Sources

by

Richard G. Wood

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

Queen's University
Kingston, Ontario, Canada

July 2015

Abstract

Optimal zero-delay coding of a finite state Markov source through quantization is considered. Building on previous literature, the existence and structure of optimal policies are studied using a stochastic control problem formulation. In the literature, the optimality of deterministic Markov coding policies (or Walrand-Varaiya type policies [20]) for infinite horizon problems has been established [11]. This work expands on this result for systems with finite source alphabets, proving the optimality of *deterministic and stationary* Markov coding policies for the infinite horizon setup. In addition, the ϵ-optimality of finite memory quantizers is established and the dependence between the memory length and ϵ is quantified. An algorithm to find the optimal policy for the finite time horizon problem is presented. Numerical results produced using this algorithm are shown.

Acknowledgements

I would most like to thank my supervisors Tamás Linder and Serdar Yüksel for their continuous enouragement, guidance and support throughout my time at Queen's University. Their interest in my success over the span of my time working with them will not be forgotten.

I am also grateful to my parents for their love and support.

Table of Contents

Chapter 1

Introduction

Zero-delay coding is a variant of the original source coding problem introduced by Shannon in 1948 [16], and expanded on in [17]. We will first discuss the original source coding problem, block coding, and rate distortion theory, moving on to zero-delay coding and an introduction to the problem considered in this thesis.

1.1 Source Coding

In [16], Shannon studied a general communication system consisting of the following components:

1. An *information source*, which produces the message or sequence of messages that is meant to be communicated over the rest of the system.

2. A *transmitter*, responsible for encoding the information source into a signal that can be transmitted over the channel.

3. A *channel*, which handles the actual communication of the signal from the transmitter to the receiver, and which can be noisy or noiseless.

4. A *receiver*, which decodes the signal that was passed over the channel and reconstructs the information source from this recieved signal.

5. A *destination*, where the message is intended to go.

Figure 1.1 contains a block diagram of the system.

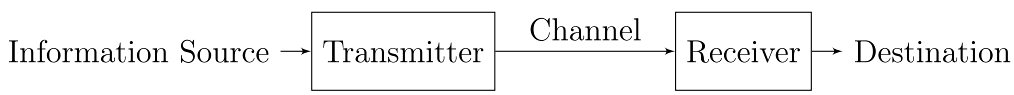

Figure 1.1: Block diagram of a general communcation system

With this communication system, three main problems are considered:

1. *Source coding* is concerned with finding methods to compress the size of the source messages to reduce the amount of data to be transmitted (through the transmission of fewer bits). Here it is assumed that the channel is noiseless.

2. *Channel coding* focuses on introducing controlled redundancy to the symbols to be transmitted across the channel to protect the message from noise in the channel.

3. *Joint source and channel coding*, which occurs when the first two problems are examined together, is the process of finding one method that both compresses the source messages and adds redundancy to protect the message being transmitted across the channel.

This thesis focuses on the source coding problem, defined formally as follows [6]. The information source $\{X_t\}_{t\geq 0}$ is assumed to be an $\mathbb{X}$-valued random process, where $\mathbb{X}$ is a discrete set. The encoder compresses the source and transmits a sequence of channel symbols $\{q_t\}_{t\geq 0}$ over the noiseless channel. Finally, the decoder reproduces the information source, where the reproduction is represented by $\{U_t\}_{t\geq 0}$, a sequence of $\mathbb{U}$-valued variables (where $|\mathbb{U}| < \infty$). The performance of the system is measured by a distortion function $d : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^+$, where $\mathbb{R}^+ := \{x \in \mathbb{R} : x \geq 0\}$ (the set of nonnegative real numbers).

When studying lossy coding systems, one is typically concerned with the transmission *rate* of an encoder and decoder, and the *distortion* of a message. We define a *rate distortion code* as follows:

Definition 1.1 (Rate Distortion Code [6]). A $(2^{RT}, T)$-*rate distortion code* encodes T source symbols $X_{[0,T-1]} := (X_0, \ldots, X_{T-1})$ at a time, and consists of an encoding function η^T and a decoding function γ^T such that

$$\eta^T : \mathbb{X}^T \to \{1, \ldots, 2^{RT}\}$$
$$\gamma^T : \{1, \ldots, 2^{RT}\} \to \mathbb{U}^T.$$

This rate distortion code has a rate of R bits per (source) symbol, as we have

$\log_2(2^{RT}) = RT$ bits per T symbols, and distortion of

$$D_T := \frac{1}{T} E\left[\sum_{t=0}^{T-1} d\left(X_t, U_t\right)\right]$$

where $(U_0, \ldots, U_{T-1}) = \gamma^T(\eta^T(X_{[0,T-1]}))$, and the expectation is taken over the distribution of $X_{[0,T-1]}$.

A common rate distortion code is a *vector quantizer*, defined below.

Definition 1.2 (Vector Quantizer)**.** A T-dimensional, M point *vector quantizer* Q^T is a mapping $Q^T : \mathbb{X}^T \to \mathsf{M}$, where $|\mathsf{M}| = M$. Q^T is determined by its quantization cells (or bins) $B_i = Q^{T^{-1}}(i) = \left\{x^T \in \mathbb{X}^T : Q^T(x^T) = i\right\}$, $i = 1, \ldots, M$, where the B_i form a partition of $\mathbb{X}^T$. An M point vector quantizer Q^T has a rate $R(Q^T) = \frac{1}{T}\log_2 M$ bits per symbol.

Definition 1.3 (Achievability of a Rate Distortion Pair [6])**.** A rate distortion pair (R, D) is said to be *achievable* if there exists a sequence of $(2^{RT}, T)$-rate distortion codes (η^T, γ^T) such that

$$\limsup_{T\to\infty} D_T \leq D.$$

With these two measurements of performance for encoders and decoders, a multitude of problems can be examined due to the inherent tradeoff between the rate and the distortion of the system, and the practical limitations that might exist in physical communication systems. For example, when considering a physical communication system, it is important to note that the cost of transmitting a bit is non-zero; there is some amount of energy required to transmit that information. Thus, if a certain level of distortion is acceptable to the system, the cheapest way to achieve that maximum distortion would be to use the encoder with the smallest rate, and vice versa, if there is a budget on the amount of data that can be transmitted, i.e. a maximum rate that can be used, one would want the to achieve the lowest distortion possible.

Definition 1.4 (Rate Distortion Function [6])**.** The *rate distortion function* $R(D)$ for a coding system is the infimum of rates R such that (R, D) is an achievable pair for a given distortion D.

Definition 1.5 (Mutual Information [6])**.** Given two discrete random variables X and Y ($\mathbb{X}$ and $\mathbb{Y}$ valued respectively) with a joint pmf $P(x, y)$, the *mutual information* $I(X;Y)$ is defined as

$$I(X;Y) \coloneqq \sum_{x\in\mathbb{X}, y\in\mathbb{Y}} P(x,y) \log \frac{P(x,y)}{P(x)P(y)}$$

For independently and identically distributed (i.i.d.) sources, the rate distortion function has an analytical expression.

Theorem 1.6 (Rate Distortion Theorem for I.I.D. Sources [6])**.** *For an i.i.d. source X with distribution $P(x)$, the rate distortion function is equal to*

$$R(D) = \min_{P_{U|X}: E[d(X,U)] \leq D} I(X;U),$$

where the minimization occurs over all conditional distributions $P_{U|X}$ such that the expected distortion between the source and the reproduction under the joint distribution $P_{U,X} = P_{U|X}P_X$ is less than or equal to D. In fact, for i.i.d. sources, for any $D > 0$

and $\delta > 0$, for T sufficiently large, then there exists a code (η^T, γ^T) with distortion

$$\frac{1}{T} E\left[\sum_{t=0}^{T-1} d\left(X_t, U_t\right)\right] < D + \delta$$

(where $(U_0, \ldots, U_{T-1}) = \gamma^T(\eta^T(X_{[0,T-1]}))$) and a rate R such that

$$R < R(D) + \delta.$$

Conversely, for any $T \geq 1$ and code (η^T, γ^T), if

$$\frac{1}{T} E\left[\sum_{t=0}^{T-1} d\left(X_t, U_t\right)\right] \leq D$$

(again where $(U_0, \ldots, U_{T-1}) = \gamma^T(\eta^T(X_{[0,T-1]}))$), then the rate of the code R satisfies

$$R \geq R(D).$$

For Markov sources, the rate distortion function is given as follows:

Theorem 1.7 (Rate Distortion Function for Markov Sources [8])**.** *For finite-state Markov sources, the rate distortion function is given by*

$$\begin{aligned} R(D) &= \lim_{T \to \infty} R_T(D) \\ R_T(D) &= \frac{1}{T} \inf_{P_{U_{[0,T-1]}|X_{[0,T-1]}} \in \mathbb{P}} I(X_{[0,T-1]}; U_{[0,T-1]}) \\ \mathbb{P} &= \left\{ P_{U_{[0,T-1]}|X_{[0,T-1]}} : E\left[\sum_{t=0}^{T-1} d(X_t, U_t)\right] \leq D \right\} \end{aligned}$$

where $P_{U_{[0,T-1]}|X_{[0,T-1]}}$ represents a conditional probability measure of the reconstructed *symbols given the source symbols, and* $I(X_{[0,T-1]}; U_{[0,T-1]})$ *is the mutual information between* $X_{[0,T-1]}$ *and* $U_{[0,T-1]}$.

In [8], Gray presented a lower bound for $R(D)$ for finite-state Markov sources as well as providing conditions on when the bound is an equality.

1.2 Zero-Delay Coding

While block coding achieves the minimum possible rate at a given distortion level, it relies on encoding blocks of data $(X_0, \ldots, X_{T-1})$, which may not be practical as the encoder has to wait until it has all T source symbols before it can encode and transmit the data. By using *zero-delay source coding*, one can eliminate this problem of waiting for data to be able to encode the source symbols. Zero-delay coding has many practical applications, including real-time control systems, audio-video systems, and sensor networks.

Definition 1.8 (Zero-Delay Source Coders)**.** A source coding system is *zero-delay* if it is "nonanticipating" and the encoding, transmission, and decoding functions occur without delay. By non-anticipating, we mean that to encode any source symbol X_t and decode the corresponding channel symbol q_t, the encoder and decoder only rely on current and past data

$$\eta_t(X_0, \ldots, X_t) = q_t,$$
$$\gamma_t(q_0, \ldots, q_t) = U_t.$$

More relaxed definitions exist, such as the definition for causal codes, see e.g. [14].

Remark 1.9. A sequence of T zero-delay encoders $\eta_{[0,T-1]}$ and decoders $\gamma_{[0,T-1]}$ can be thought of as a rate distortion code, where the encoder is of the form

$$\eta^T(X_0, \ldots, X_{T-1}) = (\eta_0(X_0), \eta_1(X_0, X_1), \ldots, \eta_{T-1}(X_0, \ldots, X_{T-1}))$$

and the decoder is similarly

$$\gamma^T(q_0, \ldots, q_{T-1}) = (\gamma_0(q_0), \gamma_1(q_0, q_1), \ldots, \gamma_{T-1}(q_0, \ldots, q_{T-1}))\,.$$

1.3 Controlled Markov Chains

Before introducing the specific problem that will be discussed in this thesis, a detour is necessary to provide some background information on *controlled Markov processes*, as the theory behind this will be important in the study of the zero-delay source coding problem presented in the thesis.

Definition 1.10 (Stochastic Kernel [9]). Let $\mathbb{W}, \mathbb{Z}$ be Borel spaces. A *stochastic kernel on* $\mathbb{W}$ *given* $\mathbb{Z}$ is a function $P(\cdot|\cdot)$ such that

1. $P(\cdot|z)$ is a probability measure on $\mathbb{W}$ for each fixed $z \in \mathbb{Z}$, and
2. $P(B|\cdot)$ is a measureable function on $\mathbb{Z}$ for each fixed Borel set $B \subset \mathbb{W}$.

In what follows, $\mathcal{P}(\mathbb{W})$ will denote the set of all probability measures on $\mathbb{W}$.

Definition 1.11 (Markov Control Model [9]). A discrete-time *Markov control model* is a five-tuple

$$(\mathbb{X}, A, \{A(x) : x \in \mathbb{X}\}, Q, c_0),$$

where

1. $\mathbb{X}$ is the *state space* of the system, or the set of all possible states of the system,

2. A is the *control space* (or action space) of the system, or the set of all admissible controls (or actions) a_t that can act on the system,

3. $\{A(x) : x \in \mathbb{X}\}$ is a subset of A consisting of all acceptable controls given the state $x \in \mathbb{X}$,

4. Q is the *transition law* of the system, a stochastic kernel on $\mathbb{X}$ given $(\mathbb{X}, A(x))$, and

5. $c_0 : \mathbb{X} \times A \to [0, \infty)$ is the *cost per time stage function* of the system, typically denoted as a function of the state and the control $c_0(x, a)$.

Thus, given a controlled Markov model and a control policy, one defines a stochastic process on $(\mathbb{X} \times A)^{\mathbb{Z}_+}$.

The $\mathbb{X}$-valued process $\{X_t\}$ is called a controlled Markov chain. Define the *admissible history* of a control model at a time $t \geq 0$ recursively with $I_0 = \mathbb{X}$ and $I_t := I_{t-1} \times A(x_{t-1}) \times \mathbb{X}$. Thus a specific history of a model $i_t \in I_t$ has the form $i_t = \{x_0, a_0, \ldots, x_{t-1}, a_{t-1}, x_t\}$.

For the purposes of this thesis, let $A(x) = A$, $\forall x \in \mathbb{X}$.

Definition 1.12 (Admissible Control Policy [9]). An *admissible control policy* $\Pi = \{\alpha_t\}_{t \geq 0}$, also called a *randomized control policy* (more simply a *control policy* or a *policy*) is a sequence of stochastic kernels on the set A given I_t. In other words, $\alpha_t : I_t \to \mathcal{P}(A)$, with α_t measureable. The set of all randomized control policies is denoted Π_A.

A randomized policy allows for both the random choice of admissible control given the current state of the system, as well as a deterministic choice of admissible control.

Definition 1.13 (Deterministic Policy [9]). A *deterministic policy* Π is a sequence of functions $\{\alpha_t\}_{t\geq 0}$, $\alpha_t : \mathbb{X}^t \to A$, that determine the control used at each time stage deterministically, i.e. $a_t = \alpha_t(x_0, \ldots, x_t)$. The set of all deterministic policies is denoted Π_D. Note that $\Pi_D \subset \Pi_A$.

Definition 1.14 (Markov Policy [9]). A *Markov policy* is a policy Π such that for each time stage the choice of control is only dependent on the current state x_t, i.e. $\Pi = \{\alpha_t\}_{t\geq 0}$ such that $\alpha_t : \mathbb{X} \to \mathcal{P}(A)$ and a_t has distribution $\alpha_t(x_t)$. The set of all Markov policies is denoted Π_M.

Definition 1.15 (Stationary Policy [9]). A *stationary policy* is a Markov policy $\Pi = \{\alpha_t\}_{t\geq 0}$ such that $\alpha_t = \alpha \ \forall t \geq 0$, where $\alpha : \mathbb{X} \to \mathcal{P}(A)$. The set of all stationary policies is denoted Π_S.

In an *optimal control problem*, a performance measure J of the system is given and the goal is to find the controls that minimize (or maximize) that measure. However this problem can be extremely complicated due to the fact that the control choice not only affects the current time stage's cost but also clearly affects the future state of the system. Also of importance is the *time horizon* T of the control problem, which is the period of time that we are concerned about minimizing (maximizing) the performance measure. The time horizon can be finite or infinite. Some common optimal control problems are as follows:

1. *Finite Horizon Average Cost Problem*: A finite horizon control problem where

the goal is to find policies that minimize the average cost

$$J_{\pi_0}(\Pi, T) := E_{\pi_0}^{\Pi}\left[\frac{1}{T}\sum_{t=0}^{T-1} c_0(X_t, a_t)\right], \tag{1.1}$$

for some $T \geq 1$.

2. *Infinite Horizon Discounted Cost Problem*: An infinite horizon control problem using discounted costs, where the goal is to find policies that minimize

$$J_{\pi_0}^{\beta}(\Pi) := \lim_{T\to\infty} E_{\pi_0}^{\Pi}\left[\sum_{t=0}^{T-1} \beta^t c_0(X_t, a_t)\right], \tag{1.2}$$

for some $\beta \in (0, 1)$.

3. *Infinite Horizon Average Cost Problem*: A more challenging infinite horizon control problem where the goal is to find policies that minimize the average cost

$$J_{\pi_0}(\Pi) := \limsup_{T\to\infty} E_{\pi_0}^{\Pi}\left[\frac{1}{T}\sum_{t=0}^{T-1} c_0(X_t, a_t)\right]. \tag{1.3}$$

In the above problems, π_0 is the distribution of the initial state variable X_0, and all of the above expectations are with respect to this distribution π_0 and under a policy Π.

Another class of systems are partially observed Markov control models, in which we cannot observe the state X_t of the system, only some (potentially noisy) observation $Y_t = g(X_t, W_t)$, where W_t is some $\mathbb{W} - \textit{valued}$, zero mean, independent and identically distributed (i.i.d.) noise source. However, these models can be transformed to a fully observed Markov control model, that is to say a Markov control model where we observe the state directly, via enlarging the state space by considering the following conditional distribution as the new state variable [23]:

$$
\begin{aligned}
\pi_{t+1}(B) &:= P(X_{t+1} \in B | y_{[0,t]}, a_{[0,t]}) \\
&= \sum_{x_{t+1} \in B} \frac{\sum_{x_t} \pi_t(x_t) P(y_t|x_t) P(a_t|\pi_{[0,t]}, a_{[0,t-1]}) Q(x_{t+1}|x_t, a_t)}{\sum_{x_t} \sum_{x_{t+1}} \pi_t(x_t) P(y_t|x_t) P(a_t|\pi_{[0,t]}, a_{[0,t-1]}) Q(x_{t+1}|x_t, a_t)} \\
&= \sum_{x_{t+1} \in B} \frac{\sum_{x_t} \pi_t(x_t) P(y_t|x_t) Q(x_{t+1}|x_t, a_t)}{\sum_{x_t} \sum_{x_{t+1}} \pi_t(x_t) P(y_t|x_t) Q(x_{t+1}|x_t, a_t)} \\
&=: F(\pi_t, a_t)(B) \qquad (1.4)
\end{aligned}
$$

A common method to solving finite horizon Markov control problems is using *dynamic programming*, which involves working backwards from the final time stage to solve for the optimal sequence of controls to use. The optimality of this algorithm is guaranteed by Bellman's principle of optimality.

Theorem 1.16 (Bellman's Principle of Optimality [9]). *Given a finite horizon T, define a sequence of functions $\{J_t(x_t)\}$ on $\mathbb{X}$ recursively such that*

$$J_T(x_T) := 0,$$

and for $0 \leq t \leq T$,

$$J_t(x_t) := \min_{a_t \in A} \left[c_0(x_t, a_t) + \sum_{x_{t+1} \in \mathbb{X}} J_{t+1}(x_{t+1}) Q(x_{t+1}|x_t, a_t) \right], \qquad (1.5)$$

then the policy $\Pi := \{a_0, \ldots, a_{T-1}\}$ is optimal with cost $J_{\pi_0}(\Pi, T) = J_0(x_0)$, where a_t is the minimizing control from Equation 1.5.

For the infinite horizon discounted cost Markov control problem, we can again use an iteration algorithm to solve for an optimal policy. This approach is commonly called the *successive approximations* method. See [9] for general conditions.

Definition 1.17 (Weak Continuity [9]). A stochastic kernel Q on $\mathbb{W}$ given $\mathbb{Z}$ is weakly continuous if the function

$$z \to \int v(w)Q(dw|z)$$

is a continuous function on $\mathbb{Z}$ whenever v is a bounded and continuous function on $\mathbb{W}$.

Definition 1.18 (Strong Continuity [9]). A stochastic kernel Q on $\mathbb{W}$ given $\mathbb{Z}$ is strongly continuous if the function

$$z \to \int v(w)Q(dw|z)$$

is a continuous function on $\mathbb{Z}$ whenever v is a bounded and measureable function on $\mathbb{W}$.

Theorem 1.19 (Iterative Solution to the Infinite Horizon Discounted Cost Problem). *For a particular $\beta \in (0,1)$, the limit of the sequence defined by*

$$J_t(x_t) = \min_{a_t \in A} \left[c_0(x_t, a_t) + \beta \sum_{x_{t-1} \in \mathbb{X}} J_{t-1}(x_{t-1}) Q(x_{t-1}|x_t, a_t) \right], \forall x_t \in \mathbb{X}$$

with $J_0(x_0) = 0$ solves the infinite horizon discounted cost problem, provided that the action set A is compact, and that one of the following hold:

1. *The one-stage cost c_0 is continuous, nonnegative, and bounded, and Q is weakly*

continuous in x_t, a_t, or

2. *The one-stage cost c_0 is continuous in a_t for every x_t, nonnegative, and bounded, and Q is strongly continuous in a_t for every x_t.*

Finally, for the infinite horizon average cost Markov control problem, we provide a brief overview of the Average Cost Optimality Equation (ACOE) below. When the ACOE holds for a policy Π, we know Π is optimal for the infinite horizon average cost problem.

Definition 1.20. The collection of functions g, h, f is a canonical triplet if for all $x \in \mathbb{X}$,

$$\begin{aligned} g(x) &= \inf_a \int g(x_{t+1}) P(dx_{t+1}|X_t = x, a_t = a) \\ g(x) + h(x) &= \inf_a \left(c(x, a) + \int h(x_{t+1}) P(dx_{t+1}|X_t = x, a_t = a) \right) \end{aligned} \tag{1.6}$$

and

$$\begin{aligned} g(x) &= \int g(x_{t+1}) P(dx_{t+1}|X_t = x, f(x)) \\ g(x) + h(x) &= c(x, f(x)) + \int h(x_{t+1}) P(dx_{t+1}|x_t = x, f(x)) \end{aligned}$$

Theorem 1.21 (Average Cost Optimality Equation [9]). *Let g, h, f be a canonical triplet. If g is a constant (in which case (1.6) is known as the Average Cost Optimality Equation or ACOE) and $\limsup_{T\to\infty} \frac{1}{T} E_{x_0}^{\Pi}[h(X_T)] = 0$ for all $x_0 \in \mathbb{X}$ and under every policy Π, then the stationary deterministic policy $\Pi^* = \{f\}$ is optimal so that*

$$g = J(x_0, \Pi^*) = \inf_{\Pi \in \Pi_A} J(x_0, \Pi)$$

where

$$J(x_0, \Pi) = \limsup_{T \to \infty} \frac{1}{T} E_{x_0}^{\Pi} \left[\sum_{t=0}^{T-1} c(X_t, a_t) \right].$$

For further details on controlled Markov processes, see [9].

Chapter 2

Problem Definition

The previous chapter provided an introduction to source coding and zero-delay coding, as well as information on controlled Markov chains. We now examine the zero-delay source coding problem studied in this thesis, and formalize the notation that will be used. We will then demonstrate how controlled Markov chains are used in the study of the problem. Finally, we will state the contributions of this thesis.

2.1 Problem Definition

This source coding problem is similar to the one discussed in [11], however this thesis only considers finite state Markov chains.

The information source $\{X_t\}_{t\geq 0}$ is assumed to be a $\mathbb{X}$-valued discrete-time Markov process, where $\mathbb{X}$ is a finite set. At each time stage, the encoder encodes the source samples and transmits the encoded versions to a receiver over a discrete noiseless channel with input and output alphabet $\mathsf{M} := \{1, 2, \ldots, M\}$, where M is a positive integer.

Formally, the encoder, or the *quantization policy*, Π is a sequence of encoder functions $\{\eta_t\}_{t\geq 0}$ with $\eta_t : \mathsf{M}^t \times (\mathbb{X})^{t+1} \to \mathsf{M}$. At a time t, the encoder transmits the M-valued message

$$q_t = \eta_t(I_t)$$

with $I_0 = X_0$, $I_t = (q_{[0,t-1]}, X_{[0,t]})$ for $t \geq 1$. The collection of all such zero-delay encoders is called the set of admissible quantization policies and is denoted by Π_A.

For fixed $q_{[0,t-1]}$ and $X_{[0,t-1]}$, as a function of X_t, the encoder

$$\eta_t(q_{[0,t-1]}, X_{[0,t-1]}, \cdot\,)$$

is a *scalar quantizer*, i.e., a mapping of $\mathbb{X}$ to the finite set M. Thus any quantization policy at each time $t \geq 0$ selects a quantizer $Q_t : \mathbb{X} \to \mathsf{M}$ based on past information $(q_{[0,t-1]}, X_{[0,t-1]})$, and then "quantizes" X_t as $q_t = Q_t(X_t)$. Upon receiving q_t, the receiver generates its reconstruction U_t, also without delay. A zero-delay receiver policy is a sequence of functions $\gamma = \{\gamma_t\}_{t\geq 0}$ of type $\gamma_t : \mathsf{M}^{t+1} \to \mathbb{U}$, where $\mathbb{U}$ denotes the finite reconstruction alphabet. Thus

$$U_t = \gamma_t(q_{[0,t]}), \qquad t \geq 0.$$

Let $\mathcal{Q}$ denote the set of all quantizers.

Theorem 2.1 (Witsenhausen [21])**.** *For the finite horizon coding problem of a Markov source, any zero-delay quantization policy $\Pi = \{\eta_t\}$ can be replaced, without any loss in performance, by a policy $\hat{\Pi} = \{\hat{\eta}_t\}$ which only uses $q_{[0,t-1]}$ and X_t to generate q_t, i.e., such that $q_t = \hat{\eta}_t(q_{[0,t-1]}, X_t)$ for all $t = 1, \ldots, T-1$.*

Let $\mathcal{P}(\mathbb{X})$ denote the space of probability measures on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ (where $\mathcal{B}(\mathbb{X})$ is the Borel σ-field over $\mathbb{X}$) endowed with the topology of weak convergence (weak topology). Given a quantization policy Π, for all $t \geq 1$ let $\pi_t \in \mathcal{P}(\mathbb{X})$ be the conditional probability defined by

$$\pi_t(A) := P(X_t \in A | q_{[0,t-1]})$$

for any set $A \subset \mathbb{X}$.

The following result is due to Walrand and Varaiya [20] who considered sources taking values from a finite set.

Theorem 2.2. *For the finite horizon coding problem of a Markov source, any zero-delay quantization policy can be replaced, without any loss in performance, by a policy which at any time* $t = 1, \ldots, T-1$ *only uses the conditional probability measure* $\pi_t = P(dx_t | q_{[0,t-1]})$ *and the state* X_t *to generate* q_t*. In other words, at time* t *such a policy uses* π_t *to select a quantizer* $Q_t : \mathbb{X} \to \mathsf{M}$ *and then* q_t *is generated as* $q_t = Q_t(X_t)$*.*

As discussed in [22], the main difference between the two structural results above is the following: in the setup of Theorem 2.1, the encoder's memory space is not fixed and keeps expanding as the encoding block length T increases. In the setup of Theorem 2.2, the memory space of an optimal encoder is fixed. More importantly, the setup of Theorem 2.2 allows one to apply the powerful theory of Markov Decision Processes on fixed state and action spaces, thus greatly facilitating the analysis.

In view of Theorem 2.2, any admissible quantization policy can be replaced by a Walrand-Varaiya type policy. We also refer to such policies as Markov policies. The class of all such policies is denoted by Π_W and is formally defined in [11] as follows.

Definition 2.3 ([11])**.** An (admissible) quantization policy $\Pi = \{\eta_t\}$ belongs to Π_W

if there exist a sequence of mappings $\{\hat{\eta}_t\}$ of the type $\hat{\eta}_t : \mathcal{P}(\mathbb{X}) \to \mathcal{Q}$ such that for $Q_t = \hat{\eta}_t(\pi_t)$ we have $q_t = Q_t(X_t) = \eta_t(I_t)$.

A policy in Π_W is called *stationary* if $\hat{\eta}_t$ does not depend on t.

Building on [22] and [11], suppose a quantizer policy $\Pi = \{\hat{\eta}_t\}$ in Π_W is used. Let $P(x_{t+1}|x_t)$ denote the transition kernel of the process $\{X_t\}$. Also note that $P(q_t|\pi_t, X_t) = 1_{\{Q_t(X_t)=q_t\}}$ with $Q_t = \hat{\eta}_t(\pi_t)$, and therefore is determined by the quantizer policy. Then, as in (1.4), standard properties of conditional probability can be used to obtain the following "filtering equation" for the evolution of π_t:

$$\begin{aligned} \pi_{t+1}(x_{t+1}) &= \frac{\sum_{x_t} \pi_t(x_t) P(q_t|\pi_t, x_t) P(x_{t+1}|x_t)}{\sum_{x_t} \sum_{x_{t+1}} \pi_t(x_t) P(q_t|\pi_t, x_t) P(x_{t+1}|x_t)} \\ &= \frac{1}{\pi_t(Q^{-1}(q_t))} \sum_{x_t \in Q^{-1}(q_t)} P(x_{t+1}|x_t)\pi_t(x_t) \end{aligned} \tag{2.1}$$

Therefore, given π_t and Q_t, π_{t+1} is conditionally independent of $(\pi_{[0,t-1]}, Q_{[0,t-1]})$. Thus $\{\pi_t\}$ can be viewed as $\mathcal{P}(\mathbb{X})$-valued controlled Markov process [9] with $\mathcal{Q}$-valued control $\{Q_t\}$ and average cost up to time $T-1$ given by

$$E^{\Pi}_{\pi_0}\left[\frac{1}{T}\sum_{t=0}^{T-1} c(\pi_t, Q_t)\right],$$

where

$$c(\pi_t, Q_t) := \sum_{i=1}^{M} \min_{u \in \mathbb{U}} \sum_{x \in Q_t^{-1}(i)} \pi_t(dx) c_0(x, u), \tag{2.2}$$

and $c_0 : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^+$, a nonnegative $\mathbb{R}$-valued distortion function [11]. In this context, Π_W corresponds to the class of deterministic Markov control policies [9].

Three specific optimization problems are considered in the thesis, where the optimal choice of encoder policy Π is the policy that minimizes some measure of cumulative distortion. For all of the problems, we assume that the encoder and decoder both know the initial distribution π_0, and $E_{\pi_0}^{\Pi,\gamma}$ denotes expectation with initial distribution π_0 for X_0, and under the quantization policy Π and receiver policy γ. The three problems are as follows:

1. *Finite Horizon Average Cost Problem*: For the finite horizon setting the goal is to minimize the average cost

$$J_{\pi_0}(\Pi, T) := \inf_{\gamma} E_{\pi_0}^{\Pi,\gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right], \tag{2.3}$$

   for some $T \geq 1$.

2. *Infinite Horizon Discounted Cost Problem*: In the infinite horizon discounted cost problem, the goal is to minimize the cumulative discounted cost

$$J_{\pi_0}^{\beta}(\Pi) := \inf_{\gamma} \lim_{T\to\infty} E_{\pi_0}^{\Pi,\gamma} \left[\sum_{t=0}^{T-1} \beta^t c(\pi_t, Q_t) \right]. \tag{2.4}$$

   for some $\beta \in (0, 1)$.

3. *Infinite Horizon Average Cost Problem*: The more challenging infinite horizon average cost problem has the objective of minimizing

$$J_{\pi_0}(\Pi) := \inf_{\gamma} \limsup_{T\to\infty} E_{\pi_0}^{\Pi,\gamma} \left[\frac{1}{T} \sum_{t=0}^{T-1} c(\pi_t, Q_t) \right]. \tag{2.5}$$

Definition 2.4 (Irreducible Markov Chain [15]). A finite state Markov chain $\{X_t\}$ is *irreducible* if it is possible to transition from any state to any other state.

Definition 2.5 (Aperiodic Markov Chain [15]). A finite state Markov chain $\{X_t\}$ is *aperiodic* if for each state i $\exists n > 0$ such that $\forall n' \geq n$,

$$P(X_{n'} = i | X_0 = i) > 0.$$

The main assumption on the Markov source $\{X_t\}$ is the following.

Assumption 2.6. $\{X_t\}$ is an irreducible and aperiodic finite state Markov chain.

2.2 Literature Review

Structural results for the finite horizon control problem described in the previous section have been developed in a number of important papers. As mentioned previously, the classic works by Witsenhausen [21] and Walrand and Varaiya [20], which use two different approaches, are of particular relevance. An extension to the more general setting of non-feedback communication was completed by Teneketzis [19], and [22] also extended these results to more general state spaces; see [11] for a detailed review.

Causal source coding was investigated in [14], establishing that for stationary memoryless sources, an optimal causal coder can either be replaced by one that is memoryless, or be replaced by a coder that is constructed by time-sharing two memoryless coders, with no loss in performance in either case. However, as noted on p. 702 of [14], the causal code restriction used is a weaker constraint than the real-time restriction. As discussed in Section 2.1, for the zero-delay problem with a finite valued, kth-order Markov source, [21] demonstrated that there exists an optimal encoder that only uses the previous k source symbols as well as the information available to the decoder. This was extended to replace the requirement of the previous channel symbols with the conditional distribution of the current source symbol given the previous channel symbols in [20]. [22] expanded these results to a partially observed setting and relaxed the condition on the source to include sources taking values in a Polish space.

[20] also discussed optimal causal coding of Markov sources with a noisy channel with feedback. Optimal causal coding of Markov sources over noisy channels without feedback was considered in [19] and [13]. Causal coding under a high-rate assumption of more general sources and stationary sources was studied in [12]. In [5], Borkar et al. studied a related problem of coding a partially observed Markov source, and obtained existence results for dynamic vector quantizers in the infinite horizon setting. It should be noted that in [5] the set of admissible quantizers was restricted to the so-called nearest neighbor quantizers, and other conditions were placed on the dynamics of the system.

In [11], Linder and Yüksel provided existence results on finite horizon and infinite horizon average-cost problems when the source is $\mathbb{R}^d$-valued. Despite the establishment of the existence of optimal quantizers for finite horizons and the existence of optimal deterministic non-stationary policies for stationary Markov sources in [11], the optimality of *stationary* and *deterministic* policies was not determined. In [11], only quantizers with convex codecells were considered (due to technical requirements), an assumption not needed in this thesis due to the finite source alphabet.

Recently [1] considered the coding of discrete independent and identically distributed (i.i.d.) sources with limited lookahead using the average cost optimality equation. [10] studied real-time joint source-channel coding of a discrete Markov source over a discrete memoryless channel with feedback. See also [4].

It is important to note that the results regarding infinite horizons do not directly follow from existing results on partially observable Markov decision processes, because in a partially observable Markov decision process, the state and the actions as well as the probability measure-valued *expanded state* and the action always constitute a *controlled Markov chain.* This is not the case in the problem considered; a quantizer and the state constructed in this thesis do not form a controlled Markov chain under an arbitrary quantizer. A detailed discussion in this aspect is present in [22].

2.3 Contributions

In view of the literature, the contributions of this work are as follows:

- Establishing the optimality (among all admissible policies) of *stationary* and *deterministic* Walrand-Varaiya type policies for zero-delay source coding for a class of Markov sources in the infinite horizon case, under both the discounted cost and average cost performance criteria.

- The establishment of ϵ-optimality of periodic zero-delay finite-length codes, and the derivation of an explicit bound on how ϵ and the length of the coding period are related.

These findings are new to our knowledge, primarily because the optimality of Witsenhausen and Walrand-Varaiya type coding schemes are based on dynamic-programming principles and thus critically depend on the finiteness of the time horizons. The findings also reveal that the increase in performance from having a large memory in the quantizers is inversely proportional with the memory of the encoder.

The thesis is structured as follows. In Chapter 3 we show that stationary Walrand-Varaiya type policies are optimal amongst all admissible policies for the finite horizon average cost problem and the infinite horizon discounted problem. In Chapter 4, we prove that stationary Walrand-Varaiya type policies are as well optimal amongst all admissible policies for the infinite horizon average cost problem, and provide results on the ϵ-optimality of finite memory policies. In Chapter 5, we present an algorithm to solve for the optimal quantization policy for the finite horizon problem (implementing the dynamic programming algorithm), and provide simulation result for the implementation. Finally in Chapter 6 we conclude the thesis and present areas for future research.

Chapter 3

The Finite Horizon Average Cost Problem and The Infinite Horizon Discounted Cost Problem

3.1 The Finite Horizon Average Cost Problem

For any quantization policy Π in Π_A and any $T \geq 1$

$$J_{\pi_0}(\Pi, T) := E_{\pi_0}^{\Pi}\left[\frac{1}{T}\sum_{t=0}^{T-1} c(\pi_t, Q_t)\right],$$

where $c(\pi_t, Q_t)$ is defined in (2.2). The following statements follow from results in [11], but also can be easily derived since in contrast to [11], there are only finitely many M-cell quantizers on $\mathbb{X}$ (as we omit the convex codecell requirement).

Theorem 3.1. *For any $T \geq 1$, there exists a policy Π in Π_W such that*

$$J_{\pi_0}(\Pi, T) = \inf_{\Pi' \in \Pi_A} J_{\pi_0}(\Pi', T). \tag{3.1}$$

Letting $J_T^T(\,\cdot\,) := 0$, $J_0^T(\pi_0) := \min_{\Pi \in \Pi_W} J_{\pi_0}(\Pi, T)$, the dynamic programming recursion

$$TJ_t^T(\pi) = \min_{Q \in \mathcal{Q}} \left(c(\pi, Q) + TE\big[J_{t+1}^T(\pi_{t+1}) | \pi_t = \pi, Q_t = Q\big] \right)$$

holds for $t = T-1, T-2, \ldots, 0$, $\pi \in \mathcal{P}(\mathbb{X})$.

Proof. By Theorem 2.2, there exists a policy Π in Π_W such that (3.1) holds. To show that the infimum is achieved, as we have a finite state space and a finite time horizon, by Theorem 1.16 we can use the dynamic programming recursion to solve for an optimal quantization policy $\Pi \in \Pi_W$. □

3.2 The Infinite Horizon Discounted Cost Problem

As discussed in Section 2.1, the goal of the infinite horizon discounted cost problem is to find policies that achieve

$$J_{\pi_0}^{\beta} := \inf_{\Pi \in \Pi_A} J_{\pi_0}^{\beta}(\Pi) \tag{3.2}$$

for some $\beta \in (0, 1)$, where

$$J_{\pi_0}^{\beta}(\Pi) = \lim_{T \to \infty} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} \beta^t c(\pi_t, Q_t) \right]. \tag{3.3}$$

As a source coding problem the discounted cost is of much less significance than the average cost. However it is important in the derivation of results for the average cost problem so it is worth examining. This result is presented formally as follows.

Theorem 3.2. *There exists an optimal deterministic quantization policy in* Π_W *among all policies in* Π_A *that solves (3.2).*

Proof. To begin the proof, observe that

$$\begin{aligned} \inf_{\Pi \in \Pi_A} \lim_{T \to \infty} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} \beta^t c(\pi_t, Q_t) \right] &\geq \limsup_{T \to \infty} \inf_{\Pi \in \Pi_A} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} \beta^t c(\pi_t, Q_t) \right] \\ &= \limsup_{T \to \infty} \inf_{\Pi \in \Pi_W} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} \beta^t c(\pi_t, Q_t) \right] \end{aligned} \tag{3.4}$$

where the equality follows from Theorem 2.2 due to the finite time horizon. Thus for each T, let Π_T denote the optimal policy in Π_W from (3.4). This sequence of policies, $\{\Pi_T\}$, can be obtained by solving the iteration algorithm

$$J_t(\pi_t) = \min_{Q_t \in \mathcal{Q}} \left[c(\pi_t, Q_t) + \beta \sum_{\pi_{t-1}} J_{t-1}(\pi_{t-1}) P(\pi_{t-1} | \pi_t, Q_t) \right]$$

with $J_0(\pi_0) = 0$. Theorem 1.19 applies, since there are finitely many quantizers for every π. So by Theorem 1.19, the sequence value functions for the policies $\{\Pi_T\}$, $\Pi_T \in \Pi_W$, i.e. $\{J_{\pi_0}(\Pi_W, T)\}$, converges to the value function of some policy $\Pi \in \Pi_W$ which is optimal amongst policies in Π_W for the infinite horizon discounted cost problem. Thus by the original set of inequalities, Π is optimal amongst all policies in Π_A. $\square$

Chapter 4

Average Cost and the ϵ-Optimality of Finite Memory Policies

4.1 Average Cost and the Average Cost Optimality Equation (ACOE)

The more challenging case is the average cost problem where one considers

$$J_{\pi_0}(\Pi) = \limsup_{T\to\infty} E_{\pi_0}^{\Pi}\left[\frac{1}{T}\sum_{t=0}^{T-1} c(\pi_t, Q_t)\right] \tag{4.1}$$

and the goal is to find an optimal policy attaining

$$J_{\pi_0} := \inf_{\Pi\in\Pi_A} J_{\pi_0}(\Pi). \tag{4.2}$$

4.2 Optimality and Approximate Optimality of policies in Π_W when the source is stationary

4.2.1 Optimality of non-stationary deterministic policies in Π_W when the source is stationary

For the infinite horizon setting the structural results in Theorems 2.1 and 2.2 are not available in the literature, as the proofs are based on dynamic programming, which starts at a finite terminal time stage and optimal policies are computed by working backwards from the end. However, as in [11], we can prove an infinite-horizon analog of Theorem 2.2 assuming that an invariant measure π^* for $\{X_t\}$ exists and $\pi_0 = \pi^*$ (which is true given Assumption 2.6).

Theorem 4.1 ([11]). *If $\{X_t\}$ starts from π^*, then there exists an optimal policy in Π_W solving the minimization problem (4.2), i.e., there exists $\Pi \in \Pi_W$ such that*

$$\limsup_{T\to\infty} E_{\pi^*}^{\Pi}\left[\frac{1}{T}\sum_{t=0}^{T-1} c(\pi_t, Q_t)\right] = J_{\pi^*}.$$

The proof of the theorem relies on a construction that pieces together policies from Π_W that on time segments of appropriately large lengths increasingly well approximate the infimum of the infinite-horizon cost achievable by policies in Π_A [11]. However for the setup considered here, we can also establish the optimality of deterministic and stationary policies. For this, we first revisit a key result which is useful in its own right given its practical implications.

4.2.2 ϵ-optimality of finite memory policies when the source is stationary

Definition 4.2 (ϵ-Optimality)**.** For the infinite horizon average cost problem, given an $\epsilon > 0$, a finite horizon policy Π over a time horizon T with average performance $J_{\pi_0}(\Pi, T)$ is *ϵ-optimal* if $J_{\pi_0}(\Pi, T) \leq J_{\pi_0} + \epsilon$, where J_{π_0} is the optimal performance for the infinite horizon average cost problem.

Define

$$J_{\pi^*}(T) := \min_{\Pi \in \Pi_A} \min_{\gamma} E_{\pi^*}^{\Pi,\gamma}\left[\frac{1}{T}\sum_{t=0}^{T-1} c(\pi_t, Q_t)\right]$$

and note that by [11], $\limsup_{T\to\infty} J_{\pi^*}(T) \leq J_{\pi^*}$. Thus there exists an *increasing* sequence of time indices $\{T_k\}$ such that for all $k = 1, 2, \ldots$

$$J_{\pi^*}(T_k) \leq J_{\pi^*} + \frac{1}{k}. \tag{4.3}$$

A key observation is that by Theorem 2.2 for all k there exists $\Pi_k = \{\hat{\eta}_t^{(k)}\} \in \Pi_W$ such that

$$J_{\pi^*}(\Pi_k, T_k) := E_{\pi^*}^{\Pi_k}\left[\frac{1}{T_k}\sum_{t=0}^{T_k-1} c(\pi_t, Q_t)\right] \leq J_{\pi^*}(T_k) + \frac{1}{k}. \tag{4.4}$$

The above shows that for every $\epsilon > 0$, there exists a finite memory encoding policy which is ϵ-optimal. This is a practically important result since most of the practical encoding schemes are finite memory schemes. In particular, for every ϵ, there exists N_ϵ such that, an encoder of the form: $\eta_t(x_t, m_t)$ with $m_t = g_t(m_{t-1}, q_{t-1})$ and $m_t \in \mathcal{N}, |\mathcal{N}| = N_\epsilon$ is ϵ-optimal. Different from [11], later in the thesis we will make this relation numerically explicit.

In the following sections, we discuss the non-stationary case.

4.3 Finite Coupling Time of Markov Processes

Before stating the final results of the thesis, we present a result that is important for the proof of the final theorems.

First note without any loss the evolution of $\{X_t\}$ can be represented by

$$X_{t+1} = F(X_t, W_t), \quad t = 0, 1, 2, \ldots \tag{4.5}$$

where $F : \mathbb{X} \times \mathbb{W} \to \mathbb{X}$ and $\{W_t\}$ where W_t is an i.i.d. $\mathbb{W}$-valued noise sequence which is independent of X_0.

Additionally, let $F_a^{-1}(b) = \{w : (a, w) \in F^{-1}(b)\}$.

Lemma 4.3. *Under Assumption 2.6, there exists a representation of the form (4.5) such that for any initial conditions $a \in \mathbb{X}$ and $b \in \mathbb{X}$, two Markov chains X'_t and X''_t, with $X'_0 = a$ and X''_t with $X''_0 = b$, driven by the same noise process W_t, we have $E[\inf\{t > 0 : X'_t = X''_t\} | X'_0 = a, X''_0 = b]$ is finite.*

Definition 4.4 (Strongly Irreducible Markov Chain)**.** A *strongly irreducible Markov chain* $\{X_t\}$ is a Markov chain that has a non-zero probability of transitioning from any state to any state in a single time stage, i.e. such that $\forall i, j \in \mathbb{X}$,

$$P(X_{t+1} = i | X_t = j) > 0.$$

In the proof of Lemma 4.3, the statement is first proven for strongly irreducible Markov chains, and then extended to general irreducible Markov chains.

A specific representation of the Markov chains is required for the subsequent analysis on the coupling time. Consider the Markov chain on the set $\{0, 1\}$ with the representation

$$X_{t+1} = F(X_t, W_t) = X_t + W_t \mod 2,$$

where W_t is a Bernoulli noise process. If we have two Markov chains X'_t, X''_t that have this representation, with $X'_0 = 0$ and $X''_0 = 1$, then clearly these two Markov chains will never couple. On the other hand, a prepresentation that allows for coupling can be constructed, as is made evident in the following proof.

Proof. Under Assumption 2.6, both Markov chains X'_t and X''_t are irreducible. First assume that both X'_t and X''_t are strongly irreducible Markov chains.

A representation that allows for the Markov chains to couple can be constructed as follows. Assume that $\mathbb{X} = \{1, \ldots, |\mathbb{X}|\}$. First, for each $x \in \mathbb{X}$ define

$$\epsilon_x := \min_{a \in \mathbb{X}} P(X_{t+1} = x | X_t = a).$$

Thus, $\forall a \in \mathbb{X}$, $P(X_{t+1} = x | X_t = a) \geq \epsilon_x$. Next, assume that the i.i.d. noise process is unformly distributed on $[0, 1]$. Now, construct the representation F as follows

$$F(X_t, W_t) = \begin{cases} 1 & \text{if } W_t \in [0, \epsilon_1) \\ 2 & \text{if } W_t \in [\epsilon_1, \epsilon_1 + \epsilon_2) \\ \vdots & \\ |\mathbb{X}| & \text{if } W_t \in [\epsilon_1 + \cdots + \epsilon_{|\mathbb{X}|-1}, \epsilon_1 + \cdots + \epsilon_{|\mathbb{X}|}) \\ G(X_t, W_t) & \text{otherwise} \end{cases}$$

where the function $G(X_t, W_t)$ dictates how the Markov chain will perform for $W_t \in$

$[\epsilon_1 + \cdots + \epsilon_{|\mathbb{X}|}, 1]$. Under this construction, these two Markov chains are guaranteed to couple if the noise sequence W_t falls within a specific range.

Now with this representation, let $\tau = \inf\{t > 0 : X'_t = X''_t\}$. As $\tau > 0$, we have

$$E[\tau | X'_0 = a, X''_0 = b] = \sum_{k=0}^{\infty} P(\tau > k | X'_0 = a, X''_0 = b) = 1 + \sum_{k=1}^{\infty} P(\tau > k | X'_0 = a, X''_0 = b).$$

Define $p = \max_{a,b \in \mathbb{X}} P(X'_i \neq X''_i | X'_{i-1} = a, X''_{i-1} = b)$, and let C be the event where $X'_0 = a$ and $X''_0 = b$.

Now, examining the conditional probability:

$$P(\tau > k | C) \tag{4.6}$$

$$= P(\cap_{i=1}^{k} (X'_i \neq X''_i) | C) \tag{4.7}$$

$$= \sum_{\substack{a_i \neq b_i \\ \forall 1 \leq i \leq k}} P(\cap_{i=1}^{k} (X'_i = a_i, X''_i = b_i) | C) \tag{4.8}$$

$$= \sum_{\substack{a_i \neq b_i \\ \forall 1 \leq i \leq k}} P(X'_k = a_k, X''_k = b_k | \cap_{i=1}^{k-1} (X'_i = a_i, X''_i = b_i), C) P(\cap_{i=1}^{k-1} (X'_i = a_i, X''_i = b_i) | C) \tag{4.9}$$

$$= \sum_{\substack{a_i \neq b_i \\ \forall 1 \leq i \leq k}} P(X'_k = a_k, X''_k = b_k | X'_{k-1} = a_{k-1}, X''_{k-1} = b_{k-1}) P(\cap_{i=1}^{k-1} (X'_i = a_i, X''_i = b_i) | C) \tag{4.10}$$

$$= \sum_{\substack{a_i \neq b_i \\ \forall 1 \leq i \leq k-1}} P(X'_k \neq X''_k | X'_{k-1} = a_{k-1}, X''_{k-1} = b_{k-1}) P(\cap_{i=1}^{k-1} (X'_i = a_i, X''_i = b_i) | C) \tag{4.11}$$

$$\leq \sum_{\substack{a_i \neq b_i \\ \forall 1 \leq i \leq k-1}} p \cdot P(\cap_{i=1}^{k-1} (X'_i = a_i, X''_i = b_i) | C) \tag{4.12}$$

$$= p \sum_{\substack{a_i \neq b_i \\ \forall 1 \leq i \leq k-1}} P(\cap_{i=1}^{k-1}(X_i' = a_i, X_i'' = b_i)|C) \tag{4.13}$$

$$\leq p^2 \sum_{\substack{a_i \neq b_i \\ \forall 1 \leq i \leq k-2}} P(\cap_{i=1}^{k-2}(X_i' = a_i, X_i'' = b_i)|C) \tag{4.14}$$

$$\vdots$$

$$\leq p^k \tag{4.15}$$

Where the ommitted steps follow the same process completed in (4.9) to (4.12). Therefore,

$$\begin{aligned} E[\tau|X_0' = a, X_0'' = b] &= 1 + \sum_{k=1}^{\infty} P(\tau > k|X_0' = a, X_0'' = b) \\ &\leq 1 + \sum_{k=1}^{\infty} p^k \\ &= \sum_{k=0}^{\infty} p^k \\ &= \frac{1}{1-p} \qquad (4.16) \\ &< \infty \end{aligned}$$

provided $p < 1$.

$$\begin{aligned} p < 1 &\iff \max_{a,b \in \mathbb{X}} P(X_i' \neq X_i''|X_{i-1}' = a, X_{i-1}'' = b) < 1 \\ &\iff 1 - \min_{a,b \in \mathbb{X}} P(X_i' = X_i''|X_{i-1}' = a, X_{i-1}'' = b) < 1 \\ &\iff \min_{a,b \in \mathbb{X}} P(X_i' = X_i''|X_{i-1}' = a, X_{i-1}'' = b) > 0 \end{aligned}$$

By the construction of F given above, it satisfies the condition that there exists $c \in \mathbb{X}$, such that $c = F(x, w)$ for all $x \in \mathbb{X}$ and $w \in B$, for some $B \subset \mathbb{W}$ with $P(B)>0$. Therefore

$$\begin{aligned} p < 1 &\Longleftrightarrow \min_{a,b\in\mathbb{X}} P(X_i' = X_i''|X_{i-1}' = a, X_{i-1}'' = b) > 0 \\ &\Longleftrightarrow \min_{a,b\in\mathbb{X}} \sum_{c\in\mathbb{X}} P(X_i' = X_i'' = c|X_{i-1}' = a, X_{i-1}'' = b) > 0 \\ &\Longleftrightarrow \min_{a,b\in\mathbb{X}} \sum_{c\in\mathbb{X}} P(F_a^{-1}(c) \cap F_b^{-1}(c)) > 0, \end{aligned} \tag{4.17}$$

which is true by the construction of F from earlier. Therefore in the case where X_t' and X_t'' are strongly irreducible, their expected coupling time is finite.

Now, assume X_t' and X_t'' are irreducible but not strongly irreducible. Then for some $n > 1, \epsilon > 0$ such that $P(X_{i+n}' = a|X_i' = b) > \epsilon$ and $P(X_{i+n}'' = a|X_i'' = b) > \epsilon$ $\forall a, b \in \mathbb{X}$. Define $Y_t' := X_{nt}'$ and $Y_t'' := X_{nt}''$, two new Markov chains that represent X_t' and X_t'' respectively sampled at every n^{th} time stage with initial conditions $Y_0' = X_0' = a$, $Y_0'' = X_0'' = b$. Thus Y_t' and Y_t'' are strongly irreducible, and therefore they couple in finite time as shown above. However, if the Y_t' and Y_t'' processes have coupled, the X_t' and X_t'' processes must have also coupled, so we have $\inf\{t > 0 : X_t' = X_t''\} \leq n(\inf\{t > 0 : Y_t' = Y_t''\})$ for all realizations of the processes. Therefore $E[\inf\{t > 0 : X_t' = X_t''\}|X_0' = a, X_0'' = b] \leq nE[\inf\{t > 0 : Y_t' = Y_t''\}|Y_0' = a, Y_0'' = b] < \infty$, and the X_t' and X_t'' processes have a finite expected coupling time. □

In the above proof, Equation (4.16) tells one that when the two processes X_t' and X_t'' are strongly irreducible, $E[\inf\{t > 0 : X_t' = X_t''\}|X_0' = a, X_0'' = b] \leq \frac{1}{1-p}$, where $p = \max_{a,b\in\mathbb{X}} P(X_i' \neq X_i''|X_{i-1}' = a, X_{i-1}'' = b)$. By replacing p with $\bar{p} =$

$1 - p = \min_{a,b \in \mathbb{X}} P(X'_i = X''_i | X'_{i-1} = a, X''_{i-1} = b)$, one can see that $E[\inf\{t > 0 : X'_t = X''_t\} | X'_0 = a, X''_0 = b] \leq \frac{1}{\bar{p}}$ is dependent on this minimum probability of transitioning from one state to another. Let K_1 be a bound on the expected coupling time $E[\inf\{t > 0 : X'_t = X''_t\} | X'_0 = a, X''_0 = b] \leq \frac{1}{\bar{p}}$, with $K_1 = \frac{1}{\bar{p}}$ for the strongly irreducible case.

When X'_t and X''_t are irreducible but not strongly irreducible, $K_1 = \frac{n}{\bar{p}_n}$, where $n = \inf\{m > 0, m \in \mathbb{Z} : P(X'_{i+m} = a | X'_i = b) > 0, P(X''_{i+m} = a | X''_i = b) > 0, \forall a, b \in \mathbb{X}\}$ and $\bar{p}_k = \min_{a,b \in \mathbb{X}} P(X'_i = X''_i | X'_{i-k} = a, X''_{i-k} = b)$. This is consistent with the strongly irreducible case, as when X'_t and X''_t are strongly irreducible, $n = 1$, and $\bar{p} = \bar{p}_1$.

Define now $\tau_{\mu_0,\zeta_0} = \min(k > 0 : X'_k = X''_k, X'_0 \sim \mu_0, X''_0 \sim \zeta_0)$.

Lemma 4.5. *For processes that satisfy Assumption 2.6, $E[\tau_{\mu_0,\zeta_0}] \leq K_1 < \infty$, where K_1 is given in Lemma 4.3.*

Proof. Let $P(x, y)$ be the joint probability measure with marginals μ_0 and ζ_0. We write

$$E[\tau_{\mu_0,\zeta_0}] = \sum_{x,y} P(x, y) E[\tau_{\delta_x,\delta_y}], \tag{4.18}$$

where δ_x denotes the probability with full mass on x. Under Assumption 2.6, for every $X'_0 = x, X''_0 = y$, the expectation of the coupling time is finite by Lemma 4.3. □

4.4 A Helper Lemma for the Subsequent Theorems

Definition 4.6 (Wasserstein Metric)**.** Let $\mathbb{X} = \{1, 2, \cdots, |\mathbb{X}|\}$ be viewed as a subset of $\mathbb{R}$. The *Wasserstein metric* of two distributions μ_0 and ζ_0 is defined as

$$W_1(\mu_0, \zeta_0) = \inf_{\mu \in \mathcal{P}(\mathbb{X}\times\mathbb{X}), \mu(x,\mathbb{X})=\mu_0(x), \mu(\mathbb{X},x)=\zeta_0(x)} \sum \mu(x,y)|x-y|.$$

Lemma 4.7. *Let $\{X'_t\}$ and $\{X''_t\}$ be two Markov chains with $X'_0 \sim \mu_0$, $X''_0 \sim \zeta_0$ (possibly dependent), and $X'_t = F(X'_{t-1}, W_{t-1})$, $X''_t = F(X''_{t-1}, W_{t-1})$ for some function F and some i.i.d. noise sequence $\{W_t\}$ with probability measure ν. Also assume both $\{X'_t\}$ and $\{X''_t\}$ satisfy Assumption 2.6. Then for all $\beta \in (0,1)$,*

$$\left| \inf_{\Pi'} E_{\mu_0}^{\Pi'}\left[\sum_{k=0}^{\infty} \beta^k c_0(X'_k, U'_k)\right] - \inf_{\Pi''} E_{\zeta_0}^{\Pi''}\left[\sum_{k=0}^{\infty} \beta^k c_0(X''_k, U''_k)\right]\right| \le 2K_2 W_1(\mu_0,\zeta_0)\|c_0\|_\infty$$

for some $K_2 < \infty$.

Proof. This follows from an argument below, which essentially builds on a related approach by V. S. Borkar [2] (see also [3] and [5]), but due to the control-free evolution of the source process and that the quantization outputs under a deterministic policy are specified once the source values are, the argument is more direct here.

The conditions will be established by, as in [5], enlarging the space of coding policies to allow for randomization at the encoder. Since for a discounted infinite horizon optimal encoding problem, optimal policies are deterministic even among possibly randomized policies, the addition of common randomness does not benefit the encoder and the decoder for optimal performance. The randomization is achieved through a parallel-path simulation argument, similar to (but different from) the one in [5].

Under the assumptions of the lemma, the simulation argument allows for the decoder outputs for X''_t to be applied as suboptimal decoder outputs for X'_t through the use of a randomized quantization procedure for X'_t. This allows for obtaining bounds on the difference between the value functions corresponding to different initial probability measures on the state process. The explicit construction of the simulation is given as follows:

1. Let $X'_{t+1} = F(X'_t, W_t)$ be a realization of the Markov chain.

2. We can write $X''_0 = G(X'_0, W'_0)$.

3. For every $X'_t = a, X'_{t+1} = b$, let $F_a^{-1}(b) = \{c : (a, c) \in F^{-1}(b)\}$. Then, generate $\hat{W}_t$ according to the simulation law $P(\hat{W}_t = \cdot | X'_t = a, X'_{t+1} = b)$ so that the following holds:

$$\nu(c) = \sum_{a,b \in \mathbb{X}} P(\hat{W}_t = c | X'_t = a, X'_{t+1} = b) P(X'_t = a) P(X'_{t+1} = b | X'_t = a) \quad (4.19)$$

   with the property that $\hat{W}_t$ and X_t are independent for any given t. An explicit construction which satisfies the relation (4.28) is to have

$$P(\hat{W}_t = c | X'_t = a, X'_{t+1} = b) = \nu(c) \frac{1_{\{c \in F_a^{-1}(b)\}}}{\nu(F_a^{-1}(b))}.$$

4. With the realized $\hat{W}_t$, generate $X''_{t+1} = F(X''_t, \hat{W}_t)$.

Note that by the construction $X'_{t+1} = F(X'_t, \hat{W}_t)$ for all t.

Lemma 4.8. *With the simulation described as above, the distribution of X_t'' is as desired, i.e., it is given by*

$$P(X_{[0,t]}'') = \zeta_0(X_0'') \prod_{i=0}^{t-1} P(X_t''|X_{t-1}'')$$

Proof. (ii) For $t = 0$, the result is correct. Note that for any i,

$$P(X_i''|X_{i-1}'') = \nu(F^{-1}_{X_{i-1}''}(X_i''))$$

$$\begin{aligned}
&P(X_1'' = b|X_0'' = a) \\
&= \sum_{d,e} P(X_1'' = b, X_1' = d, X_0' = e|X_0'' = a) \\
&= \sum_{d,e} P(W_0 \in F_a^{-1}(b), X_1' = d, X_0' = e|X_0'' = a) \\
&= \sum_{d,e} P(W_0 \in F_a^{-1}(b), W_0 \in F_e^{-1}(d)|X_0' = e, X_0'' = a \times P(X_0' = e|X_0'' = a) \\
&= \sum_{d,e} P(W_0 \in \{F_a^{-1}(b) \cap F_e^{-1}(d)\}|X_0' = e, X_0'' = a) \times P(X_0' = e|X_0'' = a) \\
&= \sum_{d,e} P(W_0 \in \{F_a^{-1}(b) \cap F_e^{-1}(d)\})P(X_0' = e|X_0'' = a) &(4.20) \\
&= \sum_{e} \left(\sum_{d} P(W_0 \in \{F_a^{-1}(b) \cap F_e^{-1}(d)\}) \right) P(X_0' = e|X_0'' = a) \\
&= \sum_{e} P(W_0 \in \{F_a^{-1}(b)\})P(X_0' = e|X_0'' = a) \\
&= \nu(W_0 \in F_a^{-1}(b)) &(4.21)
\end{aligned}$$

In (4.29) we use the fact that W_0 is independent from X_0', X_0''. It can be shown that for all t $P(X_{t+1}'' = b|X_t'' = a, X_{t-1}, \cdots, X_0) = \nu(W_t \in F_a^{-1}(b))$. □

Observe that for the process $\{X'_t\}$, we can obtain q''_t values under the optimal policy corresponding to the process with initial distribution ζ_0. The receiver applies the quantizer policy corresponding to X''_t; that is to generate U'_t we take $U'_t = \gamma''_t(q''_{[0,t]})$. This is clearly a suboptimal policy for the coding of X'_t. Thus, a *randomized encoder* and a *deterministic decoder* can be generated for the original problem. This ensures that the parallel path can be simulated, where the system is driven by the same noise process. With this construction, the applied reconstruction values will be identical.

Now, one is in a position to bound the expected costs uniformly for $\beta \in (0, 1)$. Consider the difference from the lemma

$$\left| \inf_{\Pi'} E^{\Pi'}_{\mu_0}\left[\sum_{k=0}^{\infty} \beta^k c_0(X'_k, U'_k) \right] - \inf_{\Pi''} E^{\Pi''}_{\zeta_0}\left[\sum_{k=0}^{\infty} \beta^k c_0(X''_k, U''_k) \right] \right|.$$

Suppose without loss of generality that the second term is not greater than the first one. Note that by the previous construction, the encoder for X'_t can simulate the encoder outputs for X''_t (under the initial probability measure ζ_0) with the optimal encoder Π'', and apply the decoder functions $\{\gamma''_t\}$ to compute $U''_t = \gamma''(q''_{[0,t]})$. Thus an upper bound can be obtained for the first term. Given this analysis, define now $\tau_{\mu_0,\zeta_0} = \min(k > 0 : X'_k = X''_k, X'_0 \sim \mu_0, X''_0 \sim \zeta_0)$. Thus,

$$\begin{aligned} &\left| \inf_{\Pi'} E^{\Pi'}_{\mu_0}\left[\sum_{k=0}^{\infty} \beta^k c_0(X'_k, U'_k) \right] - \inf_{\Pi''} E^{\Pi''}_{\zeta_0}\left[\sum_{k=0}^{\infty} \beta^k c_0(X''_k, U''_k) \right] \right| \\ &\quad \leq \left| E^{\Pi''}_{P(X'_0, X''_0): X'_0 \sim \mu_0, X''_0 \sim \zeta_0}\left[\sum_{k=0}^{\infty} \beta^k \big(c_0(X'_k, U''_k) - c_0(X''_k, U''_k) \big) \right] \right| \\ &\quad \leq 2E[\tau_{\mu_0,\zeta_0}] \|c_0\|_\infty \end{aligned} \tag{4.22}$$

where the last inequality follows from the coupling of the two chains.

Lemma 4.9. *For some $K_2 < \infty$, $E[\tau_{\mu_0,\bar{\mu}_0}] < K_2 W_1(\mu_0, \bar{\mu}_0)$.*

Proof. By (4.18)

$$E[\tau_{\mu_0,\bar{\mu}_0}] \leq P(x_0' \neq x_0'') \max_{x,y} E[\tau_{\delta_x,\delta_y}] \leq K_2 W_1(\mu_0, \bar{\mu}_0),$$

where the last step follows from the fact that $P(x_0' \neq x_0'') \to 0$ as $W_1(\mu_0, \bar{\mu}_0) \to 0$. □

Therefore, by the previous lemma, the statement of the theorem holds. □

4.5 Optimality of deterministic stationary quantization policies in Π_W and the Average Cost Optimality Equation (ACOE)

Theorem 4.10. *Under Assumption 2.6, for any initial distribution π_0,*

$$\begin{aligned}\inf_{\Pi \in \Pi_A} \limsup_{T \to \infty} \frac{1}{T} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right] \\ = \inf_{\Pi \in \Pi_W} \limsup_{T \to \infty} \frac{1}{T} E_{\pi^*}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right]\end{aligned}$$

where π^* is the invariant distribution.

Proof. This proof builds on an argument in *Chapter 7* of [23]. Let π_0 be the initial distribution of X_0.

Step i): Recall

$$J_{\pi_0} = \inf_{\Pi \in \Pi_A} \limsup_{T\to\infty} \frac{1}{T} E^{\Pi}_{\pi_0}\left[\sum_{t=0}^{T-1} c_0(X_t, U_t)\right]$$

Let $\{\Pi_k\}$ be a sequence of policies in Π_A such that

$$\lim_{k\to\infty} J_{\pi_0}(\Pi_k) = J_{\pi_0}.$$

Fix $n > 0$ such that

$$J_{\pi_0} \geq J_{\pi_0}(\Pi_n) - \frac{\epsilon}{3}. \tag{4.23}$$

Step ii):

Lemma 4.11 ([9])**.** *Let $\{c_t\}_{t\geq 0}$ be a sequence of nonnegative numbers. Then*

$$\begin{aligned}\liminf_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} c_t &\leq \liminf_{\beta\uparrow 1}(1-\beta)\sum_{t=0}^{\infty}\beta^t c_t \\ &\leq \limsup_{\beta\uparrow 1}(1-\beta)\sum_{t=0}^{\infty}\beta^t c_t \\ &\leq \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} c_t.\end{aligned}$$

By Lemma 4.11, $\exists\beta_\epsilon$ (close to 1) such that

$$\begin{aligned}J_{\pi_0} \geq J_{\pi_0}(\Pi_n) - \frac{\epsilon}{3} &\geq (1-\beta_\epsilon)E^{\Pi_n}_{\pi_0}\left[\sum_{t=0}^{\infty}\beta_\epsilon^t c_0(X_t, U_t)\right] - \frac{2\epsilon}{3} \\ &\geq \inf_{\Pi\in\Pi_W}(1-\beta_\epsilon)E^{\Pi}_{\pi_0}\left[\sum_{t=0}^{\infty}\beta_\epsilon^t c_0(X_t, U_t)\right] - \frac{2\epsilon}{3}\end{aligned} \tag{4.24}$$

where the final inequality follows from Theorem 3.2.

Step iii): Now we compare this to the discounted cost with the initial distribution equal to the invariant distribution π^*. For ease of interpretation, denote X'_t as the process with the condition that $X'_0 \sim \pi_0$, and X''_t as the process with the condition that $X''_0 \sim \pi^*$. Namely, we want to examine

$$\left| \inf_{\Pi \in \Pi_W} (1-\beta) E^{\Pi}_{\pi_0} \left[\sum_{t=0}^{\infty} \beta^t c_0(X'_t, U'_t) \right] - \inf_{\Pi \in \Pi_W} (1-\beta) E^{\Pi}_{\pi^*} \left[\sum_{t=0}^{\infty} \beta^t c_0(X''_t, U''_t) \right] \right| \quad (4.25)$$

Then by Lemma 4.7, $\forall \beta \in (0,1)$,

$$\left| \inf_{\Pi \in \Pi_W} (1-\beta) E^{\Pi}_{\pi_0} \left[\sum_{t=0}^{\infty} \beta^t c_0(X'_t, U'_t) \right] - \inf_{\Pi \in \Pi_W} (1-\beta) E^{\Pi}_{\pi^*} \left[\sum_{t=0}^{\infty} \beta^t c_0(X''_t, U''_t) \right] \right|$$
$$\leq 2(1-\beta) E[\tau_{\pi_0,\pi^*}] \|c_0\|_\infty \quad (4.26)$$

where $\tau_{\pi_0,\pi^*} = \min(t > 0 : X'_t = X''_t, X'_0 \sim \pi_0, X''_0 \sim \pi^*)$, and where the final inequality follows from the coupling of the two Markov chains. The above sequence of inequalities relies on three conditions to hold.

1. $\forall \Pi_n \in \Pi_W$, there exists an invariant measure ν on $\mathcal{P}(\mathbb{X}) \times \mathcal{Q}$,

2. For any stationary policy, the marginal distribution on $\mathbb{X}$ is π^*, the invariant distribution of X, and

3. The coupling argument above applies uniformly for any stationary policy.

To show 1. holds, note that if a Markov process is weak Feller on a compact space, there exists an invariant probability measure for $\{\pi_t\}$ [24]. For this case, the stochastic kernel $P(\pi_{t+1}|\pi_t, Q_t)$ is weak Feller by Lemma 11 in [11], thus 1. holds. For 2., under Assumption 2.6 an invariant distribution for X exists, and by Equation 52 in [11] it is equal to the invariant distribution of X induced by ν. Finally, 3. holds as the bound in (4.26) is independent of the policy choice.

Therefore by Lemma 4.5, as $\beta \to 1$, (4.25) goes to 0.

Thus, for $\frac{\epsilon}{3}$, $\exists \bar{\beta}$ such that $\forall \beta \geq \bar{\beta}$, (4.25) $< \frac{\epsilon}{3}$. If $\beta_\epsilon < \bar{\beta}$, then set $\beta_\epsilon = \bar{\beta}$, as moving β_ϵ closer to 1 still maintains all the existing bounds.

$$
\begin{aligned}
J_{\pi_0} &\geq \inf_{\Pi \in \Pi_W} (1-\beta_\epsilon) E^{\Pi}_{\pi_0} \left[\sum_{t=0}^{\infty} \beta_\epsilon^t c_0(X_t, U_t) \right] - \frac{2\epsilon}{3} \\
&\geq \inf_{\Pi \in \Pi_W} (1-\beta_\epsilon) E^{\Pi}_{\pi^*} \left[\sum_{t=0}^{\infty} \beta_\epsilon^t c_0(X_t, U_t) \right] - \epsilon \\
&= (1-\beta_\epsilon) E^{\Pi'}_{\pi^*} \left[\sum_{t=0}^{\infty} \beta_\epsilon^t c_0(X_t, U_t) \right] - \epsilon
\end{aligned}
$$

for some $\Pi' \in \Pi_W$ (that depends on the β_ϵ), since by Theorem 3.2, there is a policy that achieves the infimum.

Thus,

$$
\begin{aligned}
J_{\pi_0} &\geq (1-\beta_\epsilon) E^{\Pi'}_{\pi^*} \left[\sum_{t=0}^{\infty} \beta_\epsilon^t c_0(X_t, U_t) \right] - \epsilon \\
&\geq \liminf_{T \to \infty} \frac{1}{T} E^{\Pi'}_{\pi^*} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right] - \epsilon \\
&= \limsup_{T \to \infty} \frac{1}{T} E^{\Pi'}_{\pi^*} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right] - \epsilon
\end{aligned}
$$

where the final inequality follows from Lemma 4.11, and the equality follows from the fact that the system begins from the invariant measure π^*. Thus, by Theorem 4.1,

$$
J_{\pi_0} \geq \limsup_{T \to \infty} \frac{1}{T} E^{\Pi'}_{\pi^*} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right] - \epsilon
$$

$$\geq \inf_{\Pi \in \Pi_W} \limsup_{T \to \infty} \frac{1}{T} E_{\pi^*}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right] - \epsilon$$

Step iv): Thus, $\forall \epsilon > 0$, there exists an n such that (4.23) holds for $\frac{\epsilon}{3}$. Then, find a $\bar{\beta}$ such that (4.25) is $< \frac{\epsilon}{3}$. Finally, find a $\beta_\epsilon > \bar{\beta}$ such that (4.24) holds for $\frac{\epsilon}{3}$. Thus, letting $\epsilon \to 0$, there is a Walrand-Varaiya policy Π whose performance equals J_{π_0}, provided that the three conditions mentioned above hold. □

The following is one of the main results of this thesis.

Theorem 4.12. *Under Assumption 2.6, for any initial distribution* π_0*,*

$$\inf_{\Pi \in \Pi_W} \limsup_{T \to \infty} \frac{1}{T} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right]$$
$$= \inf_{\Pi \in \Pi_W} \limsup_{T \to \infty} \frac{1}{T} E_{\pi^*}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right],$$

where π^* *is the invariant distribution.*

Proof. To begin, note that by Theorem 4.10,

$$\inf_{\Pi \in \Pi_W} \limsup_{T \to \infty} \frac{1}{T} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right]$$
$$\geq \inf_{\Pi \in \Pi_W} \limsup_{T \to \infty} \frac{1}{T} E_{\pi^*}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right].$$

Now, examining the difference,

$$\left| \inf_{\Pi \in \Pi_W} \limsup_{T \to \infty} \frac{1}{T} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right] - \inf_{\Pi \in \Pi_W} \limsup_{T \to \infty} \frac{1}{T} E_{\pi^*}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right] \right|$$

by the simulation argument in Lemma 4.7, we can use the policy for the second term as a randomized policy in the first term, obtaining the same quantizer outputs for both terms. Denote this policy Π. Thus,

$$\begin{aligned}
&\left| \inf_{\Pi \in \Pi_W} \limsup_{T\to\infty} \frac{1}{T} E^{\Pi}_{\pi_0}\left[\sum_{t=0}^{T-1} c_0(X_t, U_t)\right] - \inf_{\Pi \in \Pi_W} \limsup_{T\to\infty} \frac{1}{T} E^{\Pi}_{\pi^*}\left[\sum_{t=0}^{T-1} c_0(X_t, U_t)\right] \right| \\
&\qquad \leq \left| \limsup_{T\to\infty} \frac{1}{T} E^{\Pi}_{P(X_0', X_0''): X_0' \sim \pi_0, X_0'' \sim \pi^*}\left[\sum_{t=0}^{T-1} c_0(X_t', U_t'') - c_0(X_t'', U_t'')\right] \right| \\
&\qquad \leq \limsup_{T\to\infty} \frac{1}{T} 2E[\tau_{\pi_0,\pi^*}] \|c_0\|_\infty \\
&\qquad = 0
\end{aligned}$$

since $E[\tau_{\pi_0,\pi^*}] < \infty$ by Lemma 4.5. □

Theorem 4.10 and Theorem 4.12 tell us that

$$\begin{aligned}
&\inf_{\Pi \in \Pi_A} \limsup_{T\to\infty} \frac{1}{T} E^{\Pi}_{\pi_0}\left[\sum_{t=0}^{T-1} c_0(X_t, U_t)\right] \\
&\qquad = \inf_{\Pi \in \Pi_W} \limsup_{T\to\infty} \frac{1}{T} E^{\Pi}_{\pi_0}\left[\sum_{t=0}^{T-1} c_0(X_t, U_t)\right]. \qquad (4.27)
\end{aligned}$$

The next task is to show that the infimum in (4.27) is actually a minimum. This builds on the fact that without any loss we can restrict the analysis to quantization policies in Π_W and the controlled Markov chain structure can be used. In particular, the theorem will show that there exists an optimal stationary and deterministic policy which is optimal among all policies in Π_W, and, by a consequence of Theorem 4.10 and Theorem 4.12, among Π_A as well. This results relies on the ACOE, which is only applicable under a controlled Markov chain construction.

Theorem 4.13. *Under Assumption 2.6,*

$$\inf_{\Pi \in \Pi_A} \limsup_{T \to \infty} \frac{1}{T} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right]$$
$$= \min_{\Pi \in \Pi_W} \limsup_{T \to \infty} \frac{1}{T} E_{\pi_0}^{\Pi} \left[\sum_{t=0}^{T-1} c_0(X_t, U_t) \right]$$

Furthermore, the minimizing policy is stationary.

Proof. To show that the infimum is achieved by an optimal Π, and this Π is stationary, consider the *Average Cost Optimality Equation* (or ACOE) as discussed in Section 1.3. Sufficient conditions for the ACOE to hold have been reported in [9] (see Assumptions 4.2.1 and 5.5.1 in [9]).

Assumptions 7.1.1 and 7.1.2 in [23] are more applicable to this context, since the state and the action sets are compact, weak continuity of the transition kernel is directly established from the corresponding results in [11], and the fact that the action space is finite. Examining the weak continuity of the transition kernel, as was mentioned in the proof of the previous theorem, the transition kernel $P(\pi_{t+1}|\pi_t, Q_t)$ is weak Feller, which implies weak continuity.

Thus, the boundedness and the equi-continuity of the collection of functions $V_\beta(\mu_0) = J_{\mu_0}^{\beta} - J_{\zeta_0}^{\beta}$, $\beta \in (0, 1)$ needs to be established for some arbitrary but fixed ζ_0 (with J_μ^β as defined in Equation 3.2). This follows from an argument below, which essentially builds on a related approach by V. S. Borkar [2] (see also [3] and [5]), but due to the control-free evolution of the source process and that the quantization outputs under a deterministic policy are specified once the source values are, the argument is more direct here.

The conditions will be established by, as in [5], enlarging the space of coding policies to allow for randomization at the encoder. Since for a discounted infinite horizon optimal encoding problem, optimal policies are deterministic even among possibly randomized policies, the addition of common randomness does not benefit the encoder and the decoder for optimal performance. The randomization is achieved through a parallel-path simulation argument, similar to (but different from) the one in [5].

First, recall that any finite state space Markov chain can be expressed as $X_{t+1} = F(X_t, W_t)$ for some i.i.d. noise sequence W_t and a function F. Suppose that there is a randomization device which allows for the following *parallel paths* simulation: let $X'_0 \sim \mu_0$ and $X''_0 \sim \zeta_0$ be two possibly dependent initial state variables for the path, generated through a simulation device so that both systems are driven by *identical* noise realizations, that is $X'_t = F_t(X'_0, W_{[0,t-1]})$ and $X''_t = F_t(X''_0, W_{[0,t-1]})$ for some sequence of functions F_t. What this achieves is that it allows for the decoder outputs for X''_t to be applied as suboptimal decoder outputs for X'_t through the use of a randomized quantization procedure for X'_t. This allows for obtaining bounds on the difference between the value functions corresponding to different initial probability measures on the state process. The explicit construction of the simulation is given as follows:

1. Let $X'_{t+1} = F(X'_t, W_t)$ be a realization of the Markov chain, where W_t is an i.i.d. noise process with probability measure ν.

2. Let $X'_0 \sim \mu_0$ and $X''_0 \sim \zeta_0$. We can write $X''_0 = G(X'_0, W'_0)$.

3. For every $X'_t = a, X'_{t+1} = b$, let $F_a^{-1}(b) = \{c : (a, c) \in F^{-1}(b)\}$. Then, generate $\hat{W}_t$ according to the simulation law $P(\hat{W}_t = \cdot | X'_t = a, X'_{t+1} = b)$ so that the

following holds:

$$\nu(c) = \sum_{a,b\in\mathbb{X}} P(\hat{W}_t = c | X'_t = a, X'_{t+1} = b) P(X'_t = a) P(X'_{t+1} = b | X'_t = a) \quad (4.28)$$

with the property that $\hat{W}_t$ and X_t are independent for any given t. An explicit construction which satisfies the relation (4.28) is to have

$$P(\hat{W}_t = c | X'_t = a, X'_{t+1} = b) = \nu(c) \frac{1_{\{c\in F_a^{-1}(b)\}}}{\nu(F_a^{-1}(b))}.$$

4. With the realized $\hat{W}_t$, generate $X''_{t+1} = F(X''_t, \hat{W}_t)$.

Note that by the construction $X'_{t+1} = F(X'_t, \hat{W}_t)$ for all t.

Lemma 4.14. *With the simulation described as above, the distribution of X''_t is as desired, i.e., it is given by*

$$P(X''_{[0,t]}) = \zeta_0(X''_0) \prod_{i=0}^{t-1} P(X''_t | X''_{t-1})$$

Proof. (ii) For $t = 0$, the result is correct. Note that for any i,

$$P(X''_i | X''_{i-1}) = \nu(F^{-1}_{X''_{i-1}}(X''_i))$$

$$\begin{aligned}
&P(X''_1 = b | X''_0 = a) \\
&= \sum_{d,e} P(X''_1 = b, X'_1 = d, X'_0 = e | X''_0 = a) \\
&= \sum_{d,e} P(W_0 \in F_a^{-1}(b), X'_1 = d, X'_0 = e | X''_0 = a)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{d,e} P(W_0 \in F_a^{-1}(b), W_0 \in F_e^{-1}(d) | X_0' = e, X_0'' = a \times P(X_0' = e | X_0'' = a) \\
&= \sum_{d,e} P(W_0 \in \{F_a^{-1}(b) \cap F_e^{-1}(d)\} | X_0' = e, X_0'' = a) \times P(X_0' = e | X_0'' = a) \\
&= \sum_{d,e} P(W_0 \in \{F_a^{-1}(b) \cap F_e^{-1}(d)\}) P(X_0' = e | X_0'' = a) \qquad (4.29) \\
&= \sum_{e} \left(\sum_{d} P(W_0 \in \{F_a^{-1}(b) \cap F_e^{-1}(d)\}) \right) P(X_0' = e | X_0'' = a) \\
&= \sum_{e} P(W_0 \in \{F_a^{-1}(b)\}) P(X_0' = e | X_0'' = a) \\
&= \nu(W_0 \in F_a^{-1}(b)) \qquad (4.30)
\end{aligned}
$$

In (4.29) we use the fact that W_0 is independent from X_0', X_0''. It can be shown that for all t $P(X_{t+1}'' = b | X_t'' = a, X_{t-1}, \cdots, X_0) = \nu(W_t \in F_a^{-1}(b))$. □

Observe that for the process $\{X_t'\}$, we can obtain q_t'' values under the optimal policy corresponding to the process with initial distribution ζ_0. The receiver applies the quantizer policy corresponding to X_t''; that is to generate U_t' we take $U_t' = \gamma_t''(q''_{[0,t]})$. This is clearly a suboptimal policy for the coding of X_t'. Thus, a *randomized encoder* and a *deterministic decoder* can be generated for the original problem. This ensures that the parallel path can be simulated, where the system is driven by the same noise process. With this construction, the applied reconstruction values will be identical.

Now, one is in a position to bound the expected costs uniformly for $\beta \in (0, 1)$. Consider the difference

$$
\left| \inf_{\Pi'} E_{\mu_0}^{\Pi'} \left[\sum_{k=0}^{\infty} \beta^k c_0(X_k', U_k') \right] - \inf_{\Pi''} E_{\zeta_0}^{\Pi''} \left[\sum_{k=0}^{\infty} \beta^k c_0(X_k'', U_k'') \right] \right|.
$$

Suppose without loss of generality that the second term is not greater than the first one. Note that by the previous construction, the encoder for X'_t can simulate the encoder outputs for X''_t (under the initial probability measure ζ_0) with the optimal encoder Π'', and apply the decoder functions $\{\gamma''_t\}$ to compute $U''_t = \gamma''(q''_{[0,t]})$. Thus an upper bound can be obtained for the first term. Given this analysis, define now $\tau_{\mu_0,\zeta_0} = \min(k > 0 : X'_k = X''_k, X'_0 \sim \mu_0, X''_0 \sim \zeta_0)$. Thus,

$$
\begin{aligned}
&\left| \inf_{\Pi'} E^{\Pi'}_{\mu_0}\left[\sum_{k=0}^{\infty}\beta^k c_0(X'_k, U'_k)\right] - \inf_{\Pi''} E^{\Pi''}_{\zeta_0}\left[\sum_{k=0}^{\infty}\beta^k c_0(X''_k, U''_k)\right]\right| \\
&\quad \leq \left| E^{\Pi''}_{P(X'_0,X''_0):X'_0\sim\mu_0,X''_0\sim\zeta_0}\left[\sum_{k=0}^{\infty}\beta^k\big(c_0(X'_k, U''_k) - c_0(X''_k, U''_k)\big)\right]\right| \\
&\quad \leq 2E[\tau_{\mu_0,\zeta_0}]\|c_0\|_\infty \qquad (4.31) \\
&\quad \leq 2K_1\|c_0\|_\infty \qquad (4.32)
\end{aligned}
$$

where the second last inequality follows from the coupling of the two chains, and the last inequality from Lemma 4.5.

As a result, by (4.31), the boundedness condition required for the existence of the ACOE is obtained. However, the equi-continuity of $V_\beta(\pi_0) - V_\beta(\zeta_0)$ also needs to be established for some ζ_0 over all β.

For this, consider the Wasserstein metric to select the simulation distribution P, where without loss of generality $\mathbb{X} = \{1, 2, \cdots, |\mathbb{X}|\}$ is considered, viewed as a subset of $\mathbb{R}$:

$$W_1(\mu_0, \zeta_0) = \inf_{\mu\in\mathcal{P}(\mathbb{X}\times\mathbb{X}),\mu(x,\mathbb{X})=\mu_0(x),\mu(\mathbb{X},x)=\zeta_0(x)} \sum \mu(x, y)|x - y|.$$

Lemma 4.15. *For some $K_2 < \infty$, $E[\tau_{\mu_0,\bar{\mu}_0}] < K_2W_1(\mu_0, \bar{\mu}_0)$.*

Proof. By (4.18)

$$E[\tau_{\mu_0,\bar{\mu}_0}] \leq P(x_0' \neq x_0'') \max_{x,y} E[\tau_{\delta_x,\delta_y}] \leq K_2 W_1(\mu_0, \bar{\mu}_0),$$

where the last step follows from the fact that $P(x_0' \neq x_0'') \to 0$ as $W_1(\mu_0, \bar{\mu}_0) \to 0$. □

To complete the proof Theorem 4.13, note that the above imply that $V_\beta(\mu_0) = J_{\mu_0}^\beta - J_{\zeta_0}^\beta$ is bounded from above and below, and is equi-continuous. Thus, the ACOE holds. □

4.6 ϵ-Optimality of Finite Memory Policies

Theorem 4.16. *If $X_0 \sim \pi^*$, where π^* is the invariant probability measure, then for every ϵ, there exists a finite memory, non-stationary but periodic quantization policy with period less than $\frac{2K_1\|c_0\|_\infty}{\epsilon}$ that achieves an ϵ-optimal performance, where K_1 is defined in Lemma 4.3.*

Proof. An extension to Theorem 1.21 will be needed in the proof.

Theorem 4.17 ([23])**.** *Under the condition in Theorem 1.21,*

$$\left| \frac{1}{n} E_\pi^{\Pi^*} \sum_{t=1}^{n} [c(\pi_{t-1}, Q_{t-1})] - g \right| \leq \frac{\max(|h(\pi_0)|, |E_\pi^\Pi[h(\pi_n)]|}{n} \to 0 \tag{4.33}$$

Now, note that, $J_{\pi^*}(T) \geq J_{\pi^*}$ for all T. Furthermore,

$$\frac{1}{T} E_\pi^{\Pi^*} \left[\sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] \geq J_{\pi^*}(T),$$

for any stationary and deterministic policy Π^*. Therefore, Theorem 1.21 (see also p.80 in Chapter 5 of [9]) implies, together with (4.33), the following.

$$\begin{aligned}
(J_{\pi^*}(T) - J(\pi^*)) &\leq \left| \frac{1}{T} E_\pi^{\Pi^*} \left[\sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] - J(\pi^*) \right| \\
&\leq \left| \frac{1}{T} E_\pi^{\Pi^*} \left[\sum_{t=0}^{T-1} c(\pi_t, Q_t) \right] - J(\pi^*) \right| \\
&\leq \frac{\max(|h(\pi_0)|, |E_\pi^{\Pi}[h(\pi_T)]|}{T} \\
&\leq \frac{K}{T}, \qquad (4.34)
\end{aligned}$$

where $K = \sup_{x \in \mathbb{X}} |h(x)|$, with h being an element of the canonical triplet considered in Definition 1.6.

In the vanishing discount approach, this is the limit of a converging sequence of functions $V_\beta(\pi_0) - V_\beta(\zeta_0)$ along some subsequence $\beta_k \uparrow 1$. Through (4.31) and Lemma 4.5, it follows that

$$K = \sup_\pi |h(\pi)| = \sup_{\pi_0} |V_\beta(\pi_0) - V_\beta(\zeta_0)| \leq 2K_1 \|c\|_\infty,$$

where K_1 is from Lemma 4.3. Thus, for every ϵ, we can truncate the Markov chain periodically and encode the corresponding finite horizon section with length T to arrive at ϵ-optimal policies where $\epsilon = \frac{K}{T}$.

□

Chapter 5

A Numerical Implementation for the Finite Horizon Problem

The implementation discussed below builds on the algorithm presented in [7], but removes the convex codecell restriction.

5.1 An Implementation

Given the ϵ-optimality of periodic finite memory quantizers from Section 4.2.2, in this section we present an implementation to determine the optimal zero-delay quantization policy of a stationary Markov source over a finite horizon T. The algorithm solves the dynamic programming equation by generating all possible realizations of the Markov process $\{\pi_t\}$, $0 \leq t \leq T-1$, then backward solving to find the optimal policy $\Pi \in \Pi_W$ for the finite horizon problem.

To generate all possible realizations of $\{\pi_t\}$, we start from $\pi_0 = \pi^*$, the invariant distribution of the Markov source, as this is the best estimate of the conditional probability π_0 due to the lack of information about our source yet. We rely on equation (2.1) to generate all possible π_{t+1} realizations from a given π_t and all possible (Q_t, q_t) pairs, due to the controlled Markov process property of $\{\pi_t\}$. From π_0, we find all possible π_1, and from each π_1 we can find all π_2, continuing this process until we have found all π_{T-1}. Due to the structure of the data, a tree can be used effectively use to store the data, with each level representing the possible π_t at a given time t, the parent of a node representing the π_{t-1}, and the children of a node representing all possible π_{t+1} from this π_t.

After calculating every possible realization of the $\{\pi_t\}$ process, we can backwards solve from the final time stage to determine the optimal policy. At $T-1$, the optimal quantizer is the one that minimizes the single stage expected distortion, and can be computed using the algorithm from [18]. For $0 \leq t \leq T-2$, the optimal quantizer is chosen by solving the dynamic programming recursion in (3.2), by computing the expected future cost based on the choice of Q_t and all possible q_t values, then adding the expected one stage cost for time t.

Solving the dynamic programming recursion effectively prunes the tree of all possible future states by eliminating suboptimal branches. After this has been completed, a tree is left consisting of the optimal Q_t to use for each π_t (that is, the optimal control to use for each state of our Markov process), with the children of each node corresponding to the Q_{t+1} to use given the message q_t that was produced.

One drawback of the algorithm is its computational complexity. When finding all possible realizations of the $\{\pi_t\}$ process, there are $|\mathcal{Q}||\mathsf{M}|$ possible (Q_t, q_t) pairs required to produce all possible π_{t+1}. Alternatively, this means that at the final time stage $T-1$ there are $(|\mathcal{Q}||\mathsf{M}|)^T$ probability distributions to consider.

Especially significant to the complexity is $|\mathcal{Q}|$. First, note that for every quantizer with at least one empty codecell, there exists another quantizer with no empty codecells that will perform as well as the original quantizer. Thus, we will only consider quantizers with non-empty codecells, the set of which we denote $\mathcal{Q}_{ne}$. For the finite state space case, $|\mathcal{Q}_{ne}|$ is equivalent to the number of partitions of size $|\mathsf{M}|$ of a set of size $|\mathbb{X}|$ and mapping each subset to a unique $q \in \mathsf{M} = \{1, \ldots, |\mathsf{M}|\}$. Thus $|\mathcal{Q}_{ne}| = |\mathsf{M}|! \cdot S(|\mathbb{X}|, |\mathsf{M}|)$ where $S(n,k)$ is the Stirling number of the second kind. However, in practice the specific mapping of the subsets to each element of M has no effect on performance, as the distortion of a quantizer is independent of the specific channel symbol allocation; it only depends on the input symbol x_t and the reconstructed value u_t. Thus we only need to consider one $Q \in \mathcal{Q}_{ne}$ for each partition of $\mathbb{X}$.

$|\mathcal{Q}_{ne}|$ is still a large set, and this can be mitigated by restricting to the use of quantizers with convex, non-empty codecells (the set of which we label $\mathcal{Q}_{cc}$). Based on simulation results this can cause a degradation in performance (for the $T = 1$ time horizon case it is known that the optimal quantizers have convex codecells, but it is unknown for $T > 1$). Through a combinatorial argument it can be shown that $|\mathcal{Q}_{cc}| = \binom{|\mathbb{X}|-1}{|\mathsf{M}|-1}$.

For example, Figure 5.1 demonstrates a tree with all hypothetical states for the system with settings $T = 3$, $|\mathbb{X}| = 4$, $|\mathsf{M}| = 2$, and $|\mathcal{Q}_{cc}| = 3$. However, $|\mathcal{Q}_{ne}| = 14$, so clearly there are many more states of the system if one considers all non-empty quantizers, greatly increasing the computational complexity.

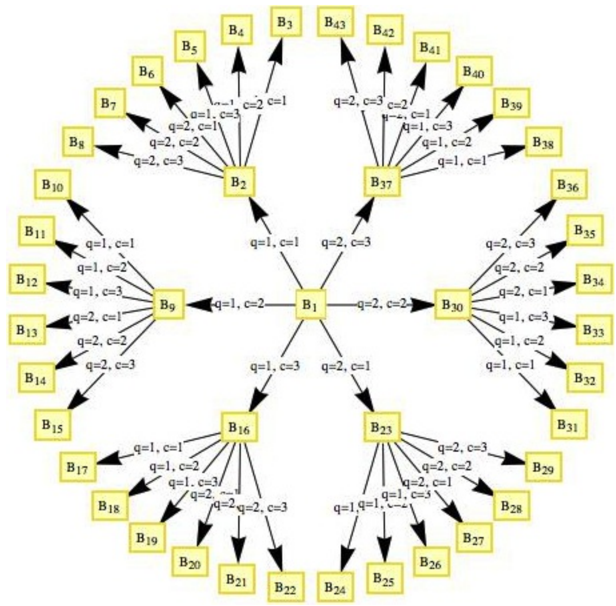

Figure 5.1: Sample tree for $T = 3$, $|\mathbb{X}| = 4, |\mathsf{M}| = 2, |\mathcal{Q}_{cc}| = 3$

5.2 Simulation Results

A simulation study was completed involving 10,000 samples being generated and encoded using the optimal policy found using the implementation above. For comparison, the optimum policy using only quantizers with convex codecells was found, and the same samples were encoded to give an idea on the change in performance due to this restriction. The mean-squared distortion measure was used to measure performance, i.e. $c_0(x, u) = (x - u)^2$, with $\mathbb{X} = \mathbb{U} = \{1, 2, \ldots, |\mathbb{X}|\} \subset \mathbb{R}$.

The system settings of the simulation are $|\mathbb{X}| = 4$ and $|\mathsf{M}| = 2$, making $|\mathcal{Q}_{ne}| = 2! \cdot S(4, 2) = 2 \cdot 7 = 14$ and $|\mathcal{Q}_{cc}| = 3$. As noted above however, we do not have to consider all the possible $Q \in \mathcal{Q}_{ne}$, just one for each partition of $\mathbb{X}$ or $S(|\mathbb{X}|, |\mathsf{M}|) = 7$

many Q's. The transition kernel for the source is given by

$$P = \begin{bmatrix} \frac{8}{10} & \frac{1}{15} & \frac{1}{15} & \frac{1}{15} \\ \frac{1}{15} & \frac{8}{10} & \frac{1}{15} & \frac{1}{15} \\ \frac{1}{15} & \frac{1}{15} & \frac{8}{10} & \frac{1}{15} \\ \frac{1}{15} & \frac{1}{15} & \frac{1}{15} & \frac{8}{10} \end{bmatrix}$$

The P matrix as defined above admits a uniform invariant distribution. In Figure 5.2, we plot the performance of the optimal zero-delay codes for various time horizon lengths (given by T), as well as the performance of the optimal zero-delay convex codecell quantizers for the same time horizons. The loss in performance (in dB) from the restriction to convex codecell quantizers is plotted in Figure 5.3.

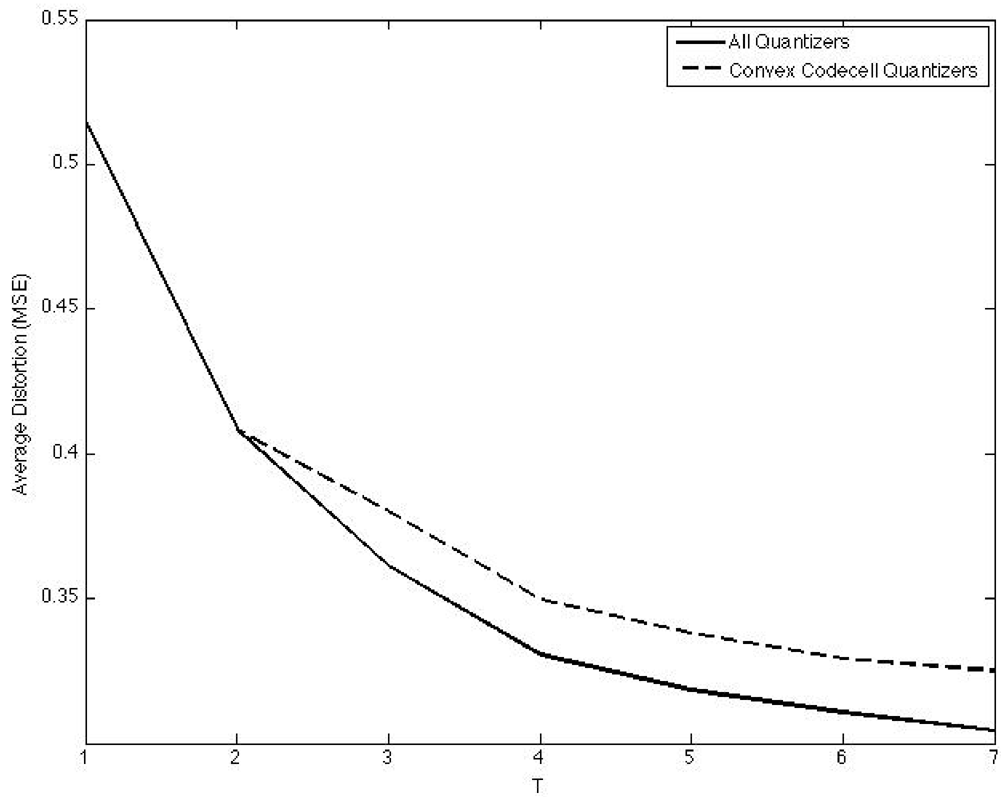

Figure 5.2: Average mean square error (MSE) for varying time horizons, $1 \leq T \leq 7$, with system settings $|\mathbb{X}| = 4$ and $|\mathsf{M}| = 2$

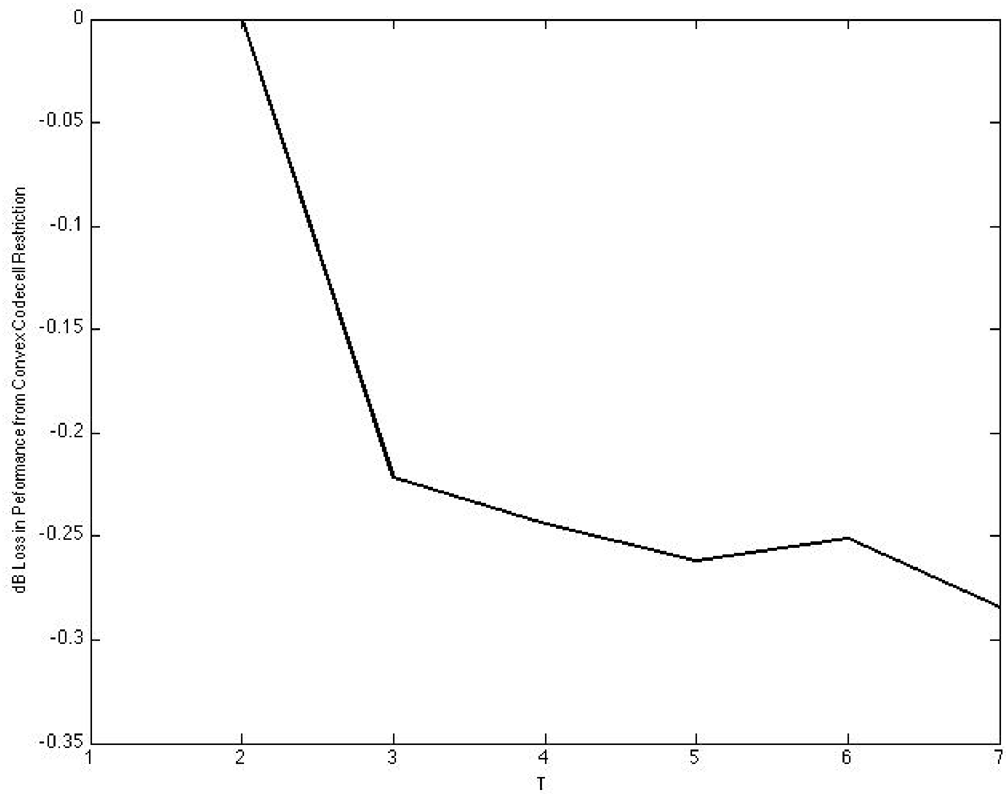

Figure 5.3: Loss in performance measured from the convex codecell quantizer restriction measured in dB for varying time horizons, $1 \leq T \leq 7$, with system settings $|\mathbb{X}| = 4$ and $|\mathsf{M}| = 2$

Chapter 6

Conclusion

Zero-delay source coding has many practical applications, justifying the importance of results centred around the problem. Due to the requirements of the system, traditional results such as rate distortion theory have limited relevance.

In this thesis, the optimality (amongst all admissible policies) of stationary and deterministic Walrand-Varaiya type policies is established for the infinite time horizon problem with both discounted and average cost measures, under the condition of a irreducible and aperiodic Markov source on a finite state space. This result builds on previous results from [21], [20], and [11].

In addition, the ϵ-optimality of periodic zero-delay finite length codes is established, with an explicit bound on the relationship between ϵ and the length of the time horizon. This result is of practical importance, as one can implement a finite horizon coding scheme, and with this result one is given a time horizon that guarantees that the performance of the finite horizon system will be within an ϵ of the performance for the optimal coder for the infinite horizon.

Further work on this problem could include extending the results to a wider class of Markov sources, namely continuous Markov sources. However, a key requirement would be a continuous version of Lemma 4.3. Additionally, the results from this thesis could be extended to cover coding problems with noisy channels with feedback.

Bibliography

[1] H. Asnani and T. Weissman. On real time coding with limited lookahead. *IEEE Transactions on Information Theory*, 59(6):3582–3606, 2013.

[2] V. S. Borkar. Average cost dynamic programming equations for controlled markov chains with partial observations. *SIAM Journal on Control and Optimization*, 39(3):673–681, 2000.

[3] V. S. Borkar. Dynamic programming for ergodic control of markov chains under partial observations: a correction. *SIAM Journal on Control and Optimization*, 45(6):2299–2304, 2007.

[4] V. S. Borkar, S. K. Mitter, A. Sahai, and S. Tatikonda. Sequential source coding: an optimization viewpoint. In *44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05.*, pages 1035–1042. IEEE, 2005.

[5] V. S. Borkar, S. K. Mitter, and S. Tatikonda. Optimal sequential vector quantization of Markov sources. *SIAM J. Control and Optimization*, 40:135–148, 2001.

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

[7] N. Daley and R. G. Wood. Real time coding with an adaptive quantizer. Undergraduate Thesis, Queen's University, April 2012.

[8] R. M. Gray. Rate distortion functions for finite-state finite-alphabet markov sources. *IEEE Transactions on Information Theory*, IT-17(2):127–134, March 1971.

[9] O. Hernandez-Lerma and J. Lasserre. *Discrete-time Markov control processes.* Springer, 1996.

[10] T. Javidi and A. Goldsmith. Dynamic joint source-channel coding with feedback. In *2013 IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 16–20. IEEE, 2013.

[11] T. Linder and S. Yüksel. On optimal zero-delay quantization of vector Markov sources. *IEEE Transactions on Information Theory*, 60:2975–5991, October 2014.

[12] T. Linder and R. Zamir. Causal coding of stationary sources and individual sequences with high resolution. *IEEE Transactions on Information Theory*, 52:662–680, February 2006.

[13] A. Mahajan and D. Teneketzis. Optimal design of sequential real-time communication systems. *IEEE Transactions on Information Theory*, 55:5317–5338, November 2009.

[14] D. L. Neuhoff and R. K. Gilbert. Causal source codes. *IEEE Transactions on Information Theory*, 28:701–713, September 1982.

[15] S. M. Ross. *Introduction to Probability Models.* Academic Press, 10th edition, 2009.

[16] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J*, XXVII(3):379–423, July 1948.

[17] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication.* University of Illinois Press, Urbana, IL, 1949.

[18] D. Sharma. Design of absolutely optimal quantizers for a wide class of distortion measures. *IEEE Transactions on Information Theory*, 24(6):693–702, 1978.

[19] D. Teneketzis. On the structure of optimal real-time encoders and decoders in noisy communication. *IEEE Transactions on Information Theory*, 52:4017–4035, September 2006.

[20] J. C. Walrand and P. Varaiya. Optimal causal coding-decoding problems. *IEEE Transactions on Information Theory*, 19:814–820, November 1983.

[21] H. S. Witsenhausen. On the structure of real-time source coders. *Bell Syst. Tech. J*, 58:1437–1451, July/August 1979.

[22] S. Yüksel. On optimal causal coding of partially observed Markov sources in single and multi-terminal settings. *IEEE Transactions on Information Theory*, 59:424–437, January 2013.

[23] S. Yüksel. Control of stochastic systems. Online course notes, May 2015.

[24] S. Yüksel and S. 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.